Research Article

Some New Properties of Poverty Measures and Deprivation Profile

Unnikrishnan Nair[1], Sunoj Sm[1]

1. Department of Statistics, Cochin University of Science and Technology, India

In the present work some new properties of the Sen-Shorrocks poverty index and the related deprivation profile are investigated. Inversion formulas for deriving the income distribution from Sen-Shorrocks index and Sen-index are derived and it is shown that while the Sen-Shorrocks index determines the income distribution uniquely, there can be more than one distribution corresponding to the Sen-index. The relationship between the deprivation profile (DP) and the Lorenz curve is exploited to characterize the income distribution, to propose parametric forms the DP and the use of DP to infer and compare patterns of poverty.

Corresponding authors: N. Unnikrishnan Nair, unnikrishnannair4@gmail.com; S. M. Sunoj, smsunoj@cusat.ac.in

1. Introduction

Beginning with the work of [1], the investigation of poverty measures has embarked on an axiomatic approach for the construction and evaluation of poverty measures that are distibution sensitive. Leaving behind some of the conventional measures like the proportion of individuals with incomes below the poverty line, quantities like the income shortfalls of the poor and inequalities among them became focal points in developing new indices. This has prompted various authors like [1], [2], [3], [4], [5] and [6] to propose different forms of poverty measures that satisfied distinct desirable properties that enabled the idetification of the poor and to understand the intensity of poverty among them. Among these the Sen-Shorrocks index, SST [6], an extension of the work of [1], is a widely used indicator of poverty that measures the incidence of poverty, its depth and inequality jointly. Futher it is invariant to a change of scale in incomes, poverty line and does not depend on the income levels of the non-poor. As a continuous function of individual incomes, it satisfies essential properties dictated by the impartiality axiom,

monotonicity axiom transfer axiom and replication invariance. Although the index SST was formulated with income distribution as the basis, it is also used extensively in the analysis of multidimensional deprivation. Some recent applications of the SST can be seen in the papers by [7], [8], [9] and [10].

The basic fabric on which most of the poverty measures including Sen and SST indices were built up is the income distribution and the observation from it are used to explain the pattern of poverty incidence. Other than defining poverty metrics using distribution functions the converse problem of determining the form of the income distribution from such metrics is a rarely visited topic in literature. One of the main objectives of the present study is to derive formulas the recovery of the income distribution from the functional form assumed by the poverty measures. Such an investigation is important in cases when the poverty index is calculated from a sample based on a specific formula, the population income distribution can be determined using our results.

In the social welfare analysis of a population, ranking of distributions per taining to economic variables like income, wealth, consumption etc are important components. The improvement of a society in terms of inequality, poverty welfare, etc are facilitated through ordering of various income profiles. Since the conclusions arrived at from the approach generally depend on the choice of the measure and the parameter required for computation, efforts were made to investigate whether the ordering of poverty measures can imply stochastic dominance criteria which provide global comparison of income distributions. Several results in this direction relating poverty ordering and stochastic dominance are well documented in [11] and [12]. The results developed in this paper have a significant contribution in this context. While determining the income distribution from the poverty measure, if it happens that more than one distribution possess the same functional form of the poverty metric, as shown in the case of the Sen index in the next section, there willl be some difficulty in fixing the correspondence between the index and the distribution. Secondly, the inversion formula expressing the distribution in terms of the index is an effective tool in establishing the dominance criterion. Further as mentioned earlier, when income model is chosen based on the form of the distribution function, instances can be seen where more than one model is appropriate by the goodness of fit criteria and hence the question of selecting the appropriate one surfaces. In such cases, our results justify the choice of the correct model as the one for which the empirical plot of the poverty index is of the same form as the index generated from the model, in view of the one-to-one correspondence established between the index and model. Finally the information about the theoretical model obtained from the index can be used for further analysis of the income distribution and all other associated metrics, than merely depending on the sample information.

