JOURNAL OF INCOME DISTRIBUTION

Volume 10, Number 1-2, Spring-Summer 2001

CONTENTS

HOW TO DECOMPOSE THE SEN-SHORROCKS-THON POVERTY INDEX: A PRACTITIONER'S GUIDE

Kuan Xu and Lars Osberg
Dalhousie University

The Sen-Shorrocks-Thon index of poverty intensity is a useful measure of poverty intensity, but the link between the index and its decompositions is not necessarily transparent. This guide documents our experience using the index and is written to assist other policy analysts to compute and understand this index and its decomposition with simple mathematics, numerical examples and geometrical figures.

Keywords: Poverty intensity, Poverty indices, Order statistics

1. Introduction

The Sen-Shorrocks-Thon index (or the SST index hereafter) is an index first proposed by Sen (1976) and later modified by Shorrocks (1995), which can also be shown to be consistent with Thon's index (1979, 1983) in the limit. While Xu (1998) examined the statistical issues relating to the SST index, Osberg and Xu (1997, 1998, 2000), in international and regional comparative empirical studies, proposed a decomposition of the SST index and a method of bootstrap inference suitable for survey data with sampling weights. The focal point of those papers was on empirical evaluation of poverty intensity and social policy and little space was provided to detailed technical issues related to the SST index and its decomposition.

The SST index and its decomposition are useful in economic research on poverty issues,[1] but the link between the SST index and its decomposition is not necessarily transparent. If the average poverty gap ratios and the Gini index of poverty gap ratios are not computed properly, the SST index decomposition will be difficult to understand or will appear not feasible. This practitioner's guide therefore explains in detail the index and its decomposition using simple mathematics, numerical examples, and geometrical figures.[2]

Direct all correspondence to: Kuan Xu, Department of Economics, Dalhousie University, Halifax, NS, Canada B3H 3J5.

The remainder of this note is organized as follows. In section 2, we discuss why the decomposition works and what are the critical conditions to pay attention to. In section 3, we present some numerical examples and geometrical interpretation of our analytical results. Finally, a conclusion is provided in section 4.

2. The SST Index and Its Decomposition

Let incomes (y_i's) for a population of size n be sorted from the smallest to the largest so that $y_1 \leq y_2 \leq \cdots \leq y_n$. Let the poverty line be z and the number of the individuals whose incomes are less than the poverty line z be q (out of n). For individuals $i = 1,2, \cdots n$, the poverty gap ratio x_i (or relative income shortfall below the poverty line) of the ith individual is defined as

$$x_i = \frac{z - y_i}{z} \qquad \text{if } \frac{z - y_i}{z} \geq 0, \tag{1}$$

$$x_i = 0 \qquad \text{if } \frac{z - y_i}{z} < 0.$$

That is, when an individual is poor, his or her poverty gap ratio is the relative income shortfall below the poverty line—he or she has a positive deprivation; when an individual is one of the nonpoor, his or her poverty gap ratio is set to zero—he or she has zero deprivation.

In Shorrocks (1995), the SST index of poverty intensity is proposed as

$$P(y;z) = \frac{1}{n^2}\sum_{i=1}^{n} (2n - 2i + 1)\frac{z - y_i}{z} \tag{2}$$

or

$$P(y;z) = \frac{1}{n^2}\sum_{i=1}^{n} (2n - 2i + 1)x_i. \tag{3}$$

Shorrocks (1995) indicates the following without giving details[3]

$$P(y;z) = \mu(x)(1 + G(x)), \tag{4}$$

where $\mu(x)$ and $G(x)$ are the average and the Gini index, respectively, for the distribution of poverty gap ratios (x is used to denote the ratios collectively).

The literature has often argued that the poverty rate (or the headcount ratio) and the average poverty gap ratio (or the income gap ratio) of the poor violate several important axioms (in particular, the Principle of Transfers, that an acceptable measure of poverty should always increase if resources are taken away from a poor person and given to a richer person) that are essential in poverty measurement.[4] However, these measures have been widely used.

Neither the poverty rate nor the poverty gap is separately a good measure of poverty, but considered together they can account for a good deal. Based on the Osberg-Xu decomposition of the SST index and international and regional empirical evidence (1997, 1998, 2000), the sum of the percentage changes in the poverty gap ratio and the average poverty gap ratio can jointly capture almost all of the percentage changes in the poverty intensity over time.

