Indirect Inference for fitting Income Distributions

Gaoyuan Tian
and
S. Rao Jammalamadaka
Department of Statistics and Applied Probability
University of California, Santa Barbara

1 Introduction

The distribution of incomes and wealth play an important role in the measurement of inequality and poverty among people as well as nations. Various methods and different parametric models for income distribution are developed in a number of articles by many economists— see e.g. Chotikapanich et al. (2007), McDonald & Xu (1995). In these papers, the Generalized Method of Moments (GMM) method is the preferred way to estimate an income distribution, which can take several parametric forms. To use GMM, for each parametric model explicit expressions for the expected values of the moments or of the functions used, are needed. In this article, we employ a general method of fitting these models, using the "indirect inference" method which allows us to estimate these quantities for a given model without needing to find explicit analytical expressions for them, and thus provides an extension of the work in Hajargasht et al. (2012).

This article is organized as follows. In Section2, we give a brief introduction to some measures of inequality including the Gini index and the Lorenz Curve (LC), because our estimation of the income distribution is based on World Bank data which provides certain values of the empirical Lorenz Curve as given in Table 2. We also review here some commonly used parametric distributions used for income. In Section 3, we describe indirect inference method as a suitable approach for these types of data sets. Theoretical properties of this estimator and a goodness-of-fit test are provided. In Section 4, we test the optimization algorithm used in our method. Also a Monte Carlo study is conducted to compare and evaluate these estimators. As a demonstration of the power of this method, in Section 5, we illustrate it by comparing the income distributions as well as inequality indices for India, China and the USA over the past 30 years. We end with a brief concluding remark in Section 6.

2 Introduction to Some Inequality Measures

2.1 Lorenz Curve

Let $x_1 \leq x_2 \leq \cdots \leq x_n$ be ordered data, say on incomes. The empirical Lorenz Curve is defined as

$$L(i/n) = s_i/s_n \tag{1}$$

where $s_i = x_1 + x_2 + \cdots + x_i, L(0) = 0, i = 0, \cdots, n$.

Let x_i denote data drawn from the distribution function $F(x)$ with mean μ. Let z_p denote the quantile corresponding to a proportion $0 \leq p \leq 1$ i.e.

$$p = F(z_p) = \int_0^{z_p} f(t)\, dt \tag{2}$$

and then the theoretical Lorenz Curve is defined

$$L(p) = \mu^{-1} \int_0^z t f(t)\, dt = \frac{\int_0^z t f(t)\, dt}{\int_0^\infty t f(t)\, dt} \tag{3}$$

The numerator sums the incomes of the bottom p proportion of the population, while the denominator sums the incomes of all the population.

Assuming that F is continuous, one may write $z = F^{-1}(p)$ and a change of variable to write the LC in a direct way:

$$L(p) = \mu^{-1} \int_0^p F^{-1}(t)\, dt \tag{4}$$

Table 1 shows LC expression for some common distributions. Notice that, for exponential distributions, LC does not depend on the scale-parameter. This property could be used for goodness of fit tests (see Gail & Gastwirth (1978)). Figure 1 compares LC for lognormal and exponential.

Distribution	CDF	Lorenz curve
Exponential	$F(x) = 1 - \exp^{-\lambda x}, x > 0$	$p + (1-p)\log(1-p)$
General Uniform	$F(x) = \frac{x-a}{\theta}, a < x < a + \theta$	$\frac{ap + \theta p^2/2}{a + \theta/2}$
Pareto	$F(x) = 1 - (a/x)^a, x > a, a > 1$	$1 - (1-p)^{(a-1)/a}$
lognormal	$F(x) = 1/2 + 1/2\,\mathrm{erf}[\frac{\log x - \mu}{\sqrt{2}\sigma}]$	$\Phi(\Phi^{-1}(p) - \sigma)$

Table 1: Lorenz Curve for some distributions

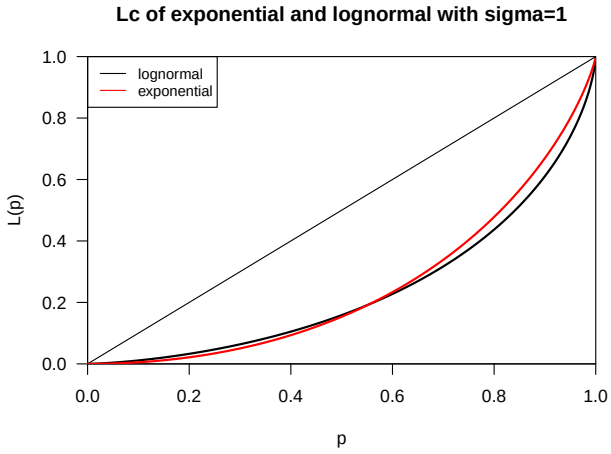

Figure 1: Lorenz Curve of lognormal and exponential

2.2 Gini Index

Gini index is a number between 0 and 1 which gives information about the income inequality of a country, and is the most commonly used measure of inequality. It is also a U-statistic widely used in goodness of fit tests. Jammalamadaka & Goria (2004) introduced a test of goodness of fit based on Gini index of spacings. Recently, Noughabi (2014) introduced a general test of goodness of fit based on the Gini index of data. One way to define Gini index is through expected mean difference.