[6] while introducing the modifications of the [1] index, also presented the novel concept of deprivation profile (DP) associated with SST. It was shown that twice the area under the profile of a normalized income variable is equivalent to SST thereby giving a geometric interpretation to SST. Thus result provides a connection between deprivation profile and the poverty measure discussed in this work. In a later work [13] gave some additional properties of DP, including a stochastic order involving DP's to compare the poverty patterns. The relationship DP has with Lorenz curve opens up several other interesting properties that could be exploited in poverty analysis. Motivated by this we have derived several properties of DP. These include determination of income distribution from DP, implications of dominance of income distributions and the DP ordering, parametric representations of DP by sorting out the conditions under which an arbitrary function can be a DP. A probability distribution based on DP derived in this context summarises some aspects of deprivation on the population.

The present work consists of five sections. After the introductory section we move to Section 2 to discuss the SST index and its properties. This is followed by a discussion on the DP in Section 3. Finally in Section 4, we propose some funcitonal forms of DP and illustrate their application to real data.

2. Sen-Shorrocks index

It is assumed throughout this paper that Y is a continuous non-negative random variable representing the income in a population with distribution function $F(y)$ and survival function $\overline{F}(y) = 1 - F(y)$. Also $y = z$ is chosen as the poverty line so that any individual in the population with income less than z is considered poor. The Sen-Shorrocks index, defined in [6] has the form

$$P_1(y, z) = \mu(x)(1 + G(x)) \tag{2.1}$$

where x is the realization of the normalized random variable $X = \max\left(\frac{z-Y}{z}, 0\right)$, $\mu(x)$ and $G(x)$ are the mean and Gini index of X respectively. We have

$$\mu(x) = \int_0^z \frac{z-y}{z} dF(y) \tag{2.2}$$

and

$$G(x) = 1 - \frac{2}{\mu(x)} \int_0^z \frac{z-y}{z} F(y) dF(y) \tag{2.3}$$

so that

$$P_1(y, z) = \frac{2}{z} \int_0^z (z-y)(1 - F(y) dF(y)). \tag{2.4}$$

Equation (2.4) expresses SST in terms of the income distribution. It also known that $P_1(y,z)$ admits a decomposition of the form

$$P_1(y,z) = H\left[(2-H)I + H(1-I)G_p\right]$$

where $H = F(z)$, the head count ratio, I is the income gap ratio,

$$I(z) = \frac{1}{F(z)}\int_0^z \frac{z-y}{z} dF(y)$$

and G is the Gini index for the poor,

$$G_p(z) = \frac{2}{\mu(z)}\int_0^z yF(y)dF(y) - 1.$$

Equation (2.4) indicates that the index $P_1(y,z)$ can be calculated directly from the income distribution $F(y)$ and hence the properties of P_1 can be directly linked to F. A simplified version of (2.4), which does not seem to have appeared in literature, can be obtained from (2.4) in a straightforward manner as

$$P_1(y,z) = 1 - \frac{1}{z}\int_0^z \overline{F}^2(y)dy. \tag{2.5}$$

The first question we consider is a formula that expresses the income distribution F in terms of $P_1(y,z)$, so that the income distribution can be derived from the knowledge of the functional form of $P_1(y,z)$. The income distribution $F(y)$ can be recovered from the SST through

$$F(z) = 1 - \left[1 - P_1(y,z) - zP_1'(y,z)\right]^{\frac{1}{2}}. \tag{2.6}$$

By writing (2.5) as

$$\left[1 - P_1(y,z)\right] z = \int_0^z \overline{F}^2(y)dy. \tag{2.7}$$

and differentiating (2.7).

Distribution functions and quantile functions $Q(u) = \inf[x|F(x) \geq u],\ 0 \leq u \leq 1$, are two equivalent ways of specifying a distribution. When $F(x)$ is continuous and strictly increasing, $F(x) = u$ if and only if $x = Q(u)$ and $dx = q(u)du$ or $du = f(x)dx$, where $q(u) = \frac{dQ}{du}$, is the quantile density function and $f(x)$ is the probability density function. Hence the results in (2.5) and (2.6) when written in terms of quantile functions have the form in the next proposition. These equations enable the calculation of the income quantile function from the poverty quantile function, providing an alternative methodology.

