

Parametric and nonparametric income distribution estimators in CGE micro-simulation modeling

Dorothee Boccanfuso*, Patrick Richard*, and Luc Savard*

*Université de Sherbrooke - GRÉDI

June, 2012 (*Preliminary version*)

Abstract

We consider the issue of income distribution modeling in the context of poverty analysis based on computable general equilibrium micro-simulation models. Specifically, we study the situation where a poverty index of the FGT class is used to measure the impact of a given policy on the level of poverty in a population, which usually implies computing the chosen index in the pre and post simulation samples. The empirical distribution function (EDF) is by far the most commonly used estimator in practice. It is, however, not the only available consistent estimator and there may be situations in which a different estimator would be able to provide more accurate results.

An alternative that solves this problem is to use a smooth estimator of the population income distribution. Broadly speaking, two types of such estimators are available: parametric and nonparametric ones. In the first case, one has to choose a particular parametric form for the distribution function and estimate its parameters. The main drawback of this approach is the difficulty associated with the selection of the functional form, which must be done so as to balance finite sample bias and variance. The nonparametric approach sidesteps this functional form issue by using kernel density estimators that only impose mild restrictions on the distribution function. This is obviously an important advantage, but its cost is that the accuracy of these estimators typically depends to a large extent on the bandwidth used in the kernel function. Another advantage of the nonparametric kernel approach is that it nests the EDF as a special case.

We propose to extend the work of Boccanfuso et al. [2008] in two ways. First, we consider a larger set of parametric functions, including the 5 parameter generalized beta distribution and some of its special cases. Second, we use non-parametric kernel estimators and study their accuracy under different bandwidth selection schemes. Lastly, we provide Monte Carlo comparisons of the accuracy of these methods with the widely used EDF.

1 Introduction

Traditionally, computable general equilibrium (CGE) models have been used to simulate the impact of exogenous shocks or policy changes on the socio-economic system. In the late seventies, some modellers [Adelman and Robinson, 1978, Taylor and Lysy, 1979] attempted to use CGE models for distributional analysis, but limited the exercise to decomposing households in the CGE model into income quintiles. Later, in the early nineties, Janvry et al. [1991] analysed the impact of structural adjustment programs and were the first to use the Foster, Greer and Thorbecke (FGT) metric [Foster et al., 1984] to measure poverty changes in the Ivory Coast. Because CGE models are calibrated on the basis of a social accounting matrix (SAM) for a reference period characterized by a set of consistent initial conditions, the SAM does not contain any information on the income distribution of socioeconomic household groups. Therefore, conventional CGEs can only simulate the effect of a scenario on the representative households specified in the model. Using this representative household approach implies a very strong assumption that the distribution of income within groups of households is exogenous to the model. This method can then lead to misleading conclusions, as demonstrated in Savard [2005]. As income and expenditure surveys become more readily available, it becomes relatively easy to integrate large sample of households or entire survey samples into a SAM to calibrate the CGE model in order to take into account intra-group distributional changes. In this context, two main approaches have been used to link macro reforms to changes in income distribution and poverty.¹

The integrated multi-household (IMH) approach first proposed by Decaluwé et al. [1999], consists in including all households from the household survey in a CGE model.² However, it can raise some difficulties at the implementation and resolution stage such as data reconciliation [Hertel and Reimer., 2004], each household account needs to be balanced and aggregated to the level found in the social accounting matrix. Another problem is related to the challenges of finding a numerical solution, as raised by Chen and Ravallion [2004]. Finally, the modeller is constrained to use functional forms that respect the standard conditions of general equilibrium. The second approach is referred to as the CGE micro-simulation sequential method (MSS) and was formalized in Chen and Ravallion [2004]. The general idea of the MSS approach is that a CGE module feeds market

¹For an interesting review and discussion on the value of the CGE macro-micro approach to analyse poverty and inequality impact, the reader can consult Hertel and Reimer. [2004] and Bourguignon and Spadaro [2006].

²Some authors refer to this approach as a CGE micro-simulation application.

and factor price changes into a micro-simulation household model. The approach offers more flexibility with respect to household behaviour modelling, as standard constraints imposed by CGE models need not be respected. To summarize, the RH approach is inappropriate if household data is available—as the variance of the income distribution is held constant (no intra-group distributional impact)—but it is easier to implement; the IMH approach is theoretically sound but is more data intensive and can present challenges at the resolution stage; and the MSS approach is more flexible. In this article, we use the MSS approach applied to Mali.

Researchers' interest in income distribution began at the end of the nineteenth century. One of the objectives was to provide a mathematical description of the size of income distribution necessary to approximate the “true” distribution. It was initially believed that incomes are normally distributed but Pareto [1897] empirically demonstrated that incomes are lognormally distributed and that the skewness to the right had a flat tail, meaning unequal distribution. Since Pareto, various functional forms have been proposed that can be grouped into three main categories. The first category consists of forms describing an income distribution generated by a stochastic process [Champernowne, 1953a,b, Rutherford, 1955, Parker, 1999]. The major criticism raised against this group is that its supporters only consider the theoretical properties of income variables and neglect the empirical aspect. The second category consists of the functional forms that provide a good fit to empirical data but have no theoretical basis [Salem and Mount, 1974, McDonald, 1984]. In the last group, functional forms are found as a solution to specified differential equations such that the theoretical foundation is developed on the basis of empirical evidence [Pareto, 1897, Singh and Maddala, 1976, Dagum, 2008]. The current literature offers many alternative forms of probability density functions to approximate “true” income distribution. It is acknowledged, however, that even if lognormal and Pareto distributions are easy to estimate and interpret, other distributions such as displaced lognormal or beta, can improve the fit. The Pareto distribution appears to be appropriate only for the upper tail of income distribution where the fit over the whole range of incomes is limited. This limit seems to be a rule for all two-parameter income distributions. The lognormal income distribution proposed by Gibrat [1931] and further examined by Aitchinson and Brown [1957] seems to correctly fit at the lower income levels but is considered to adequately fit income distribution for a fairly homogeneous population. Salem and Mount [1974] provide empirical evidence that gamma distributions have a better fit than lognormal distributions. Champernowne [1953a,b] suggested a three-parameter distribution, which provides a better fit than the two-parameter one. Furthermore, according to McDonald [1984], McDonald and Xu [1995], Bordley et al. [1996], Bandourian

et al. [2003], Reed and Wu [2008] even if beta distributions are flexible and can take a variety of shapes, it is a two-parameter distribution, and its accuracy for fitting data is limited.

Over the past few years, these researchers, among others, have contributed to generalizing the beta function. This more complex model seems to reflect more appropriately the impact of economic fluctuations. Better fits may be obtained with two members of the Burr family: Singh and Maddala [1976] (Burr 12), Dagum [2008] (Burr 3). Singh–Maddala is a generalization of Pareto and Weibull distributions, and in terms of goodness-of-fit, this model outperforms both lognormal and gamma distributions in light of the US income data application done by Salem and Mount [1974]. Dagum [2008] proposed a theoretical description based on regular income-elasticity characteristics observed in income distribution. Tadikamalla [1980] and Kleiber [1996] compares the Singh–Maddala distribution with the Dagum. He demonstrates that the Dagum almost systematically exhibit a better fit. Bandourian et al. [2003] show that the Dagum is clearly the best fitting three parameter models. Comparing the ability of eleven probability distribution functions to fit income data for 23 countries over time, the Dagum outperforms three parameter distributions in 84% of the time. However, Boccanfuso et al. [2008] find that the Singh-Maddala give so good results as Dagum but conclude that there is no single “best fitting” functional form for all groups and that the most flexible ones appear to be more efficient in most cases.