Definition 1. *Gini* $:= \dfrac{E|X-Y|}{2 \cdot E(X)}$ *where* X, Y *are two random points drawn independently from the distribution* F*.*

The sample version could be written in the following way:

$$Gini(S) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2(n-1) \sum_{i=1}^{n} x_i} \tag{5}$$

It could also be calculated via LC (Gastwirth 1972):

$$G(t) = 2 \cdot \int_0^1 (t - L(t))\, dt \tag{6}$$

2.3 Some Popular Parametric models for income

The income distribution is heavily positively skewed and has a long right tail. The popular income distribution models include Generalized Beta-2 distribution, Gamma distribution and the lognormal distribution.

Generalized Beta-2 distribution (1.7) is widely used for modeling income distribution. Beta-2 ($a = 1$), Singh-Maddala ($p = 1$), Dagum ($q = 1$) and Generalized gamma ($q \to \infty$) are special cases of Generalized beta-2 distribution (see McDonald & Xu (1995)).

$$f(x; a, b, p, q) = \frac{ax^{ap-1}}{b^{ap}B(p, q)(1 + (x/b)^a)^{p+q}}, x > 0 \tag{7}$$

Lognormal distribution (1.8) is another popular income distribution model, its pdf could be derived from $log(X) = Y$ which has a normal distribution.

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(log(x)-\mu)^2}{2\sigma^2}}, x > 0, \sigma > 0 \tag{8}$$

Many alternate models exist, but as Cowell (1995) says, the more complicated four parameters densities are not particularly good choices. Their parameters are hard to interpret and may have an over-fitting problem. He argues in favor of lognormal and gamma density which has two parameters. Among the distribution with two parameters, the Pareto density is nice for modeling high incomes while gamma and lognormal are nice for modeling middle range incomes. In this article, lognormal distribution is chosen for illustrative purposes.

3 Indirect Inference Method

3.1 Indirect Inference Framework

C.Gourierou & Renault (1993) first introduced indirect inference as a simulation based method for estimating the parameters of an extensive class of models. This method is particularly important when, the likelihood function is analytically intractable or considerably difficult to evaluate. Indirect inference method greatly simplifies the estimation procedure because all we need is to simulate the required moments or functions for a candidate model.

The auxiliary parameter vector, denoted by $\boldsymbol{\pi}(\theta)$, which is a function of $\boldsymbol{\theta}$, and has an easy-to-compute empirical estimator $\hat{\boldsymbol{\pi}}$. The relationship between $\hat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}(\boldsymbol{\theta})$ is not required to be explicit, compared with the generalized method of moment (GMM) proposed by Hansen (1982). In general, an estimator of $\boldsymbol{\theta}$ could be defined by the solution of the following optimization problem:

$$\underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\operatorname{argmin}} \ (\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}(\boldsymbol{\theta}))^T \boldsymbol{\Omega} (\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}(\boldsymbol{\theta}))$$

where $\boldsymbol{\Theta}$ is the parameter space, and $\boldsymbol{\Omega}$ is a positive definite weight matrix. The idea here is to find the parameter vector $\boldsymbol{\theta}$ such that $\hat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}(\boldsymbol{\theta})$ are as close as

possible. If $\boldsymbol{\pi}(\boldsymbol{\theta})$ could be calculated for given $\boldsymbol{\theta}$, either as an explicit function or estimated through a software, then the estimator $\boldsymbol{\theta}$ could be obtained by a standard optimization algorithm. Otherwise, the estimator could be approximated by parametric bootstrap as follows:

Step 1 H samples of sample size N is simulated from $F_{\boldsymbol{\theta}}$.

Step 2 For each sample h, $h = 1, 2, \ldots, H$, its $\boldsymbol{\pi}^{*h}(\boldsymbol{\theta})$ is calculated based on its empirical distribution function.

Step 3 $\boldsymbol{\pi}(\boldsymbol{\theta})$ could be approximated by $\boldsymbol{\pi}^*(\boldsymbol{\theta}) = \dfrac{1}{H}\sum_{h=1}^{H} \boldsymbol{\pi}^{*h}(\boldsymbol{\theta})$.

Thus the indirect inference estimator $\hat{\boldsymbol{\theta}}$ is defined as follows:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}} \ (\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}^*(\boldsymbol{\theta}))^T \boldsymbol{\Omega} (\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}^*(\boldsymbol{\theta}))$$

3.2 Lorenz Curve based Indirect Inference

Here we wish to determine the distribution based on values of the empirical Lorenz curve. The auxiliary estimator $\hat{\boldsymbol{\pi}}$ is set to be the sample mean and 9 points on empirical LC in Table 1.3: $\hat{\boldsymbol{\pi}} = (\bar{X}, \hat{L}(0.1), \cdots, \hat{L}(0.9))^T$. Thus the auxiliary parameters corresponds to the theoretical mean and 9 points on theoretical LC implied by lognormal distribution. The optimal choice of the weight matrix is given by $\boldsymbol{\Omega} = (Var(\hat{\boldsymbol{\pi}}))^{-1}$, thereby ensuring that the estimator $\hat{\boldsymbol{\theta}}$ is asymptotically efficient. However, $(Var(\hat{\boldsymbol{\pi}}))^{-1}$ is a function of $\boldsymbol{\theta}$, we obtain an estimate for this matrix is through a two-step procedure (similar to the two-step GMM) as follows:

Step 1 $\boldsymbol{\Omega} = \boldsymbol{I}$ is used with $\boldsymbol{I}$ denoting the identity matrix, to solve the optimization problem and obtain the initial estimate $\boldsymbol{\theta}_1$.

