Economics Bulletin

Volume 30, Issue 2

Tricks with the lorenz curve

Subramanian Sreenivasan
Madras Institute of Development Studies

Abstract

This note develops, for the Gini coefficient of inequality, a very simple generalization that directly incorporates judgments on 'relative inter-group inequality aversion' by making the inequality measure sensitive to the skewness of the Lorenz curve. The resulting family of inequality indices can be seen as complements to the Gini coefficient: some members of the family reflect 'left-leaning', and others 'right-leaning', distributional values relative to the 'centrist' position assumed by Gini.

This note is based on the text of an Endowment Lecture delivered by the author in honour of Professor R. Elango, at the Annamalai University in Tamil Nadu, India, on February 1st 2010. It owes much to the late Professor M. N. Murthy, who suggested, several years ago in conversation with the author, the inequality measure mentioned as a 'special case' in these notes: his suggestion has played a crucial role in motivating the entire text. For this, and his customary kindness in commenting on an earlier draft (which goes back many years), many grateful thanks are owed. Thanks are also due to William Neilson for helpful suggestions, to D. Jayaraj for constructive discussions on the subject of this note, and to A. Arivazhagan for the graphics. For providing me with an occasion to return to a problem which I had set aside for years, I must thank N. Ramagopal. The usual disclaimer applies.

Citation: Subramanian Sreenivasan, (2010) "Tricks with the lorenz curve", *Economics Bulletin*, Vol. 30 no.2 pp. 1594-1602.
Submitted: Apr 15 2010. **Published:** June 02, 2010.

TRICKS WITH THE LORENZ CURVE[1]

1

The aim of this note is to present an extremely simple generalization of the Gini coefficient of inequality – a generalization that displays sensitivity to the skewness of the Lorenz curve. The accent is on the word 'simple', and, indeed, the indices advanced here may well attract the charge, even, of simple-mindedness. The emphasis will be entirely on the intuitive plausibility of a certain line of reasoning, and there can be nothing of interest here for the scholar who insists that an impeccable axiomatic rationalization is an indispensable accompaniment of any proposed real-valued measure of inequality. There are no axiomatics in this essay; what there is, it seems to me, is an intriguing little curiosum, which may be worth the effort of deeper investigation. A mitigating factor for the unrefined simplicity of this note resides in the fact - as anybody who has had anything to do with Lorenz curves will appreciate - that it is difficult to resist the temptation of getting up to tricks of one kind of another in the presence of the seemingly infinite possibilities offered up by the curve.

2

A well-known shortcoming of the Gini coefficient of inequality (G) is its inability to take account of the skewness of the Lorenz curve in assessing the extent of inequality in a distribution. Specifically, G will pronounce as equally unequal all distributions for which the areas enclosed by their respective Lorenz curves and the diagonal of the unit square are the same. This raises the question: should inequality among the relatively poorer income groups – as would be reflected in a Lorenz curve skewed toward (1,1) of the unit square - be regarded as more pernicious than inequality among the relatively richer groups – as would be reflected in a Lorenz curve skewed toward (0,0)? A 'left-leaning' ideology would probably uphold this point of view, since in a Lorenz curve skewed toward (1,1) (that is to say, a curve which 'bulges at the bottom'), the income-share of the poorer income groups is lower than in a Lorenz curve skewed toward (0,0) (that is to say, a curve which 'bulges at the top') and which encloses the same area between itself and the diagonal as does the former curve. A 'right-leaning' ideology would uphold the converse point of view, and a 'centrist' ideology would regard inequality among the poorer income groups to be neither better nor worse than inequality among the richer groups[2].

Amartya Sen (1973; p.36) puts the matter thus:

[1] Readers will recognize that the title has been lifted from earlier work by Gorman (1976).

[2] The terms 'left-leaning', 'centrist', and 'right-leaning' are reminiscent of Serge-Christophe Kolm's 'leftist', 'centrist', and 'rightist' designations for inequality measures. The contexts of use, however, are entirely different, Kolm's concerns being with the relative merits of the so-called 'scale invariance' and 'translation invariance' properties of inequality indices - see S-C. Kolm (1976a and 1976b). The sense in which the terms 'left', 'right' and 'centre' are employed in this note has, in fact, much to do with what Kolm (*op. cit.*) has called the 'principle of diminishing transfers'.