Proposition 2.1. *When Y has quantile function $Q(u)$.*

i. $p_1(u) = P_1\left(u, Q(u)\right) = 1 - \frac{1}{Q(u)}\int_0^u (1-p)^2 q(p)dp$ *and*

ii. $Q(u) = \exp\left[\int_0^u \frac{p_1'}{1-p_1-(1-p)^2}\,dp\right]$.

Proof. Equation (i) follows by substituting $z = Q(u)$ in (2.5). To prove the second identity we use the transformation $z = Q(u)$ in (2.7) to get

$$Q(u)\left(1 - p_1(u)\right) = \int_0^u (1-p)^2 q(p)dp.$$

Differentiating and simplifying

$$\frac{q(u)}{Q(u)} = \frac{p_1'(u)}{1 - p_1(u) - (1-u)^2}$$

from which (ii) is obtained. □

The importance of the above result lies in the fact that they enable the unique determination of F when the functional form of $P_1(y,z)$ or $p_1(u)$ are ascertained from empirical data. Also when two models are found fitting the data in black box modelling, one can choose the correct one by appealing to the form of P_1 or p_1. Though Proposition 2.1 is valid for all distributions, it is particularly useful when the income model is a quantile function that does not have closed form distribution function to analytically make use of (2.5) and (2.6). The role and applications of such quantile functions in income analysis can be found in [14], [15], [16], [17] and [18].

Apart from calculating the SST for empirical data, if we look at $P_1(y,z)$ as a function of the poverty line, several interesting behaviour can be noticed.

1. The extent of the increase in P_1 depends on the income distribution. For example, in the case of the Pareto distribution, $\overline{F}(y) = y^{-\alpha}, y > 0$ we have

$$P_1(y,z) = 1 + \frac{2\alpha}{(1-2\alpha)z} + \frac{1}{(1-2\alpha)z^{2\alpha}}, \alpha > 1,\ 1 - 2\alpha < 0.$$

However, for the power distribution $F(x) = \left(\frac{x}{\alpha}\right)^\beta,\ 0 \le x \le \alpha, \alpha, \beta > 0$ (see [19] for the application of the power distribution to describe the lower tail of the incomes of poor households)

$$P_1(y,z) = \frac{2z^\beta}{(\beta+1)\alpha^\beta} - \frac{z^{2\beta}}{(2\beta+1)\alpha^{2\beta}}.$$

For the Pareto and power distributions, P_1 is increasing in z, but it is concave for the former and convex for the latter.

2. When the income distribution is specified by a quantile function with no closed form distribution function the above reasoning fails and Proposition 2.1 is needed. For example, when X has distribution

$$Q(u) = \alpha \log u + 2\beta u + \gamma,\ 0 \leq u \leq 1, \alpha, \beta, \gamma > 0$$

When u tends to 0^+, by usual convention $\log 0$ is taken as zero and the support is $(Q(0), Q(1)) = (\gamma, 2\beta + \gamma)$, on taking $\log 0 = 0$. [17] have used this model for the income data of the California state. We then have $q(u) = \frac{\alpha}{u} + 2\beta$ and the quantile-based SST is

$$p_1(u) = \frac{6\alpha(1+\beta) + \beta(4-\alpha)\alpha u + 4\beta u^3}{6(\alpha \log u + 2\beta u + \gamma)}.$$

3. Proposition 2.2 permits us the ranking of income distributions based on their poverty measures. Let X_1 and X_2 be the income random variables with distribution functions $F_1(y)$ and $F_2(y)$ with SST's denoted by $P_1(y,z)$ and $P_2(y,z)$. Recall that (a) F_1 has first degree stochastic dominance over F_2, written as $X_1 D_1 X_2$ if $\overline{F}_1(y) > \overline{F}_2(y)$ for all $y > 0$, (ii) X_1 has less SST than X_2, written as $X_1 <_{SST} X_2$ if $P_1(y, z < P_2(y,z))$.

