Economics Bulletin

Volume 39, Issue 3

Further tricks with the Lorenz curve

Sreenivasan Subramanian
Previously: Madras Institute of Development Studies, Currently: Independent Scholar

Abstract

The present note is the third in a series that deals with simple manipulations of the Lorenz curve intended to cast light on this or that aspect of the measurement of inequality. In this note, a procedure is outlined for ‘adjusting' the Gini coefficient of inequality for the skewness of the Lorenz curve, and for similarly ‘adjusting' the so-called ‘Kolkata' index of inequality, as well as the Lorenz curve itself, in such a way that the inequality orderings of distributions induced by these adjusted measures and the adjusted Lorenz curve are equivalent.

The author is indebted to an anonymous referee for helpful suggestions, and to A Arivazhagan for help with the graphics.
Citation: Sreenivasan Subramanian, (2019) "Further tricks with the Lorenz curve", *Economics Bulletin*, Volume 39, Issue 3, pages 1677-1686
Contact: Sreenivasan Subramanian - ssubramanianecon@gmail.com.
Submitted: April 23, 2019. **Published:** July 07, 2019.

1. Introduction

This is the third in a sequence of papers of which the first was titled 'Tricks with the Lorenz Curve' (Subramanian, 2010), and the second, 'More Tricks with the Lorenz Curve' (Subramanian, 2013). Undoubtedly, the sequence threatens to become a serial killer (if it is not already that), but the good news is that the prospect of additional 'Tricks' may be naturally limited by the difficulty of finding titles for them. The present instalment is, sadly, triggered by the same impulse as the previous two versions, namely, that '… it is difficult to resist the temptation of getting up to tricks of one kind or another in the presence of the seemingly infinite possibilities offered up by the [Lorenz] curve' (Subramanian, 2010; p. 1595). As with the earlier papers, the focus of the present one will be on a severely pragmatic approach to measurement. My objective is to advance an easily comprehended inequality index, the construction of which reflects a set of long-standing concerns in the relevant literature. This approach relies on an intuitively plausible line of reasoning rather than on sophisticated axiomatics or powerful general propositions. Thus, the accent in the paper, throughout, is on a simple and direct approach to the problem, which it shares with its two predecessors. This concludes the sub-section on apologies and caveats.

The full title of the paper should read: 'Further Tricks with the Lorenz Curve: Or What To Do With It If It Is Found Crossing Another One.' That is to say, the concern is with the well-recognized problem of ranking intersecting Lorenz curves, a problem that is mediated by differential sensitivity to income transfers at the upper and lower ends of a distribution, and one which has been dealt with in considerable depth by authors such as Kolm (1976), Shorrocks and Foster (1987), and Aaberge (2009). The notion of 'transfer-sensitivity' of an inequality index is here interpreted quite simply in terms of the skewness of the Lorenz curve, and how the Gini coefficient of inequality might be 'corrected' or 'adjusted' for the skewness of the curve.

2. Basic Concepts

I shall draw substantially on Kakwani (1980) and Subramanian (2010) in setting out the basic concepts and notation for this paper. We shall let *x* stand for a random variable designating income. It is distributed over the interval $[0,\infty)$, and the mean of the distribution is denoted by μ. The density function of *x* (the proportion of the population with income *x*) is $f(x)$; $F(x)$ is the cumulative density function (the cumulative proportion of the population with incomes not exceeding *x*); and $L(F(x))$ is the first-moment distribution function (the cumulative share in income of the population with incomes not exceeding *x*):