> Can it be asserted that our judgment of the extent of inequality will not vary according to whether the people involved are generally poor or generally rich? Some have taken the view that our concern with inequality increases as a society gets prosperous since the society can 'afford' to be inequality-conscious. Others have asserted that the poorer an economy, the more 'disastrous' the consequences of inequality, so that inequality measures should be sharper for low average income. This is a fairly complex question and is bedeviled by a mixture of positive and normative considerations. The view that for poorer economies inequality measures must be themselves sharper can be contrasted with the view that greater *importance* must be attached to any given inequality measure if the economy is poorer. The former incorporates the value in question into the measure of inequality itself, while the latter brings it in through the evaluation of the relative importance of a given measure at different levels of average income.

My own concern here will be with the first of the two types of exercise that Sen alludes to at the conclusion of the quoted passage. I shall advance a specific approach to the incorporation, into the inequality measure itself, of alternative values relating to whether or not the measure should be sharper for poorer societies. This approach proposes a family of inequality measures, $G(\lambda)$, where λ is in the nature of an 'indicator of inter-group inequality aversion'. For parametric variation in λ (within bounds that will be discussed later), $G(\cdot)$ will be seen to exhibit varying degrees of sensitivity to the skewness of the Lorenz curve[3].

3

There are different notational ways of representing the Lorenz curve, and in setting out the preliminary formalities, it is of great assistance to draw on the work of Nanak Kakwani (1980b). One can begin by letting x stand for a random variable designating (say) income, distributed, with mean μ, over the interval $[0, \bar{x}]$. $f(x)$ is the density function of x (the proportion of the population with income x), $F(x)$ the cumulative density function (the cumulative proportion of the population with incomes not exceeding x), and $F_1(x)$ the first-moment distribution function (the cumulative share in income of the population with incomes not exceeding x):

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

$$F_1(x) = \left(\frac{1}{\mu}\right)\int_0^x yf(y)dy;$$

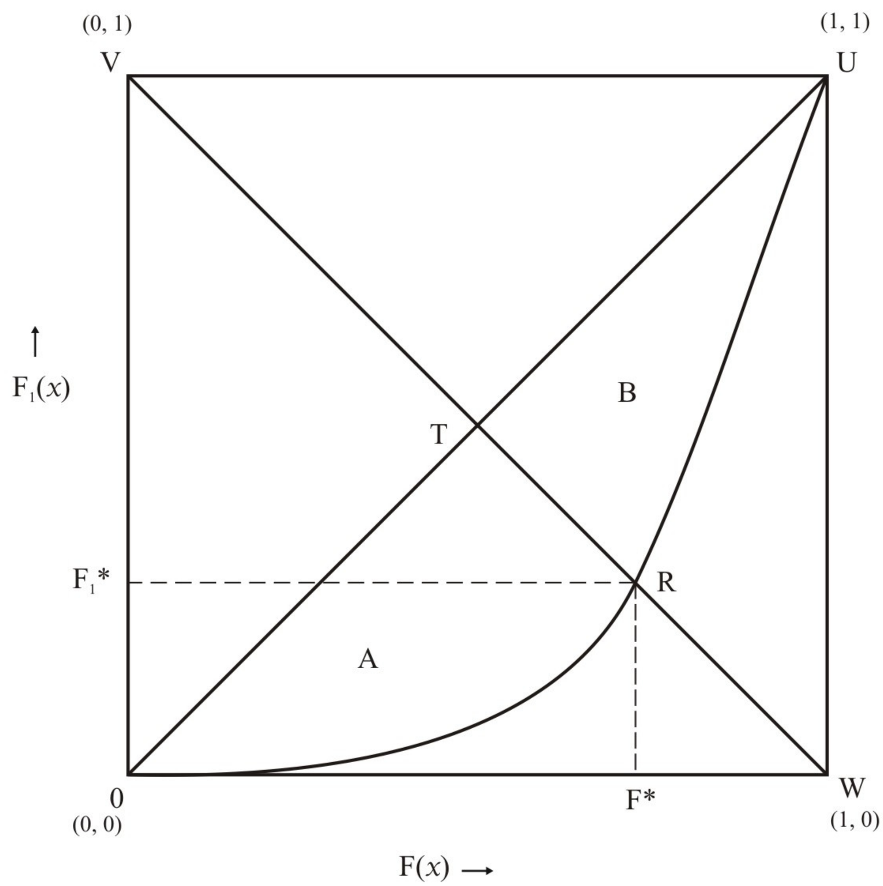
$Lim_{x\to 0}F(x) = Lim_{x\to 0}F_1(x) = 0$; and $Lim_{x\to \bar{x}}F(x) = Lim_{x\to \bar{x}}F_1(x) = 1$.