Step 2 The weighting matrix is estimated with $\hat{\boldsymbol{\Omega}} = (Var(\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_1)))^{-1}$.

The expression of $Var(\hat{\boldsymbol{\pi}}(\boldsymbol{\theta}_1))$ is hard to derive and hence evaluated through parametric bootstrap with simulations using the initial estimate $\boldsymbol{\theta}_1$. The optimization algorithm used in this case is the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm which is an iterative method that solves non-linear optimization problems. This estimation procedure could be described by the genetic algorithm shown in Figure 2. $\boldsymbol{\pi}(\boldsymbol{\theta})$ could be calculated given a reasonable initial value of $\boldsymbol{\theta}$. Thereafter an iterative process is triggered to search the optimal θ until some convergence criteria are satisfied.

Having defined our proposed estimator and described the procedure to obtain the estimator in practice, the next section studies the asymptotic and robustness properties of this estimator.

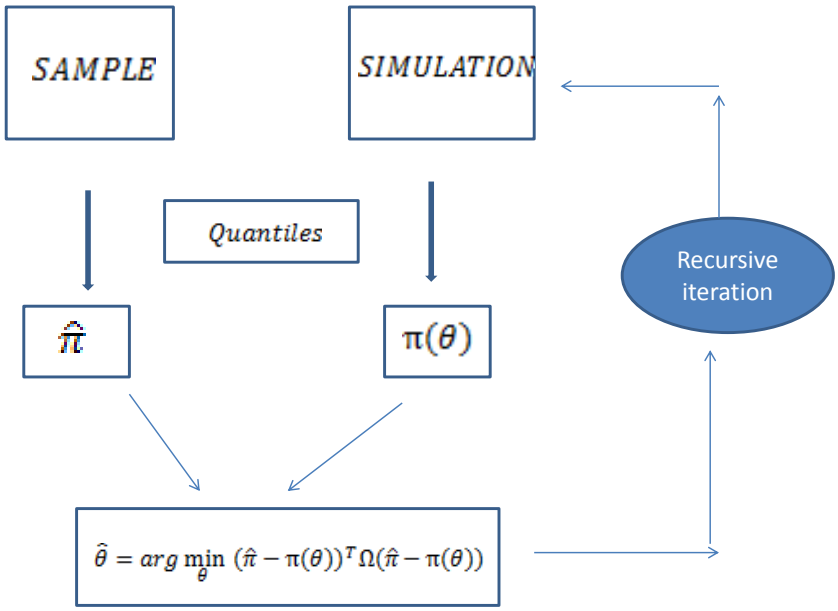

Figure 2: Estimation Algorithm

3.3 Theoretical Properties

To study the asymptotic properties we make use of the existing results and conditions for indirect estimators given in C.Gourierou & Renault (1993). Denoting $\boldsymbol{\theta}_0$ as the true parameter vector, let us first investigate the conditions which ensure the consistency and asymptotic normality of the proposed estimator in our case:

(A1) $\boldsymbol{\xi}_n = \sqrt{n}(\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}(\boldsymbol{\theta}_0)) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{V})$ where $\boldsymbol{V} = \lim_{n\to\infty} Var(\boldsymbol{\xi}_n)$

(A2) There is a unique $\boldsymbol{\theta}_0$ such that auxiliary estimator equals the auxiliary parameter: $\hat{\boldsymbol{\pi}} = \boldsymbol{\pi}(\boldsymbol{\theta_0}) \Rightarrow \boldsymbol{\theta} = \boldsymbol{\theta}_0$

(A3) If $\boldsymbol{\Omega}$ is estimated by $\hat{\boldsymbol{\Omega}}$, then $\hat{\boldsymbol{\Omega}} \xrightarrow{P} \boldsymbol{\Omega}$, where $\boldsymbol{\Omega} > 0$

(A4) $\boldsymbol{\pi}(\boldsymbol{\theta})$ is a differentiable function with $\mathbf{D}(\boldsymbol{\theta}) = \partial \boldsymbol{\pi}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}^T$.

(A5) The matrix $\mathbf{D}^T(\boldsymbol{\theta}) \boldsymbol{\Omega} \mathbf{D}(\boldsymbol{\theta})$ is full rank.

(A6) $\boldsymbol{\Theta}$ is compact.

(A7) The choice of the initial value of $\boldsymbol{\theta}$ is independent of the estimation algorithm.

