

The “Deciles Implied Inequality Indices” for Measuring Income Distribution

Bijan Bidabad¹

Abstract

Different ways have been proposed to measure income inequality; there is no best way to calculate the inequality index that expresses income distribution as it is. Popular inequality indices provide information about some points on the distribution function and analyze the inequality of income without reference to the amount of the budget needed to improve the income distribution.

In this paper, we propose a set of “Deciles Implied Inequality Indices”. By using this index, we can show how much transfer payment is needed as a redistribution policy to achieve a desired income distribution consistent with the perceived economic goals of the society. That is, we try to find a fiscal-compensation-based index for reducing inequality.

By using “Deciles Implied Inequality Indices”, we may measure how much income (tax and subsidy) may be redistributed to reach the targeted income distribution policy.

Keywords: Income distribution, Inequality Index, Lorenz curve, Concentration Surface, Income Redistribution.

JEL: C63

Introduction

For better distribution of income and welfare, the government might redistribute income through tax, subsidy, or transfer payments to achieve equity and promote the utility of every member of the society.

Popular inequality indices provide comparative information about the income or wealth distribution without responding to the question of what the amount of the money needed to improve the distribution. In this paper, we define a series of income inequality indices; namely, the “Deciles Implied Inequality Indices” to show how much transfer payments or redistribution policies are needed to achieve a desired distribution of income in accordance of the economic equity goal of the policy-maker.

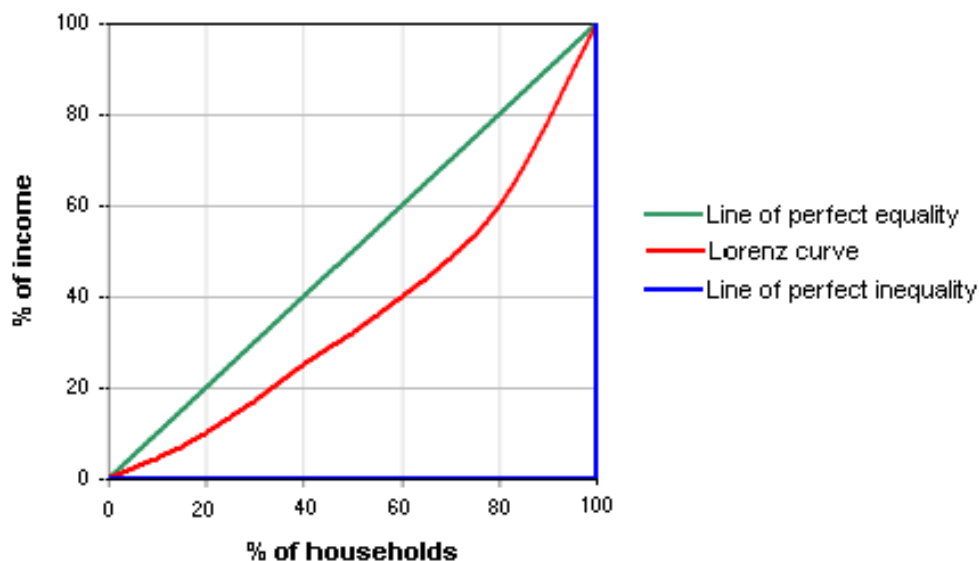
There are different measures of income inequality (Atkinson (1970); Cowell (1977); Sen (1973)). Popular measures are such as the Gini coefficient; the decile ratio; the proportions of total income earned by the bottom 50%, 60%, and 70% of the households; the Robin Hood index; the Atkinson index; and Theil's entropy measure.

The Lorenz curve may illustrate the degree of income inequality. The Lorenz curve is a simple case of the Concentration Surfaces in statistics that are used to show the shape and distribution of inequality in data. In other words, the Lorenz Curve is a representation of the cumulative income distribution function. It

¹ (B.A., M.Sc., Ph.D., Post-Doc.) Professor of Economics and Chief Economic Advisor to Bank Melli Iran.
<http://www.bidabad.ir> bjan@bidabad.ir

explains what portion of the total income is received by the bottom percentages of the households. The percentages of households are shown on the horizontal axis, and the percentages of earned incomes are plotted on the vertical axis as Figure 1 shows it graphically.