The Lorenz curve is simply the functional relational between $F_1(x)$ and $F(x)$ and is drawn in Figure 1: as plotted in the unit square, the curve for an unequal distribution, typically, would be an increasing and strictly convex one, running from (0,0) to (1,1) of

[3] For alternative generalizations of the Gini coefficient, see Kakwani (1980a); Donaldson and Weymark (1980); and Yitzhaki (1983).

the square; for an equal distribution, the Lorenz curve would coincide with the 45° line, or diagonal of the unit square, which is the 'line of equality'.

Figure 1 : The Lorenz Curve



Note : $G^M = G + (A - B)$.

Consider a measure of central tendency – call it $x*$ - such that the poorest $F(x*)$ proportion of the population earn $(1-F(x*))$ proportion of the total income. It is convenient to regard $x*$ in the light of a certain distinguished 'relative' poverty line which separates the relatively poorer segment of the population from its relatively richer segment. $x*$, clearly, is the income level corresponding to which the Lorenz curve (see Figure 1) intersects the diagonal drawn from (0,1) to (1,0) of the unit square (this diagonal is referred to by Kakwani 1980b as the 'alternative diagonal'). In what follows, I shall simplify the notation somewhat: the cumulative proportion of the population with incomes no higher than $x*$, $F(x*)$, will be written as just $F*$, while the cumulative income share of the poorest F^* proportion of the population, $F_1(x^*)$, will be written as just $F_1{}^*$.

The Gini coefficient of inequality in the distribution of income is given by the following well-known expression (again see Kakwani 1980b for a derivation):

$G = 1 - 2\int_0^{\bar{x}} F_1(x) f(x) dx$, or, equivalently,

$$G = 1 - 2[\int_0^{x^*} F_1(x) f(x) dx + \int_{x^*}^{\bar{x}} F_1(x) f(x) dx]. \tag{1}$$

Letting G_1 and G_2 stand for the Gini coefficients of inequality in the distribution of income among, respectively, those with incomes not exceeding x^* and those with incomes exceeding x^*, it is routine to note that

$G_1 = 1 - 2\int_0^{x^*} F_1(x) f(x) dx / F^* F_1{}^*$, so that

$$2\int_0^{x^*} F_1(x) f(x) dx / F^* F_1{}^* = (1 - G_1) F^* F_1{}^*; \tag{2}$$

and

$G_2 = 1 - 2[\int_{x^*}^{\bar{x}} F_1(x) f(x) dx - F^*(1 - F^*)]/(1 - F^*)(1 - F_1{}^*)$, so that

$$2\int_{x^*}^{\bar{x}} F_1(x) f(x) dx = (1 - G_2)(1 - F^*)(1 - F_1{}^*) + 2F^*(1 - F^*). \tag{3}$$

From (2) and (3), and making use of the fact that since $(F^*, F_1{}^*)$ is a point on the alternative diagonal of the unit square, so that $F^* + F_1{}^* = 1$, one has:

$$2\int_0^{x^*} F_1(x) f(x) dx = F^*(1 - F^*)(1 - G_1); \tag{4}$$

and

$$2\int_{x^*}^{\bar{x}} F_1(x) f(x) dx = F^*(1 - F^*)(1 - G_2) + 2(1 - F^*)^2. \tag{5}$$

