# AN IMPLIED INCOME INEQUALITY INDEX USING $L_1$ NORM ESTIMATION OF LORENZ CURVE

Hamid Shahrestani, Ohio University-Chillicothe
Bijan Bidabad, Islamic Azad University, Science and Research Branch, Tehran

**ABSTRACT**

*Distribution of income is among the most important issues in welfare economics. Economic literature provides different ways to measure income inequality. Most common inequality indices provide information about the income distribution and analyze the inequality of income allocation without any reference to the amount of money needed to reduce the income inequality.In this paper, we design a model to estimate the Lorenz curve function parameters. By this approach, any census summary data can be used to measure the distribution of income. We also introduce a new fiscal-compensation-based index for reduction of the degree of inequality. Using this index, we show how much transfer payment is needed to achieve the desired distribution of income consistent with the perceived economic goals of the society.*

**JEL:** C63

**KEYWORDS:** Income distribution, Inequality index, Lorenz curve.

## INTRODUCTION

Estimation of the Lorenz curve is a challenge and is coupled with some difficulties. To estimate, first we need to define an appropriate functional form that can accept different curvatures. Moreover, to generate the necessary data set for estimation of the corresponding parameters, a large scale of computation of sample income data is inevitable. In section (A) of this paper, we introduce a shortcut and use the probability density function of population income to estimate the Lorenz function parameters. We develop the continuous $L_1$ norm smoothing method to estimate the regression parameters. We use two different probability density functions: (a) Pareto density distribution function that is integrable and (b) log-normal function that is more suitable for a wider range of income but is not integrable. Most inequality indices concentrate on the statistical aspect of the distribution of income. That is, they generally analyze the distribution without inferring the amount of funds needed to correct the income inequality. In section (B) of this paper, we introduce an implied inequality index. We identify an income inequality index to show how much of a transfer payment is needed to achieve the desired distribution of income consistent with the perceived economic equality goals of the policy-maker.

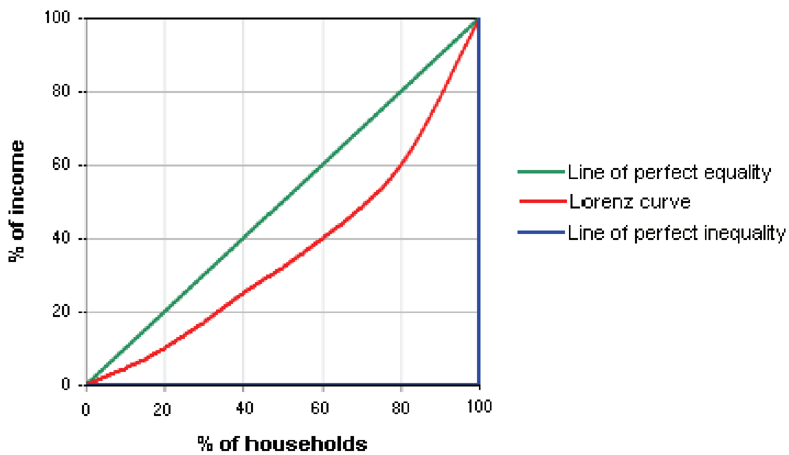
## REVIEW OF LITERATURE

The degree of income inequality can be illustrated with a Lorenz curve. The Lorenz curve is a graphical representation of the cumulative income distribution function. It shows what portion of the total income **y** is received by the bottom percentages of households. The percentages of households are plotted on the horizontal axis, and the percentages of earned incomes are plotted on the vertical axis, as Figure 1 shows.

The income inequality shown by the Lorenz curve can be measured by the Gini ratio. This coefficient is a measure to express distribution inequality. It is defined as a ratio between 0 and 1. Its numerator is the area between the Lorenz curve and the diagonal line, which represents the uniform income distribution. The denominator is the area under the uniform distribution line. Since the skewness of income distribution is persistently exhibited for different populations, the Lorenz curve becomes a method to analyze the skew distributions. There is also a relation between the area under the Lorenz curve and the corresponding probability distribution function. It is discussed that Pearsonian family distributions are rival functions to explain income distribution. See, Kendall and Stuart (1977).

When the probability distribution function is known, we may find the corresponding Lorenz curve as well as the Gini coefficient. Estimation of the Lorenz curve is associated with some difficulties. For this estimation, we should define an appropriate functional form that can accept different curvatures. In this regard, we need to have access to a large data set. Even though estimation of the corresponding parameters requires an extensive computational procedure, nevertheless, it leads to a low degree of significance for the estimated parameters. See, Bidabad and Bidabad (1989).

Figure 1: Lorenz Curve




In section (A), we use the probability density function of population income to estimate the Lorenz function parameters and concentrate on two rival probability density functions of Pareto and log-normal. The Pareto density function is integrable, and from this, we may uniquely derive the corresponding Lorenz function. However, the log-normal density function, which is more suitable for a wider income range than the Pareto distribution and is a better fit for higher income level -- see, Cramer (1973), Singh and Maddala (1976), Salem and Mount (1974) -- is not integrable. In this case, we need to define a general Lorenz curve functional form and apply the $L_1$ norm smoothing to estimate the relevant parameters. Using the approximation method for the discrete case, Bidabad (1989) has solved the $L_1$ norm estimation problems of linear one and two parameter models. He has also proposed certain directions for a continuous smoothing case. In our paper, this method is applied to estimate the Lorenz functional form introduced by Gupta (1984) and Bidabad and Bidabad (1989).

To have a solid understanding of guiding principles of income distribution, it is not enough to know the Lorenz curve and be familiar with the conventional inequality indices. The redistribution policies must enumerate specific budget guidelines to promote a more equal income distribution in the society. Economic literature provides different ways to measure income inequality (Atkinson, 1970; Cowell, 1977; Sen, 1973). Some of the most commonly used measures include: the Gini coefficient; the decile ratio; the proportions of total income earned by the bottom 50%, 60%, and 70% of households; the Robin Hood index; the Atkinson index; and Theil's entropy measure. The Gini is calculated as the ratio of the area between the Lorenz curve and the 45° line, to the whole area below the 45° line. Kakwani (1980) is a recalculation of the Gini coefficient and measures the length of the Lorenz curve. The Robin Hood index is equivalent to the maximum vertical distance between the Lorenz curve and the line of equal incomes. The Atkinson (1970) index is one of the few inequality measures that explicitly incorporate normative judgments about social welfare. It is derived by calculating the so-called equity-sensitive average income, which is defined as that level of per capita income, which if enjoyed by everyone, would make total welfare equal to the total welfare generated by the actual income distribution. Theil's (1967) entropy measure is derived from the notion of entropy in information theory.