McDonald [1984] develops two four-parameter generalizations of the beta distribution, called the generalized beta of the first and second kind (GB1 and GB2). As shown by McDonald and Xu [1995], every two and three parameter distributions previously used to model income distributions are special cases of either the GB1 or the GB2. Empirically, Bandourian et al. [2003] find that the GB2 distribution fits the data better than the GB1 and any three parameter distribution in 44% of 82 cases considered.

In the literature, other four parameter distributions have been used to estimate income distribution. Among these are the type II Dagum [Dagum, 2008] and double Pareto-lognormal (dPIN) proposed by Reed and Jorgensen [2004] but the authors of these papers have not demonstrated that these distribution that they were a better choice. They generally find that more flexible distributions (the ones with more parameters) exhibit better estimations. Many authors went further to propose five parameters distribution such as McDonald and Xu [1995] who proposed the generalized beta (GB) from which one can derive the GB1 and GB2 under certain assumptions. Bandourian et al. [2003], Dastrup et al. [2007],

McDonald and Ransom [2008] all illustrate that this distribution performs well without being able to systematically arrive at the conclusion that it is the best distribution. Reed [2007] as well as Reed and Wu [2008] proposed the Generalized Normal-Laplace (GNL), which is a generalization of the five parameter double Pareto-lognormal (dPIN) to estimate the income distribution of households in nine countries. They find that the GNL is more appropriate compared to other distributions including the GB of McDonald and Xu [1995]. However, Reed and Wu [2008] conclude that without the reduced form for the density function and the cumulative for this function, it is technically binding. They even question the usefulness of using five parameter functions of model income distribution. More recently, Kaniadakis [2002, 2005] developed the three parameters κ -generalized distribution, a non-Gaussian distributions with asymptotic power-law tails, more flexible to build statistical models [Clementi et al., 2010]. Clementi et al. [2010] find that in a satisfactory number of cases the performance of this distribution is not to be considered inferior from the statistical point of view to that of the Singh-Maddala, Dagum and GB2 distributions. Given that a great deal of literature has been produced on modeling income distribution over the past 60 years, we believe that it has not been fully exploited in the context of CGE micro-simulation analysis.

Janvry et al. [1991] applied Pareto distributions to model the income distribution of different subgroups in Ecuador, the best suited to represent the income distribution for higher income groups. Chia et al. [1994] and Montaud [2003] used lognormal distribution for groups in the Ivory Coast and Burkina Faso respectively. Dervis et al. [1991] also chose this two parameters distribution assuming the change in mean is equal to the change in income of the representative household generated by the CGE model and use it to calculate a new theoretical variance for this function. The lognormal distribution should be more appropriate for a high concentration of low income groups. Adelman and Robinson [1978] also performed a statistical test on lognormal that proved unsatisfactory in a few cases when testing for skewness and kurtosis. To solve this problem, they eliminated a socio-economic group by aggregation. Decaluwé et al. [2009], Stifel and Thorbecke [2003], Agénor et al. [2005], Bazlul et al. [2006] and Oum [2009] among others, applied the beta distribution arguing that the income distribution modeling approach and statistical literature provide evidence that income distribution may be represented by more appropriate and flexible functional forms [Bordley et al., 1996]. Boccanfuso et al. [2008] verify whether the results were affected by the choice of various income distribution functions used in the CGE context such as lognormal, gamma or beta. They have illustrated that using more flexible functional forms to model

income distribution is more appropriate compared to what is presently used in the CGE modeling literature. But there is no single “best fitting” functional form for all groups. Furthermore, the greater the degree of household disaggregation, the more attention should be paid to the choice of income distribution functional form and estimation/computation method if one is to provide accurate poverty analysis. An other approach used to model income distribution in this context was first proposed by Cockburn [2002] and others after him. In his CGE-IMH model applied to Nepal, he used an empirical method to compare the income distribution of different household groups. However, Boccanfuso et al. [2008] showed that when sample sizes are relatively small, this empirical approach is found not to be as sensitive as the continuous functional forms.

However, beyond the choice of the distribution through the scope of its flexibility characterized by its number of parameters to estimate, some authors reveal another problem for parametric estimation of distributions. Dastrup et al. [2007] find that the best fitting parametric distribution depends on the definition of income (earnings, total income, disposable income, expenditures,...) and may change following a policy (ex: tax increases). This means that whenever attempting to fit a parametric distribution to a sample of observed income variables, one is forced to estimate several different specifications, which may be computationally cumbersome, as it proves to be with our Malian data set. Clementi et al. [2010] expressed the same observation in their conclusion.

Furthermore, even when maximum likelihood estimation proceeds well, one must face some difficulties at the model selection step. For instance, the GB1 and GB2 distributions are not nested, which renders testing somewhat more complicated than simply using a likelihood ratio statistic. Indeed, even when the distributions are nested the null hypothesis may impose a set of constraints which lie on the border of the parametric space. For instance, a test of the GB2 against the GB imposes that $c = 1$ under H_0 while $0 \leq c \leq 1$ under H_1 . It is well known that common statistics such as the likelihood ratio do not follow their usual χ^2 distribution under this type of null hypothesis (see e.g. Gouriéroux et al. [1982]) and that the bootstrap fails to deliver valid inferences [Andrews, 2000]. Other, more complicated, inference methods such as subsampling are available, but their finite sample properties often are somewhat poor.

Considering the importance of modeling the impact of price variations on income distribution to analyze the impact of poverty changes in the context of CGE microsimulation modeling, the difficulties associated with the choice of the “good distribution” that can change according to the characteristics of the target groups

based on the effects of simulations as is shown in Boccanfuso et al. [2008] or again given a welfare indicator as in Dastrup et al. [2007], Clementi et al. [2010], and the limits of the inference method raised earlier, we propose to compare poverty analysis given different parametric distributions (Singh-Maddala, GB2 and GB) with results from empirical FGT and finally with non parametric methods in the CGE microsimulation context. Indeed to circumvent the problem of selecting and distribution, we investigate the usefulness of non parametric estimation our paper. Interestingly, this approach is commonly used in the literature, under certain conditions, it has not been experimented to our knowledge in the CGE microsimulation context.

The rest of this paper is structured such that we first describe the estimation methods (section 2). In the third section, after a brief presentation of the CGE model and the Malian data used, we describe the estimation results (parametric and non parametric) as well as poverty impact (FGT0). In the next section (4), we perform Monte Carlo simulation to identify the best estimates for poverty variations and finally we provide concluding remarks and highlight the contribution of the paper.

2 Maximum likelihood and nonparametric CDF estimation

Maximum likelihood estimation of parametric distribution using modern software is very easy but should not be regarded as an automated process. Because the likelihood function is generally too complicated to allow for its maximum to be found analytically, maximum likelihood relies on numerical search algorithms. The basic principle of these algorithms is always the same: as a first step, they must be provided with a starting value. From this starting value, they search in which direction the likelihood function seems to be higher, then decide how far to go in this direction and repeat the process from the point where they landed unless this point is judged to be a probable maximum. Some algorithms work better than others in different situations, so this choice may be of importance. Usually, algorithms of the Newton type work quite well in econometrics and are therefore used by virtually every major econometric software, see Davidson and MacKinnon [2004] (chap. 6), for an accessible discussion.

Even with an efficient search algorithm, maximum likelihood estimation requires the investigator to make several often crucial decisions, the most important of which being the choice of the values at which the search algorithm is to start.