Proposition 2.2.

$$\text{If } X_1 D_1 X_2 \text{ then } X_1 <_{SST} X_2.$$

Proof.

$$\begin{aligned} X_1 D_1 X_2 \quad & \Longrightarrow \overline{F}_1(y) > \overline{F}_2(y) \text{ for all } y > 0 \\ & \Longrightarrow \frac{1}{z}\int_0^z \overline{F}_1^2(y)dy \geq \frac{1}{z}\int_0^z \overline{F}_2^2(y)dy \text{ for all } z \in (0, z^{\max}) \\ & \Longrightarrow 1 - \frac{1}{z}\int_0^z \overline{F}_1^2(y)dy \leq 1 - \frac{1}{z}\int_0^z \overline{F}_2^2(y)dy \\ & \Longrightarrow X_1 <_{SST} X_2. \end{aligned}$$

□

The converse of the above proposition need not be true. In the next proposition, we give the additional conditions required to show that the SST order (lesser SST index) implies the dominance order (X_2 dominates the random variable X_1).

Proposition 2.3. *If* $\frac{\int_0^z \overline{F}_1^2(y)dy}{\int_0^z \overline{F}_2^2(y)dy}$ *is increasing for all* $z \in (0, z^{\max})$*, then* $X_1 <_{SST} X_2 \Longrightarrow X_1 D_1 X_2$.

Proof.

$\frac{\int_0^z \overline{F}_1^2(y)dy}{\int_0^z \overline{F}_2^2(y)dy}$ is increasing $\Longrightarrow \overline{F}_1^2(z)\int_0^z \overline{F}_2^2(y)dy - \overline{F}_2^2(z)\int_0^z \overline{F}_1^2(y)dy \geq 0$

$$\implies \frac{\overline{F}_1^2(z)}{\overline{F}_2^2(z)} > \frac{\int_0^z \overline{F}_1^2(y)dy}{\int_0^z \overline{F}_2^2(y)dy} \text{for all~} z \in (0, z^{\max}). \tag{2.8}$$

Since $X_1 <_{SST} X_2$ we have

$$\begin{aligned}
&1 - \frac{1}{z}\int_0^z \overline{F}_1^2(y)dy < 1 - \frac{1}{z}\int_0^z \overline{F}_2^2(y)dy \\
\implies &\int_0^z \overline{F}_2^2(y)dy < \int_0^z \overline{F}_1^2(y)dy \\
\implies &\frac{\int_0^z \overline{F}_1^2(y)dy}{\int_0^z \overline{F}_2^2(y)dy} > 1 \implies \frac{\overline{F}_1^2(z)}{\overline{F}_2^2(z)} > 1 \text{ by (2.8)} \\
\implies &\overline{F}_1(z) > \overline{F}_2(z) \implies X_1 D_1 X_2.
\end{aligned}$$

□

Remark 2.1. For a positive integer r, F_1 is less than in rth degree stochastic dominance denoted by $X_1 D_r X_2, r = 1, 2, \cdots$, if $F_{1,r} \leq F_{2,r}$ for all x in $[a, b] \subset [0, \infty)$ where $F_{i,r} = \int_0^x F_{i,r-1}(t)dt, i = 1, 2$. It is known that $F_1 D_{r-1} F_2 \implies F_1 D_r F_2, r \geq 2$. Thus $X_1 <_{SST} X_2$ implies all higher order dominances and their implications established in literature.