Obviously, there is no single "best" measure of the degree of income inequality. Some measures (e.g., the Atkinson index) are more bottom-sensitive than others, i.e., more strongly correlated with the extent of poverty. The measures perform differently under various types of income transfers. For instance, the Gini is much less sensitive to income transfers between households if they lie near the middle of the income distribution compared to the tails. The Robin Hood index is insensitive with respect to income transfers between households on the same side of the mean income. While there are alternative methods, there is no best way to calculate the inequality index, especially for fiscal purposes. That is, they generally analyze the distribution without inferring the amount of funds needed to correct income inequality.

In section (B), we introduce an implied inequality index, which satisfies the needs of the above mentioned policy implications. We identify an index that can be used to reduce the degree of inequality. We show how to use any income or census summary data (i.e., the average and the median income) to measure the distribution of income and calculate the amount of money needed to be levied on the upper income group and then transferred to the lower income group to promote a better income distribution within the society.

## DEVELOPMENT OF CONTINUOUS L1 NORM FUNCTIONS

In this section, we develop the continuous $L_1$ norm smoothing method to estimate the regression parameters. First, we explain linear *one* and *two* parameter models respectively. Next, we estimate the Lorenz curve parameters. To estimate; we use two different probability density functions: (a) Pareto density distribution function that is integrable and (b) log-normal function that is more suitable for a wider range of income but is not integrable. Finally, we have a hypothetical numerical example for calculation of the parameters of the Lorenz curve.

Generally, $L_p$ norm of a function f(x) (see, Rice and White, 1964) is defined by,

$$\|f(x)\|_p == \int_{x\in I}\left(|f(x)|^P dx\right)^{\frac{1}{p}} \qquad (1)$$

Where "I" is a closed bounded set. The $L_1$ norm of f(x) is simply written as,

$$\|f(x)\|_1 = \int_{x\varepsilon I} |f(x)|dx \qquad (2)$$

Suppose the non-stochastic function $f(x,\boldsymbol{\beta})$ and the stochastic disturbance term <u>u</u> form y(x) as follows:

$$y(x) = f(x, \boldsymbol{\beta}) + u \qquad (3)$$

Where $\boldsymbol{\beta}$ is unknown parameters vector. Rewriting <u>u</u> as the residual of $y(x)-f(x,\boldsymbol{\beta})$, for $L_1$ norm approximation of "$\boldsymbol{\beta}$", we should find "$\boldsymbol{\beta}$" vector such that the $L_1$ norm of "u" is minimum. That is,

$$\text{Min: } S= \|u\|_1 = \|y(x)-f(x,\boldsymbol{\beta})\|_1 = \int_{x\varepsilon I} |y(x)-f(x,\boldsymbol{\beta})|dx \qquad (4)$$

**B**

<u>Liniar One Parameter Continuous L1 Norm Smoothing</u>

Redefine $f(x,\boldsymbol{\beta})$ as $\beta x$ and y(x) as the following linear function,

$$y(x) = \beta x + u \qquad (5)$$

Where "**β**" is a single (non-vector) parameter. Expression (4) reduces to:

$$\underset{\beta}{\text{Min:}}\ S = \|u\|_1 = \|y(x) - \beta x\|_1 = \int_{x \varepsilon I} |y(x) - \beta x| dx \tag{6}$$

The discrete analogue of (6) is solved by Bidabad (1988, 1989). In those papers, he proposed applying discrete and regular derivatives to the discrete problem by using a slack variable "t" as a point to distinguish negative and positive residuals. A similar approach is used here to minimize (6). To do so in this case, certain Lipschitz conditions are imposed on the functions involved (see, Usow, 1967a). Rewrite (6) as follows:

$$\underset{\beta}{\text{Min:}}\ S = \int_{x \varepsilon I} |x|\ |y(x)/x - \beta| dx \tag{7}$$

Let us define "I" as a closed interval [0,1]. The procedure may be applied to other intervals with no major problem (see, Hobby and Rice, 1965, Kripke and Rivlin, 1965, Usow, 1967a). To minimize this function we should first remove the absolute value sign of the expression after the integral sign. Since "x" belongs to a closed interval "I", both functions, y(x) (which is a linear function of "x") and y(x)/x are smooth and continuous. And since y(x)/x is a uniformly increasing or decreasing function of "x", a value of tεI can be found to have the following properties:

$$\begin{aligned} y(x)/x &< \beta \quad \text{if } x < t \\ y(x)/x &= \beta \quad \text{if } x = t \\ y(x)/x &> \beta \quad \text{if } x > t \end{aligned} \tag{8}$$

The value of the slack variable "t" actually is the border of negative and positive residuals. If the value of "t" were known, when x=t, we could calculate the optimal value of "β". Nevertheless, neither "t" nor "β" are known. To solve, according to (8), we can rewrite (7) as two separate definite integrals with different upper and lower bounds.

$$\underset{\beta}{Min}\ S = -\int_0^t |x|\ (y(x)/x - \beta)dx + \int_t^1 |x|\ (y(x)/x - \beta)dx \tag{9}$$

Decomposition of (7) into (8) has been done by use of the slack variable "t". Since both "β" and "t" are unknown, to solve (9), we partially differentiate it with respect to "t" and "β".

$$\frac{\partial S}{\partial \beta} = \int_0^t |x|dx - \int_t^1 |x|dx = 0 \tag{10}$$

And using the Liebniz rule to differentiate the integrals with respect to their variable bounds "t", yields,

$$\frac{\partial S}{\partial t} = -|t|\left[\frac{y(t)}{t} - \beta\right] - |t|\left[\frac{y(t)}{t} - \beta\right] = 0 \tag{11}$$

Since "x" belongs to [0,1], equation (10) can be written as,

$$\int_0^t x dx - \int_t^1 x dx = 0 \tag{12}$$

Or,

$$\tfrac{1}{2} t^2 - \tfrac{1}{2} + \tfrac{1}{2} t^2 = 0 \tag{13}$$

This yields,

$$t = \sqrt{2}/2 \tag{14}$$

Substituting for "t" in equation (11), yields,

$$\beta = \frac{y(\sqrt{2}/2)}{\sqrt{2}/2} \quad (15)$$

Liniar Two Parameters Continuous L1 Norm Smoothing

Given that y(t) is function y(x) evaluated at x=t. Value of "β" given by (15) is the optimal solution of (6). The above procedure is in fact a generalization of the Laplace weighted median for continuous case. Before applying this to the Lorenz curve, let us develop the procedure for the linear two parameters model.