Figure 1: Lorenz Curve



In Figure 1, the diagonal line presents the uniform distribution line. That is the distribution of income is perfectly equal. The kinked blue line is presenting full inequality. The income inequality shown by the Redline is the Lorenz curve that can be measured and expressed by the Gini ratio as a statistic. The Gini coefficient is the ratio of the area between the Lorenz curve and the angle bisector 45° line, to the total area below the 45° line. This coefficient expresses the distribution inequality just by one figure. Gini ratio is calculated by a ratio that its numerator is the area between the Lorenz curve and the diagonal line, and the denominator is the area under the uniform distribution line. This ratio is between 0 and 1.

Kakwani (1980) by some recalculation of the Gini coefficient measures the length of the Lorenz curve as an inequality index. 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 wether enjoyed by everybody would make total welfare precisely equal to the overall well-being generated by the actual income distribution. Theil (1967) entropy measure derives from the notion of entropy in information theory.

There is no single "best" measure of income inequality. Some measures such as the Atkinson index are more bottom-sensitive than others are; i.e., more strongly correlated with the extent of poverty. The measures perform differently under various types of income transfers. For instance, the Gini coefficient 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 concerning income transfers between households on the same side of the mean income, and so on. While there are alternative methods, there is no best way to calculate the inequality index, mainly concentrating on fiscal view. That is, they generally analyze the distribution without inferring about the amount of fund needed to correct income inequality.

Since the skewness of income distribution is persistently exhibited for different statistical populations, the Lorenz curve becomes a method to analyze the skew distributions. Therefore, the Pearsonian family distributions are rival functions to explain income distribution. There is also a relation between the area under the Lorenz curve and the corresponding probability distribution function (see, Kendall and Stuart (1977)). That is, when the probability distribution function is known, we may find the corresponding Lorenz curve and Gini coefficient as well.

Estimation of the Lorenz curve has some difficulties. For this estimation, we should define an appropriate functional form that can accept different curvatures. There is another problem, that is, to create a large-size dataset for estimating the corresponding parameters of the Lorenz curve, a large amount of computation on raw sample income data is inevitable. These problems, despite their computational difficulties, make the significance of the estimated parameters weak (see, Bidabad and Bidabad (1989)). To avoid this, Bidabad (1989a,b) proposed the estimation of the functional form of the Lorenz curve by defining and using continuous L_1 norm smoothing. The continuous L_1 norm estimation problems of linear one and two parameter models were solved by Bidabad (1988a,b).

The viewpoint of this paper is to introduce a set of “Deciles Implied Inequality Indices”, which satisfies the policy implications needs. These indices can be used for reducing the degrees of inequality. We show how to use Lorenz Curve to measure the distribution of income and to calculate the amount of money needed to be levied on rich and then transferred to the poor to promote income distribution of the society. However, some more mathematical sophistication to estimate the Lorenz Curve regarding “Implied Inequality Index” was given by Bidabad and Shahrestani (2008). In that paper, we introduced a shortcut to use the probability density function of population income to estimate the Lorenz function parameters. That is the continuous L_1 norm smoothing method used to estimate the regression parameters of the Lorenz curves corresponding to two different probability density functions: (a) log-normal function that is more suitable for a broader range of income but is not integrable and (b) Pareto density distribution function that is integrable.

Lorenz Curve

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

$$\left(P(V|V \leq v), \frac{E(V|V \leq v)}{E(V)} \right) \quad v \in R \quad (1)$$

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

$$\left(\int_{-\infty}^v f(w)dw, \frac{\int_{-\infty}^v wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} \right) \equiv (x(v), y(x(v))) \quad (2)$$

Taguchi (1972a,b,c,73,81,83,87,88) multiplies the second element of (1) by $P(V|V \leq v)$ which is not correct; his definition of (2) is equivalent to ours. We denote (2) by ordered pair $(x(v), y(x(v)))$ where $x(v)$ and $y(x(v))$ are its elements. "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. We may use the form introduced by Gupta (1984) and a modified version of Bidabad and Bidabad (1989), which benefits from certain properties as the explicit function of the Lorenz curve.

Gupta (1984) proposed the functional form,

$$y = xA^{x-1} \quad A > 1 \quad (3)$$

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

$$y = x^B A^{x-1} \quad B \geq 1, A \geq 1 \quad (4)$$

To estimate the above functions we need discrete data from the population, to construct relevant x and y vectors to estimate "A" of (3) or "A" and "B" of (4).