Because of the importance of the SST index decomposition, it is necessary to examine equation (4) as the basis of further decomposition. First, the average of poverty gap ratios

$$\mu(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5}$$

in equation (4) can be decomposed into a product of the poverty rate

$$H = \frac{q}{n}, \tag{6}$$

and the average poverty gap ratio (or the income gap ratio) of the poor[5]

$$I = \frac{1}{q}\sum_{i=1}^{n} x_i. \tag{7}$$

Second, the measure of equality of poverty gap ratios for the population is $(1 + G(x))$ in which the Gini index $G(x)$ is defined as:[6]

$$G(x) = 1 - \sum_{i=1}^{n} \left(\frac{1}{n}\right)\left(\frac{\sum_{j=1}^{i} x_{(j)} + \sum_{\substack{i\geq 2 \\ j=2}}^{i} x_{(j-1)}}{n\mu(x)}\right) \tag{8}$$

or

$$G(x) = 1 - \frac{1}{n^2\mu(x)}\sum_{i=1}^{n} (2n - 2i + 1)x_{(i)}$$

with the x_i's being sorted in increasing order[7]

$$x_{(i)} \leq x_{(2)} \leq \cdots \leq x_{(n)}. \tag{9}$$

Thus, equation (4) can be rewritten as

$$P(y;\, z) = HI(1+\mathrm{G}(x)). \tag{10}$$

This decomposition expresses poverty intensity as a product of the poverty rate, the average poverty gap ratio of the poor, and the measure of the inequality of the poverty gap ratios of the population. This decomposition also allows economists to view the sources of a change in the measure of poverty intensity in terms of changes in the poverty rate, the average poverty gap ratio, and the (in)equality of the poverty gap ratios. More specifically, taking the natural logarithm of both sides of equation (10) gives

$$\ln P(y;z) = \ln H + \ln I + \ln(1+\mathrm{G}(x)), \tag{11}$$

where $\ln\,(1+G(x))$ is an approximate of $G(x)$ based on the first-order Taylor series expansion. Let $\Delta A = A - A_{-1}$, where A_{-1} is the A measured at the previous period. This is applicable to any first-order difference. Then equation (11) can be transformed into

$$\Delta \ln\ P(y;z) = \Delta \ln\ \ H + \Delta \ln\ I + \Delta \ln\,(1 + G(x)), \tag{12}$$

where $\Delta\ln\,(1+G(x))$ is an approximation of $\Delta G(x)$. Equation (12) indicates that the percentage change in $P(y;z)$ is the sum of the percentage changes in H, I, and $G(x)$.

To explain how to derive equation (4) [or (10)] using equation (8), we take the following four steps.

First, discuss the income sequence y (y_i's), the unsorted poverty gap ratio sequence x'(x_i's), and the sorted poverty gap ratio sequence x ($x_{(i)}$'s). Note that in the income sequence y (y_i's) incomes are arranged in increasing order, i.e.,

$$y_1 \le y_2 \le \cdots \le y_n. \tag{13}$$

As poverty gap ratios x_i's are given by equation (1), poverty gap ratios in the unsorted poverty gap ratio sequence x' (x_i's) are arranged in decreasing order, i.e.,

$$x_1 \ge x_2 \ge \ldots \ge x_q \ge x_{q+1} = \cdots = x_n = 0. \tag{14}$$

Note that the nonpoor have zero poverty gap ratios. Poverty gap ratios in the sorted poverty gap ratio sequence x ($x_{(i)}$'s) are arranged in increasing order, i.e.,

$$0 = x_{(1)} = \cdots = x_{(n-q)} \le x_{(n-q+1)} \le \cdots \le x_{(n-1)} \le x_{(n)}. \tag{15}$$

Second, note that sorting data in sequence x' to get sequence x does not affect the absolute value but alters the sign of the Gini index; that is

$$G(x) = -G(x'). \tag{16}$$

The detailed explanation for equation (16) is given in Appendix A.

Third, given the above results equation (4), therefore equation (10), is valid.

$$P(y,z) = \mu(x)(1-G(x')), \tag{17}$$

is valid because

$$\begin{aligned} \mu(x)(1-G(x')) &= \mu(x)\left(1-1+\tfrac{1}{n^2\mu(x)}\sum_{i=1}^{n}(2n-2i+1)x_i\right) \\ &= \tfrac{1}{n^2}\sum_{i=1}^{n}(2n-2i+1)x_i \\ &= P(y;z). \end{aligned} \tag{18}$$

Finally, it can be concluded that equation (17) can be written as equations (4) and (10) because of equation (16).

3 Numerical Examples and Geometric Explanations

In this section, we illustrate three examples. Example One is used to show that $G(x) = -G(x')$. Example Two is designed to show the SST index decomposition $P(y;z) = HI(1 + G(x))$. Example Three is used to show that the arc portion of the deprivation profile is an inverted Lorenz curve of poverty gap ratios of the poor geometrically. Appendix B presents a GAUSS program that computes the SST index and its decomposed components. The GAUSS program shows the computations for Examples Two and Three.