To apply the above technique to the linear two parameters model, rewrite (4) as,

$$\underset{\alpha,\beta}{\text{Min}}: S=\|u\|_1=\|y(x)-\alpha-\beta x\|_1=\int_{x\varepsilon I} |y(x)-\alpha-\beta x|dx \quad (16)$$

Where, "α" and "β" are two single (non-vector) unknown parameters, and y(x) and "x" are as before. According to Rice (1964c), let f(α*,β*,x) interpolates y(x) at the set of canonical points $\{x_i;i=1,2\}$, if y(x) is such that: y(x)-f(α*,β*,x) changes sign at these $x_i$'s and at no other points in [0,1], then f(α*,β*,x) is the best $L_1$ norm approximation to y(x) (see also, Usow, 1967a). With the help of this rule, if we denote these two points to $t_1$ and $t_2$, we can rewrite (16) for I=[0,1] as,

$$S=\int_0^{t_1} [y(x)-\alpha-\beta x]dx - \int_{t_1}^{t_2} [y(x)-\alpha-\beta x]\, dx + \int_{t_2}^{1} [y(x)-\alpha-\beta x]\, dx \quad (17)$$

Since $t_1$ and $t_2$ are also unknowns, we should minimize S with respect to α, β, $t_1$ and $t_2$. Taking a partial derivative of (17) using Liebniz' rule with respect to these variables and equating them to zero, we will have,

$$\frac{\partial S}{\partial \alpha} = -\int_0^{t_1} dx + \int_{t_1}^{t_2} dx - \int_{t_2}^{t_1} dx = 0 \quad (18)$$

$$\frac{\partial S}{\partial \beta} = -\int_0^{t_1} dx + \int_{t_1}^{t_2} dx - \int_{t_2}^{t_1} dx = 0 \quad (19)$$

$$\frac{\partial S}{\partial t_1} = 2[y(t_1) - \alpha - \beta t_1] = 0 \quad (20)$$

$$\frac{\partial S}{\partial t_2} = -2[y(t_2) - \alpha - \beta t_2] = 0 \quad (21)$$

Equations (18) through (21) may be solved simultaneously for α, β, $t_1$ and $t_2$. Thus, we have the following system of equations,

$$2t_2 - 2t_1 - 1 = 0 \quad (22)$$
$$t_2^2 - t_1^2 - ½ = 0 \quad (23)$$
$$y(t_1) - \alpha - \beta t_1 = 0 \quad (24)$$
$$y(t_2) - \alpha - \beta t_2 = 0 \quad (25)$$

The solutions are:

$$t_1=1/4 \quad (26)$$
$$t_2=3/4 \quad (27)$$
$$\alpha = y(3/4)-(3/4)\beta = y(1/4)-(1/4)\beta \quad (28)$$
$$\beta = 2[y(3/4)-y(1/4)] \quad (29)$$

This procedure may be expanded to include "m" unknown parameters. Some computational methods for solving the different cases of "m" parameters models are investigated by Ptak (1958), Rice and White (1964), Rice (1964a, 1964b, 1964c, 1969, 1985), Usow (1967a), Lazarski (1975a, 1975b, 1975c, 1977) (see also, Hobby and Rice, 1965, Kripke and Rivlin, 1965, Watson, 1981). Now, let us have a look at the Lorenz curve and its proposed functional forms.

Continuous $L_1$ Norm Smoothing Of Lorenz Curve

The Lorenz curve for a random variable with probability density function f(v) may be defined as an ordered pair.

$$(P(V|V \le v), \frac{E(V|V \le v)}{E(V)}) \qquad v \in R \tag{30}$$

For a continuous density function f(v), (30) can be written as,

$$(\int_{-\infty}^{v} f(w)dw, \frac{\int_{-\infty}^{v} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw}) \equiv (x(v), y(x(v))) \tag{31}$$

Taguchi (1972a, 1972b, 1972c, 1973, 1981, 1983, 1987, and 1988) multiplies the second element of (30) by P(V|V≤v); his definition of (31) is equivalent to ours. We denote (31) by the ordered pair (x(v),y(x(v))) where x(v) and y(x(v)) are its elements. Next, "x" is a function which maps "v" to x(v) and "y" is a function which maps x(v) to y(x(v)). The function y(x(v)) is simply the Lorenz curve function. For the explicit function for the Lorenz curve, we use the form introduced by Gupta (1984) and a modified version, which benefits from certain properties. Gupta (1984) proposed the functional form,

$$y = xA^{x-1} \qquad A>1 \tag{32}$$

The modified version of Bidabad and Bidabad (1989) suggests the following functional form:

$$y = x^B A^{x-1} \qquad B \ge 1, A \ge 1 \tag{33}$$

To estimate "A" of (32) or "A" and "B" of (33), we need discrete data from the population to construct x and **y** vectors. On the other hand, if the probability distribution of income is known, we can estimate the Lorenz curve by using the continuous $L_1$ norm smoothing method for continuous functions. To estimate the Lorenz curve parameters when the probability density function of income is known and integrable, we can find the functional relationship between the two elements of (31) by simple mathematical derivation. However, when integrals of (31) are not obtainable, we will follow another procedure.