If the algorithm is initiated too far from the global maximum, it might converge to a local one or not converge at all. A simple way to minimize the incidence of this problem is to run the search from different starting values. Another aspect which warrants vigilance is the fact that the algorithm may stop for any of several reasons: a gradient sufficiently close to 0, very similar successive parameter values, maximum number of iterations reached or numerical over or underflow. It is unlikely that a maximum has been reached if termination occurs because of any but the first reason. It is therefore important that the investigator check the reason for the termination before considering the terminal point as a maximum. Even if termination occurs as a result of a very small gradient, it is always possible that the algorithm has converged to a local maximum, so that the estimated vector of parameter does not globally maximize the likelihood function. Some methods, such as simulated annealing, ensure convergence to a global maximum. They are however much more difficult to apply than Newton-type algorithms and are not typically included in standard econometric packages.

Nonparametric kernel estimation requires less intervention by the investigator. The kernel CDF estimator of a distribution F at a point x is

$$\hat{F}_h(x) = \frac{1}{n} \sum_{t=1}^n K\left(\frac{x - x_t}{h}\right), \quad (1)$$

where K is a cumulative kernel, which consists of a cumulative density function usually symmetric around 0, and where h is the bandwidth. In essence, the only two necessary choices are those of a cumulative kernel function, which determines the shape of decrease of the weights given to observations away from the point of interest, and a bandwidth, which determines the rate of this decrease. The former is typically of little importance while the latter usually has a great influence on the accuracy of the resulting estimation. In any given case, the optimal bandwidth, that is the bandwidth which yields the most accurate estimates, depends on the characteristics of the data generating process, which is always unknown.

It is easy to see that if $h \rightarrow \infty$, then $\left(\frac{x-x_t}{h}\right) \rightarrow 0$ which implies, by the symmetry of $K(\cdot)$, that $\hat{F}_h(x) = 0.5$ for all x . Then, $\text{Var}\left(\hat{F}_h(x)\right) = 0$ but its bias is quite large. If, on the other hand, $h \rightarrow 0$, then $\left(\frac{x-x_t}{h}\right) \rightarrow +\infty$ or $-\infty$ depending on the sign of $x - x_t$. It therefore follows that $K\left(\frac{x-x_t}{h}\right) \rightarrow I(x \geq x_t)$, where $I(z)$ is the indicator function which equals 1 if the argument z is true and 0 otherwise. Thus, $\hat{F}_h(x)$ converges to the EDF, which means that it has a small bias and a very large variance. The former case is referred to as oversmoothing while the latter is called undersmoothing. Techniques to select h are all based on minimizing some

functional of $\hat{F}_h(x)$'s mean squared error, which is a combination of its bias and variance.

Several methods for the choice of h have been proposed. Rule-of-thumb selection methods are the easiest ones to use and take the general form $h = a_0 n^b$, where n is the sample size and a and b are parameters. It can be shown that, if one wishes to estimate a PDF, then the optimal value of b is $-1/5$ while if one estimates a CDF, then b should be $-1/3$. The optimal value of a depends on the true distribution of the data. In particular, a should be chosen to be function of the spread of the distribution. When this is close to Gaussian, the standard deviation is a good measure of the spread. If the distribution has thick tails however, the standard deviation tends to overestimate the spread and one might prefer to use the interquartile range (IQR), that is the difference between the 0.75^{th} and 0.25^{th} quantiles of the data. See Azzalini [1981], Wand and Jones [1995] or Q.Li and Racine [2006] for proofs of these results and Davidson and MacKinnon [2004] (chap 15) for an accessible discussion.

Another selection method for h which uses more of the information contained in the sample is least squares cross validation. This is based on the principle of selecting the value of h which minimizes the integrated squared error of $\hat{F}_h(x)$ given the sample. Bowman et al. [1998] propose using h that minimizes

$$CV_F(h) = \frac{1}{n} \sum_{t=1}^n \int \left[1(x_t \leq x) - \hat{F}_{-t}(x) \right]^2 dx, \quad (2)$$

where

$$\hat{F}_{-t}(x) = \frac{1}{n-1} \sum_{j \neq t}^n K \left(\frac{x - x_j}{h} \right).$$

Notice that $\hat{F}_{-t}(x)$ simply is the kernel CDF estimate with bandwidth h computed by excluding observation t . This exclusion is necessary because otherwise the function (2) would converge monotonically to 0 as $h \rightarrow 0$, see Q.Li and Racine [2006] (chap 1), for a discussion of this point. Bowman et al. [1998] show that the value of h that minimizes (2) converges in probability to the asymptotically optimal choice. The main drawback of cross-validation methods is the computational time necessary to minimize (2), which usually is considerable.

3 Estimation with actual data and CGE

In this section, we consider the accuracy of maximum likelihood estimation under different distributional assumptions as well as that of nonparametric kernel estimation using different bandwidth selection methods with actual income data and income data resulting from four CGE simulations. To illustrate we used the model developed by Boccanfuso and Savard [2011]. The social accounting matrix (SAM) is decomposed into 10 production sectors, seven of which are tradable and three of which are non-tradable based on 2001 data. The CGE module includes four agents, namely an aggregate household, the government, firms, and the rest of the world. We have also included savings and investment accounts. All 4,966 households of the Malian household survey “Enquête malienne d’évaluation de la pauvreté” (EMEP - 2001) have been integrated into a household micro-simulation model.

Using a Cobb-Douglas production function (for value-added), we assume producers have a cost-minimizing behaviour constrained by this production function. Value-added is a combination of capital and labour and is related to intermediate consumption with a fixed-share assumption. Capital is assumed to be fixed, which generates a branch-specific return on capital. The returns on capital combined with the wage provide for ten factor payments. These factors payments are the main source of the heterogeneous impact on household income changes. The government collects its revenues through income taxes (imposed on households and firms), goods and services taxes, import duties, and transfers from other agents (the rest of the world). It spends this revenue by paying subsidies and by producing public services. Household income is composed of wage payment, capital payments, dividends, and transfers from other agents (households and remittances from abroad). Household expenditure is derived from maximizing a Cobb-Douglas utility function under budgetary constraint. The income tax rate corresponds to the effective tax rate and not the administrative rate. The welfare indicator used at the household level is the real income. The nominal income is deflated by a household-specific consumer price index. This approach is different than the endogenous poverty line approach proposed by Decaluwé et al. [2005], as it captures a household-specific price effect of the simulation based on each household consumption structure. As for closure rules, we also assume that government saving is exogenous and that total investment is endogenous, since its level is determined by the level of savings (domestic and foreign). The total labour supply is assumed fixed and workers can move from one sector to another following a simulation. Hence, there is no endogenous unemployment in the model. The current account

balance (CAB) is exogenous, as is the nominal exchange rate, which also serves as the numeraire. The price index (GDP deflator) is endogenous and allows for clearing the CAB. We assume in a standard fashion that Mali is a small open economy with the Armington [1969] assumption for the demand of imported goods: imperfect substitution with constant elasticity of substitution function (CES) and constant elasticity of transformation function (CET) to model the export supply.

The distributional analysis is performed from an output originating from the micro-simulation household module but including 4,966 households. We transmit price changes and factor payment changes from the CGE module to the micro module in a top-down fashion. With these new prices, the module computes the new incomes and household-specific price indices. The price indices are a function of each household expenditure structure. The output of the micro module is a vector of new real income. We identified six household groups according to their geographical location (Bamako, other urban areas and rural areas) and the gender of the head of household. Most of the population lives in rural areas, with 70.92% for Mali. The Table 1 presents the size and the proportion of each group.