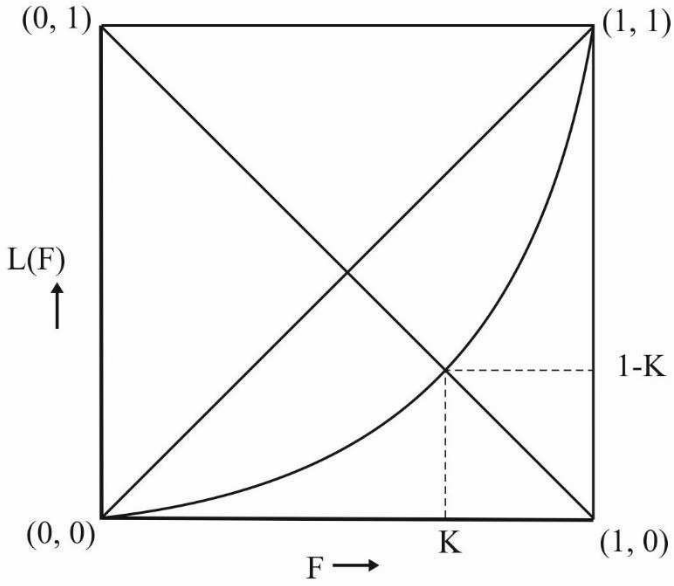
$$F(x) = \int_0^x f(y)dy;$$

$$L(F(x)) = \left(\frac{1}{\mu}\right)\int_0^x yf(y)dy;$$

$$Lim_{x\to 0}F(x) = Lim_{x\to 0}L(F(x)) = 0; \text{ and } Lim_{x\to\infty}F(x) = Lim_{x\to\infty}L(F(x)) = 1.$$

The Lorenz curve is simply the plot of the income share of the poorest $F(x)$ proportion of the population against the poorest $F(x)$ proportion of the population, for all values of x. As drawn in the unit square of Figure 1, the Lorenz curve for an unequal distribution, typically, would be an increasing and strictly convex curve, running from (0,0) to (1,1) of the square; and for an equal distribution, the Lorenz curve would coincide with the 'line of equality', which is the diagonal of the unit square.

Figure 1: The Lorenz Curve



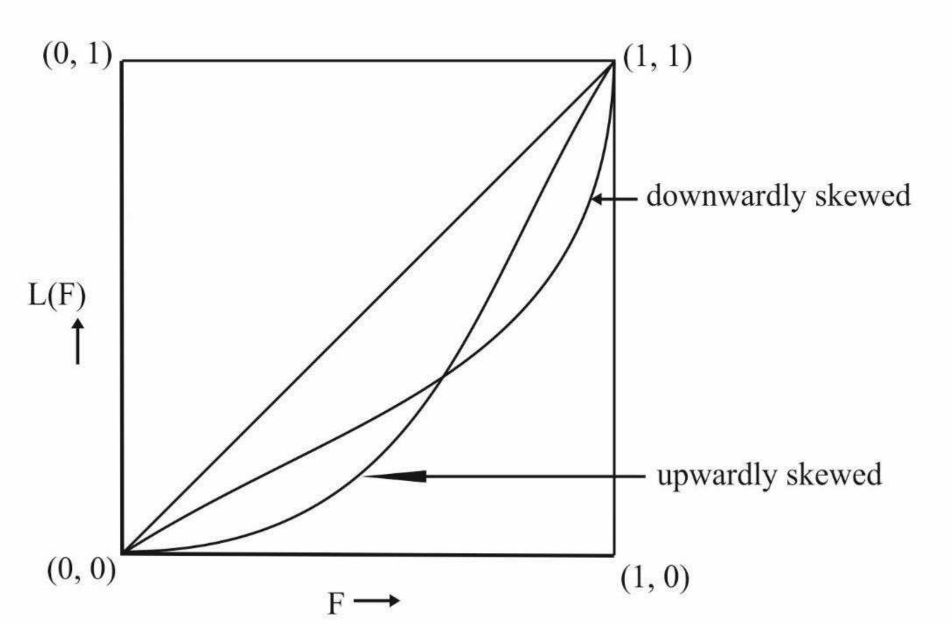
Define x^* to be a measure of central tendency such that the poorest $F(x^*)$ proportion of the population earn $(1-F(x^*))$ proportion of the total income. $F(x^*)$ has been seen in the light of a measure of inequality in its own right, one which has been called the *Kolkata Index* (K) by Chatterjee et al. (2017); accordingly, the quantity $F(x^*)$ will henceforth simply be referred to as K in this paper. It is easy to see that x^* is that level of income corresponding to which the Lorenz curve (see Figure 1) intersects the diagonal drawn from (0,1) to (1,0) of the unit square (a diagonal which Kakwani, 1980, refers to as the 'alternative diagonal').