Suppose that the income of a society is distributed using the probability density function f(w). This density function may be a skewed function, such as Pareto or log-normal, as follows:

$$f(w) = \theta k^{\theta} w^{-\theta-1}, \qquad w,k>0,\ \theta>0 \tag{34}$$

$$f(w) = [1/w\sigma\sqrt{(2\pi)}]\exp\{-[\ln(w)-\mu]^2/2\sigma^2\}, \qquad w\varepsilon(0,\infty),\ \mu\varepsilon(-\infty,+\infty),\ \sigma>0 \tag{35}$$

These two distributions are known as good candidates for representing distribution of personal income. In the case of the Pareto density function of (34), we can simply derive the Lorenz curve function as follows:

Let F(w) denote the Pareto distribution function:

$$F(w) = 1-(k/w)^{\theta} \tag{36}$$

With mean equal to,

$$E(w) = \theta^k/(\theta-1), \quad \theta>1 \tag{37}$$

If we find the function y as stated by (31) as a function of x, the Lorenz function will be derived. Rearrange the terms of (31) as,

$$x(v) = \int_{-\infty}^{v} f(w)dw \quad (38)$$

$$y(x(v)) = [1/E(W)]\int_{-\infty}^{v} wf(w)dw \quad (39)$$

Substituting Pareto distribution function,

$$x(v) = F(v) = 1-(k/v)^{\theta} \quad (40)$$

$$y(x(v)) = [(\theta-1)/\theta^{k}] \int_{k}^{v} w\theta k^{\theta} w^{-\theta-1} dw \quad (41)$$

Or $y(x(v)) = 1-(k/v)^{\theta-1}$ (42)

By solving (40) for "v" and substituting in (42), the Lorenz curve for Pareto distribution is derived as,

$$y = 1-(1-x)^{(\theta-1)/\theta} \quad (43)$$

For log-normal distribution, we proceed as follows:
According to (30) and (31) independent and dependent variables of (32) and (33) may be written as,

$$x(v) = \int_{o}^{v} f(w)dw \quad (44)$$

$$y(x(v)) = [1/E(W)]\int_{0}^{v} wf(w)dw \quad (45)$$

Substitute (44) and (45) in (32) and include the random error term u, we will have,

$$[1/E(w)]\int_{0}^{v} wf(w)dw = \int_{0}^{v} f(w)dw.\, A^{\int_{0}^{v} f(w)dw-1}.e^{u} \quad (46)$$

Or,

$$y(x) = x\, A^{x-1} e^{u} \quad (47)$$

Similarly for the model (35),

$$[1/E(w)]\int_{0}^{v} wf(w)dw = \left\{\int_{0}^{v} f(w)dw\right\}^{B}.\, A^{\int_{0}^{v} f(w)dw-1}.e^{u} \quad (48)$$

Or,

$$y(x) = x^{B} A^{x-1} e^{u} \quad (49)$$

Taking the natural logarithm of (47) and (49), gives,

$\ln y(x) = \ln x + (x-1)\ln A + u$ (50)

$\ln y(x) = B.\ln x + (x-1)\ln A + u$ (51)

With respect to the properties of a Lorenz curve and the probability density function of f(w) and equations (46) to (49), it can be seen that x belongs to the interval [0,1]. Thus, the $L_1$ norm objective function for minimizing (50) or (51) is given by

$$\underset{A}{Min}: S = \int_{0}^{1} |u| dx \quad (52)$$

Or,

$$\underset{A}{Min}: S = \int_{0}^{1} |\ln y(x) - \ln x - (x-1)\ln A| dx \quad (53)$$

Or,

$$\underset{A}{Min}: S = \int_{0}^{1} |x-1| \; |[\ln y(x) - \ln]/(x-1) - \ln A| dx \quad (54)$$

By a technique similar to the one used by (9), we can rewrite (54) as,

$$\underset{A}{Min}: S = \int_0^t |x-1| \{[\ln y(x) - \ln x]/(x-1) - \ln A\} dx - \int_t^1 |x-1| \{[\ln y(x) - \ln]/(x-1) - \ln A\} dx \quad (55)$$

Since, 0≤x≤1 we have

$$\underset{A}{Min}: S = \int_0^1 [\ln y(x) - \ln x - (x-1)\ln A] dx + \int_t^1 [\ln y(x) - \ln x - (x-1)\ln A] dx \quad (56)$$

Differentiate (56) partially with respect to "t" and "A"

$$\frac{\partial S}{\partial A} = +\int_0^1 [(x-1)/A] dx - \int_t^1 [(x-1)/A] dx = 0 \quad (57)$$

$$\frac{\partial S}{\partial t} = -2[\ln y(t) - \ln t - (t-1)\ln A] = 0 \quad (58)$$

From equation (57) we have,

$$t = 1 \pm \sqrt{2}/2 \quad (59)$$

Since "t" should belong to the interval [0,1], we accept

$$t = 1 - \sqrt{2}/2 \quad (60)$$

Substitute (60) in (58), and solve for "A", gives the $L_1$ norm estimation for "A" equal to

$$A = \left[ \frac{1-\sqrt{2}/2}{y(1-\sqrt{2}/2)} \right]^{\sqrt{2}} \quad (61)$$

Now, let us apply this procedure to another Lorenz curve functional form of (33), as redefined by (51). Rewrite the L1 norm objective function (52) for the model (51),

$$\underset{A,B}{Min}: S = \int_0^1 |\ln y(x) - B\ln x - (x-1)\ln A| dx \quad (62)$$

Or,

$$\underset{A,B}{Min}: S = \int_0^1 |x-1| |[\ln y(x)]/(x-1) - (\ln x)/(x-1) - \ln A| dx \quad (63)$$

The objective function (63) is similar to (16). Thus, by a similar procedure to those of (17) through (29), we can write "S" as,

$$\begin{aligned} \underset{A,B}{Min}: S &= \int_0^{t_1} |x-1| \{[\ln y(x)]/(x-1) - (\ln x)/(x-1)\ln A\} dx \\ &- \int_{t_1}^{t_2} |x-1| \{[\ln y(x)]/(x-1) - (\ln x)/(x-1) - \ln A\} dx \\ &+ \int_{t_1}^{1} |x-1| \{[ly(x)]/(x-1) - (\ln x)/(x-1) - \ln A\} dx \end{aligned} \quad (64)$$

Since 0≤x≤1, (64) reduces to

$$\begin{aligned} \underset{A,B}{Min}: S &= -\int_0^{t_1} [\ln y(x) - B\ln x - (x-1)\ln A] dx + \int_{t_1}^{t_2} [\ln y(x) - B\ln x - (x-1)\ln A] dx \\ &- \int_{t_2}^{1} [\ln y(x) - B\ln - (x-1)\ln A] dx \end{aligned} \quad (65)$$

Differentiate "S" partially with respect to "A", "B", $t_1$ and $t_2$,

$$\frac{\partial S}{\partial A} = \frac{1}{A}\left[ \int_0^{t_1} (x-1)dx - \int_{t_1}^{t_2} (x-1)dx + \int_{t_2}^{1} (x-1)dx \right] = 0 \quad (66)$$

$$\frac{\partial S}{\partial B} = \int_0^{t_1} \ln(x)dx - \int_{t_1}^{t_2} \ln(x)dx + \int_{t_2}^{1} \ln(x)dx = 0 \quad (67)$$

$$\frac{\partial S}{\partial t_1} = -2\{\ln[y(t_1)] - B\ln(t_1) - (t_1 - 1)\ln(A)\} = 0 \quad (68)$$

$$\frac{\partial S}{t_2} = 2\{\ln[y(t_2)] - B\ln(t_2) - (t_2 - 1)\ln(A)\} = 0 \quad (69)$$