Making the appropriate substitutions from (4) and (5) into (1) yields:

$G = 1 - [F^*(1 - F^*)(1 - G_1) + F^*(1 - F^*)(1 - G_2) + 2(1 - F^*)^2]$

which, upon simplification, can be written as

$$G = 1 - 2(1 - F^*)[1 - F^*\left(\frac{1}{2}G_1 + \frac{1}{2}G_2\right)]. \tag{6}$$

Writing G in the form of (6) has the advantage of bringing sharply into focus the fact that Gini can be written as a weighted sum of G_1 and G_2, with the weights on G_1 and G_2 being identical. This paves the way for what one could see to be a very natural generalization of G to a (parametrized) family of measures that we may call $G(\lambda)$, and which is given by

$$G(\lambda) = 1 - 2(1 - F^*)[1 - F^*\left(\lambda G_1 + (1 - \lambda) G_2\right)], \lambda \in [0,1]. \tag{7}$$

λ, appearing on the Right Hand Side of (7), may be regarded as an indicator of 'relative inter-group inequality aversion', the groups in question being constituted by, respectively, the relatively poor population and the relatively rich population. In principle, one could let λ take any value in the closed interval [0,1]. For $\lambda \in [0,1/2)$, we would have a continuum of (decreasingly) 'right-leaning' inequality measures. For $\lambda = 1/2$, we would have a 'centrist' inequality measure – which is, precisely, the Gini coefficient. For $\lambda \in (1/2,1]$, we would have a continuum of (increasingly) 'left-leaning' inequality measures. In assessing overall inequality, λ and $(1-\lambda)$ are the relative weights assigned to inequality among the 'poorer' and the 'richer' groups respectively, the two groups, to recall, being separated from one another by the 'relative poverty line' x^*. $\lambda = 1$ would correspond to Rawls' maximin rule (only the claim of the worse-off of the two groups matters), and $\lambda = 0$ would correspond to an extreme anti-Rawlsian 'maximax' rule (only the claim of the better-off of the two groups matters). In general, an 'egalitarian' bias is injected by confining λ to the interval $(1/2,1]$. Note that, for $\lambda > 1/2$,
$G(\lambda) - G = 2F^*(1-F^*)(\lambda - 1/2)(G_1 - G_2) >,<$ or $= 0$ according as $G_1 >,<$ or $= G_2$: $G(\lambda)$ exceeds G if the Lorenz curve is skewed toward (1,1) of the unit square, is less than G if the Lorenz curve is skewed toward (0,0), and coincides with G if the Lorenz curve is symmetric. This property realizes the intuitive appropriateness (if one is left-leaning!) of penalizing a distribution for which the Lorenz curve is skewed toward (1,1) of the unit square relative to a distribution for which it is skewed toward (0,0). From a purely pragmatic point of view, an advantage with the proposed generalization of Gini is that the policy maker may not find it terribly hard to 'understand the meaning' of λ: λ is, after all, in the nature of a straightforward proportional weight placed on the extent of inequality among the 'poorer' income groups[4].

Additionally, there is a particular value of λ - ¾ as it happens - which yields up a very simple, intuitively clear, and visually appealing left-leaning variant of the Gini coefficient, which is greater than, equal to, or less than G according as the Lorenz curve is skewed toward (1,1), symmetric, or skewed toward (0,0). The index $G(3/4)$ may be called G^M, the superscript 'M' standing for (Professor M. N.) Murthy, who proposed the index to the author in personal conversation. The interesting feature about the index G^M is that it is very simply given (I desist here from providing a demonstration) by: $G^M = G + (A - B)$ where – see Figure 1 - A is the area enclosed by the diagonal of the unit square and the Lorenz curve, marked ORT, to the left of the alternative diagonal, and B is the area enclosed by the diagonal of the unit square and the Lorenz curve, marked TRU, to the right of the alternative diagonal.

