

## **COMPARISON BETWEEN LORENZ CURVE AND BUTLER-MCDONALD MEASURE FOR INCOME AND CONSUMPTION INEQUALITY**

**T. H. ABOUELMAGD<sup>1</sup> and A. A. E. AHMED<sup>2</sup>**

<sup>1</sup>Benha University  
Egypt

<sup>2</sup>Helwan University  
Egypt

<sup>1,2</sup>Taibah University  
Saudi Arabia  
e-mail: [ahmed63feb@gmail.com](mailto:ahmed63feb@gmail.com)

### **Abstract**

The Lorenz curve is a tool used to represent income distributions: It tells us which proportion of total income is in the hands of a given percentage of population. While, Butler and McDonald [4] and [5] presented a very broad family of distributions for measuring inequality based on incomplete moments. Lorenz curve can be considered an example of incomplete moment of the income distribution for measuring inequality, where Lorenz curve is the first member of this family; it is obtained from this family, when  $r$  is equal one. In this article, we study the features of Lorenz curve and the higher orders of Butler-McDonald family and compare among them. In addition, we study Lorenz curve and Butler-McDonald measures under Pareto distribution. Finally, we use a real data of Egyptian income, and consumption to prove that Lorenz curve gives more accurate estimation of income and consumption inequality than higher orders of Butler-McDonald family.

---

2010 Mathematics Subject Classification: 47Jxx, 47J10, 47J35.

Keywords and phrases: Lorenz curve, Gini coefficient, Bulter-McDonald measure, income and consumption inequality, and Pareto distribution.

Received December 21, 2015

## 1. Introduction

Equality of income distribution is found when every income unit receives its proportional share of the total income. The income units may be further defined as individual, head of family, and a consumption unit; see, for example, Cowell [6], Lambert and Lanza [14].

Inequality may be defined as any deviation from equality. Thus, if any person received less than his proportionate share of the aggregate income, the distribution would be unequal, see, for example, Schutz [15] and Zheng and Formby [20]. The Lorenz curve can be defined as the relationship between the cumulative proportion of income units and the cumulative proportion of income received when units are arranged in ascending order of their income; see, for example, Cowell [7], Sen [16], and Shahateet [17]. Also, the Lorenz curve has been used as a graphical device to represent size distribution of income and wealth. In addition, the Lorenz curve is not limited to a specific type of population or variable that is distributed among that population; see Aaberge [1], Bellú and Liberati [3], and Kakwani [12]. In the other words, Lorenz curve can be used to measure income, expenditure, and other variables.

The Lorenz curve can be represented by a function  $L(F)$ , where  $F$  is the horizontal axis and  $L$  is the vertical axis. The points in the Lorenz curve are indexed in non-decreasing order that mean Lorenz curve is linear function; see Daniel [8], Frosini [9], and Lakner and Milanovic [13].

The Lorenz curve can measure the inequality in any society by the distance between the equality line and Lorenz curve. If the Lorenz curve is farther away from the line of perfect equality, then that case is considered to have less equality than a case with a curve nearer to the equality line; see Gastwirth [10] and Zenga [19]. There are two definitions to Lorenz curve; the first one was presented by Lorenz in 1905 in order to compare and analyze inequalities of wealth in a country

during different epochs, or in different countries during the same epoch. The second definition of Lorenz curve was presented by Gastwirth [11] which applies to discrete as well as to continuous variables.

### 1.1. The first definition of Lorenz curve

Let a vector of income  $X$  (a positive random variable) from a continuous distribution with cumulative distribution function (cdf)  $F(x) = F$ , density function  $f(x)$ ,  $x(F) = F^{-1}(x)$  quantile function and let  $X_{i:n}$  denote the corresponding order statistics for a general distribution function  $F(x)$ .