The Gini coefficient of inequality G for a distribution can be written in a number of ways. One of these is as a normalized, weighted sum of the deviations of all incomes from the mean income, the weights being the cumulative proportions of the population with incomes exceeding each relevant income:

$$G = (2/\mu)\int_0^\infty (\mu - x)[1-F(x)]f(x)dx. \tag{1}$$

As is well-known, the Gini coefficient will rank two intersecting Lorenz curves indifferently so long as the areas enclosed by the two curves between themselves and the diagonal of the unit square are the same—without regard for whether the curves are skewed toward (1,1) or (0,0) of the unit square (see Figure 2). We shall call these two types of Lorenz curve 'upwardly skewed' and 'downwardly skewed' respectively, unlike a symmetric Lorenz curve, which, on the left side of the alternative diagonal drawn from (0,1) to (1,0) of the unit square, is a mirror image of the portion of the curve on the right side of this diagonal.

Figure 2: Intersecting Lorenz Curves with the Same Gini Value



An important property of an inequality index is that of 'transfer sensitivity', which essentially requires that the downward impact on inequality of a progressive ('richer-to-poorer') transfer of income should be more pronounced the lower down the income distribution one travels. Typically, the Gini coefficient of inequality (exactly like the squared coefficient of variation) fails the transfer sensitivity property—being equally responsive, as it is, to income transfers of a given amount between incomes a fixed distance apart, or a fixed number of individuals apart. One way of interpreting the property of transfer sensitivity of an inequality measure in terms of the skewness of the Lorenz curve is the following. For two intersecting Lorenz curves with the same Gini-value (G) of inequality, transfer sensitivity should demand that the upwardly skewed curve should be penalized vis-à-vis the downwardly skewed curve. The reference standard is provided by the symmetric Lorenz curve: a 'skewness-adjusted' Gini coefficient—call it G^*—might be expected to display the 'right' sort of transfer-sensitivity whenever G^* is greater than G for an upwardly skewed Lorenz curve, equal to G for a symmetric Lorenz curve, and less than G for a downwardly skewed Lorenz curve.

In deriving the index G^*, reference will be made to the distinguished measure of central tendency earlier denoted by x^*, which, to recall, is simply that level of income for which the poorest Kth fraction of the population receives a (1-K)th fraction of the total income. The quantities x^* and K are defined in Subramanian (2010), where K is written as $F(x^*)$, and further discussed in Lambert and Subramanian (2015). As mentioned earlier, Chatterjee et al. (2017) advance the merits of K as a well-defined measure of inequality which is christened the Kolkata Index, and whose relationship with the Gini coefficient is studied in that paper.

In the following section, I shall advance a procedure for ranking intersecting Lorenz curves, in terms of the interplay between the quantities G (the Gini coefficient), G^* (a skewness-adjusted version of G), s (an indicator of skewness of the Lorenz curve), x^* (a measure of central tendency), K (the Kolkata Index of inequality), K^* (a measure derived from K), and L^* (a pseudo-Lorenz curve derived from the actual Lorenz curve L).