The Sen index is given by

$$P(y, z) = F(z)\left[I(z + (1 - I(z))G_p(z))\right].$$

Substituting for $I(z)$ and $G_p(z)$ from (2.5) and (2.6)

$$\begin{aligned}
P(y, z) \quad &= \frac{2}{zF(z)}\int_0^z (z - y)(F(z) - F(y))dF(y) \\
&= \frac{2}{zF(z)}\left[\int_0^z dF(y) - z\int_0^z F(y)dF(y) - F(z)\int_0^z ydF(y) + \int_0^z yF(y)dF(y)\right] \\
&= \frac{2}{zF(z)}\left[F(z)\int_0^z F(y)dy - \int_0^z F^2(y)dy\right].
\end{aligned} \tag{2.9}$$

Setting $z = Q(u)$ and $y = Q(t)$, the quantile version of (2.9) is

$$p(u) = p(Q(u)) = \frac{2}{uQ(u)}\left[u\int_0^u tq(t)dt - \int_0^u t^2 q(t)dt\right]. \tag{2.10}$$

The income distribution is to be obtained by solving for F from (2.10) or $Q(u)$ from (2.10). We have

$$uQ(u)p(u) = 2u\int_0^u tq(t)dt - 2\int_0^u t^2 q(t)dt.$$

Differentiating twice with respect to u yields

$$up(u)\frac{d^2Q}{du^2} + 2\left(up'(u) + p(u) - u\right)\frac{dQ}{du} + \left(up''(u) + 2p'(u)Q\right) = 0 \tag{2.11}$$

a second order differential equation with variable coefficients. The general solution for $Q(u)$ from the differential equation is difficult to find. But for particular forms of $p(u)$ solutions exist, but it may not be unique, so that corresponding to a Sen index, there can be more than one income distribution. We show this by an example.

Example 2.1. Let X be distributed in the power form with $F(y) = \left(\frac{y}{\alpha}\right)^{\beta},\ 0 \leq y \leq \alpha$ or $Q(u) = \alpha u^{\frac{1}{\beta}}$. From (2.10),

$$p(u) = \frac{2\beta u}{(1+\beta)(1+2\beta)},$$

a linear function of u. We now show that there are two income distributions with this$p(u)$. We can write $p(u) = au; a = \frac{2\beta}{(1+\beta)(1+2\beta)} = \frac{1}{3}$, say. We have to find Q from (2.11). To solve the resulting differential equation, assume that $Q(u) = u^m$ be a solution. The corresponding auxillary equation is

$$2am(m-1)u^{m-2} + 2(2a-1)mu^{m-1} + 2au^m = 0$$

or

$$am(m-1) + 2(2a-1)m + 2a = 0$$

giving

$$m = \frac{2 - 3a \pm \sqrt{a^2 - 12a + 4}}{2a}.$$

With $a = \frac{1}{3}$ the solutions are $m = 1, 2$. Hence $Q(u) = u^{\frac{1}{2}}$ or u. There are two power distributions, $F(x) = x$ and $F(x) = x^2$ which has the same Sen index, $p(u) = \frac{1}{3}$.

The Sen index does not share the uniqueness property of the income distribution of the Sen-Shorrocks index. This makes the ordering of $p(u)$ for X_1 and X_2 giving more than one set of F_{X_1} and F_{X_2} or equivalently $Q_{X_1}(u)$ and $Q_{X_2}(u)$.

3. Deprivation profiles

Deprivation profiles are often a method of depicting various aspects of the patterns in which poverty prevails in different populatons. As a measure of indicating the extent of deprivation, these profiles can also be employed to compare and decide which population is more poor than another. In this context [6] identified deprivation by considering the distribution $F(x)$ of incomes, through the formula

$$B(u) = \mu_X \left(1 - L_X(1-u)\right) \tag{3.1}$$

with mean income as μ_X and L_X as the Lorenz curve of X. In terms of $F(x)$ we can write (3.1) as

$$B(u) = \int_{Q_X(1-u)}^{1} x dF(x) = \int_{1-u}^{1} Q_X(p) dp. \tag{3.2}$$

The objective of the present section is to infer several properties of $B(u)$ that do not seem to have appeared in literature and throws light on several new results concerning the income distribution F as well as those of B.

(a) Writing $B^*(u) = \mu_X^{-1} B(u)$, it is seen that $B^*(0) = 0$, $B^*(1) = 1$, B^* is continuous and increasing, so that it is a distribution function of some non-negative random variable V, with support $[0, 1]$. It is a summary of the poverty profile of Y, and will be called the poverty distribution. It has mean

$$\mu_* = E(V) = \int_0^1 (1 - B^*(u))\, du = \frac{1 - G_p}{2}$$

where G_p is the Gini index of the poor. It is easy to see that the mean of the poverty distribution increases when the income inequality indicated by G_p decreases and as such can be thought of as an index to compare the incidence of poverty in two populations.