Finally, a swift and simple example may help to illustrate the concerns of this note, and also possibly assist in giving content to the distinction between 'left-leaning'

[4] In a minor aside, it may be noted that this relative 'interpretation-friendliness' is perhaps slightly in contrast with the status that obtains in many standard formulations of inequality aversion, in which the magnitude of this quantity is sought to be captured by the value of the exponent in a power function. It *is* a bit hard to imagine the average lay policy-maker feeling entirely comfortable about making informed judgments on the degree of convexity s/he wishes, through restrictions on the value of the exponent, to impart to the curvature of an underlying social welfare function!

and 'right-leaning' inequality judgments. Imagine a population of 1 million persons with a mean income of 2 Rupees, and let x_i be the income of the *i*th poorest person. Consider two distributions **x** and **y** respectively, such that, in distribution **x**:

$x_i = 0 \forall i = 1,...,500,000;$

$\quad = 4 \forall i = 500,001,...,1000,000;$

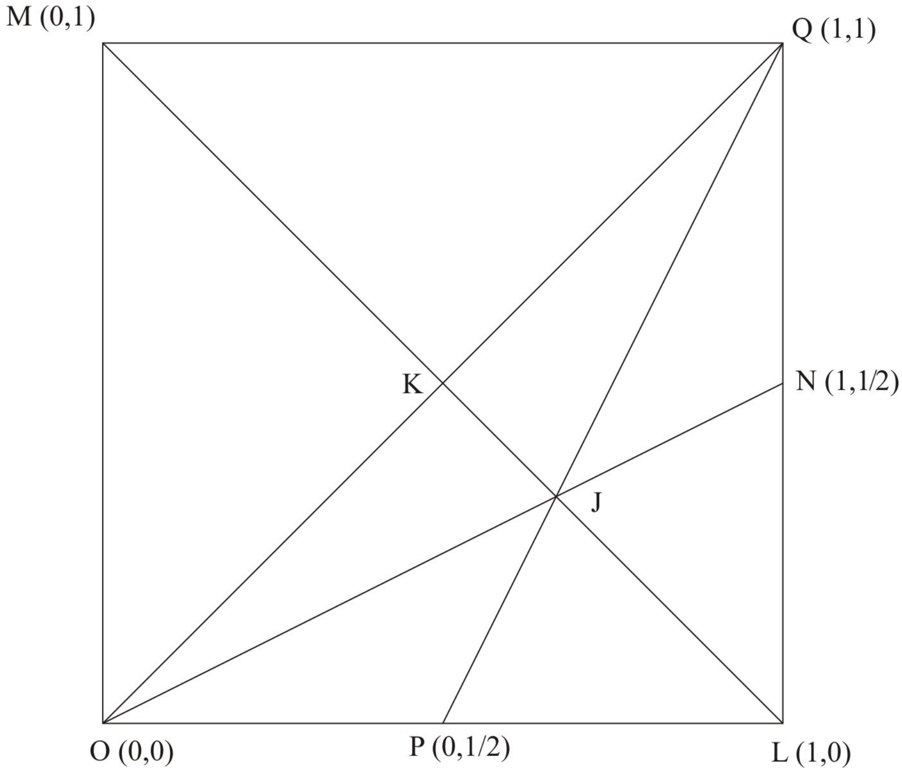
and in distribution **y**:

$x_i = 1999,900/999,999 (\cong 1.9999) \forall i = 1,...,999,999;$

$\quad = 100$ for $i = 1000,000$.