3. Adjusting the Gini Coefficient for Skewness of the Lorenz Curve

Basic to the measurement of the skewness of a Lorenz curve is the measure of central tendency x^*. To see what is involved, notice first that the slope of the Lorenz curve at any point on it corresponding to an income level of x is given by x/μ (see Kakwani, 1980), so that at the point on the Lorenz curve corresponding to the mean income, the slope of the curve is unity. When the Lorenz curve is symmetric, it is easy to see that x^* coincides with μ, so that the slope of the Lorenz curve at the point of its intersection with the alternative diagonal is precisely one. For an upwardly skewed Lorenz curve, x^* is clearly greater than μ while, for a downwardly skewed Lorenz curve, x^* is smaller than μ (see Figure 2). The proportionate deviation of x^* from μ suggests itself as a naturally plausible measure s of the skewness of the Lorenz curve, given by:

$$s = 1 - x^*/\mu. \tag{2}$$

From what has been said before, s is simply the difference between the slope of the Lorenz curve at the point corresponding to the mean income and its slope at the point corresponding to the income level x^*: s is zero for a symmetric Lorenz curve, it is positive for a downwardly skewed Lorenz curve, and it is negative for an upwardly skewed Lorenz curve.

Returning to the expression for the Gini coefficient furnished by equation (1), one can see that the inability of this inequality measure to take account of the Lorenz curve's skewness has to do with taking income deviations from the mean income μ rather than from the income level x^*. What if, instead, we were to measure inequality as a normalized weighted sum of all income deviations from x^*, the weights, again, being the cumulative proportions of the population with incomes in excess of each x? Call the resulting inequality measure G^*, so that, by definition,

$$G^* = (2/\mu)\int_0^\infty (x^* - x)(1 - F(x)) f(x)dx,$$

which can be written as

$$G^* = (2/\mu)\int_0^\infty (\mu - x)(1 - F(x)) f(x)dx - (2/\mu)\int_0^\infty (\mu - x^*)(1 - F(x)) f(x)dx,$$

which, in turn, can be simplified to read:

$$G^* = (2/\mu)\int_0^\infty (\mu - x)(1 - F(x)) f(x)dx - (1 - x^*/\mu),$$

so that, in view of the expressions for G and s provided by equations (1) and (2) respectively, we can now write:

$$G^* = G - s. \tag{3}$$

Since s is zero for a symmetric Lorenz curve, positive for a downwardly skewed Lorenz curve, and negative for an upwardly skewed Lorenz curve, equation (3) assures us that G* coincides with G for a symmetric Lorenz curve, is less than G for a downwardly skewed Lorenz curve, and exceeds G for an upwardly skewed Lorenz curve, thus meeting the requirement of rewarding a Lorenz curve for which the income-shares of the relatively poor population are higher (that is, a Lorenz curve which is downwardly skewed), and penalizing a Lorenz curve for which the income-shares of the relatively poor population are lower (that is, a Lorenz curve which is upwardly skewed)—see Figure 2.

Before proceeding further, a word on Lorenz-dominance. A distribution F will be said to Lorenz-dominate a distribution F'—written $L(F) \succ L(F')$—if and only if the Lorenz curve for F lies somewhere inside the Lorenz curve of F', and nowhere outside it. What can we say about the ordering of distributions by the Lorenz-dominance criterion and by the real-valued measures G and K? It is not hard to see that we can have a situation in which for a pair of distributions F and F', it is the case that $L(F) \succ L(F')$ and $G(F) < G(F')$, but $\neg(K(F) < K(F'))$; or a situation in which the Lorenz curves for F and F' intersect (so that the distributions cannot be ordered by the Lorenz-dominance relation $\succ$), but $G(F) = G(F')$ and $K(F) \neq (K(F')$; and so on. Is there a means by which the inequality orderings of distributions by the Lorenz-dominance relation and by real-valued measures similar to G and K can be made to be identical? Some middle-school geometry and algebra are deployed toward this end in what follows.