If the two random variables have, by a possible transformation, the same mean then

$$\begin{aligned} X_1 D_1 X_2 \qquad & \Rightarrow \int_{1-u}^{1} Q_1(p) dp \geq \int_{1-u}^{1} Q_2(p) dp \Rightarrow \int_0^1 Q_1(p) dp \geq \int_0^1 Q_2(p) dp \\ & \Rightarrow E(X_1) \geq E(X_2) = \mu_{X_1} \geq \mu_{X_2} \end{aligned}$$

where μ_{X_i} is the mean of $X_i, i = 1, 2$. The last inequality means from (a) above that $\frac{1-G_1}{2} \geq \frac{1-G_2}{2}$ or $G_1 \leq G_2$. This would mean that of two income distributions have the same mean

$$X_1 D_1 X_2 \Rightarrow G_1 \leq G_2.$$

While modelling data wth quantile funtions, when a candidate model is not found to be appropriate, it is common to add or multiply the chosen with another quantile function to find a suitable model. Sometimes two data sets may be combined toether for specific analysis and the sum of the quantile functions of the form $Q_1(u) + Q_2(u)$ may be the resulting model which is again another quantile function. This property of quantile functions is not shared when distribution functions are considered to represent the distribution. From the above discussion it follows that if X_1 and X_2 are incomes with quantile functions Q_1 and Q_2 with profiles B_1 and B_2, then the deprivation profile of $Q_1 + Q_2$ is $B_1 + B_2$. in particular if $X = Y + \alpha$, then $B_X(u) = B_Y + u\alpha$.

(b) The income distribution F is determined upto a scale factor from $B(u)$ by the formula $Q(u) = \mu B'(1-u)$ or $F(u) = Q^{-1}(u),\ 0 \leq u \leq 1$. Note that the same $Q(u)$ will result even if every income is multiplied by a constant K.

(c) Let X_1 and X_2 be two income random variables with distributions F_1 and F_2 and quantile functions Q_1 and Q_2 respectively. When lower deprivation measures are preferred, F_1 is deprivation dominates F_2, if $B_{F_2}(u) \geq B_{F_1}(u)$ for all u in $[0,1]$ [13] or equivalently

$$\int_{1-u}^{1} Q_{X_2}(p)dp \geq \int_{1-u}^{1} Q_{X_1}(p)dp, p \in [0,1]$$

It follows that, if the income distribution of X_2 dominates that of X_1, then F_1 is deprivation dominant over F_2.

However $B_{F_1} \geq B_{F_2}$ does not necessaily imply $X_2 D_1 X_1$. A condition for this to hold is that if $\frac{B_{F_1}(u)}{B_{F_2}(u)}$ is increasing in u then $B_{F_2}(u) \geq B_{F_1}(u) \Rightarrow X_2 D_1 X_1$.

The above properties add to the importance of DP as a poverty measure on its own merit, just as the role of Lorenz curve in income inequality. Hence it is worthwhile to construct DPs of income distributions. To enable this we identify the necessary and sufficient conditions for a function to be a DP.

Proposition 3.1. *Let $A : [0,1] \to \mathbb{R}$ be an arbitrary twice differentiable function. Then it will represent a DP if and only if (i) $A(0) = 0$, (ii) $A(1) = \mu > 0$, (iii) $A'(u) \geq 0$ and $A''(u) \leq 0$.*

Proof. Let $A(u)$ be a function on $[0,1]$ satisfying the above properties. Then $A(u)$ being differentiable it is continuous and there exists a non-negative continuous function $C(x)$ such that $A(u) = \int_{1-u}^{1} C(x)dx$.