Implied-Inequality-Index

Most inequality indices are concentrated on the statistical aspect of the income distribution. That is, they generally analyze the distribution without inferring about the amount of fund needed to correct income inequality. In this section, we will introduce an inequality index, which shows how much money should be transferred from the upper-income group to the lower group to achieve the desired distribution of income. Bidabad and Shahrestani (2010) introduced an implied inequality index using L_1 norm estimation of the Lorenz curve. Here we extend this index to be more qualified for income redistribution policies and implications.

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

$$\int_{-\infty}^v wf(w)dw = \int_v^{+\infty} wf(w)dw \quad (5)$$

By definition, we have:

$$\mu = \int_{-\infty}^{+\infty} wf(w)dw = \int_{-\infty}^v wf(w)dw + \int_v^{+\infty} wf(w)dw \quad (6)$$

That is:

$$\int_{-\infty}^v wf(w)dw = \mu/2 \quad (7)$$

On the other hand:

$$\frac{\int_{-\infty}^v wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} = 1/2 \quad (8)$$

According to (2) this is a point on the Lorenz curve with the following ordered pair:

$$\left(\int_{-\infty}^v wf(w)dw, 1/2 \right) \quad (9)$$

Thus, we define implied-inequality-index (iii) as $\int_{-\infty}^v f(w)dw$ when v satisfies (5). That is,

$$iii = \int_{-\infty}^v f(w)dw \quad \text{when } v \text{ satisfies } \frac{\int_{-\infty}^v wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} = 1/2 \quad (10)$$

To find iii, (7) should be solved for v and its value be replaced in (10). As iii approaches $1/2$, 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 (left) high inequality. The values of iii less than $1/2$, the iii has no economic implication for income distribution.

Let us define the cost of equalization as:

$$C = [(iii - \frac{1}{2}) \times N \times \mu] \quad (11)$$

This expression means that to equalize the distribution of income without changing the average income of the society, the amount of C in terms of money should be transferred from higher income earner to the 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:

$$\text{Relative cost of equalization} = [(iii - \frac{1}{2}) \times N \times \mu] / (N \times \mu) = (iii - \frac{1}{2}) \quad (12)$$

Numerical Example

To illustrate, the following table 1 of income distribution for a hypothetical society is used. Consider a society of 400 households with a total income of the society equal to \$2000 where 280 poor income earners receive half of it (\$1000) and 120 richer earn another 50% (\$1000) of the society's income. These values can be simply understood from table 1. At the half of total income of the society (\$1000), the bottom 70% of the population earns 50% of society's income, and 30% of the top of the population earn other 50% of the total income of the society. According to table 1, we have:

$N = 400$	(Number of households)
$\mu = 2000/400 = 5$	(Average income)
$\mu_{\text{lower}} = 1000/280 = 3.57$	(Average income of lower category)
$\mu_{\text{upper}} = 1000/120 = 8.33$,	(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 total tax of \$400 from the top 30% of the population and transfer it to the lower 70% of the income earners, the average income of both groups will be the same:

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

$$\text{Relative cost of equalization} = 0.7 - 0.5 = 0.2 = \%20$$

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

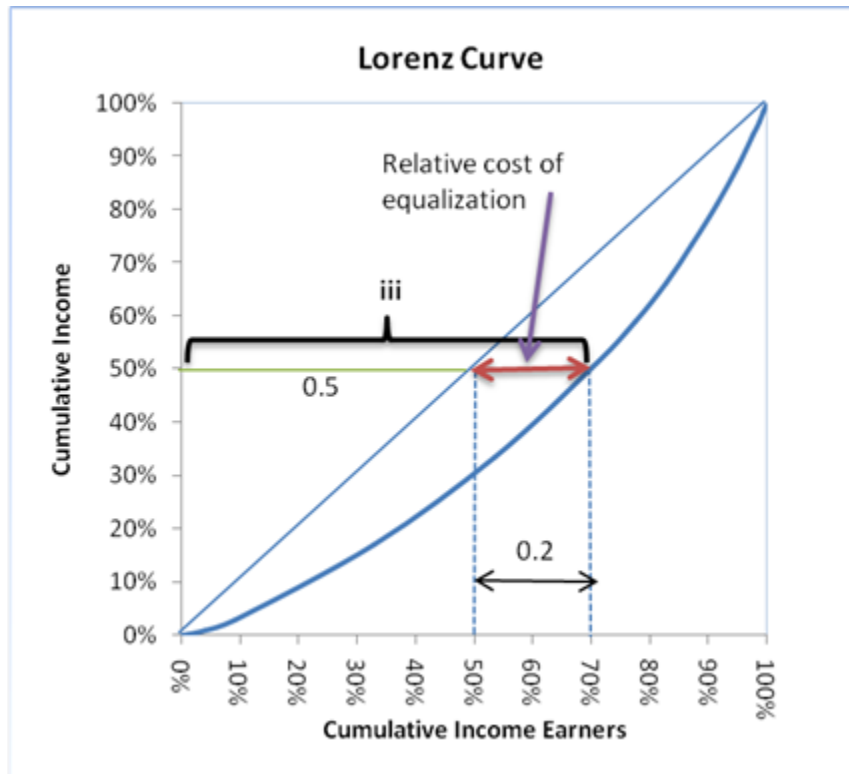
According to table 1, we may depict the iii and the relative cost of equalization on the Lorenz curve as follows. This index is illustrated by using columns (4) and (9) of table 1.