3.1 *Example One*

This example is used to show the impact of data sorting on the sign and value of the Gini index. First of all, it is useful to show geometrically and numerically that $G(x) = -G(x')$. Assume that the hypothetical data are $x_{(1)} = 0$, $x_{(2)} = 1$, $x_{(3)} = 2$ and $n = 3$.[8] Let us look at the problem geometrically. It is well-known that the Gini index can be defined as Figure 1 as the ratio of Area A to Area $A + B$, i.e.,

$$\begin{aligned} G(x) &= \tfrac{A}{A+B} \\ &= \tfrac{2A}{2(A+B)} \\ &= 2A \\ &= 1-2B \end{aligned} \tag{19}$$

As shown in Figure 2, for the data, Area B is the sum of the area of a small triangle 1/18, the area of a square 1/9, and the area of a large triangle 1/9. Thus, according to equation (19),

Figure 1

Lorenz Curve and the Gini Index

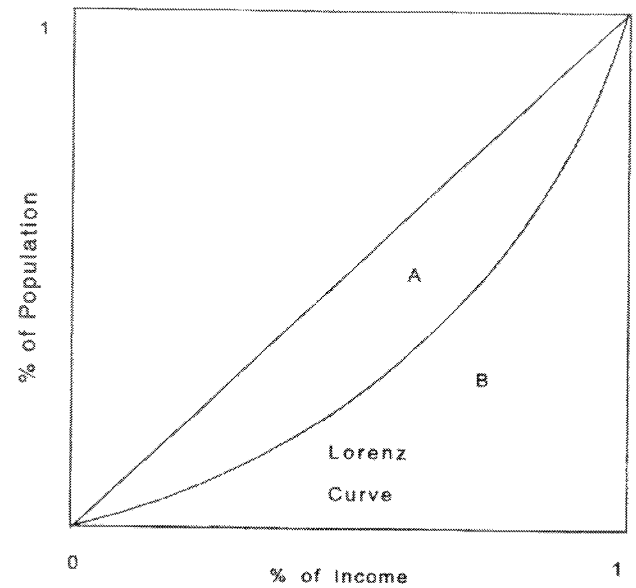
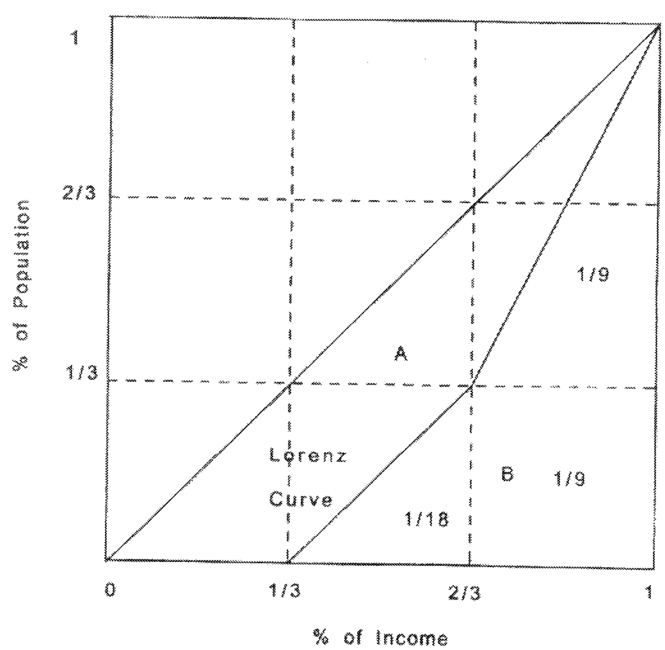


Figure 2

Numerical Example One—Area B for the Gini Index



$$\begin{aligned} G(x) &= 1-2\left(\tfrac{1}{18}+\tfrac{1}{9}+\tfrac{1}{9}\right) \\ &= 1-2\tfrac{5}{18} \\ &= \tfrac{4}{9}. \end{aligned} \tag{20}$$

Alternatively, one can compute $G(x)$ numerically using equation (8) and the same data as follows.

$$\begin{aligned} G(x) &= 1-\left(\tfrac{1}{3}\right)\left(\tfrac{1}{3}\right)((0)+(1+0)+(3+1)) \\ &= 1-2\tfrac{5}{18} \\ &= \tfrac{4}{9} \end{aligned} \tag{21}$$

Clearly, the above two approaches give the same result.

If the x_i's are sorted in decreasing order so that $x_1 = 2$, $x_2 = 1$, and $x_3 = 0$, $G(')$ can be computed using equation (8):

$$\begin{aligned} G(x') &= 1-\left(\tfrac{1}{3}\right)\left(\tfrac{1}{3}\right)((2)+(3+2)+(3+3)) \\ &= 1-\tfrac{13}{9} \\ &= -\tfrac{4}{9}. \end{aligned} \tag{22}$$

Thus, $G(x) = -G(x')$. This example shows that the way in which the data are sorted will not change the absolute value of the Gini index but will alter its sign.