Table 1: Statistics of the household groups

		# Obs	% Sample	% Malian households
Mali		4966	100%	100%
Urban (with Bamako) Men	Bamako Men	454	9.14%	10.16%
	Other urban Men	1151	23.18%	14.56%
Urban (with Bamako) women	Bamako Women	52	1.05%	1.00%
	Other urban Women	183	3.69%	3.36%
Rural Men		2950	59.40%	65.77%
Rural Women		176	3.54%	5.15%

Source: EMEP, 2001

Let us recall the main objective of this paper that is on one hand to see if we can identify the best parametric distribution among the ones used in the literature given different sample size of target groups and on the other hand we aim to compare the robustness of results obtained with the non parametric approach in the context of CGE microsimulation. To achieve these objectives, we used two simulations presented in Boccanfuso and Savard [2011]. These two simulations which reproduce the impact of the food crisis, linked to the increase in international food prices. In the first simulation a 70% generalized increase food prices (all

types of agricultural goods) is performed. For the second simulation, it focuses on the increase in cereal prices excluding cotton (since Mali is an important cotton exporter).

3.1 Parametric estimation

As table 1 shows, our analysis is based on a sample of 4,966 observations and a number of subsamples representing different strata of the population. Sample sizes range from 52 for Bamako households headed by a women to 4,966 for the country as a whole. We will first take a look at the parametric estimation results. The main issue there is to choose which distribution fits the data best in order to obtain an accurate poverty rate estimate. Table 2 shows the distribution chosen by likelihood ratio tests based on standard χ^2 5% critical values for each subsample. Because of the problems mentioned above associated with these tests, it is likely that the critical values are not always valid so that these inferences may not be as accurate as they should.

Furthermore, because they are not nested, it is not always possible to make a clear choice between the Singh-Maddala and Dagum distributions, hence in some cases where their loglikelihood are almost identical, we report that both distributions appear to fit the data equally well. In some other cases involving tests of the GB2 against the GB, the LR statistic is very close to the $\chi^2(1)$ critical value. When this happens, table 2 reports the P value of this test.

Table 2: Parametric distribution selected by LR tests

	Reference	Simulation 1	Simulation 2
Mali	GB	GB	GB
Urban	GB2 ($P = 0.0540$)	GB ($P = 0.0058$)	GB ($P = 0.0060$)
Rural	GB ($P = 0.0488$)	GB2 ($P = 0.1252$)	GB ($P = 0.0460$)
Rural men	GB2 ($P = 0.0580$)	GB2 ($P = 0.1797$)	GB2 ($P = 0.0540$)
Rural women	DM or SM	DM or SM	DM or SM
Urban men	DM or SM	DM or SM	DM or SM
Urban women	GB	GB	GB
Bamako men	DM	DM	GB
Bamako women	DM or SM	DM or SM	DM or SM
Non-Bamako men	SM	SM	SM
Non-Bamako women	DM or SM	DM or SM	DM or SM

Note: P values in parenthesis

Starting with the actual data, we find that LR tests select each of our 4 distri-

butions at least once, but that there often is considerable ambiguity. While the GB seems to be appropriate for the population as a whole (the LR statistics of the GB2 against the GB is 58.73), it appears to be over-parametrized for at least six population sub-groups. Depending on how strict one wants to be about the 5% level of the test, the GB may or may not be the best choice to model the distribution among urban and rural households and among rural households with a men at their head.

This sort of ambiguity also happens between the Dagum and Singh-Maddala distributions and can sometimes have an important impact on the analysis based on these estimations. As an illustration, let us consider the population subgroup consisting of Bamako households with a female head, of which our sample contains 52 observations. Table 2 indicates that both the Singh-Maddala and Dagum distributions provide appropriate representations of income distributions. Their estimated head count poverty ratio are 13.9 and 13.5 respectively. Meanwhile, the headcount ratio estimated by the GB distribution is 15.8. Thus, assuming that the Singh-Maddala and Dagum distributions are indeed appropriate, the use of an over-parametrized distribution such as the GB may lead to substantial distortions in the poverty ratio estimates. Therefore, a modeler performing an analysis involving several population subgroups must be very careful not to impose a single functional form to every groups.

One must also be careful not to systematically use the same distribution for a given subgroup while performing an impact analysis for a policy because the policy may change the distribution. Our data provide one clear example of this with the case of Bamako households headed by a man before and after simulation 2, where the Dagum is selected over the GB and GB2 distributions with the original data, but soundly rejected in favour of the GB even at 1% with the simulated data. Table 3 shows the LR statistics for this population subgroup. Other selection criteria such as choosing the distribution with the smallest sum of squared errors run into the same kind of troubles.

3.2 Nonparametric estimation

One of the purposes of this paper is to see whether nonparametric kernel distribution estimators can provide accurate estimates of income distribution functions that can conveniently be used for policy impact analysis. This method has the advantages of not requiring that the modeler specify a parametric function for the

Table 3: LR tests statistics for Bamako households headed by a men

	Original data	Simulation 1	Simulation 2
LR (GB2 vs GB)	1.2817	3.9684**	11.4112***
LR (SM vs GB)	7.5824***	9.7431***	17.9644***
LR (SM vs GB2)	6.3007***	5.7747**	6.5532***
LR (DM vs GB)	1.4600	3.8284*	11.5837***
LR (DM vs GB2)	0.1799	0.1401	0.17256

5% critical values for $\chi^2(1)$ and $\chi^2(2)$ distributions are 3.84 and 5.99 respectively.
 ***, ** and * denote rejection of the nested distribution at 1%, 5% and 10% levels respectively.

data’s distribution and of not requiring any numerical optimization algorithm, except, perhaps, a cross-validation search for the optimal bandwidth, which usually is much simpler than maximum likelihood parametric distribution estimation.

The kernel nonparametric estimators we used were all based on a standard normal gaussian kernel. This was done for simplicity’s sake and may actually not be optimal. For one thing, Serfling [1980] showed that, in a wide range of situations, the Epanechnikov kernel, which is quite different from the gaussian kernel, minimises the kernel density estimator’s integrated mean squared error. Also, the fact that income data is nonnegative is not properly being taken into account by usual kernels such as the gaussian or the Epanechnikov. Indeed, these kernels give positive weights to negative income values, which are outside the support of the data. This creates a bias, aptly called boundary bias, which can be eliminated by using asymmetric kernels that assign 0 weights to every values outside the data’s support. Bouezmarni and Scaillet [2005] provide some theoretical properties of nonparametric density estimation based on asymmetric kernels and use a Brazilian data set to show that they may indeed improve upon usual nonparametric estimators. Because using these kernel estimators is somewhat more complicated than using gaussian ones, we do not consider them here. Finally, it is well known that smoothed nonparametric kernel distribution estimators are biased, a flaw they compensate by a reduced variance compared to the unsmoothed empirical density estimator. It is possible to decrease this bias by using higher order kernels. Briefly, standard kernels such as the gaussian and Epanechnikov are called second order kernels because their second moment is their smallest nonzero one. A m order kernel is one for which the smallest nonzero moment is its m^{th} . It can be shown that using a m order kernel with $m > 2$ decreases the bias. This however comes at the cost of a greater computational burden and with the undesirable feature that the density estimate is no longer guaranteed to be positive at every point of

its support. To keep the analysis as simple as possible, we do not consider higher order kernels. Interested readers may consult Q.Li and Racine [2006] for further details.

No matter what type of kernel is used, for any given data set or any underlying distribution, the choice of bandwidth is of much greater importance than that of the kernel function. As discussed by Azzalini [1981] and Wand and Jones [1995], $h_1 = 1.587\hat{\sigma}n^{-1/3}$ and $h_2 = 1.3\hat{\sigma}n^{-1/3}$ are optimal choices when the data has been generated by a distribution resembling the normal and when one is interested in the whole CDF or its tails respectively. Because income data distribution are usually quite different from the normal, it maybe expect that the cross-validation procedure Bowman et al. [1998] would provide more accurate estimates.