Theorem 1. *(C.Gourierou & Renault 1993) Under the conditions of (A1)-(A7) and the other usual regularity conditions, our indirect estimator is asymptotically normal, when H is fixed and n goes to infinity:*

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \overset{\mathcal{D}}{\longmapsto} \mathcal{N}(\mathbf{0}, \boldsymbol{\Lambda})$$

with $\boldsymbol{\Lambda} = (1 + \frac{1}{H})\boldsymbol{\Gamma}\mathbf{V}\boldsymbol{\Gamma}^T$ *where* $\boldsymbol{\Gamma} = \left(\mathbf{D}^T(\boldsymbol{\theta}_0)\boldsymbol{\Omega}\mathbf{D}(\boldsymbol{\theta}_0)\right)^{-1}\mathbf{D}^T(\boldsymbol{\theta}_0)\boldsymbol{\Omega}$.

This theorem provides the asymptotic normality of the estimator $\hat{\boldsymbol{\theta}}$ by that of auxiliary estimator $\hat{\boldsymbol{\pi}}$. Since the asymptotic normality is obtained, the consistency property follows. Notice, the factor $(1 + \frac{1}{H})$ distinguish the asymptotic variance of indirect inference with that of GMM: when H goes to infinity, they have the same expression. H is set to be 100 in this article. Now, a bit of explanation about these conditions (A1) to (A7).

For Condition (A1), we need to prove the asymptotic normality of $\hat{\boldsymbol{\pi}} = (\bar{X}, \hat{L}(0.1), \cdots, \hat{L}(0.9))^T$. By central limit theorem, $\bar{X}$ is asymptotic normal. Under some mild conditions, Goldie (1977) proved the weak convergence of the Lorenz process $l_n(\boldsymbol{p}) = \sqrt{n}[L_n(\boldsymbol{p}) - L(\boldsymbol{p})], 0 \leq \boldsymbol{p} \leq 1$, to a Gaussian process if $L(\boldsymbol{p})$ is continuous at the empirical points. Thereafter the asymptotic normality of $\hat{\boldsymbol{\pi}} = (\bar{X}, \hat{L}(0.1), \cdots, \hat{L}(0.9))^T$ is established.

Condition (A2) is often called the "global identifiability" problem in econometrics and is often hard to prove and such, is assumed in many cases. In condition (A3), our 2-step matrix $\hat{\boldsymbol{\Omega}}$ is estimated through the 2-step GMM procedure described above and thus is consistent. The rest of the conditions are standard conditions for indirect estimators such as the one put forward in this paper. We therefore have that the estimator $\hat{\boldsymbol{\theta}}$ proposed here is consistent and asymptotically normal.

3.4 Data

The data comes from the Website of the World Bank, and takes the form of summary statistics including mean income, measures of inequality and 9 points on the empirical LC. In Table 2, the poverty line is the minimum level of income deemed adequate in a particular country. The head-count ratio is the proportion of a population lives below the poverty line. The first part of Table 2 shows the data in the following way: the first 10% of the population owns 1.7% of the total income, the second 10% of the population owns 3.4% of the total income, etc. Since the sum of these 10 numbers equals 1, only the numbers of the first 9

USA 's Income share by deciles(%)										
Year	lowest	2nd	3rd	4th	5th	6th	7th	8th	9th	highest
2010	1.70	3.40	4.56	5.73	7.00	8.44	10.19	12.52	16.25	30.19
USA's poverty index										
Year		mean($/month)			pov.line		headcount(%)		Gini index(%)	
2010		1917.38			1.90		1.00		41.06	

Table 2: Original Data

By deciles(%)									
p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\hat{L}(p)$	1.70	5.10	9.66	15.39	22.39	30.83	41.02	53.54	69.77

Table 3: Transformed Data (cumulative share)

groups need to be included in the moment conditions. The cumulation of these 9 numbers yields the 9 points on the empirical LC $\hat{L}(\boldsymbol{p})$ in Table 3.

With our indirect inference estimator $\hat{\boldsymbol{\theta}}$, $\hat{L}(\boldsymbol{p})$ and $L(\boldsymbol{p}, \hat{\boldsymbol{\theta}})$ are compared as shown at Table 1.4 . This table contains lognormal and gamma, and could be extended for more models to assess the goodness of fit.

3.5 Goodness of Fit Analysis

Here the test statistics J_n can be used to test the validity of the assumed income distribution.

$$J_n = n\left(\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}(\hat{\boldsymbol{\theta}})\right)^T \hat{\boldsymbol{\Omega}} \left(\hat{\boldsymbol{\pi}} - \boldsymbol{\pi}(\hat{\boldsymbol{\theta}})\right) \xrightarrow{D} \chi^2_{M-K} \tag{9}$$

where M is the dimension of auxiliary parameters, K is the number of parameters in the parametric model. $\hat{\boldsymbol{\Omega}}$ is the above two-step weight matrix. n is assumed to be 10000 . The parametric models chosen here are lognormal and gamma. The test results are counterintuitive: the p-value of this test for both models are 0.0000.(Hajargasht et al. 2012)