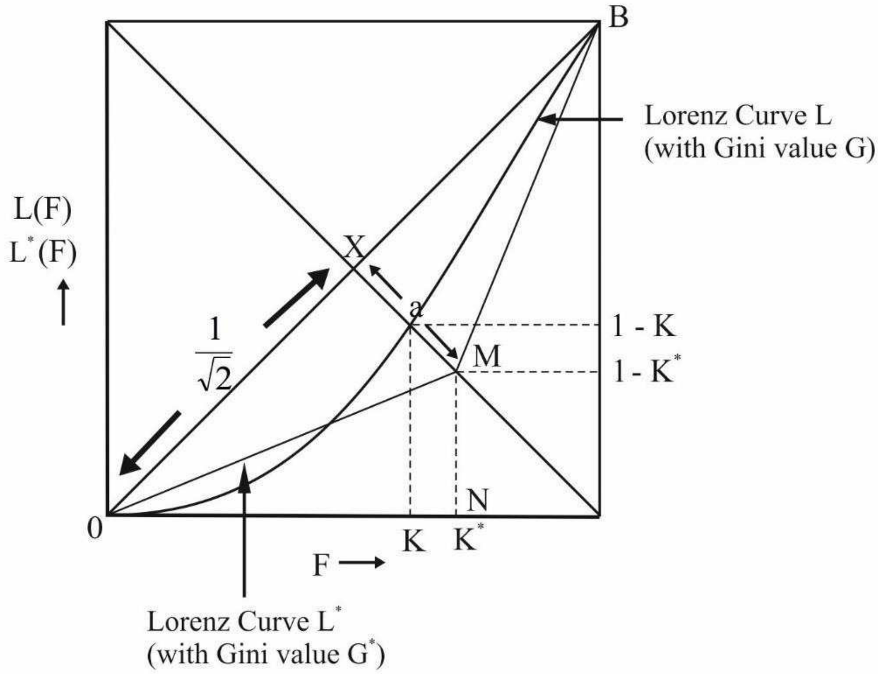
We note first that, if F^{-1} stands for the generalized inverse of the distribution function F, the relationship between the measures G*, G and K revealed by equation (3) can also be written as:

$$G^* = G - [(\mu - F^{-1}(K)) / \mu]. \tag{4}$$

(($F^{-1}(K)$ is, of course, the income level x*.)

Now consider the following construction. Given Figure 1, where we have a Lorenz curve L with a Gini-value of G and a Kolkata Index-value of K, it is in principle a simple matter, given equation (4), to compute the value of the index G*. In Figure 3 we draw a symmetric Lorenz curve with two linear arms as depicted in the illustration, such that the Gini-value of this Lorenz curve is G*. The value of the Gini coefficient is, of course, twice the area enclosed by the Lorenz curve and the diagonal of the unit square. By construction, the Lorenz curve OMB in Figure 3 has been drawn such that its corresponding Gini value is G*; and since the Lorenz curve is symmetric, the value of G* must be 4(Area OMX) [since G* = 2(Area OMB) and Area OMX = (1/2)(Area OMB)]. But the area of the triangle OMX (given by (1/2)[base$\times$altitude)) is one-half of the product of the length OX along the diagonal of the unit square and the length of the line-segment, marked a, along the alternative diagonal. Noting that the length of the line segment OX is $(1/\sqrt{2})$, while that of the line segment XM is a, it follows that $(1/2)(1/\sqrt{2})a = G^*/4,$ whence $a = G^*/\sqrt{2}$.

Figure 3: The Pseudo-Lorenz Curve L*



To summarize, given the actual Lorenz curve L with a Gini value of G, we compute the value of G*, and draw a *pseudo-Lorenz curve* L* by means of the following construction: we mark off the point M on the alternative diagonal of the unit square by measuring along it a length $a = G^*/\sqrt{2}$ from the point of intersection X of the two diagonals of the Lorenz curve to the point M on the alternative diagonal. The coordinates of the point M are denoted $(K^*, 1-K^*)$. The pseudo-Lorenz curve L* can be seen to be a symmetrically distributed two-part 'linearized' equivalent of the actual Lorenz curve L.