3.2 *Example Two*

This example is used to show the deprivation profile and the SST index decomposition. Let the income data sorted in increasing order be $y_1 = 3$, $y_2 = 9$, $y_3 = 11$, and $y_4 = 15$ and the poverty line z be 10. Using equation (1), we can get $x_1 = 0.7$, $x_2 = 0.1$, $x_3 = 0$, and $x_4 = 0$. Using equation (3) we can compute $P(y;z)$

$$\begin{aligned} P(y;z) &= \tfrac{1}{n^2}\sum_{i=1}^{n}(2n-2i+1))x_i \\ &= \tfrac{1}{16}(7(0.7)+5(0.1)) \\ &= 0.3375. \end{aligned} \tag{23}$$

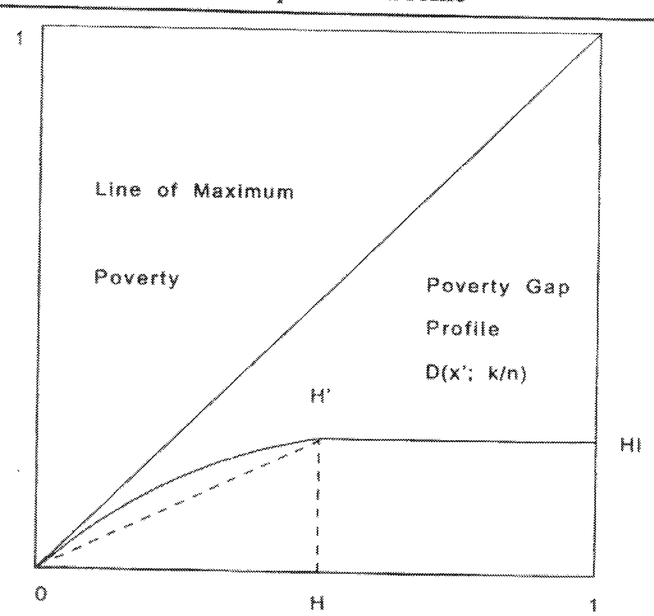
Like the Gini index, the SST index also has a simple geometric interpretation. The deprivation profile for discrete relative deprivation measures[9] $D(.,.)$ is a function of k/n for a given data set x':

$$D\left(x';\tfrac{k}{n}\right)=\tfrac{1}{n}\sum_{i=1}^{k} x_i,$$

where x' refers the sequence of x_i's arranged in *decreasing* order.[10] As can be seen in Figure 3, the deprivation profile curve starts from the origin reaching H' and, after that point, it becomes horizontal. The point H indicates the poverty rate while HI gives the average poverty ratio of the population. The average poverty ratio of the poor, I, is represented by the slope of dotted line OH' because $HI/H = I$. The arc OH' is similar to a Lorenz curve. The degree that the

Figure 3

The Deprivation Profile



arc OH' departs from the dotted OH' implies the degree of inequality of deprivation measures. The SST index can be viewed as the ratio of the area under the poverty gap profile to the area under the line of maximum poverty (the 45 degree line) as shown in Figure 3.

For the same data set, the poverty rate is $H = 0.5$, the average poverty gap ratio of the poor is $I = 0.4$, and the average poverty gap ratios of the population is $\mu(x) = HI = 0.2$. Based on the sorted x_i's ($x_{(1)} = 0$, $x_{(2)} = 0$, $x_{(3)} = 0.1$, and $x_{(4)} = 0.7$), the Gini index is given by

$$\begin{aligned} G(x) &= 1 - \sum_{i=1}^{n}\left(\tfrac{1}{n}\right)\left(\frac{\sum_{j=1}^{i} x_{(j)} + \sum_{\substack{i\geq 2\\ j=2}}^{i} x_{(j-1)}}{n\mu(x)}\right) \\ &= 1 - \tfrac{1}{4}\left(\tfrac{0.1+0.9}{0.8}\right) \\ &= 0.6875. \end{aligned} \quad (24)$$

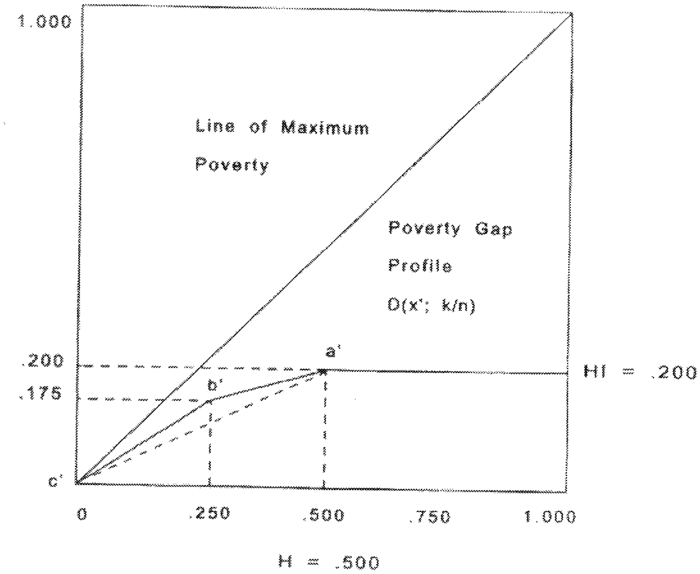
Thus,

$$\begin{aligned} P(y;z) &= HI(1+G(x)) \\ &= (0.2)(1+0.6875) \\ &= 0.3375. \end{aligned} \quad (25)$$

The results from equations (23) and (25) are identical. Thus, the SST index computed directly using equation (2) is equivalent to that computed indirectly by the product of its three decomposed components. In Figure 4, the deprivation profile curve starts from the origin reaching a' and, after that point, it