The accuracy of the different estimation procedures is measured using criteria similar to those used by Boccanfuso et al. [2008]. These are the sum of squared errors (SSE), the sum of absolute errors (SAE) and, in the maximum likelihood cases, the maximized log-likelihood value.

$$SSE = \sum_{t=1}^n \left(n_i/n - p_i(\hat{\theta}) \right)^2,$$

$$SAE = \sum_{t=1}^n | n_i/n - p_i(\hat{\theta}) |,$$

where n is the sample size and

$$p_i(\hat{\theta}) = \int_{-\infty}^{x_i} \hat{f}(x)dx,$$

$$n_i = \sum_{t=1}^{x_i} I(x_t < x_i),$$

where $\hat{f}(x)$ is an estimate of the PDF $f(x)$. In simpler terms, p_i is the estimated weight of $f(x)$ to the left of the point x_i while n_i/n is the fraction of the data in the actual sample to the left of that point, that is, the EDF evaluated at x_i .

Table 4 reports some results from estimating a number of distributions from our Malian data set. We chose to focus on the different estimates' SSE and SAE as a basis of comparison, but we also report FGT0 indices since they are of special interest to us (Table 5). It must be noted here that one can trivially obtain a nonparametric estimator with zero SSE and SAE by setting $h = 0$, which yields

Table 4: SSE of estimated distributions with original Malian data

	GB	GB2	SM	Dagum	K h_{mse}	K h_{imse}	K h_{cv}
Mali	0.041	0.169	0.133	0.535	0.025	0.051	0.001
Urban	0.019	0.039	0.070	0.118	0.028	0.057	0.064
Rural	0.088	0.106	0.243	1.064	0.068	0.129	0.0049
Rural men	0.019	0.043	0.071	0.131	0.026	0.051	0.006
Rural women	0.657	0.604	0.612	0.552	2.582	4.297	0.758
Urban men	0.996	0.911	1.090	0.847	1.370	2.548	0.091
Urban women	0.067	0.109	0.234	0.978	0.073	0.140	0.065
Bamako men	1.153	1.097	1.146	0.843	0.704	1.286	0.132
Bamako women	6.668	4.302	4.307	4.289	8.358	14.73	0.651
Non-Bamako men	0.045	0.051	0.072	0.450	0.069	0.134	0.099
Non-Bamako women	1.561	1.505	1.571	1.499	1.578	2.859	0.153

the EDF. Because minimizing the either the SSE or the SAE is not the purpose of any of our parametric or nonparametric estimator, we consider this measure appealing.

Table 5: FGT0 indices with original Malian data

	GB	GB2	SM	Dagum	K h_{mse}	K h_{imse}	K h_{cv}	EDF
Mali	0.5199	0.51257	0.51040	0.50233	0.51343	0.51202	0.51628	0.51651
Urban	0.2601	0.25969	0.25925	0.26374	0.26152	0.26246	0.26145	0.26467
Rural	0.6398	0.64252	0.63505	0.63285	0.63749	0.63538	0.64094	0.64427
Rural men	0.6531	0.65643	0.64993	0.64740	0.65073	0.64879	0.65348	0.65728
Rural women	0.4067	0.40587	0.40410	0.40770	0.39739	0.39447	0.40741	0.42613
Urban men	0.2532	0.25256	0.25064	0.25312	0.25912	0.26716	0.26461	0.24255
Urban women	0.2617	0.26078	0.26023	0.26435	0.26441	0.26517	0.23976	0.26791
Bamako men	0.1680	0.16789	0.17778	0.16905	0.18681	0.19309	0.17854	0.18061
Bamako women	0.1577	0.13500	0.13903	0.13501	0.20722	0.22648	0.16413	0.17307
Non-Bamako men	0.2980	0.29857	0.29597	0.30380	0.29958	0.30009	0.29982	0.30060
Non-Bamako women	0.2856	0.28418	0.28253	0.28654	0.28255	0.29102	0.26041	0.26229

4 Monte Carlo Simulation

The interest of the present paper centers on the estimation of poverty incidence within a population based on an observed sample. Let z be the poverty line and F be the actual cumulative distribution of households incomes in the economy. The true population poverty rate is then

$$PR_0 = F(z). \quad (3)$$

Let $\hat{F}(x)$ denote a consistent estimator of $F(x)$. Then, a consistent estimator of PR_0 is

$$\hat{P}R = \hat{F}(z).$$

The simplest consistent estimator of $F(x)$ is the EDF, which yields an estimator of PR_0 that simply is the fraction of households with income below z over the total number of households:

$$\hat{P}R_{EDF} = \hat{F}_{EDF}(z) = \frac{1}{n} \sum_{t=1}^n I(y_t < z)$$

where $I()$ is the indicator function. This estimator is the most widely utilized in practice and is often the default in popular statistical softwares such as DASP. One weakness of the EDF is that it has a rather large variance, and this characteristic carries over to $\hat{P}R_{EDF}$. To illustrate this, consider a sample of 100 observations with 20 observations below the poverty line, so that $\hat{P}R_{EDF} = 0.2000$. Suppose then that an additional observation is added to the sample and consider the impact of this observation on $\hat{P}R_{EDF}$. If the new observation is above z , then $\hat{P}R_{EDF}$ becomes $20/101 = 0.1980$, a variation of 1%. If on the other hand the new observation is below z , then $\hat{P}R_{EDF}$ becomes $21/101 = 0.2079$, a variation of nearly 4%. Such large changes on account of a single observation are evidently not a desirable feature.

Smooth estimators such as those provided by maximum likelihood and nonparametric kernels reduce this problem by trading some variance for more bias. The question is then how much bias is acceptable for a given reduction of variance? The usual answer to this question is that the best estimator is the one which minimizes the mean squared error, defined here as

$$MSE(\hat{P}R) = E \left(\hat{P}R - PR_0 \right)^2. \quad (4)$$

Of course, the expectation in (4) cannot be computed analytically because we do not generally know the distribution of $\hat{P}R$. It can nevertheless be approximated with arbitrary accuracy through Monte Carlo experiments. Suppose we draw M samples from F and let $\hat{P}R_i$ denote the poverty rate estimated from \hat{F} in the i^{th} sample. Then, we can approximate $MSE(\hat{P}R)$ by

$$M\hat{S}E(\hat{P}R) = \frac{1}{M} \sum_{i=1}^M \left(\hat{P}R_i - PR_0 \right)^2. \quad (5)$$

For a large enough M , a law of large numbers implies that $M\hat{SE}(\hat{PR}) \xrightarrow{p} MSE(\hat{PR})$. It must be stressed that expression (5) cannot be computed for any given sample because PR_0 depends on the true characteristics of the data's generating process, which is not known. This expression can only be used to numerically derive some theoretical, finite sample, properties of different estimators of the poverty rate. This is what we do in this section.

To carry out the Monte Carlo simulations necessary to approximate (5), one must first choose a distribution from which to draw artificial samples. In all that follows, samples are drawn from a Singh-Maddala distribution, which is a special case of the GB distribution with $c = p = 1$. The main reason for this choice is our desire to control the thickness of the right tail of the distribution from which we generate data, which may be a crucial parameter. With the SM distribution, the tail thickness is given by the product of q and a . A second reason is that it is somewhat easier to draw random numbers from the SM than from the GB2 or GB distributions (the SM is easily solved for y). Finally, as we saw in Table 2 and as was found by several authors [Brachmann et al., 1996] the SM distribution fits actual data quite well.

In order to draw artificial samples, values for the parameters a , b and q must be chosen. The parameter b determines the scale of the distribution, and is of little importance, so we have set it to 3, which is close to the values we found for our different Malian population subgroups. The product of a and q determines the thickness of the distribution's right tail. The lower this product is, the thicker the tail is and for $aq \leq 2$ and $aq \leq 1$, the variance and mean, respectively, are infinite. Estimation with our Malian data set yielded tail indexes ranging from slightly over 2 to about 3.5, depending on the population sub-group considered. Our first set of experiments uses samples drawn with a tail index of 2, which thus represents a case slightly more challenging than, yet not too far remote from, what we found with the Malian data. Other samples were drawn with tail indexes of 3 and 1 and will be discussed later.