USA 2010									
p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\hat{L}(\boldsymbol{p})$	1.70	5.10	9.66	15.39	22.39	30.83	41.02	53.54	69.77
$L(\boldsymbol{p}, \hat{\boldsymbol{\theta}})(lognormal)$	2.15	5.63	10.14	15.63	22.31	30.59	40.66	52.98	69.17
$L(\boldsymbol{p}, \hat{\boldsymbol{\theta}})(Gamma)$	1.29	4.22	8.56	14.34	21.55	30.49	41.39	54.99	72.39

Table 4: Goodness of Fit Assessment

4 Simulation Study

4.1 Numerical Optimization

The default optimization algorithm used in R is Broyden-Fletcher-Goldfarb-Shanno(BFGS) algorithm. Similar to Newton's method, it is an iterative method for solving non-linear optimization problems. In this case, the parameter space for σ is $(0, \infty)$. Since it has a lower bound, sometimes this optimization algorithm breaks down when searching for points larger than but close to 0. Instead, we would estimate the parameters $\boldsymbol{\theta} = (\theta_1, \theta_2)$, where $(\mu, \sigma) = (\theta_1, \exp(\theta_2))$. The estimated parameter $\hat{\sigma}$ approximately equals $\log(\hat{\theta_2})$.

Here we want to verify that the estimated point is the local minimum. The true parameters $\boldsymbol{\theta} = (4.8276, -0.4963)$ is obtained from the estimate value of data in Table 2. The data (9 points on Lorenz curve and mean) is simulated from lognormal distribution with above parameters with sample size $N = 1000$. The estimated value $\hat{\boldsymbol{\theta}} = (4.8381, -0.4515)$. It has a local minimum as can be seen from Figures 3 and 4.

4.2 Monte Carlo Study

Suppose we only have the 9 points on the LC, sample median and sample mean. For lognormal distribution, the mean $EX = \exp(\mu + \sigma^2/2)$, Median $m = \exp(\mu)$. By setting these equal to their empirical parts, a method of moment estimators are given by:

$$\hat{\mu} = \log(m), \hat{\sigma} = \sqrt{2(\log(\bar{x}) - \log(m))} \tag{10}$$

Suppose the true parameters $(\mu, \sigma) = (4.8276, \exp(-0.4963))$. Box-plots to compare these two estimators are obtained by Monte Carlo study with sample size $N = 1000$ and Monte Carlo replication $B = 1000$ in Figure 5 and Figure 6. Our indirect inference method has smaller variance especially for σ.

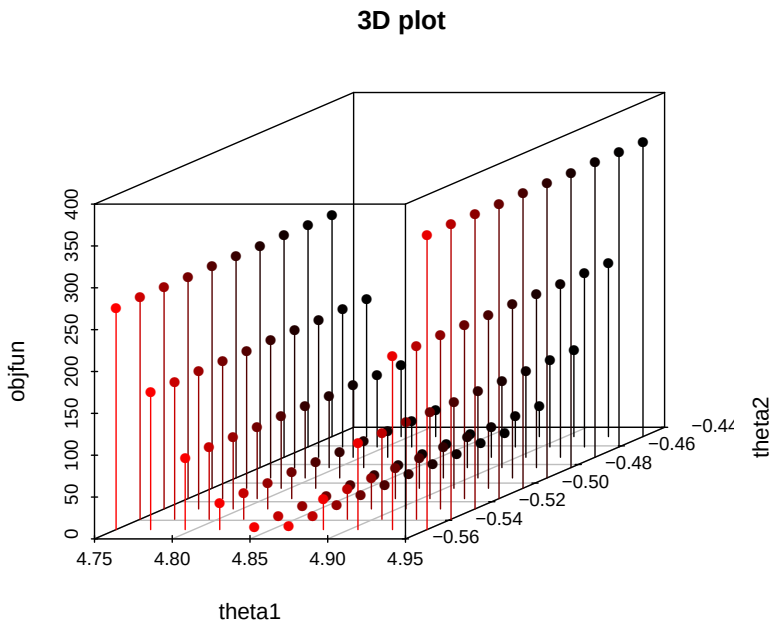

Figure 3: Objective function vs (θ_1, θ_2)