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%		48	1933	96.7%	
13	3	398	0.8%	99.5%		39	1972	98.6%	
14	2	400	0.5%	100.0%	120	28	2000	100.0%	1000
	400		100%		400	2000			2000

In table 1, the column (1) depicts dollar values of income categories, and the 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. The column (10) shows half of the total income of society.

Figure 3: Implied inequality index iii



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

Deciles-Implied-Inequality-Indices

In this section, we will introduce deciles implied inequality indices, which can be simply generalized to quantiles or percentiles indices. They show how much money should be transferred from the upper-income deciles to the lower deciles or quantiles or percentiles groups to achieve the desired distribution of income.

As before, suppose there is a personal income v at which a decile or quantile of the total income of the population belongs to those who have less than v , and the other income belongs to those who have a higher income than v . That is for the D_i (i^{th} decile, $i=1, 2, \dots, 10$):

$$\int_{-\infty}^{v_1} wf(w)dw = \int_{v_1}^{v_2} wf(w)dw = \dots = \int_{v_9}^{+\infty} wf(w)dw \quad (13)$$

By definition, we have:

$$\mu = \int_{-\infty}^{+\infty} wf(w)dw = \int_{-\infty}^{v_1} wf(w)dw + \int_{v_1}^{v_2} wf(w)dw + \dots + \int_{v_9}^{+\infty} wf(w)dw \quad (14)$$

That is:

$$\begin{aligned} \int_{-\infty}^{v_1} wf(w)dw &= \mu/10 \\ \int_{v_1}^{v_2} wf(w)dw &= 2\mu/10 \\ &\vdots \\ \int_{v_9}^{+\infty} wf(w)dw &= 10\mu/10 \end{aligned} \quad (15)$$

On the other hand:

$$\begin{aligned} \frac{\int_{-\infty}^{v_1} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} &= 1/10 \\ \frac{\int_{v_1}^{v_2} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} &= 1/10 \\ &\vdots \\ \frac{\int_{v_9}^{+\infty} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} &= 1/10 \end{aligned} \quad (16)$$

According to (2) these points are on the Lorenz curve with the following ordered pairs:

$$\left(\int_{-\infty}^{v_i} wf(w)dw, i/10 \right), \quad i = 1, 2, \dots, 10 \quad (17)$$

Thus, we define decile-implied-inequality-index (d_{iii}) as $\int_{-\infty}^{v_i} f(w)dw$ when v_i satisfies (13). That is,

$$d_{i,iii} = \int_{-\infty}^{v_i} f(w)dw \quad \text{when } v \text{ satisfies } \frac{\int_{-\infty}^{v_i} wf(w)dw}{\int_{-\infty}^{+\infty} wf(w)dw} = i/10 \quad (18)$$

To find $d_{i,iii}$, (17) should be solved for v_i , and its value be replaced in (18). As $d_{i,iii}$ approaches $i/10$, distribution becomes more symmetric. If $d_{i,iii}$ tends to 1, distribution tends to be fully right-skewed, indicating high (right) inequality and as $d_{i,iii}$ tends to 0, distribution tends to be left-skewed, and distribution tends to left high inequality. The values of $d_{i,iii}$ less than $i/10$, however, have no economic implication for income distribution. Let us define the cost of equalization policy as:

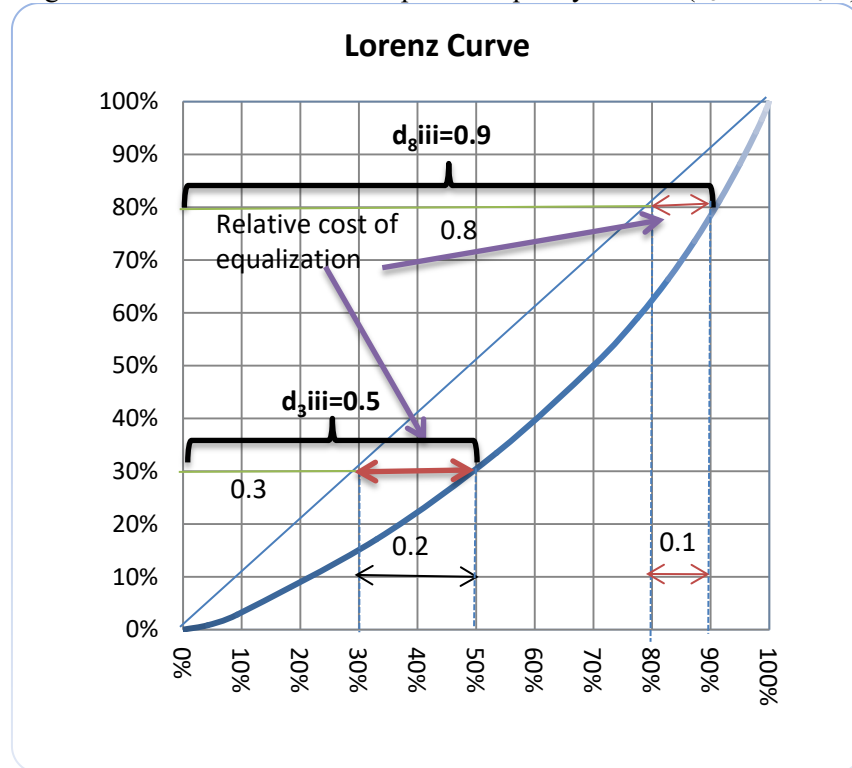
$$C_i = [d_{i,iii} - i/10] \times N \times \mu \quad (19)$$

The above expression means that to equalize the distribution of income without changing the average income of the society, the amount of C_i should be transferred from higher income earner to lower income earner (lower than the i^{th} decile), 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:

$$\text{Relative decile cost of equalization} = [(d_{i,iii} - i/10) \times N \times \mu] / (N \times \mu) = (d_{i,iii} - i/10) \quad (20)$$

Figure 4. The deciles 3 and 8 Implied Inequality Indices ($d_{3,iii}$ and $d_{8,iii}$)



This figure depicts the two deciles (3 and 8) of Decile Implied Inequality Indices and the relative cost of equalization as corresponding parts of the Lorenz curve

Conclusion

To have a better understanding and policy arrangements about the equality of income distribution, it is not enough to know the traditional inequality indices. The redistribution policies or income distribution monitoring need to deal with specific budget guidelines to promote the society to a better distributional position. While there are alternative methods for presenting income distribution, there is no best way to present the inequality index, mainly concentrating on fiscal viewpoints. That is, they generally analyze the statistical distribution without inferring about the amount of fund needed to promote income equality. The view of this paper is to introduce a set of Deciles Implied Inequality Indices, which satisfies these policy implications needs. The designed Deciles Implied Inequality Indices are good fiscal guidepost for monitoring or performing equalization policies of society's income.

References

Atkinson A. B. (1970), On the measurement of inequality, *Journal of Economic Theory* 1970; 2: 244-263.

Bidabad, Bijan; Behrouz Bidabad (1989), Functional form for estimating the Lorenz curve, Australasian Econometric meeting, Australian National University, Australia, 1989. *American Finance & Banking Review*, 4(1), 17-21, 2019.

<https://www.cribfb.com/journal/index.php/amfbr/article/view/286>

<http://www.bidabad.ir/doc/functional-form-lorenz.pdf>

<http://www.bidabad.ir/doc/functional-form-lorenz.pptx>

DOI: 10.1313/amfbr.v4i1.286

Bidabad Bijan (1989a) "Discrete and continuous L_1 norm regressions, proposition of discrete approximation algorithms and continuous smoothing of concentration surface", Ph.D. thesis, Islamic Azad University, Tehran, Iran.