Our first concern is to find out whether, in our chosen setting, any given estimation method yields the most accurate poverty rate estimates for all possible poverty rate values. Our measure of accuracy is the mean squared error of the estimates, which we compute as

$$MSE(\hat{PR}) = \frac{1}{M} \sum_{j=1}^M (P\hat{R}_j - PR_0)^2,$$

where M is the number of Monte Carlo samples used and PR_0 is the true poverty rate. Figures 1 to 3 show the estimated MSE of the poverty rate estimated by the EDF, the GB2 distribution and the nonparametric kernel using the gaussian reference rules of thumbs and cross validation for samples of 25, 100 and 1000 observations. Obviously, every estimator gains some accuracy as the sample size increases, which is a consequence of the fact that they are all consistent.

It is quite clear from these figures that the EDF estimator, which is the default one in softwares such as DASP, is not optimal. Indeed, the GB2 estimator has lower MSE for all poverty rates considered. Thus, provided one has access to a convenient maximum likelihood estimation package, one appears to always be better off using the GB2 than the EDF.

It is also clear that using nonparametric kernel estimators with gaussian-optimal rules of thumbs is not at all a good idea because their MSE depends greatly on the actual poverty rate and is often much larger than the others. This evidently is due to the fact that the Singh-Maddala distribution is very different from the normal.

Not surprisingly, because it does not require the underlying distribution to be normal, kernel estimation with crossvalidation yields much better results. For small samples, it is almost always more accurate than the EDF and often also outperforms the parametric estimator. It is, however, more sensitive to the actual poverty rate than the parametric estimator. This may be due to the fact that the crossvalidation algorithm is designed to minimize the integrated MSE over the whole distribution, and not at the poverty line. There may therefore be room for improvement by modifying the crossvalidation algorithm so that it gives more weight to points in a close neighborhood of the poverty line. This will be explored in a later version of this paper.

Figures 4 and 5 explore the impact of the thickness of the tail of the distribution on the MSE of nonparametric (with crossvalidation) and parametric poverty rate estimates. It seems that both estimators are fairly robust to all but very thick tails. In this latter case, the parametric estimator seems to be preferable for low poverty rates while the nonparametric one seems to be preferable for other rates. Given that distributions with no moments at all (that is, with a tail index of 1) do not appear to be common for income data, we may conclude that both type of estimators appear to be quite robust to the type of tails usually encountered in practice.

5 Conclusion

This paper extends from the findings of Boccanfuso et al. [2008] and aims to verify if more flexible income distribution functions are systematically preferred over less flexible ones to perform poverty analysis in the context of CGE microsimulation framework. This issue is important since this framework of analysis commonly uses variable size of household groups for the distributional impact analysis. In the context of an empirical analysis, we find that among the four parametric distributions of 4 to 5 parameters, each is the preferred option at least once and that the preferred function can change for the same group after a simulation. Moreover, our results reveal that the poverty analysis can be affected by the choice of the distribution.

Hence, our results would lead us to think that non parametric approach and more specifically with the estimation based on the cross validation provides an interesting alternative to the empiric and parametric approaches. This finding is even stronger when working with target groups of a small size. However, caveats should apply and our results should be considered to be a basis for further investigation. Among possible extension to this work we could highlight the asymmetric kernel that attributes no weight to negative income as oppose to the Gaussian kernel. Moreover, the estimation bias of our results could be reduced with the use of a kernel functions of order superior to two. Finally, an extension could generalize the approach with applications to other poverty and inequality indices. From an economic policy perspective, our results illustrates the importance of considering the sensitivity of poverty analysis to the choice of methodology used by the modeler when using the empiric and less flexible function forms to represent the income distribution.

6 Annexes

Table 6: DASP Results

		Bamako	Other Urban Areas	Urban	Rural	Mali
	Poverty line	149 037	146 068	146 398	119 623	157 920
		Men head household				
Initial	FGT0	18,06%	30,06%	26,79%	65,73%	59,97%
Sim 1	Δ % FGT0	<i>-13,41%</i>	<i>-14,74%</i>	<i>-15,81%</i>	<i>-14,29%</i>	<i>-9,97%</i>
Sim 2	Δ % FGT0	0,00%	2,02%	0,93%	0,62%	0,60%
Sim 3	Δ % FGT0	<i>-24,39%</i>	<i>-25,72%</i>	<i>-26,51%</i>	<i>-19,55%</i>	<i>-13,97%</i>
		Women head household				
Initial	FGT0	17,31%	26,23%	24,26%	42,61%	
Sim 1	Δ % FGT0	-11,11%	-2,08%	-3,51%	-10,67%	
Sim 2	Δ % FGT0	0,00%	0,00%	0,00%	0,00%	
Sim 3	Δ % FGT0	-11,11%	<i>-12,50%</i>	<i>-12,28%</i>	<i>-16,00%</i>	

Source: EMEP 2001

Italic : significant at 5%

Figure 1. Mean Squared Error (n=25)

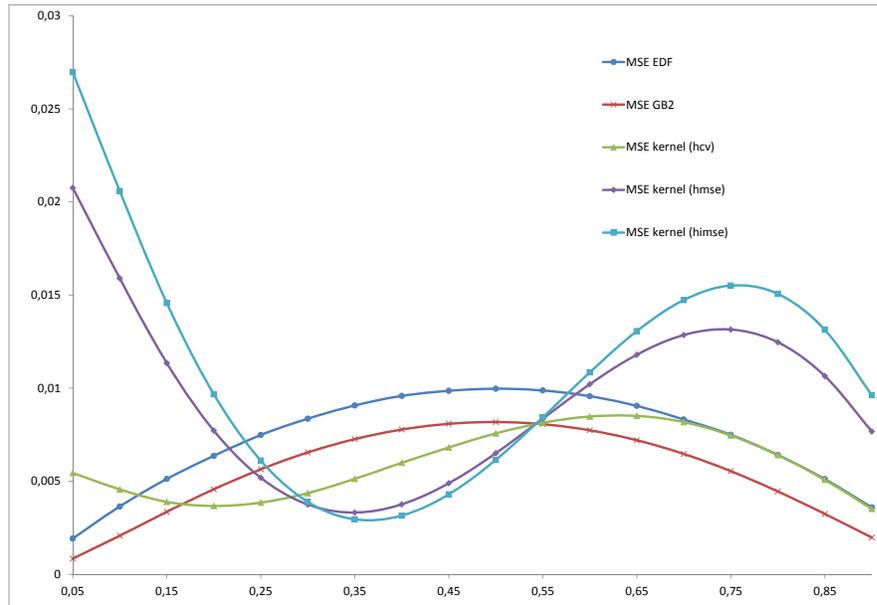


Figure 2. Mean Squared Error (n=100)