Given an actual Lorenz curve L with a Gini value of G and a Kolkata Index-value of K, we have corrected for the skewness of the Lorenz curve by transforming the triple $\langle L, G, K \rangle$ into an 'equivalent' triple $\langle L^*, G^*, K^* \rangle$. Given the fact that the Lorenz curve L* has been both 'symmetrified' and 'linearised', it is clear that exactly and only one Lorenz curve L* can be drawn through any given point on the alternative diagonal of the Lorenz curve. It follows that two L* Lorenz curves can never intersect: the ordering induced by L* is a *complete* ordering, unlike the partial ordering induced by L. Since, by construction, the length a of the line segment XM in Figure 3 is an increasing function of G*, and K* is an increasing function of a, K* is also an increasing function of G*: it follows that the ranking of distributions by the Lorenz curve L* and by the real-valued measures G* and K* must always be identical. That is, for all distributions F and F':

$$L^*(F) \succ L^*(F') \leftrightarrow G^*(F) < G^*(F') \leftrightarrow K^*(F) < K^*(F').$$

4. Conclusion with an Empirical Example

Kakwani (1980) has provided very useful empirical information on the equations of the Lorenz curves for the distribution of income in fifty countries. The equation of the Lorenz curve is specified in terms of a transformed (π, η) coordinate system, given by $\eta = r\pi^{\alpha}(\sqrt{2} - \pi)^{\beta}$, where r, α and β are parameters estimated by regression, and π and η are derived from the coordinates of the Lorenz curve through the relationships $\pi = [F + L(F)]/\sqrt{2}$ and $\eta = [F - L(F)]/\sqrt{2}$ respectively (see Kakwani and Podder, 1976). Of interest to us are the downwardly skewed Lorenz curve for Thailand's 1970 urban household distribution and the upwardly skewed Lorenz curve for Yugoslavia's 1968 household distribution of income. The parameter estimates of the transformed coordinate system for Thailand and Yugoslavia respectively, are given by: Thailand: $r = .377, \alpha = .924, \beta = .912$; and Yugoslavia: $r = .315, \alpha = .823, \beta = .862$ (the details are available in Table 17.1 of Kakwani, 1980). Using these parameter estimates, we can generate data on selected coordinates of the Lorenz curve for each of Thailand and Yugoslavia, as is done in Table 1 below:

Table 1: Data on Selected Coordinates of the Lorenz Curves of Household Income Distribution for Urban Thailand (1970) [T] and Yugoslavia (1968) [Y]

F_T	0	.17	.29	.44	.54	.64	.74	.82	.87	.96	1.0
$L(F_T)$	0	.06	.11	.20	.26	.36	.46	.58	.73	.84	1.0
F_Y	0	.155	.29	.41	.52	.62	.72	.81	.88	.95	1.0
$L(F_Y)$	0	.045	.11	.19	.28	.38	.48	.59	.72	.85	1.0

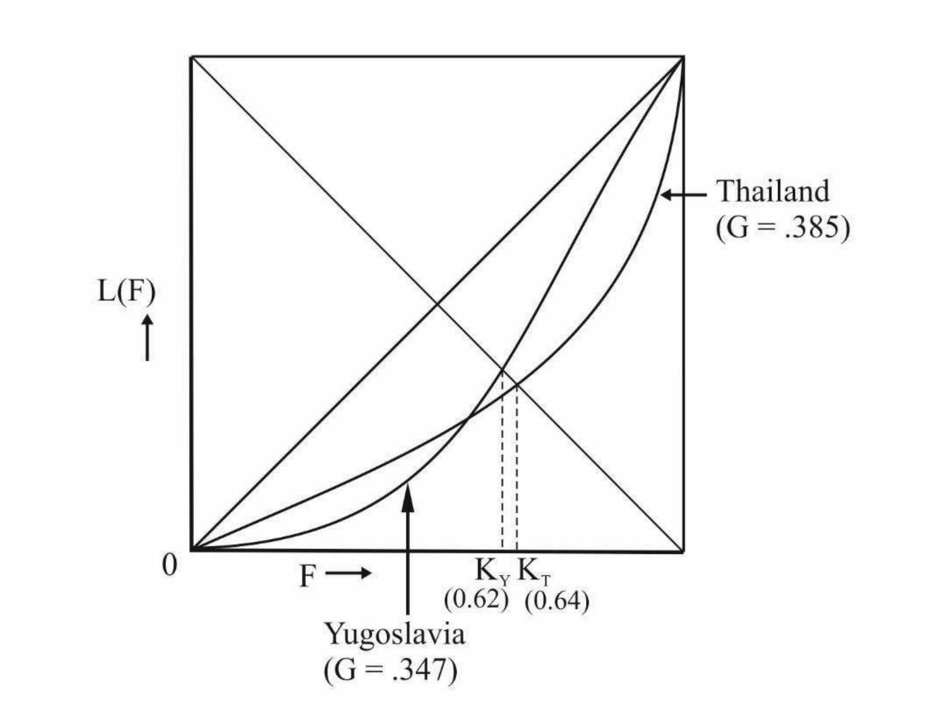
Source: Computations based on the parameter values of r, α and β available in Table 17.1 of Kakwani (1980).