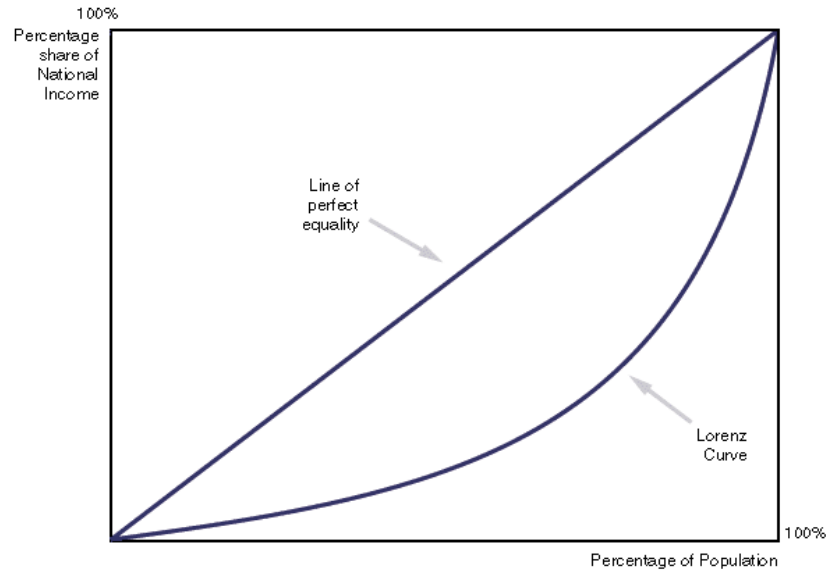
$$F(x) = \int_0^x f(x)dx,$$

where  $F(x)$  can be interpreted as the proportion of units having an income less than or equal to  $x$ ;  $F(x)$  varies from 0 to 1. Furthermore, if the mean  $\mu$  of the distribution exists, the first-moment distribution function of  $X$  is defined as

$$F_1(x) = \frac{1}{\mu} \int_0^x x f(x)dx.$$

Also,  $F_1(x)$  varies from 0 to 1 and  $F_1(x)$  can be interpreted as the proportional share of total income of the units having income less than or equal of  $x$ .

The Lorenz curve ( $L(p)$ ) is the relationship between the variables  $F(x)$  and  $F_1(x)$  and is obtained by inverting functions  $F(x)$  and  $F_1(x)$ , and eliminating  $x$  if the functions are invertible. Alternatively, the curve can be plotted by generating the values of  $F(x)$  and  $F_1(x)$  by considering the arbitrary values of  $x$ . Also, the straight line  $F_1 = F$  is called the egalitarian line (equality line).



**Figure 1.** Lorenz curve.

If the curve coincides with the equality line, it is implied that each unit receives the same income; this is the case of perfect equality of incomes. In the case of perfect inequality of incomes, it implies that all the income is received by only one unit in the population.

### 1.2. The second definition of Lorenz curve

Gastwirth [11] provided an alternative definition of the Lorenz curve, which is expressed in terms of the inverse of the distribution function

$$F^{-1}(t) = \text{Min}[x : F(x) \geq t].$$

This definition ensures the existence of  $x$  for all values of  $F(x)$ , even though in the case of a discrete distribution corresponding to some values of  $F(x)$ ,  $x$  is not exist. Also, if the density function is continuous, this definition of inverse is identical to the usual definition of an inverse function, because  $x = F^{-1}(t)$  and  $f(x)dx = dt$ , where  $t$  varies from zero

to unity when  $x$  varies from zero to infinity. The Lorenz curve can be rewritten as

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

where  $0 \leq p \leq 1$ .

### 1.3. The disadvantage of the Lorenz curve; see Bellù [2]

(i) The Lorenz curve cannot rise above the line of perfect equality and cannot sink below the  $X$ -axis.

(ii) The Lorenz curve is not defined if the mean of probability distribution is zero or infinite.

## 2. Butler-McDonald Measure

Butler and McDonald [4] presented four interdistributional Lorenz curves based on normalized incomplete moments. Also, they defined Lorenz curve, Gini, and Pietra measures of inequality across two populations.

The  $h$ -th partial moment given by  $y < x$  for density function  $f(y)$  is defined as

$$I(x; h) = \int_0^x y^h f(y)dy.$$

From which it follows that

$$E(y^h) = \lim_{x \rightarrow \infty} I(x; h),$$

and

where  $F(y)$  denotes the cumulative distribution function for  $y$ . The fraction of  $h$ -th moment of  $y$ , corresponding to  $y < x$  is given by

$$M_{BM}^r = \frac{\int_0^x y^h f(y) dy}{E(y^h)} = \frac{I(x; h)}{E(y^h)}.$$

They called this the normalized incomplete moment which represent the proportion of each respective moment that is accounted for by income levels less than  $x$ . They presented Lorenz curve as a plot of  $\phi(x; 0)$  and  $\phi(x; 1)$ , depicted in Figure 1. The  $45^\circ$  diagonal is the line of complete equality in which all individuals in the population have the same income. The Lorenz curve indicates the fraction of total income held by different fractions of the population and this curve is concave to the  $45^\circ$  diagonal.