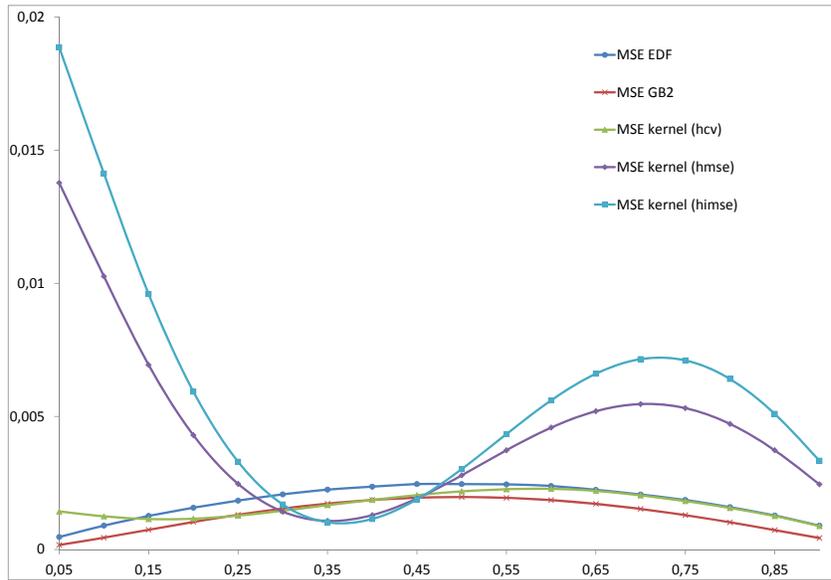


Figure 3. Mean Squared Error (n=1000)

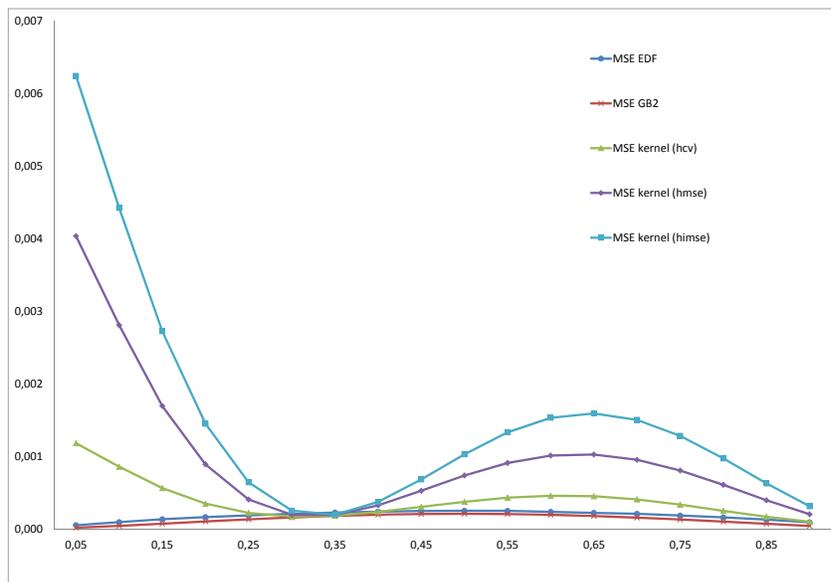


Figure 4. MSE nonparametric (n=100)

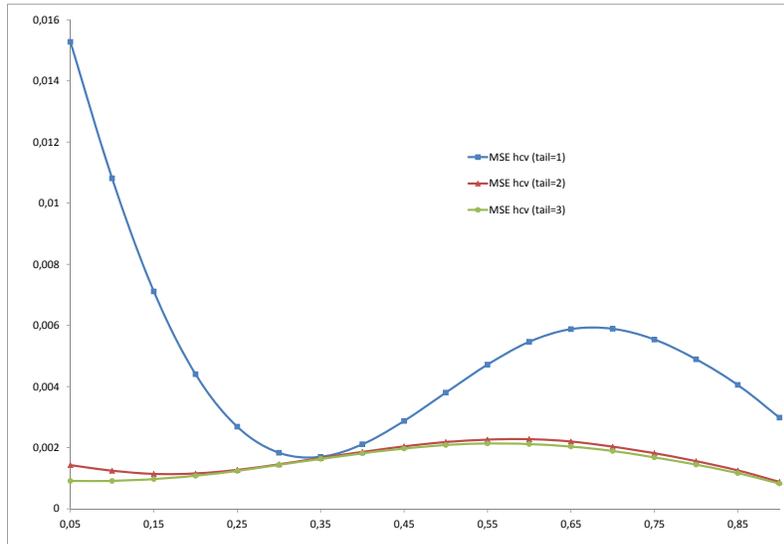
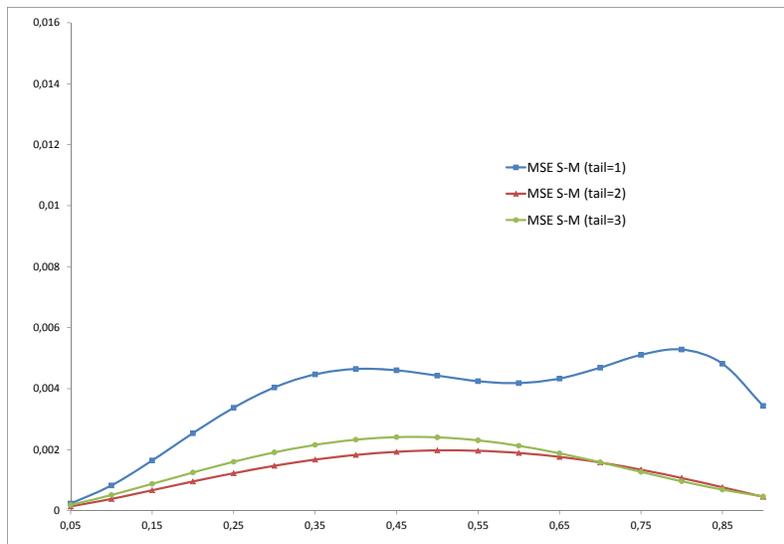


Figure 5. MSE parametric (n=100)



Bibliography

- I. Adelman and S. Robinson. *Income Distribution Policy: A Computable General Equilibrium Model of South Korea*. Stanford University Press, 1978.
- P. Agénor, M. Grimm, and D. Chen. Macro models and household survey data: Linkages for poverty and distributional analysis. mimeo Manchester University, 2005.
- J. Aitchinson and J.A. C. Brown. *The lognormal distribution*. Cambridge University Press, Cambridge, 1957.
- D. Andrews. Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, 68(2):399–405, 2000.
- P. Armington. A theory of demand for products distinguished by place of production (une théorie de la demande de produits différenciés d’après leur origine). *Staff Papers - International Monetary Fund*, 16(1):pp. 159–178, 1969.
- A. Azzalini. A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika*, 68(1):326–328, 1981.
- R. Bandourian, J. B. McDonald, and R. S. Turley. Income distributions: An intertemporal comparison over countries. *Estadística*, 55:135–152, 2003.
- K. Bazlul, M. Mujeri, and S. Raihan. Welfare and poverty impacts of tariff reforms in bangladesh: a general equilibrium approach. Working Papers MPIA 2006-05, PEP-MPIA, 2006.
- D. Boccanfuso and L. Savard. The food crisis and its impacts on poverty in senegal and mali: Crossed destinies. *Development Policy Review*, 29(2):211–247, 03 2011. URL <http://ideas.repec.org/a/bla/devpol/v29y2011i2p211-247.html>.
- D. Boccanfuso, B. Decaluwé, and L. Savard. Poverty, income distribution and cge micro-simulation modeling: Does the functional

- form of distribution matter? *Journal of Economic Inequality*, 6:149–184, 2008. ISSN 1569-1721.
- R. Bordley, J. B. McDonald, and A. Mantrala. Something new, something old: Parametric models for the size distribution of income. *Journal of Income Distribution*, 6(1):97–102, 1996.
- T. Bouezmarni and O. Scaillet. Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data. *Econometric Theory*, 21(02):390–412, April 2005.
- F. Bourguignon and A. Spadaro. Microsimulation as a tool for evaluating redistribution policies. *Journal of Economic Inequality*, 4(1):77–106, 2006.
- A. W. Bowman, M. C. Jones, and I. Gijbels. Testing monotonicity of regression. *Journal of Computational and Graphical Statistics*, 7(4):489–500, December?Winter 1998. URL <http://oro.open.ac.uk/24028/>.
- K. Brachmann, A. Stich, and M. Trede. Evaluating parametric income distribution models. *Allgemeine Statistisches Archiv*, 80:285–298, 1996.
- D. Champernowne. A model for income distribution. *Economic Journal*, 53(1):318–351, 1953a.
- D. Champernowne. The graduation of income distribution. *Econometrica*, 20(4):591–615, 1953b.
- S. Chen and M. Ravallion. Welfare impacts of china’s accession to the world trade organization. *The World Bank Economic Review*, 18(1):29–57, 2004.
- N. Chia, S. Wahba S., and J. Whalley. Poverty-reducing targeting programmes: A general equilibrium approach. *Journal of African Economies*, 3(2):309–38, 1994.
- F. Clementi, M. Gallegati, and G. Kaniadakis. A model of personal income distribution with application to italian data. *Empirical Economics*, 39(2):559–591, 2010.