The above system of simultaneous equations can be solved for the unknowns: $t_1$, $t_2$, "A" and "B". Equation (66) is reduced to,

$$t_1^2 - t_2^2 - 2(t_1 - t_2) - 1/2 = 0 \quad (70)$$

Equation (67) can be written as,

$$t_1(\ln t_1 - 1) - t_2(\ln t_2 - 1) - 1/2 = 0 \quad (71)$$

Calculate $t_1$ from (70) as,

$$t_1 = 1 \pm \sqrt{(t_2^2 - 2t_2 + 3/2)} \quad (72)$$

Since $0 \le t_1 \le 1$ we accept,

$$t_1 = 1 - \sqrt{(t_2^2 - 2t_2 + 3/2} \quad (73)$$

Substitute $t_1$ from (73) into (71), and rearrange the terms, gives,

$$\ln \frac{\left[1 - \sqrt{(t_2^2 - 2t_2 + 3/2)}\right]^{\left[1 - \sqrt{(t_2^2 - 2t_2 + 3/2)}\right]}}{t_2^{t_2}} + t_2 - 3/2 + \sqrt{(t_2^2 - 2t_2 + 3/2)} = 0 \quad (74)$$

We can compute the root of equation (74) by a numerical algorithm. For five digits decimal point, we have:

$$t_2 = 0.40442 \quad (75)$$

Value of $t_1$ is derived by substituting $t_2$ into (73),

$$t_1 = 0.07549 \quad (76)$$

Values of "B" and "A" are computed from (68) and (69) using $t_2$ and $t_1$ given by (75) and (76). Thus,

$$B = \frac{(t_2 - 1)\ln y(t_1) - (t_1 - 1)\ln y(t_2)}{(t_2 - 1)\ln(t_1) - (t_1 - 1)\ln(t_2)} \quad (77)$$

Or,

$$B = -0.84857 \ln[y(0.07549)] + 1.31722 \ln[y(0.40442)] \quad (78)$$

And,

$$A = [y(0.07549)]^{1.28986}[y(0.40442)]^{-3.68126} \quad (79)$$

Now, let us describe how equation (61) for the model (32) and equations (78) and (79) for the model (33) can be used to estimate the parameters of the Lorenz curve when the probability distribution function is known. For the model (32) we should solve (44) for "v" such that,

$$x(v) = \int_0^v wf(w)dw = 1-\sqrt{2}/2 = 0.29293 \tag{80}$$

By substituting this value of "v" into (45), value of $y(1-\sqrt{2}/2)$ is computed. This value is used to compute the parameter "A" given by (61) for model (32).
Similarly, for the model (33) we will find two values for "v" by solving:

$$x(v) = \int_0^v wf(w)dw = 0.07549 \tag{81}$$

&

$$x(v) = \int_0^v wf(w)dw = 0.40442 \tag{82}$$

These values of "v" are substituted in (45) to find y(0.07549) and y(0.40442). These values of "y" are used to compute the parameters of the model (33) by substituting them into (78) and (79).The computation of related definite integrals of x(v) defined by (80), (81) and (82), can be done by using the appropriate numerical methods.

Numerical Example

Suppose the sample mean and median of income distribution of the society are given. For calculation of the parameters of the Lorenz curve, the following notations have been coded for MathCAD 11.
Assume that the sample mean of income distribution of the society is: $60,000.
Assume that the sample median of income distribution of the society is: $40,000.

The standard deviation can be calculated as: $\sigma := \sqrt{2\cdot\ln\left(\frac{\text{Mean}}{\text{Med}}\right)}$

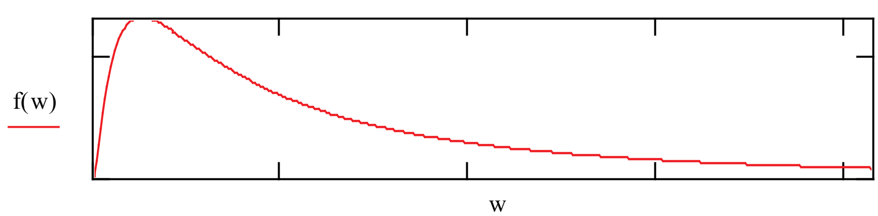
And, $\mu = \ln(\text{Med})$ such that, $\mu = 10.5966$, $\sigma = .9005$

Calculation of log-normal density function parameters based on sample mean and median is

log-normal Probability Density Function: $f(w) := \left(\frac{1}{w\cdot\sigma\cdot\sqrt{2\cdot\pi}}\right)\cdot\exp\left[\frac{-(\ln(w)-\mu)^2}{2\cdot\sigma^2}\right]$

For selective range for log-normal plot: $w := 10^{-5}, \frac{\text{Mean}}{200} .. 2\cdot\text{Mean}$

Figure 2: log-normal plot




Precision Tolerance level TOL :=0.00001

TOL value can be changed for more accurate solutions (less TOL = higher precision).

For equation (45) we have: $y(v) := \left(\frac{1}{\text{Mean}}\right)\cdot\int_0^v w\cdot f(w)\,dw$

For equation (44) we have: $x(v) := \int_{0.00001}^{v} f(w)\,dw$

Calculation for Gupta Model:

Initial guess for v: $v := 20000$ It might be changed for faster convergence and less iterations.

For (60): $t_0 := 1 - \frac{\sqrt{2}}{2}$

Calculating v for (80): $v := \mathrm{root}(x(v) - t_0, v)$, $v = 27136.6437$

y(t) $y(v) = 0.04208$ $z_0 := y(v)$

For (61): estimated A: $A := \left(\frac{t_0}{z_0}\right)^{\sqrt{2}}$ $A = 15.54768$

For (53): $S := \int_0^1 \left|\ln(z_0) - \ln(t_0) - (t_0 - 1)\cdot\ln(A)\right|\,dx$

Sum of absolute residuals: $S = 0$

Range of variable for plotting the Lorenz curves: $X := 0, 0.005 .. 1$

Gupta Lorenz curve: $Y(X) := X\cdot A^{X-1}$

Calculation of Gini coefficient: $Gini := 1 - 2\cdot\int_0^1 Y(X)\,dX$ $Gini = 0.51967$

Calculation for Bidabad Model:

For (76): $t_1 := 0.07549$

Initial guess for v: $v := 8000$ It might be changed for faster convergence and less iterations.