The Pietra measure ( $P$ ) associated with a single population is twice the area of the largest triangle that can be inscribed between the  $45^\circ$  line of complete inequality and Lorenz curve. They expressed  $P$  in terms of incomplete moments,

$$P = \phi(\mu; 0) - \phi(\mu; 1),$$

which is the difference between the fraction of the population having income less than the mean ( $\mu$ ) and the fraction of total income held by that group. Also, the Pietra measure is  $E|x - \mu| / \mu$ , see Gastwirth [10].

Butler and McDonald [4] presented the Gini coefficient as a weighted average of incomplete moments

$$Gini = \int_0^\infty [(x / \mu)\phi(x; 0) - \phi(x; 1)]f(x) dx.$$

Butler and McDonald [4] extended the notion of a Lorenz curve to examine the distance between the black and white income distribution. They plotted the most natural interdistributional Lorenz curve between black and white income distribution which the black incomplete moment,  $\phi_b(x; 0)$ , against the white incomplete moment,  $\phi_w(x; 0)$ . If the distribution

for whites is first-order dominant over blacks, then the Lorenz curve will lie below the 45° diagonal.

They defined four cases for interdistributional Pietra index for blacks and whites income distribution.

(1) The first Lorenz curve with a natural interpretation is obtained by plotting  $(\phi_b(x; 0)$  and  $\phi_w(x; 0)$ )

$$P(0, 0) = \phi_b(\mu_w; 0) - \phi_w(\mu_b; 0),$$

which is the difference between the fraction of blacks with income less than the mean income level for whites and the fractions of whites with incomes less than the mean income for blacks.

(2) The second Lorenz curve is obtained by plotting  $(\phi_b(x; 1)$  and  $\phi_w(x; 1)$ )

$$P(1, 1) = \phi_b(\mu_w; 1) - \phi_w(\mu_b; 1),$$

which is the difference between the fraction of total income held by blacks with incomes less than the mean income for whites and the fraction of total income held by whites with incomes less than the mean income for blacks.

(3) The third Lorenz curve is obtained by plotting  $(\phi_b(x; 0)$  and  $\phi_w(x; 1)$ )

$$P(0, 1) = \phi_b(\mu_w; 0) - \phi_w(\mu_b; 1).$$

(4) The fourth Lorenz curve is obtained by plotting  $(\phi_b(x; 1)$  and  $\phi_w(x; 0)$ )

$$P(1, 0) = \phi_b(\mu_w; 1) - \phi_w(\mu_b; 0).$$

The last two were presented in a paper on black and white racial discrimination by Smith and Welch [18].

### 3. Estimation

When using the sample income data  $Y_1, Y_2, \dots, Y_n$  and the empirical distribution function, the Lorenz curve can be estimated as follows:

$$L(\widehat{p}) = \frac{\sum_{i=1}^n I(Y_i < x) Y_i}{\sum_{i=1}^n Y_i},$$

and Butler-McDonald measure can be estimated as follows:

$$\widehat{M}_{BM}^r = \frac{\sum_{i=1}^n I(Y_i^r < x) Y_i^r}{\sum_{i=1}^n Y_i^r}.$$

### 4. Studying Lorenz Curve and Butler-McDonald Measure under Pareto Distribution

If  $X_1, X_2, \dots, X_n$  are a random variables from Pareto distribution with density function  $pdf = ak^a x^{-(a+1)}$ , cumulative distribution function  $F = 1 - \left(\frac{k}{x}\right)^a$ , quantile function  $X(F) = k(1 - F)^{\frac{-1}{a}}$ ,  $r$ -th noncentral moments

$$\mu_r = ak^r (a - r)^{-1} \text{ and } \mu_X^{(h,r)} = \int_0^\infty x F^r dF.$$

Then we can derive Lorenz curve by using Pareto distribution

$$L(p) = \frac{1}{\int_0^1 x dF(x)} \int_0^{y_r} x dF(x),$$

which when using the indicator function gives



$$L(p) = \frac{1}{\int_0^1 x(F) dF} \int_0^1 I(x \leq y_r) x dF,$$

or equivalently,

$$L(p) = \frac{1}{\int_0^1 k(1-F)^{\frac{-1}{a}} dF} \int_0^1 I(k(1-F)^{\frac{-1}{a}} \leq y_r) k(1-F)^{\frac{-1}{a}} dF,$$

$$L(p) = \frac{1}{\int_0^1 \left( (1-F)^{\frac{-1}{a}} \right) dF} \int_0^1 I\left( F \leq 1 - \left( \frac{k}{y_r} \right)^a \right) \left( (1-F)^{\frac{-1}{a}} \right) dF,$$

$$L(p) = \frac{1}{\int_0^1 \left( (1-F)^{\frac{-1}{a}} \right) F^0 dF} \int_0^{1 - \left( \frac{k}{y_r} \right)^a} \left( (1-F)^{\frac{-1}{a}} \right) F^0 dF.$$