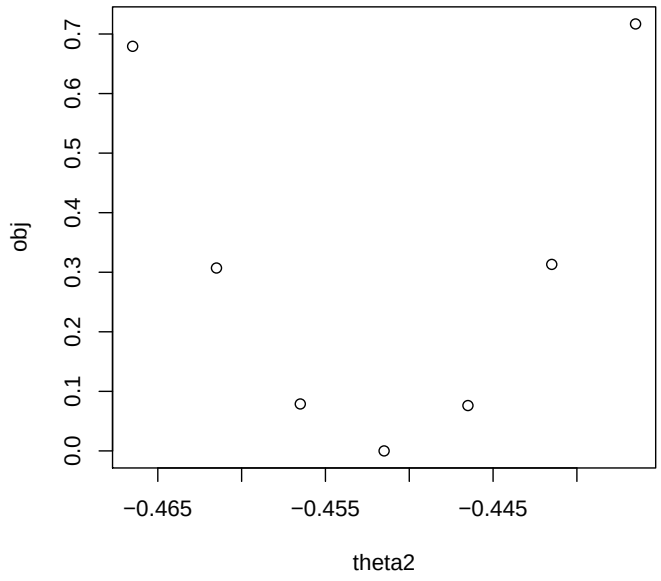

Figure 4: Objective function vs θ_2

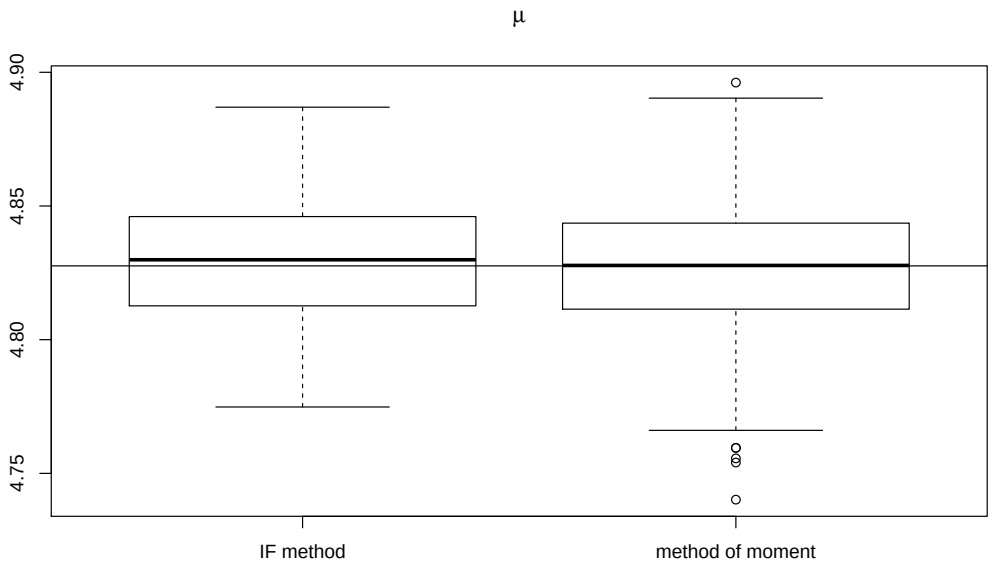

Figure 5: Boxplot of $\hat{\mu}$

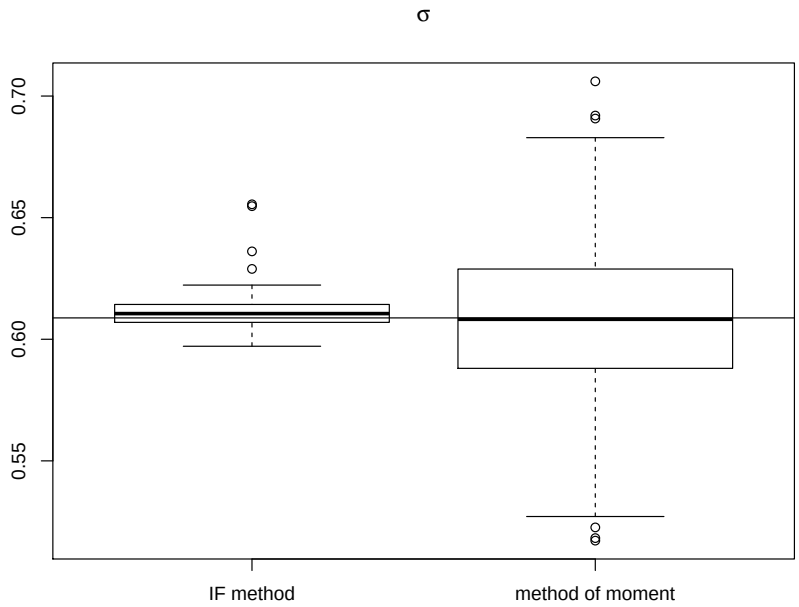

Figure 6: Boxplot of $\hat{\sigma}$

5 Case Study

Greenwood & Jovanovic (1990) found a positive correlation between growth and income inequality in a cross-section of international data. Here we are interested to see the change of India's income distribution and inequality in the past 30 years. In addition, income distribution and inequality of India and China are compared with USA, the largest economy.

5.1 Data

Data is collected every 3 years by the World Bank. It takes the form of summary statistics as shown at Table 5 eg. for India.

India (Urban)'s Income share by deciles(%)										
Year	lowest	2nd	3rd	4th	5th	6th	7th	8th	9th	highest
2010	2.92	4.04	4.87	5.76	6.76	7.95	9.45	11.49	15.01	31.75
1983	3.59	4.61	5.57	6.51	7.50	8.60	9.94	11.75	14.80	27.12

Poverty Index				
Year	mean($/month)	pov.line($/day)	headcount(%)	Gini index(%)
2010	129.75	1.90	19.85	39.35
1983	89.36	1.90	34.20	33.33

Table 5: Income inequality of India: 1983 v.s 2010

5.2 Some Results

With the 9 points on the empirical LC, a smooth empirical LC is estimated by the non-parametric spline technique in R. The income distributions are assumed to be lognormal and are estimated by above indirect inference method. The results are illustrated from Figure 7 to Figure 11.