Table 2: The Values of Some Inequality Measures for Thailand and Yugoslavia based on the Information in Table 1

Inequality Indicator/Country	*G*	*G**	*K*	*K**
Thailand	0.385	0.379	0.641	0.690
Yugoslavia	0.347	0.366	0.624	0.684
Ratio of Thai Indicator to Yugoslav Indicator	1.11	1.04	1.03	1.01

Source: *G*-values are from Table 17.1 of Kakwani (1980); all other values have been computed from data available in the same Table.

Figure 4: Lorenz Curves for Urban Thailand (1968) and Yugoslavia (1970) Household Income Distributions

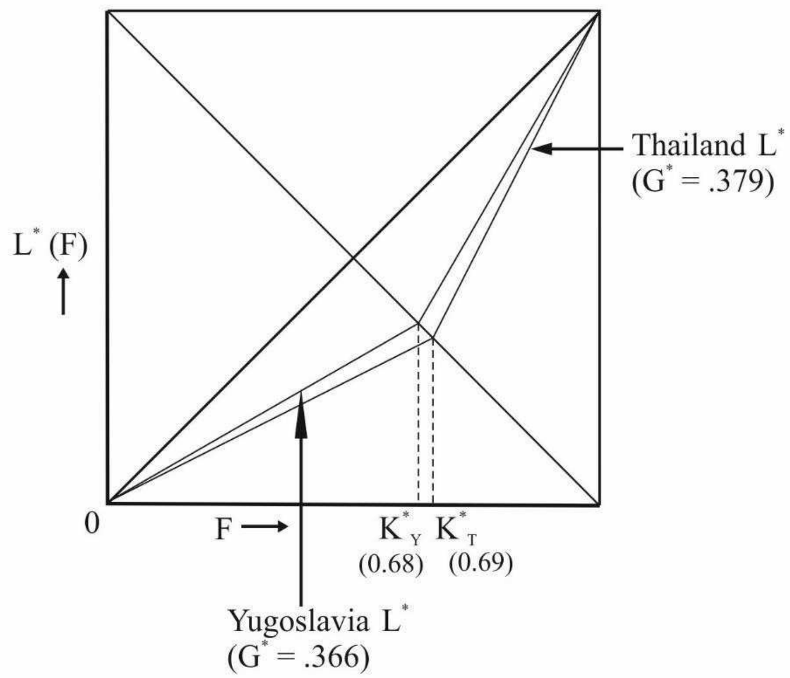


From the data provided in Table 1, the Lorenz curves for Thailand and Yugoslavia have been plotted in Figure 4, from which one can see that the Thai Lorenz curve is downwardly skewed and the Yugoslavian curve is upwardly skewed, with the two curves intersecting each other once. Further, the Kolkata Index, *K*, for Thailand, is 0.64, suggesting that the poorest 64 per cent of the population account for just 36 per cent of the income, while the K-value for Yugoslavia is smaller, at 0.62, suggesting that the poorest 62 per cent of the population own 38 per cent of the total income. Additionally, from estimates available in Table 17.1 of Kakwani (1980), we find that the Gini coefficients for Thailand and Yugoslavia are 0.385 and 0.347 respectively. Table 2 provides additional information on the values of the indices G^* and K^* respectively.