It is a simple matter to see that the Lorenz curve for distribution **x** – see Figure 2 – is given by the curve OPQ (where P is the point (1/2,0)), while the Lorenz curve for distribution **y** is given by the curve ONQ (where N is the point (1,1/2)). Clearly, the Lorenz curve for **x** is skewed toward (1,1) of the unit square, while the Lorenz curve for **y** is skewed toward (0,0). Noting that the Gini coefficient of inequality is twice the area enclosed by the Lorenz curve and the diagonal of the unit square, it follows that the Gini coefficients for the distributions **x** and **y** are given, respectively, by G(**x**) = 2AreaOPQ = 2(AreaOPJK + AreaKJQ), and G(**y**) = 2AreaONQ = 2(AreaOJK + AreaKJNQ). It is straightforward to see that AreaOPJK = AreaKJNQ $(\equiv \Delta_1, say)$, while AreaKJQ = AreaOJK $(\equiv \Delta_2, say)$. That is, the Gini coefficients for the two distributions are identically the same, at, say, G. Now the index G^M for the two distributions will be given, repectively, by G^M(**x**) = G + AreaOPJK – AreaKJQ = $G + (\Delta_1 - \Delta_2)$, and G^M(**y**) = G + AreaOJK – AreaKJNQ = G – (AreaKJNQ – AreaOJK) = $G - (\Delta_1 - \Delta_2)$, so that if we designate $(\Delta_1 - \Delta_2)(> 0)$ by Δ, we have: G^M(**x**) = $G + \Delta$ > G^M(**y**) = $G - \Delta$: the 'left-leaning' inequality index G^M will penalize the distribution (**x**) for which the Lorenz curve is skewed toward (1,1) of the unit square vis-à-vis the distribution (**y**) for which the Lorenz curve is skewed toward (0,0). Why might this be a reasonable outcome? To see this, consider the following circumstantial details which, while they may not be strictly necessary for the argument, may yet assist in comprehending its thrust. Suppose 1 Rupee is a poverty line such that those with incomes below this line are certified as being absolutely impoverished. Notice that in distribution **y** *no person is poor*, though a vast majority of the population are very much worse off than a single extremely rich individual. In contrast, 50 per cent of the population in distribution **x** are not just poor *but wholly destitute*, even though the majority of the non-poor in **x** are better off than in **y**. If our focus is on the poorer population and our differential sympathy resides with this section of the population, then it would be natural to deem **x** as a worse distribution than **y**. This, precisely, is the inequality judgment incorporated in the 'left-leaning' measure G^M.

Figure 2: Distributions with Skewed Lorenz Curves sharing the Same Gini Coefficient



Note : The Lorenz curve OPQ is skewed toward (1,1), while the curve ONQ is skewed toward (0,0). Both distributions share the same value of the Gini coefficient.

4

It remains to summarize and conclude. In this note, I have explored an elementary approach to sensitizing an inequality measure to the skewness of the Lorenz curve. This approach develops a simple and natural generalization $G(\lambda)$ of the Gini coefficient of inequality G. While Gini is insensitive to whether inequality is more pronounced at the lower or the upper end of the distribution, its generalized form $G(\lambda)$ permits, through selection of appropriate ('right-leaning'/'centrist'/'left-leaning') values for the parameter λ, judgments on 'relative inter-group inequality aversion' to be directly incorporated in the inequality measure. A special case of $G(\lambda)$ has been proposed. This is the index G^M. G^M, which is an uncomplicated measure that makes a direct visual appeal to the beholder, penalizes a Lorenz curve skewed toward (1,1) of the unit square relative to a curve skewed toward (0,0).

REFERENCES

Donaldson, D. and J. A. Weymark (1980) "A Single-Parameter Generalization of the Gini Indices of Inequality", *Journal of Economic Theory*, 22(1), 67-86.

Gorman, W. M. (1976) "Tricks with Utility Functions" in *Essays in Economic Analysis* by M. Artis and R. Nobay, Eds., Cambridge University Press: Cambridge.

Kakwani, N. C. (1980a) "On a Class of Poverty Measures", *Econometrica*, 48(2), 437-446.

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

Kolm, S. C. (1976a) "Unequal Inequalities: I", *Journal of Economic Theory*, 12(3), 416-442.

Kolm, S. C. (1976b) "Unequal Inequalities: II", *Journal of Economic Theory*, 13(1), 82-111.

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

Yitzhaki, S. (1983) "On an Extension of the Gini Inequality Index", *International Economic Review*, 24(3), 617-628.