- J. Cockburn. Trade liberalisation and poverty in nepal: A computable general equilibrium micro simulation analysis. Technical report, 2002.
- C. Dagum. A new model of personal income distribution: Specification and estimation. In Duangkamon Chotikapanich, editor, *Modeling Income Distributions and Lorenz Curves*, volume 5 of *Economic Studies in Inequality, Social Exclusion and Well-Being*, pages 3–25. Springer New York, 2008. ISBN 978-0-387-72796-7.
- S. Dastrup, R. Hartshorn, and J. McDonald. The impact of taxes and transfer payments on the distribution of income: A parametric comparison. *Journal of Economic Inequality*, 5(3):353–369, 2007.
- R. Davidson and J.G. MacKinnon. *Econometric Theory and Methods*. Oxford University Press, 2004. URL <http://books.google.ca/books?id=vqkap8AwAy0C>.
- B. Decaluwé, J.C. Dumont, and L. Savard. How to measure poverty and inequality in general equilibrium framework. CREFA Cahier de recherche 9920, 1999.
- B. Decaluwé, L. Savard, and E. Thorbecke. General equilibrium approach for poverty analysis: With an application to cameroon. *African Development Review*, 17(2):213–243, 2005. ISSN 1467-8268. doi: 10.1111/j.1017-6772.2005.00113.x. URL <http://dx.doi.org/10.1111/j.1017-6772.2005.00113.x>.
- B. Decaluwé, A. Patry, L. Savard, and E. Thorbecke. Poverty analysis within a general equilibrium framework. In A.K. Fosu, G.M. Mwabu, and E. Thorbecke, editors, *Poverty in Africa: Analytical and Policy Perspectives*, pages 127–183. University of Nairobi Press, 2009. ISBN 9789966846624. URL <http://books.google.ca/books?id=6Rn1KcNWTEwC>.
- K. Dervis, J. de Melo, and S. Robinson. *General Equilibrium Models for Development Policy*. Cambridge University Press, 1991. ISBN 978-0821312742.
- J. Foster, J. Greer, and E. Thorbecke. A class of decomposable poverty measures. *Econometrica*, 52(3):761–66, 1984.

- R. Gibrat. *Les Inégalités Economiques*. Sirely, Paris, 1931.
- C. Gourieroux, A. Holly, and A. Monfort. Likelihood ratio test, wald test, and kuhn-tucker test in linear models with inequality constraints on the regression parameters. *Econometrica*, 50(1):63–80, 1982.
- T. Hertel and J. Reimer. Predicting the poverty impacts of trade reform. World Bank Policy Research Working Paper 3444, World Bank, Washington, 2004.
- A. De Janvry, E. Sadoulet, and A. Fargeix. *Adjustment and Equity in Ecuador*. Adjustment and Equity in Developing Countries. Development Centre of the Organisation for Economic Cooperation and Development, 1991. ISBN 9789264135390. URL <http://books.google.ca/books?id=Ic228DByU0cC>.
- G. Kaniadakis. Statistical mechanics in the context of special relativity. *Phys. Rev. E*, 66:056125, Nov 2002.
- G. Kaniadakis. Statistical mechanics in the context of special relativity. ii. *Phys. Rev. E*, 72:036108, 2005.
- C. Kleiber. Dagum vs. singh-maddala income distributions. *Economics Letters*, 53(3):265 – 268, 1996.
- J. McDonald. Some generalized functions for the size distribution of income. *Econometrica*, 52(3):647–63, May 1984.
- J. McDonald and M. Ransom. The generalized beta distribution as a model for the distribution of income: Estimation of related measures of inequality. In Duangkamon Chotikapanich, editor, *Modeling Income Distributions and Lorenz Curves*, volume 5 of *Economic Studies in Inequality, Social Exclusion and Well-Being*, pages 147–166. Springer New York, 2008. ISBN 978-0-387-72796-7.
- J. McDonald and Y. Xu. A generalization of the beta distribution with applications. *Journal of Econometrics*, 69(2):427–428, 1995.
- J.-M. Montaud. Dotations en capital et pauvreté des ménages au burkina faso : une analyse en Équilibre général calculable. *Revue d'économie du développement*, 17(1):43–71, 2003.

- S. Oum. Income distribution and poverty in a cge framework: A proposed methodology. Technical report, 2009.
- V. Pareto. *Cours d'économie publique*. Rouge, Lausanne, 1897.
- S. Parker. The generalised beta as a model for the distribution of earnings. *Economics Letters*, 62(2):197 – 200, 1999.
- Q.Li and J. S. Racine. *Nonparametric Econometrics: Theory and Practice*, volume 1 of *Economics Books*. Princeton University Press, 2006. URL <http://ideas.repec.org/b/pup/pbooks/8355.html>.
- W. Reed. Brownian–laplace motion and its use in financial modelling. *Communications in Statistics - Theory and Methods*, 36(3):473–484, 2007.
- W. Reed and M. Jorgensen. The double pareto-lognormal distribution—a new parametric model for size distributions. *Communications in Statistics - Theory and Methods*, 33(8):1733–1753, 2004.
- W. Reed and F. Wu. New four- and five-parameter models for income distributions. In Duangkamon Chotikapanich, editor, *Modeling Income Distributions and Lorenz Curves*, volume 5 of *Economic Studies in Inequality, Social Exclusion and Well-Being*, pages 211–223. Springer New York, 2008. ISBN 978-0-387-72796-7.
- R. Rutherford. Income distribution a new model. *Econometrica*, 23: 277–294, 1955.
- A. Salem and T. Mount. A convenient descriptive model of income distribution: The gamma density. *Econometrica*, 42(6):pp. 1115–1127, 1974.
- L. Savard. Poverty and inequality analysis within a cge framework: A comparative analysis of the representative agent and micro-simulation approaches. *Development Policy Review*, 23(3):313–332, 2005.
- R. J. Serfling. *Approximation Theorems for Mathematical Statistics*. Wiley, New York, NY, 1980.
- S. Singh and G. Maddala. A function for size distribution of incomes. *Econometrica*, 44(5):pp. 963–970, 1976.

- D. Stifel and E. Thorbecke. A dual-dual cge model of an archetype african economy: trade reform, migration and poverty. *Journal of Policy Modeling*, 25(3):207–235, April 2003.
- P. Tadikamalla. A look at the burr and related distributions. *International Statistical Review / Revue Internationale de Statistique*, 48(3):pp. 337–344, 1980.
- L. Taylor and F. Lysy. Vanishing income redistributions : Keynesian clues about model surprises in the short run. *Journal of Development Economics*, 6(1):11–29, 1979.
- P. Wand and C. Jones. *Kernel Smoothing*. Monographs on Statistics and Applied Probability. Chapman & Hall, 1995. ISBN 9780412552700. URL <http://books.google.ca/books?id=GT00i5yE008C>.