Notice from Table 2 that because the Thailand Lorenz curve is downwardly skewed, it is rewarded by a G^* value (0.379) which is less than its *G* value; and because the Yugoslav Lorenz curve is upwardly skewed, it is penalized by a G^* value (0.366) which is greater than its *G* value. It can be verified, through an appropriate application of Pythagoras' Theorem on right-angled triangles to the triangles OXM and OMN in Figure 3, that K^* and G^* are connected by the relation $K^* = (1+G^*)/2$. It turns out that the starred values of *G* and *K*, as the last row of Table 2 reveals, bring Thailand and Yugoslavia closer together in terms of inequality: this is the outcome of correcting for the skewness of the Lorenz curve.

Finally, the Kolkata Index, as we have seen, conveys an intuitively appealing interpretation of inequality in terms of the poorest *p*th fraction of the population earning a $(1-p)$th fraction of the total income. Unfortunately, however, the index *K* concentrates on a single point on the Lorenz curve (the point where the latter intersects the alternative diagonal of the unit square), whereas the Gini coefficient *G* takes the entire Lorenz distribution into account, making for a situation in which the inequality orderings of distributions by G and K, when the underlying Lorenz curves intersect, need not be identical. This problem is however overcome in the case of the indices G^* and K^* which will always rank distributions in a mutually identical fashion, and also identically with the Lorenz dominance orderings of the corresponding L^* curves. The L^* curves for Thailand and Yugoslavia are drawn in Figure 5. The fact that G^* and K^* rank distributions identically leaves us free to rank them by either indicator, and in particular by the indicator K^* which conveys the pithy information that once we have corrected for the skewness of the Lorenz curve, it is 'as if' the poorest 69 per cent of the urban Thai population earn 31 per cent of the total income—not vastly differently than in Yugoslavia where it is 'as if' the poorest 68 per cent of the population earn 32 per cent of the total income.

Figure 5: The L*-Curves for the Thailand and Yugoslavia Distributions of Figure 4



REFERENCES

Aaberge, R. (2009) 'Ranking Intersecting Lorenz Curves', *Social Choice and Welfare* **33(2)**, 235-259.

Chatterjee, A., A. Ghosh and B. K. Chakrabarti (2017) 'Socio-Economic Inequality: Relationship between Gini and Kolkata Indices', *Physica A* **466**, 583-595.

Kakwani, N. C. (1980) *Income Inequality and Poverty: Methods of Estimation and Policy Application*, Oxford University Press: New York.

Kakwani, N. C. and N. Podder (1976) 'Efficient Estimation of the Lorenz Curve and Associated Inequality Measures from Grouped Observations', *Econometrica* **44(1)**, 137-148.

Kolm, S-Ch. (1976) 'Unequal Inequalities II', *Journal of Economic Theory* **13(1)**, 82-111.

Lambert, P. and S. Subramanian (2015) 'Shaikh and Ragab's "Incomes of the Vast Majority": Some Additions and Extensions', *Sankhya (Series B)* **77(2)**, 359-370.

Shorrocks, A. F. and J. E. Foster (1987) 'Transfer Sensitive Inequality Measures', *Review of Economic Studies* **54(3)**, 485-497.

Subramanian, S. (2010) 'Tricks with the Lorenz Curve', *Economics Bulletin* **30(2)**, 1594-1602.

Subramanian, S. (2015) 'More Tricks with the Lorenz Curve', *Economics Bulletin* **35(1)**, 580-589.