Calculating v for (81): $v := \mathrm{root}(x(v) - t_1, v)$ $v = 9464.04318$

y(0.07549) $y(v) = 0.00442$ $z_1 := y(v)$

For (75): $t_2 := 0.40442$

Initial guess for v: $v := 27000$ It might be changed for faster convergence and less iterations.

Calculating v for (82): $v := \mathrm{root}(x(v) - t_2, v)$ $v = 38826.25803$

y(0.40442): $y(v) = 0.07722$ $z_2 := y(v)$

For (79): $A := (z_1)^{1.28986}\cdot(z_2)^{-3.68126}$

For (78): $B := -0.84857\cdot\ln(z_1) + 1.31722\cdot\ln(z_2)$

Estimated A and B: $A = 11.41481$ $B = 1.22709$

For (62): $S := \int_0^1 \left|\ln(z_1) - B\cdot\ln(t_1) - (t_1 - 1)\cdot\ln(A)\right|\,dx$

Sum of absolute residuals: $S = 0.00002$

Range of variable for plotting the Lorenz curves: $X := 0, 0.005 .. 1$

Modified Lorenz curve: $Y(X) := X^{B} \cdot A^{X-1}$

Calculation of Gini coefficient: $Gini := 1 - 2 \cdot \int_0^1 Y(X)\,dX$ $Gini = 0.51834$

## INTRODUCTION OF IMPLIED-INEQUALITY-INDEX

In this section we introduce a new income inequality index. The redistribution policies must enumerate specific budget guidelines to promote a more equal income distribution in the society. Using our calculated implied-inequality-index, we show how much of a transfer payment is needed to achieve the desired distribution of income consistent with the perceived economic equality goals of the policy-maker. While there are alternative methods; there is no best way to calculate the inequality index .Most inequality indices concentrate on the statistical aspect of the distribution of income. That is, they generally analyze the distribution without inferring the amount of funds needed to correct the income inequality. First, we define our index. Second, we use a hypothetical numerical example to show how much money should be transferred from the upper income group to the lower income group to achieve the desired distribution of income.

Definition Of The Inequality Index

Suppose there is a personal income υ at which half of the total income of the population belongs to those who have an income less than υ, and the other half of the income belongs to those who have a higher income than υ. That is:

$$\int_{-\infty}^{v} wf(w)dw = \int_{v}^{+\infty} wf(w)dw \tag{83}$$

By definition, we have:

$$\mu = \int_{-\infty}^{+\infty} wf(w)dw = \int_{-\infty}^{v} wf(w)dw + \int_{v}^{+\infty} wf(w) \tag{84}$$

That is:

$$\int_{-\infty}^{v} wf(w)dw = \mu/2 \tag{85}$$

On the other hand:

$$\frac{\int_{-\infty}^{v} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} = 1/2 \tag{86}$$

According to (31) this is a point on the Lorenz curve with the following ordered pair:

$$(\int_{-\infty}^{v} wf(w)dw, 1/2) \tag{87}$$

Thus, we define implied-inequality-index (iii) as $\int_{-\infty}^{v} f(w)dw$ when υ satisfies (83). That is,

$$iii = \int_{-\infty}^{v} f(w)dw \qquad \text{when v satisfies} \qquad \frac{\int_{-\infty}^{v} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} = 1/2 \tag{88}$$

To find iii, (85) should be solved for υ and its value be replaced in (88). As iii approaches ½, distribution becomes more symmetric. If iii tends to 1, distribution tends to be fully right-skewed, indicating high (right) inequality and, as iii tends to 0, distribution tends to be left-skewed and distribution tends to high (left) inequality. The values of iii less than ½, however, have no economic implication for income distribution. Let us define the cost of equalization as:

$$C= [iii-½]\times N\times \mu \qquad (89)$$

The above expression means that to equalize the distribution of income without changing the average income of the society, the amount of C should be transferred from higher income earner to lower income earner, where N and μ are the population size and average income of the society. We may normalize this index by dividing the equalization cost by total income of the society and find an inter-societies comparable index. That is:

$$\text{Relative cost of equalization} = [(iii-½)\times N\times \mu] / (N\times \mu) = (iii-½) \qquad (90)$$

<u>Numerical Example</u>

To illustrate, the following table of income distribution for a hypothetical society is used. Consider a society of 400 households with total income of the society equal to $2,000, where 70% of the population (280 households) receives only half of the total income, and the remaining 30% (120 households) receive the other half. According to Table 1, we have:

| | |
|---|---|
| N = 400 | (Number of households) |
| $\upsilon = \mu = 2000/400 = 5$ | (Average income) |
| $\mu_{lower} = 1000/280 = 3.57$ | (Average income of lower category) |
| $\mu_{upper} = 1000/120 = 8.33$ | (Average income of upper category) |
| iii = 280/400 = 0.7 | (implied inequality index |
| $C = (0.7-0.5) \times 400\times 5 = \$400$ | (Cost of equalization) |

That is, if we collect a total tax of $400 from the top 30% of the population and transfer it to the lower 70% of income earners, the average income of both groups will be the same:

$$(1000+400)/280 = (1000-400)/120 = 5$$

Relative cost of equalization = 0.7-0.5 = 0.2

That is, the cost of such equalization is 20% of the total income of the society.