Figure 4

Numerical Example Two—A Deprivation Profile

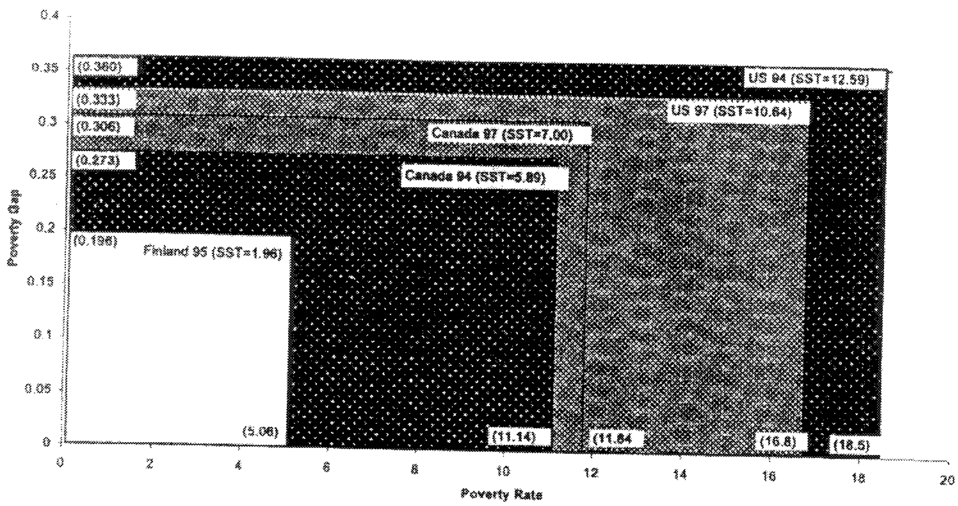


becomes horizontal. As expected, the point $H = 0.5$ gives the poverty rate while $HI = 0.2$ gives the average poverty ratio of the population. The average poverty ratio of the poor, I, is represented by the slope of dotted line $c'a'$ because $HI/H = I = 0.4$. The arc $c'b'a'HI$ is similar to a Lorenz curve which will be explained further in Example Three. The SST index can be viewed as the ratio of the area under the poverty gap profile (0.16875) to the area under the line of maximum poverty (0.5), that is 0.3375.

Using the international income survey data, Osberg and Xu (1997, 1998, 2000) found that the percentage changes in $(1 + G(x))$ are relatively small over time. Therefore, percentage changes in the SST index over time can be captured largely by percentage changes in the poverty rate (or rate) and average poverty gap ratio (or gap). Because of this, Osberg (2000) proposed the concept of "poverty box" which uses the horizontal sides of a "poverty box" to represent the rate and the vertical sides of the box to represent the gap. In Figure 4, the poverty box is represented by a box formed by the horizontal line from the origin to 0.5 and the vertical line from the origin to 0.2. Using this box can facilitate communication about poverty research. For example, by placing a poverty box of one country over that of another for the same year or aligning poverty boxes at the origin for one country over time, one can illustrate cross-country differences between countries or over-time changes of a country in two major dimensions of poverty. In Figure 5, changes in poverty in the United States and Canada from 1994 to 1997 are illustrated with reference to the poverty rate and gap in Finland in 1995. As can be seen in Figure 5, Canada's poverty was worsened when that of the United States has been improved over the period but both countries experienced more poverty than Finland did.

Figure 5

The Poverty Box: intensity in Canada, the United States, and Fineland



Source: Osberg (2000). Poverty Line = 1/2 Median After-Tax Equivalent Income

3.3 *Example Three*

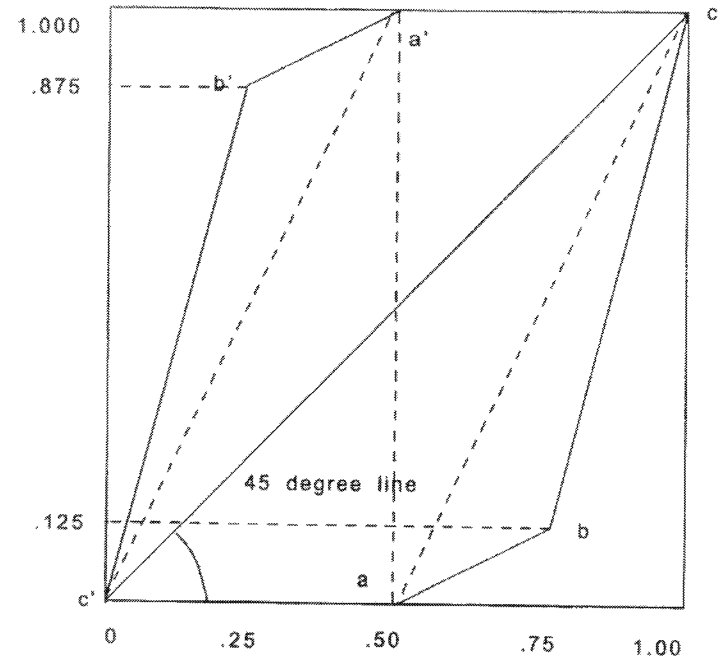
While the poverty box is used for communicating two major dimensions of poverty, this example is used to show that the arc portion of the deprivation profile is an inverted Lorenz curve and that the curvature of the arc represents the level of inequality among the poor. Using the same data set given in Example Two, draw one Lorenz curve and one inverted Lorenz curve for deprivation measures in Figure 6. The Lorenz curve is for sequence x—that is, the $x_{(i)}$'s arranged in increasing order; see the solid curve $c'abc$ in Figure 6. The inverted Lorenz curve is for sequence x'—that is, the x_i's arranged in decreasing order; see the solid curve $c'b'a'c$ in Figure 6.