<http://www.bidabad.ir/doc/L1-norm-thesis-en.pdf>

Bidabad Bijan (1989b) "Discrete and continuous L_1 norm regressions, proposition of discrete approximation algorithms and continuous smoothing of concentration surface", Ph.D. thesis, Islamic Azad University, Tehran, Iran. Persian translation.

<http://www.bidabad.ir/doc/L1-norm-thesis-fa.pdf>

Bidabad, Bijan (2019) Continuous L_1 norm estimation of Lorenz curve. *Bangladesh Journal of Multidisciplinary Scientific Research*, Vol. 1, no. 1, PP. 41-49, 2019.

<http://www.bidabad.ir/doc/l1-article4.pdf>

DOI: 10.1313/BJMSR.v1i1.314

Bidabad, Bijan (2019) Estimating Lorenz curve for Iran by using continuous L_1 norm estimation, *Economics and Management Journal*, Islamic Azad University, No. 19, winter 1993, pp. 83-101.

Reprinted: *International Journal of Marketing Research Innovation*, 3(1), 11-21, 2019.

<https://www.cribfb.com/journal/index.php/ijmri/article/view/322>

<http://www.bidabad.ir/doc/iraninc-l1.pdf>

DOI: 10.1313/ijmri.v3i1.322

Bidabad, Bijan (2019) USA Income distribution counter-business-cyclical trend (Estimating Lorenz curve using continuous L_1 norm estimation). First meeting of the Society for the Study of Economic Inequality (ECINEQ), Palma de Mallorca, Spain, July 20-22, 2005.

<http://www.bidabad.ir/doc/estimating-lorenz-us.pdf>

- Bidabad, Bijan, Hamid Shahrestani (2010) An implied inequality index using L_1 norm estimation of Lorenz curve. Global Conference on Business and Finance Proceedings. Mercedes Jalbert, managing editor, ISSN 1931-0285 CD, ISSN 1941-9589 Online, Volume 3, Number 2, 2008, The Institute for Business and Finance Research, Ramada Plaza Herradura, San Jose, Costa Rica, May 28-31, 2008, pp. 148-163. Global Journal of Business Research, Vol. 4, No. 1, 2010, pp.29-45.
<http://www.bidabad.ir/doc/L1-Implied-inequality-index-4.pdf>
- 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*, 52, 1313-1314.
- 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*, 48, 1063-64.
- Kakwani, N.C. & Podder, N (1976) "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations." *Econometrica* 44, 137-148.
- Kendall, M. & Stuart, A. (1977) "The advanced theory of statistics", Vol.1, Charles Griffin & Co., London.
- Mills, Jeffrey & Sourush, A. (1997) "Statistical Inference via Bootstrapping for Measures of Inequality", *Journal of Applied Economics* 12:133-150.
- Rasche, R.H., Gaffney, J., Koo, A.Y.C. & Obst, N. (1980) "Functional forms for estimating the Lorenz curve", *Econometrica*, 48, 1061-1062.
- Rice, J.R. (1985) "Numerical methods, software, and analysis", McGraw-Hill, ch. 11.
- Rice, J.R., White, J.S. (1964) "Norms for smoothing and estimation", *SIAM Rev.*, 6, 243-256.
- Salem, A.B.Z. & Mount, T.D. (1974) "A convenient descriptive model of income distribution: the gamma density", *Econometrica*, 42, 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*, 44, 963-970.
- Slotje, 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.*, 24(2), 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.*, 24(3), 599-619.

Taguchi, T. (1972c) “Concentration polyhedron, two-dimensional concentration coefficient for discrete type distribution and some new correlation coefficients etc.”, *The Inst. of Stat. Math.*, 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.*, 25(1), 215-237.

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

Taguchi, T. (1978) “On a generalization of Gaussian distribution”, *Ann. of the Inst. of Stat. Math.*, 30(2), A, 211-242.

Taguchi, T. (1981) “On a multiple Gini's coefficient and some concentrative regressions”, *Metron*, Vol. XXXIX (1-2), 5-98.”.

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

Xu, Kuan (2004). “How has the literature on Gini’s index evolved in the past 80 years?”, <http://economics.dal.ca/RePEc/dal/wprarch/howgini.pdf>