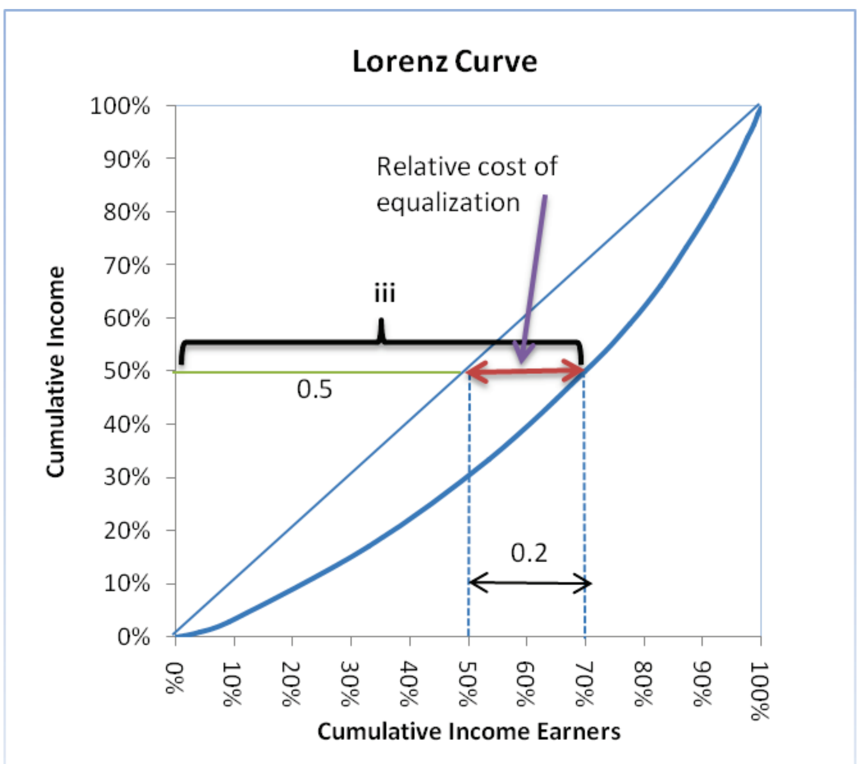
In Figure 3, we show the implied inequality index (iii) and the relative cost of equalization on the Lorenz curve. This is depicted by using columns (4) and (9) of Table 1.

Table 1: Income Distribution for a Hypothetical Society

| Income w | Frequency f | Cumulative Frequency F | Relative Frequency | Relative Cumulative Frequency | Half Income Earner | w . f (1)*(2) | Cumulative Income | Relative Cumulative Income | Half Income |
|---|---|---|---|---|---|---|---|---|---|
| ($) | (Numbers) | (Numbers) | (%) | (%) | (Numbers) | ($) | ($) | ($) | ($) |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| 1 | 17 | 17 | 4.3% | 4.3% | | 17 | 17 | 0.9% | |
| 2 | 20 | 37 | 5.0% | 9.3% | | 40 | 57 | 2.9% | |
| 3 | 95 | 132 | 23.8% | 33.0% | | 285 | 342 | 17.1% | |
| 4 | 82 | 214 | 20.5% | 53.5% | | 328 | 670 | 33.5% | |
| 5 | 66 | 280 | 16.5% | 70.0% | 280 | 330 | 1000 | 50.0% | 1000 |
| 6 | 30 | 310 | 7.5% | 77.5% | | 180 | 1180 | 59.0% | |
| 7 | 21 | 331 | 5.3% | 82.8% | | 147 | 1327 | 66.4% | |
| 8 | 18 | 349 | 4.5% | 87.3% | | 144 | 1471 | 73.6% | |
| 9 | 17 | 366 | 4.3% | 91.5% | | 153 | 1624 | 81.2% | |
| 10 | 14 | 380 | 3.5% | 95.0% | | 140 | 1764 | 88.2% | |
| 11 | 11 | 391 | 2.8% | 97.8% | | 121 | 1885 | 94.3% | |
| 12 | 4 | 395 | 1.0% | 98.8% | 120 | 48 | 1933 | 96.7% | |
| 13 | 3 | 398 | 0.8% | 99.5% | | 39 | 1972 | 98.6% | |
| 14 | 2 | 400 | 0.5% | 100.0% | | 28 | 2000 | 100.0% | 1000 |
| | 400 | | 100% | | 400 | 2000 | | | 2000 |

*In Table 1, column (1) depicts dollar values of income categories and column (2) shows the number or frequencies of households in each income category of column (1). Columns (3), (4) and (5) are for cumulative frequencies, relative frequencies and relative cumulative frequencies. Column (6) shows the number of lower and higher income earners. Column (7) shows the multiplication of the paired elements of the columns (1) and (2). Column (8) cumulates (7) and (9) shows the relative cumulative income. Column (10) shows half of the total income of the society.*

Figure 3: Implied Inequality Index iii




*This figure depicts the information of Table 1. The implied inequality index (iii) and relative cost of equalization are shown as corresponding parts of a Lorenz curv*

## CONCLUSION

Estimation of the Lorenz curve is fraught with difficulty. To circumvent this, we tried to estimate the functional form of the Lorenz curve by using continuous information. We employed the probability density function of population income to estimate the Lorenz function parameters by using the continuous L1 norm smoothing method. To have a better understanding of policy arrangements regarding the inequality of income distribution, it is not enough to know the conventional inequality indices. The redistribution policies need to deal with specific budget guidelines in order to lead society toward a position of greater equality. Obviously, there is no single "best" measure of income inequality. While there are alternative methods, there is no best way to calculate the inequality index, especially when concentrating on the fiscal view. That is, the existing methods generally analyze the distribution without inferring about the amount of funds needed to correct income inequality. In this paper, we introduced an implied inequality index, which satisfies these policy implications. We designed an implied inequality index as a fiscal guidepost to improve income distribution. In this paper, we only used the 50% point as the benchmark for our policy objective. This model can be extended to use quantiles or deciles points as equalization policy objectives. To do so, we need to derive the necessary formula for the model. These developments will improve the policy applications of the derived indices.

## REFERENCES

Atkinson, A. B. (1970) "On the Measurement of Inequality," *Journal of Economic Theory* vol. 2, p. 244-263.

Bidabad, B. (1988) "A Proposed Algorithm for Least Absolute Error Estimation & Multiple Regression," *Proceedings of the Seminar of Mathematical Analysis,* Shiraz University, Iran, p. 24-50

Bidabad, B. & B. (1989) "Functional Form for Estimating the Lorenz Curve*," The Australian Meeting of the Econometric Society,* Canberra, Australia.

Bidabad, B. (1989) "Discrete and Continuous $L_1$Norm Regressions," *Doctorate Dissertation*, Islamic Azad University, Tehran, Iran

Cowell, F.A. (1977) "Measuring Inequality," *Oxford: Philip Allan.*

Cramer, J.S. (1973) "Empirical Econometrics," *North-Holland, Amsterdam*

Gupta, M.R. (1984) "Functional forms for estimating the Lorenz curve*," Econometrica* vol. 52, p. 1313-1314

Hobby, C.R. Rice, J.R. (1965) *"*A Moment Problem in $L_1$ Approximation," *Proc. Amer. Math*. Soc., vol. 16, p. 665-670

Kakwani, N.C. (1980) "Income Inequality and Poverty," *New York, Oxford University Press.*

Kakwani, N.C. (1980) *"* Functional Forms for Estimating the Lorenz Curve: a reply," *Econometrica*, vol. 48, p.1063-64.