From the Lorenz curve, the Gini index is valued at $G(x) = 0.6875$. This number is exactly the ratio of the area between the 45 degree line and the Lorenz curve $c'abc$ (0.34375) to the lower triangle of the unit square (0.5). Based on geometry, it can also be interpreted as the ratio of the area above the Lorenz curve segment abc (0.34375) to the right-hand-side half of the unit square (0.5). Because of symmetry, $G(x) = 0.6875$ can also be interpreted as the ratio of the area below the inverted Lorenz curve segment $c'b'a'$ (0.34375) to the left-hand-side half of the unit square (0.5). It is the last interpretation that helps in interpreting the inequality of deprivation measures.

The inverted Lorenz curve $c'b'a'c$ in Figure 6 can be mapped to the deprivation profile $c'b'a\ HI$ in Figure 4 by changing the height from the unity to 0.2. Computing the SST index as the ratio of the area under the deprivation profile to the half of the unit square in Figure 4 is equivalent to multiplying $(1 + G(x))$ by HI. The unity in $(1 + G(x))$ can be viewed as the ratio of the right-hand-side

Figure 6

Numrical Example Three—Lorenz Curves for Deprivation Measures x (c'abc) and x' (c'b'a'c)



half of the unit square in Figure 6 to itself. $G(x)$ in term $(1 + G(x))$ can be viewed as the ratio of the area below the inverted Lorenz curve segment $c'b'a'$ in Figure 6 to the left-hand-side half of the unit square. Thus, $(1 + G(x))$ can be interpreted as the ratio of the area under the Lorenz curve $c'b'a'c$ in Figure 6 to the area of the unit square. To map this ratio to the SST index $P(y;z)$—the ratio of the area under the deprivation profile to the area under the line of maximum poverty in Figure 4, $(1 + G(x))$ needs to be multiplied by HI.

4 Concluding Remarks

This note is written to provide a practitioner's guide on how to use the Sen-Shorrocks-Thon index decomposition and geometric explanations. Although the basis for such decomposition was provided in Shorrocks (1995) and the further decomposition was made and used in Osberg and Xu (1997, 1998, 2000), in our communicating with other policy analysts we have found that the link between the SST index and its decomposition is not necessarily transparent. Since the ultimate value of a theoretical work is its usefulness to policy analysts, we write this note to make the SST index and its decomposition more accessible through simple mathematics, numerical examples, and geometrical figures.

Acknowledgments

We would like to thank the Editor and anonymous referees for their constructive comments and suggestions which lead to a substantial revision of an earlier version of this paper. We would like to thank Carla Farmer and Heather Lennox for secretarial assistance.

Notes

1. J. Myles and G. Picot (2000) use the SST index and its decomposition proposed by Osberg and Xu (1997) in research at Statistics Canada. Our communication with them and others has convinced us that a practitioner's guide would be useful to policy analysts.
2. For a more comprehensive theoretical survey of the Sen family of poverty indices, see Xu and Osberg (2000).
3. For theoretical economists, this may be alright. But for practitioners, some further elaboration is clearly necessary.
4. The empirical counterpart of this debate includes the empirical research of Rodgers and Rodgers (1991) and Osberg and Xu (1997, 1998, 2000).
5. One may alternatively adopt the steps taken in equation (6) in Osberg and Xu (1997, 2000).
6. Note that $G(x)$ is a measure of inequality while $1+G(x)$ is a measure of equality. Also note the following facts are used for two different expressions of the Gini index $G(x)$:

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{i} x_j &= x_1 + (x_1 + x_2) + (x_1 + x_2 + x_3) + (x_1 + x_2 + \ldots x_n)\\ &= nx_1 + (n-1)x_2 + (n-2)x_3 + \ldots + x_n\\ &= \sum_{i=1}^{n} (n-i+1)x_i,\end{aligned}$$

and

$$\begin{aligned}\sum_{i=2}^{n}\sum_{\substack{i\geq 1\\ j=2}}^{i} x_j &= x_1 + (x_1 + x_2) + (x_1 + x_2 + x_3) + (x_1 + x_2 + \cdots + x_{n-1})\\ &= (n-1)x_1 + (n-2)x_2 + \cdots + x_{n-1}\\ &= \sum_{i=1}^{n-1} (n-i)x_i.\end{aligned}$$

7. It should be noted that if one arranges x_i's in decreasing order rather than increasing order in computation, the right-hand and left-hand sides of equation (4) are *not* equal.
8. We use integers here because they allow the easy illustration of our computation geometrically.
9. Xu and Osberg (1998) focused on absolute deprivation measures and deprivation dominance tests.
10. This is corresponding to

$$D(F;p) = \int_{F^{-1}(1-p)}^{\infty} x dF(x),$$

where F is the underlying distribution function for x in the support and p is the probability taking a value from zero to one. Alternatively,

$$D(F;p) = \int_{1-p}^{1} F^{-1}(w) dw,$$

where w takes a value from zero to one.