Although India's Gini index slightly increased in the last 30 years, the population proportion of low income class decreases (Figure 8). The population proportion of low income class of India is significantly larger than USA'S, but smaller than China's (Figure 9 and Figure 10).

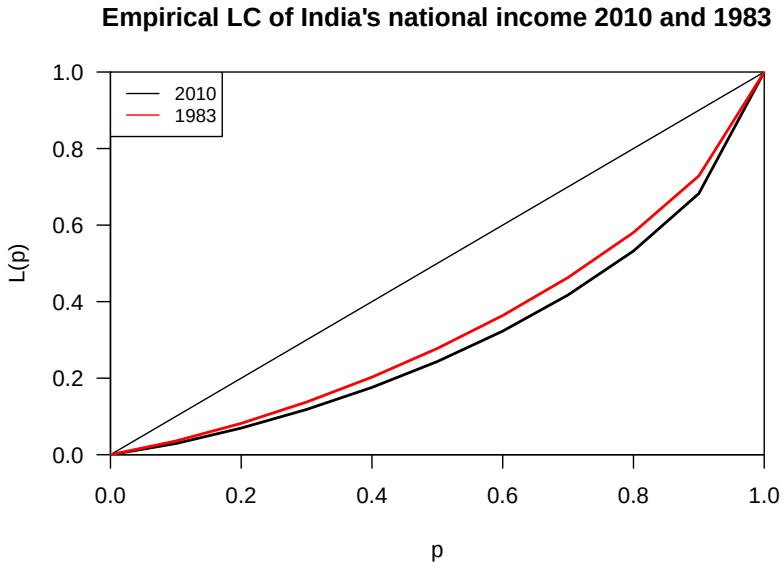

Figure 7: Lorenz Curve of India 1983 vs 2010

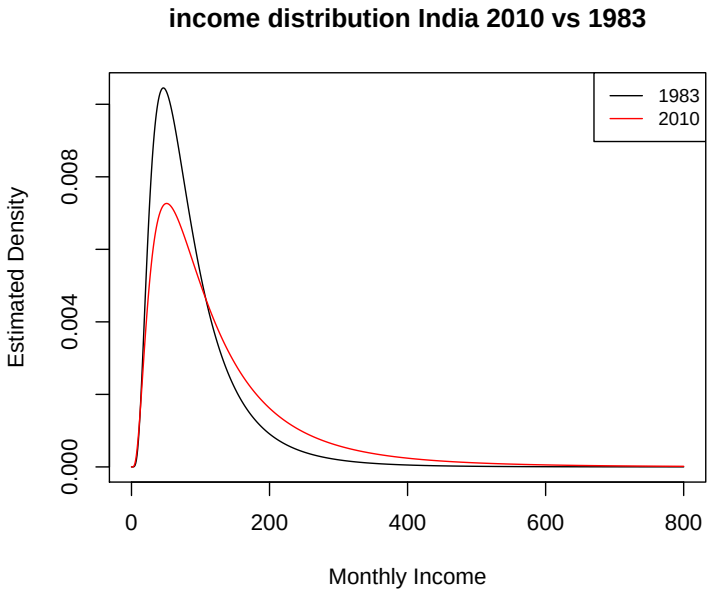

Figure 8: Income distribution of India 2010 vs 1983

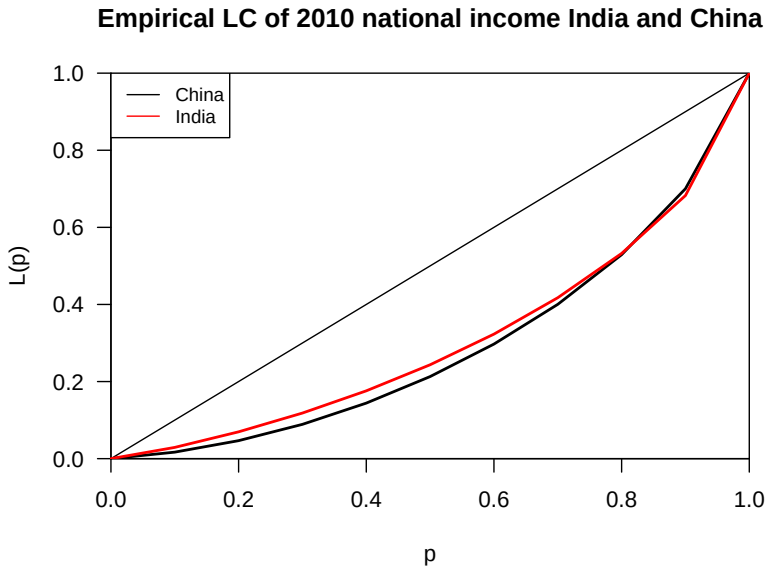

Figure 9: Lorenz Curve of 2010 India vs China

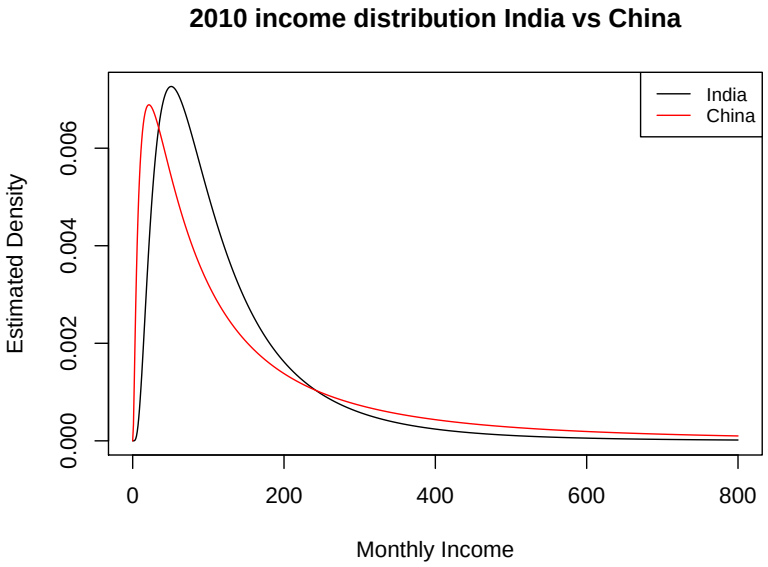

Figure 10: Income distribution of 2010 India vs China

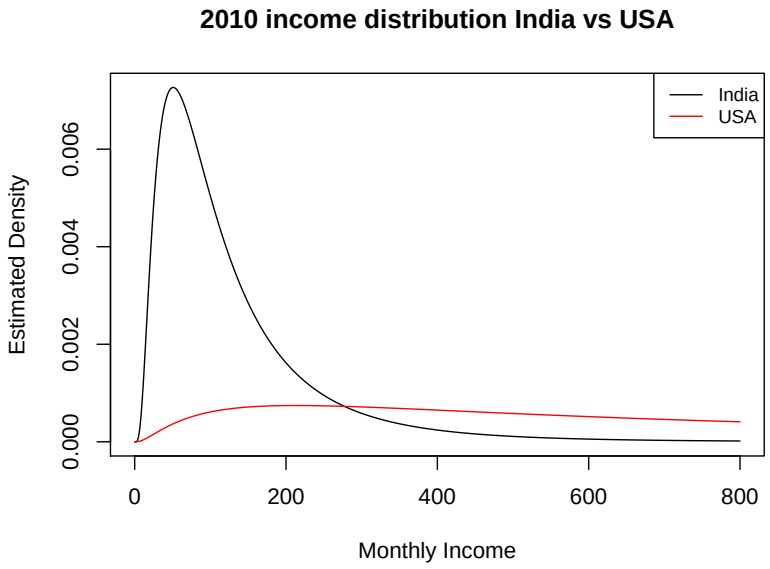

Figure 11: Income distribution of 2010 India vs USA

6 Conclusions

We develop a practical estimation framework based on indirect inference for analyzing income distributions and income inequality given a limited amount of data. This simulation based method is very flexible, namely that the parametric model and/or the auxiliary parameters are adjusted adaptively.

References

C.Gourierou, A. M. & Renault, E. (1993). Indirect inference, *Journal of Applied Econometrics* **8**: 1–10.

Chotikapanich, D. et al. (2007). Estimating and combining national income distribution using limited data, *Journal of Business & Economic statistics* **25**: 97–109.

Cowell, F. (1995). *Measuring Inequality*, Harvester Wheatsheaf, Hemel Hempstead.