And by using Karl Pearson's incomplete beta function. We find

$$L(p) = \frac{I\left( \frac{-1}{a} + 1, 1; 1 - \left( \frac{k}{\beta\left( \frac{-1}{a} + 1, 1 \right)} \right)^a \right)}{\beta\left( \frac{-1}{a} + 1, 1 \right)}.$$

While Butler-McDonald measure can be driven under Pareto distribution as follows:

$$\begin{aligned} M_{BM}^r &= \frac{1}{\int_0^1 x^r dF(x)} \int_0^{y_r} x^r dF(x). \\ &= \frac{1}{\int_0^1 x^r dF} \int_0^1 I(x^r \leq y_r) x^r dF \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\int_0^1 (x(F))^r dF} \int_0^1 I(x^r \leq y_r) (x(F))^r dF \\
&= \frac{1}{\int_0^1 \left(k(1-F)^{\frac{-1}{a}}\right)^r dG} \int_0^1 I\left(\left(k(1-F)^{\frac{-1}{a}}\right)^r \leq y_r\right) \left(k(1-F)^{\frac{-1}{a}}\right)^r dF, \\
M_{BM}^r &= \frac{1}{\int_0^1 \left((1-F)^{\frac{-1}{a}}\right)^r dF} \int_0^1 I\left(F \leq 1 - \left(\frac{k^r}{y_r}\right)^{\frac{a}{r}}\right) \left((1-F)^{\frac{-1}{a}}\right)^r dF.
\end{aligned}$$

Thus,

$$M_{BM}^r = \frac{1}{\int_0^1 (1-F)^{\frac{-r}{a}} F^0 dF} \int_0^{1 - \left(\frac{k^r}{y_r}\right)^{\frac{a}{r}}} (1-F)^{\frac{-r}{a}} F^0 dF.$$

In addition, by using Karl Pearson's incomplete beta function. We find that

$$M_{BM}^r = \frac{I\left(\frac{-r}{a} + 1, 1; 1 - \left(\frac{k^r}{\beta\left(\frac{-r}{a} + 1, 1\right)}\right)^{\frac{a}{r}}\right)}{\beta\left(\frac{-r}{a} + 1, 1\right)}.$$

### 5. Empirical Application: Income and Earnings Inequality in Egypt 2009 and 2011

Using income data from the survey of national income and individuals incomes in Egypt in 2009 and 2011, we consider the applicability to compute and compare among Butler-McDonald family. The data set was obtained from the Central Agency for Public mobilization and statistics who takes a sample to study incomes and

consumptions of individuals every two years regularly. The sample used in the survey is the simple random sample from overall Egypt in 2009 and 2011. This sample represents all Governorates and the Central Agency approximately gave us the data of incomes and consumptions of 12000 individuals of the survey in 2009 and gave us the data of incomes and consumptions of 3600 individuals of the survey in 2011. Also, these data contain all details about incomes and consumptions of the individuals. For example, the individual income consists of salary and wages, agrarian income, and inherited income etc. Also, the individual consumption consists of the expenditure on cloths, food, and drinks etc. The inequality measures will be computed for total income and total consumption. Therefore, the computation of inequality measures will be for a sample consists of 3000 individual for total income and total consumption in 2009 and 2011.

We have estimated income inequality by using Butler-McDonald measures in 2009 and 2011 in Table 1. Also, we used these measures to estimate the consumption inequality for individual's consumption in 2009 and 2011 in Table 2.