References

Myles, J. and G. Picot, 2000, Poverty Indices and Policy Analysis, *Review of Income and Wealth 46*, 161-179.

Osberg, L., 2000, Poverty in Canada and the USA: Measurement, Trends and Implications, *Canadian Journal of Economics 33*, 847-877.

Osberg, L. and K. Xu, 1997, International Comparisons of Poverty Intensity: Index Decomposition and Bootstrap Inference, Department of Economics, Working Paper 97-03, Dalhousie University, Halifax, Canada.

Osberg, L. and K. Xu, 1998, Poverty Intensity - How Well Do Canadian Provinces Compare? *Canadian Public Policy 25*, 1-17.

Osberg, L. and K. Xu, 2000, International Comparisons of Poverty Intensity: Index Decomposition and Bootstrap Inference, Journal of Human Resources 35, 51-81. Errata, Journal of Human Resources, 35(3).

Rodgers, J.L. and J.R. Rodgers, 1991, Measuring the Intensity of Poverty Among Subpopulations: Applications to the United States, *Journal of Human Resources 26*, 338-361.

Sen, A.K., 1976, Poverty: An Ordinal Approach to Measurement, *Econometrica 44*, 219-231.

Shorrocks, A.F., 1995, Revisiting the Sen Poverty Index, *Econometrica 63*, 1225-1230.

Thon, D., 1979, On Measuring Poverty, *Review of Income and Wealth 25*, 429-440.

Thon, D., 1983, A Poverty Measure, *The Indian Economic Journal 30*, 55-70.

Xu, K., 1998, The Statistical Inference for the Sen-Shorrocks-Thon Index of Poverty Intensity, *Journal of Income Distribution 8*, 143-152.

Xu, K. and L. Osberg, 1998, A Distribution-Free Test for Deprivation Dominance, *Econometric Reviews 17*, 415-429.

Xu, K. and L. Osberg, 2000, The Social Welfare Implications, Decomposability, and Geometry of the Sen Family of Poverty Indices, Department of Economics, Dalhousie University, Halifax, Canada; forthcoming in *Canadian Journal of Economics*.

Appendix A
The Impact of Data Sorting on the Sign and Value of the Gini Index

For sequences x' and x, the following correspondence relations hold for all i:

$$x_{(i)} = x_{n-i+1}, \quad x_i = x_{(n-i+1)}. \tag{26}$$

This technical identity will be used in the following. Sorting data in sequence x' to get sequence x does not affect the absolute value but alters the sign of the Gini index; that is

$$G(x) = -G(x') \tag{27}$$

where the Gini index $G(x)$ is a function of the sorted sequence x($x_{(i)}$'s) and the Gini index $G(x')$ is a function of the unsorted sequence x' (x_i's). More specifically, the Gini index of the sorted sequence $G(x)$ is given by

$$G(x) = 1 - \sum_{i=1}^{n}\left(\tfrac{1}{n}\right)\left(\frac{\sum_{j=1}^{i} x_{(j)} + \sum_{\substack{i\geq 2\\ j=2}}^{i} x_{(j-1)}}{n\mu(x)}\right), \tag{28}$$

or

$$G(x) = 1 - \tfrac{1}{n^2\mu(x)} \sum_{i=1}^{n} (2n - 2i + 1)x_{(i)}, \tag{29}$$

while the Gini index of the unsorted sequence $G(x')$ is given by

$$G(x') = 1 - \sum_{i=1}^{n}\left(\tfrac{1}{n}\right)\left(\frac{\sum_{j=1}^{i} x_j + \sum_{\substack{i\geq 2\\ j=2}}^{i} x_{j-1}}{n\mu(x)}\right), \tag{30}$$

or

$$G(x') = 1 - \tfrac{1}{n^2\mu(x)} \sum_{i=1}^{n} (2n - 2i + 1)x_i. \tag{31}$$

Because of equation (26) and $\Sigma_i^n x_{(i)} = \Sigma_i^n x_i$, equation (31) can be written as

$$G(x') = 1 - \sum_{i=1}^{n}\left(\tfrac{1}{n\Sigma_i^n x_{(i)}}\right)\left((2n-1)x_{(n)} + \cdots + 5x_{(3)} + 3x_{(2)} + x_{(1)}\right). \tag{32}$$