Kakwani, N.C. & Podder, N (1976) "Efficient Estimation of the Lorenz Curve and Associated Inequality Measures from Grouped Observations," *Econometrica*, vol. 44, p. 137-148

Kendall, M. & Stuart, A. (1977) “The Advanced Theory of Statistics,” *Charles Griffin &Co.,* vol. 1, London.

Kripke, B.R. & Rivlin, T.J. (1965) “Approximation in the Metric of $L_1$(X,u),” *Trans. Amer. Math. Soc.*, vol. 119, p. 101-22

Lazarski, E. (1975a) “Approximation of Continuous Functions in the Space $L_1$,” *Automatika*, vol. 487, p. 85-93

Lazarski, E. (1975b) “The Approximation of the Continuous Function by the Polynomials of Power Functions in $L_1$ Space,” *Automatika*, vol. 487, p. 95-106

Lazarski, E. (1975c) “On the Necessary Conditions of the Uniqueness of Approximation by the Polynomials of Power Functions in $L_1$ Space,” *Automatika*, vol. 487, p. 107-117

Lazarski, E. (1977) “Approximation of Continuous Functions by Exponential Polynomials in the $L_1$ Space,” *Automatika*, vol. 598, p. 82-87

Mills, Jeffrey & Sourush, A. (1997) “Statistical Inference via Bootstrapping for Measures of Inequality,” *Journal of Applied Economics,* vol. 12, p.133-150

Ptak, V. (1958) “On Approximation of Continuous Functions in the Metric $\int_a^b|x(t)|dt$,” *Czechoslovak Math.* Journal, vol. 8(83), p. 267-273

Rasche, R.H. Gaffney, J. Koo, A.Y.C. & Obst, N. (1980) “Functional Forms for Estimating the Lorenz Curve,” *Econometrica*, vol. 48, p. 1061-1062

Rice, J.R. (1964a) “On Computation of $L_1$ Approximations by Exponentials, Rationales, and Other Functions,” *Math Comp,* vol.18, p. 390-396

Rice, J.R. (1964b) “On Nonlinear $L_1$ Approximation,” *Arch. Rational Mech. Anal,* vol. 17, p. 61-66

Rice, J.R. (1964c) “The Approximation of Functions, vol. I, Linear Theory,” *Reading Mass: Addison-Wesley*

Rice, J.R. (1969) “The Approximation of Functions, vol. II, Linear Theory,” *Reading Mass: Addison-Wesley*

Rice, J.R. (1985) “Numerical Methods, Software, and Analysis,” *McGraw-Hill*

Rice, J.R., White, J.S. (1964) “Norms for Smoothing and Estimation,” *SIAM Rev.*, vol. 6, p. 243-256

Salem, A.B.Z. & Mount, T.D. (1974) “A Convenient Descriptive Model of Income Distribution: the Gamma Density,” *Econometrica,* vol. 42, p. 1115-1127

Sen A. (1973) "On Economic Inequality," *Oxford: Oxford University Press*

Singh, S.K. & Maddala, G.S. (1976) “A Function for the Size Distribution of Income,” *Econometrica*, vol.44, p.963-970

Slottje, D.J. (1989) "The Structure of Earnings and the Measurement of Income Inequality in the U.S.," *North-Holland Publishing Company,* Amsterdam

Taguchi, T. (1972a) "On the Two-dimensional Concentration Surface and Extensions of Concentration Coefficient and Pareto Distribution to the Two Dimensional Case-I," *Annals of the Inst. of Stat. Math.,* vol. 24(2), p.355-381

Taguchi, T. (1972b) "On the Two-dimensional Concentration Surface and Extensions of Concentration Coefficient and Pareto Distribution to the Two Dimensional Case-II," *Annals of the Inst. of Stat. Math.*, vol.24 (3), p. 599-619

Taguchi, T. (1972c) "Concentration Polyhedron, Two Dimensional Concentration Coefficient for Discrete Type Distribution and Some New Correlation Coefficients," *The Inst. of Stat. Math*., p. 77-115

Taguchi, T. (1973) "On the Two-dimensional Concentration Surface and Extensions of Concentration Coefficient and Pareto Distribution to the Two Dimensional Case-III," *Annals of the Inst. of Stat. Math*., vol.25 (1), p.215-237

Taguchi T. (1974) "On Fechner's thesis and statistics with norm p," *Ann. of the Inst. of Stat. Math*., vol. 26(2), p.175-193

Taguchi, T. (1978) "On a Generalization of Gaussian Distribution," *Ann. of the Inst. of Stat. Math*., vol.30(2), A, p.211-242

Taguchi, T. (1981) "On a Multiple Gini's Coefficient and Some Concentrative Regressions," *Metron,* vol. XXXIX (1-2), p.5-98

Theil, H. (1967) "Economic and Information Theory," *Amsterdam: North Holland*

Usow, K.H. (1967a) "On $L_1$ approximation: computation for continuous functions and continuous dependence", *SIAM J. of Numerical Anal*., vol.4, p.70-88

Watson, G.A. (1981) "An Algorithm for Linear $L_1$ Approximation of Continuous Functions," *SIMA J. of Numerical Anal.*, vol.1, p.157-167

**BIOGRAPHY**

Hamid Shahrestani is an Associate Professor of Economics at Ohio University-Chillicothe. He received his B.A. in Economics from Concordia University in Canada, M.A. and Ph.D. in Economics from Western Michigan University and University of Cincinnati, respectively. In the past 30 years, he has been an active consultant in the private sector as well as a teacher and academic researcher. He can be reached by email at: shahrest@ohio.edu

Bijan Bidabad is a Visiting Professor of Economics at Islamic Azad University, Science and Research Branch, Tehran, Iran. He received his Ph.D. from Azad University and is an independent economic researcher specializing in Econometrics and Mathematical Modeling. He is the Former Director of Foreign Exchange and International Research Department at the Monetary and Banking Research Academy of Central Bank of Iran; He can be reached by email at: bijan@bidabad.com