Gail, M. H. & Gastwirth, J. L. (1978). A scale-free goodness of fit test for the exponential distribution based on the lorenz curve, *Journal of the American Statistical Association* **73**: 787–793.

Gastwirth, J. L. (1972). The estimation of the lorenz curve and gini index, *The Review of Economics and statistics* **54**: 306–316.

Goldie, C. M. (1977). Convergence theorems for empirical lorenz curve and their inverses, *Advances in Applied Probability* **9**: 765–791.

Greenwood, J. & Jovanovic, B. (1990). Financial development, growth, and the distribution of income, *The Journal of Political Economy* **98**: 1076–1107.

Hajargasht, G. et al. (2012). Inference for income distribution using grouped data, *Journal of Business & Economic statistics* **30:4**: 563–575.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators, *Econometrica* **50**: 1029–1054.

Jammalamadaka, S. R. & Goria, M. (2004). A test of goodness of fit based on gini's index of spacing, *Statistics & probability letters* **68**: 177–187.

Kuznets, S. (1995). Economic growth and income inequality, *American Economic Review* **45**.

McDonald, J. & Xu, Y. (1995). A generalization of the beta distribution with applications, *Journal of Econometrics* **66**: 133–152.

M.Ravallion (1995). Growth and poverty:evidence for developing countries in the 1980s, *Economics letters* **48**: 411–417.

Noughabi, H. (2014). A test of goodness of fit based on gini index, *Journal of the turkish statistical association* **7**: 23–32.