The second term within the summation sign on the right-hand side of equation (32) can be rewritten as

$$\begin{aligned} & (2n-1)x_{(n)}+\cdots+5x_{(3)}+3x_{(2)}+x_{(1)} \\ = & \left(2nx_{(n)}-x_{(n)}\right)+\cdots+\left(2nx_{(3)}-(2n-5)x_{(3)}\right) \\ & +\left(2nx_{(2)}-(2n-3)x_{(2)}\right)+\left(2nx_{(1)}-(2n-1)x_{(1)}\right) \\ = & \ 2n\sum_{i}^{n}x_{(i)}-\left((2n-1)x_{(1)}+\cdots+5x_{(n-2)}+3x_{(n-1)}+x_{(n)}\right). \end{aligned} \tag{33}$$

Substituting the last line in equation (33) into equation (32) gives

$$\begin{aligned} G(x') = & \ 1-\sum_{i=1}^{n}\left(\frac{1}{n\sum_{i}^{n}x_{(i)}}\right) \\ & \times\left(2n\sum_{i}^{n}x_{(i)}-\left((2n-1)x_{(1)}+\cdots+5x_{(n-2)}+3x_{(n-1)}+x_{(n)}\right)\right) \\ = & \ 1-2+\sum_{i=1}^{n}\left(\frac{1}{n\sum_{i}^{n}x_{(i)}}\right) \\ & \times\left((2n-1)x_{(1)}+\cdots+5x_{(n-2)}+3x_{(n-1)}+x_{(n)}\right) \\ = & \ -1+\sum_{i=1}^{1}\left(\frac{1}{n\sum_{i}^{n}x_{(i)}}\right) \\ & \times\left((2n-1)x_{(1)}+\cdots+5x_{(n-2)}+3x_{(n-1)}+x_{(n)}\right) \\ = & \ -G(x). \end{aligned} \tag{34}$$

Equation (34) also implies

$$G(x) = -G(x'). \tag{35}$$

The last result indicate that if one wants to compute the Gini index using the data sorted in *increasing* order, he or she can get the same by multiplying -1 to the Gini index using the data sorted in *decreasing* order.

Appendix B
The GAUSS program

The following a GAUSS program that computes the SST index and its decomposed components. Since the main purpose is to show the computation procedure, this program is written in lower level commands.

```
/*
* File name: sstindex.prg
* Purpose:      This program is written to compute the SST index and its de-
composition proposed by Osberg and Xu (1997).
*
* Input:
*               y =     income vector;
*
* Intermediate Variables:
*               n =     the number of individuals;
*               z =     the poverty line set at the half of the median income;
*               x =     the vector of poverty gap ratio in decreasing order;
*               q =     the number of the poor individuals whose incomes are
*                       below the poverty line z;
*               xs =    the vector of poverty gap ratio in increasing order;
*
* Outputs:
*               sst1=   the Sen-Shorrocks-Thon index (or the SST index) of
*                       poverty intensity---Shorrocks equation;
*               sst2 =  the Sen-Shorrocks-Thon index (or the SST index) of
*                       poverty intensity---Osberg-Xu decomposition;
*               rate =  the poverty rate;
*               gap =   the average poverty gap ratio of the poor;
*               gini =  the Gini index.
*
* Remarks:
*               sst1 and sst2 are identical for the same data set.
*
*/
@ set the output file @
output file=sstindex.out reset;
@ read in income data @
y={3,9,11,15};
@ sort income data in increasing order @
y=sortc(y,1);
/*
@ find median @
ymed=median(y);
@ define the poverty line z @
z=.5*ymed;
```

```
*/
z=10;
@ compute the relative poverty gap ratio @
x=(z-y)./z;
@ figure out the number of the poor, q @
@ set negative x[i] to zero @
n=rows(y);
q=0;
i=1;
do while i le n;
if x[i] ge 0; q=q+1; endif;
if x[i] lt 0; x[i] = 0; endif;
i=i+1;
endo;
@ compute the SST index, sst1 @
sst1=(1/n^2)*sumc( (2*n-2*seqa(1,1,n)+1).*x );
@ compute the poverty rate, rate @
rate=q/n;
@ compute the poverty gap, gap @
gap=(1/q)*sumc(x);
@ compute the Gini index of x, gini @
@ sort x in increasing order @
xs=sortc(x,1);
@ set the do-loop counter to 1 @
i=1;
@ set the numerator sum, num @
num=0;
do while i le n;
if i lt 2;
 num=num+xs[1:i];
else;
 num=num+sumc(xs[1:i])+sumc(xs[1:i-1]);
endif;
i=i+1;
endo;
@ compute the Gini index, gini @
gini=1-( num/(n*sumc(xs)) );
@ the SST index, sst2, computed from rate, gap, and 1+gini @
sst2=rate*gap*(1+gini);
"Output:";
"gini" gini;
"rate" rate;
"gap " gap;
"sst1" sst1;
"sst2" sst2;
```

The computation results are:

```
Output:
gini  0.68750000
rate  0.50000000
gap   0.40000000
sst1  0.33750000
sst2  0.33750000
```