**Table 1.** Income inequality of 2009 and 2011

$r$	$M_{BM}^r$ in 2009	$M_{BM}^r$ in 2011
1	0.4266285	0.4423661
2	0.2892503	0.3154167
3	0.1435348	0.1331508
4	0.06090147	0.03773861
5	0.02488659	0.01300343

**Table 2.** Consumption inequality of 2009 and 2011

$r$	$M_{BM}^r$ in 2009	$M_{BM}^r$ in 2011
1	0.4357518	0.4512436
2	0.2933374	0.3234504
3	0.1481554	0.1237589
4	0.06480088	0.0345481
5	0.0288145	0.01096196

**A number of points are noteworthy from the results of Tables 1 and 2**

(1) To begin with, in Table 1, the sensitivity for Butler-McDonald measure decreases when  $r$ -th increase, which means that Lorenz curve is more sensitive than the higher order of Butler-McDonald measure. Where the income inequality is 0.4266285 when  $r = 1$ , but it begins to decrease when  $r = 2$  until reaches 0.02488659 when  $r = 5$ .

(2) The measures of income inequality, in Table 1, show that the income inequality ranges from 0.4266285 in 2009 and 0.4423661 at 2011, when  $r = 1$ , these values can be explained as a low inequality in the population.

(3) The measures values are range between 0 and 1, which mean that these measures satisfy the range axiom.

(4) The results in Table 2 confirm the results concluded in Table 1 about the sensitivity, range, and values of inequality.

## 6. Conclusion

The Lorenz curve can be defined as the relationship between the cumulative proportion of income units and the cumulative proportion of income received when units are arranged in ascending order of their income. Butler-McDonald [4] and [5] presented a general family of distributions for measuring inequality using incomplete moments. On the

other hand, Lorenz curve is obtained from Butler-McDonald family of inequality distributions, when  $r$  is equal one. In this article, we study Lorenz curve and Butler-McDonald measures under Pareto distribution. By using a real data application of Egyptian incomes and consumptions in 2009 and 2011 to compare among Lorenz curve and Butler-McDonald family, we found that Lorenz curve gives more accurate results for all kinds of inequality, such as income and consumption.

### References

- [1] R. Aaberge, Characterizations of Lorenz curves and income distributions, *Social Choice and Welfare* (forthcoming).
- [2] L. G. Bellù, *The Lorenz Curve*, EASYPOL, 2005.
- [3] L. G. Bellù and P. Liberati, Basic poverty measures, EASYPOL, 2005.
- [4] R. J. Butler and J. B. McDonald, Interdistributional income inequality, *J. of Business and Economic Statistics* 5 (1987), 13-18.
- [5] R. J. Butler and J. B. McDonald, Using incomplete moments to measure inequality, *J. of Econometrics* 18 (1989), 109-119.
- [6] F. A. Cowell, On the structure of additive inequality measures. *Rev. Econ. Stud.* 47 (1980), 521-531.
- [7] F. A. Cowell, *Measuring Inequality*, (2nd Edition), Harvester Wheatsheaf, Hamel Hempstead, 1995.
- [8] J. Daniel, *How to Generate a Lorenz Curve and Gini Coefficient*, Seattle Washington, 2009.
- [9] B. Frosini, Inequality measures for histograms, *Statistica* 65 (2005), 27-40.
- [10] J. Gastwirth, The estimation of the Lorenz curve and Gini index, *Rev. Econ. Stat.* 54 (1972), 306-316.
- [11] J. L. Gastwirth, A general definition of the Lorenz curve, *Econometrica* 39 (1971), 1037-1039.
- [12] N. Kakwani, *Income Inequality and Poverty*, Oxford University Press, London, 1980.
- [13] C. Lakner and B. Milanovic, Global income distribution: From the fall of the Berlin wall to the great recession, *World Bank Working Paper No 6719*, (2013).
- [14] P. J. Lambert and G. Lanza, The effect on inequality of changing one or two incomes, *Economics Department Working Paper, No. 2003, 15*, University of Oregon, (2003).
- [15] R. R. Schutz, On the measurement of income inequality, *American Economic Review* 41 (1951), 107-122.
- [16] A. Sen, *On Economic Inequality*, Clarendon, Oxford, UK, 1997.

- [17] M. I. Shahateet, How serious regional economic inequality in Jordan? Evidence from two national household surveys, *American Journal of Applied Sciences* 3 (2006), 1735-1744.
- [18] J. P. Smith and F. R. Welch, Racial discrimination- a human capital perspective, Unpublished paper delivered at the three Worlds of Labor Economics Conference, Snowbird, UT, October, 1985.
- [19] M. Zenga, Inequality curve and inequality index based on the ratios between lower and upper arithmetic means, *Stat. Appl.* 6 (2007), 137-151.
- [20] B. Zheng and J. Formby, Inequality Orderings, Stochastic Dominance and Statistical Inference, *Journal Bus. Econ. Stat.* 18 (2000), 490-488.