$$A'(u) = -C(1-u) \geq 0 \text{and} A''(u) = \frac{dC(1-u)}{du},$$

showing $\frac{dC(1-u)}{du} \leq 0$ or $C(1-u)$ is decreasing which is equivalent to $C(u)$ increasing. Thus $C(u)$ is the quantile function of a non-negative random variable and $A(u)$ is its DP. Conversely if $A(u)$ is the DP of a quantile function $Q(u)$, then from $A(u) = \int_{1-u}^{1} Q(p)dp$ it is easy to see that (i), (ii) and (iii) hold. □

Some popular measures of income inequality are related to the DP, for instance, the Pietra index defined as

$$I = \frac{1}{\mu} \int_{0}^{F(\mu)} (\mu - Q(p))\, dp, \mu = E(X).$$

Then I can be expressed via DP as

$$\begin{aligned} I &= \int_0^{F(\mu)} (1 - B'(1-t))\, dt \\ &= F(\mu) + B\left(\overline{F}(\mu)\right) - 1. \end{aligned}$$

When F is symmetric, $I = B\left(\frac{1}{2}\right) - \frac{1}{2}$. The relationships with Lorenz curve and Gini index has already been pointed out.

4. Represntation of deprivation profiles

As in the case of Lorenz curves, the deprivation profiles $B(u)$ also do not have simple functional forms for many standard distributions used as income models. Hence it is advantageous to use simple parametric functional forms that satisfy the properties of DP mentioned in Proposition 3.1. In this section we present functional forms of $A(u)$ by using the parametric Lorenz curves that are established in literature and the relationship (3.1). The corresponding income distribution represented by its quantile function, the poverty distribution and its mean μ are also presented in each case for easy reference.

(i) [20]:

$$L(p) = p\exp\left[-\delta(1-p)\right], \delta > 0,$$

$$B(u) = \mu[1 - (1-u)e^{-\delta u}]; B^*(u) = (1-u)e^{-\delta u},$$

$$Q(u) = e^{-\delta(1-u)}(1+\delta u).$$

There is no closed form distribution function. The random variable X has support $\left(\mu e^{-\delta}, \mu(1+\delta)\right)$.

(ii) [4]:

$$L(u) = u\left(1 - \delta(1-u)^{1/2}\right), \delta \in (0,1).$$

We have

$$B(u) = \mu\left[1 - (1-u)(1-\delta u^{1/2})\right]; B^*(u) = (1-u)\left(1 - \delta u^{1/2}\right); \mu = \frac{15+8\delta}{30},$$

$$Q(u) = \mu\left[1 - \frac{\delta(2-3u)}{2(1-u)^{1/2}}\right],$$

there is no closed from distribution function.

(iii) [21]:

$$L(u) = \frac{(1-\theta)^2 u}{(1+\theta)^2 - 4u},\ 0 < \theta < 1,$$

$$B^*(u)=\frac{(1+\theta)^2u}{(1+\theta)^2-4\theta(1-u)},$$

$$Q(u)=\frac{\mu(1-\theta^2)^2}{(1+\theta)^2-4\theta u};\mu_*=\frac{(1-\theta)^2}{8\theta^2}\log\frac{1+\theta}{1-\theta},$$

and

$$F(x)=\frac{(1+\theta)^2}{4\theta}\left[1-\frac{(1-\theta)^2x}{\mu}\right],\frac{\mu}{(1+\theta)^2}\le x\le\frac{\mu}{(1-\theta)^2}.$$

(iv) [22]:

$$L(u)=ua^{u-1},a>1,$$

$$B^*(u)=(1-u)a^{-u},$$

$$Q(u)=a^{u-1}\left(1+u\log a\right),\ 0\le u\le 1,$$

no closed form distribution function, $\mu_*=(\log a)^{-1}+(a-1)\log a$.

(v) [23]:

$$L(u)=u\left[1-(1-u)^\theta\right],\ 0\le\theta\le 1,$$

$$B(u)=\mu\left[1-(1-u)\left(1-u^\theta\right)\right];B^*(u)=(1-u)\left(1-u^\theta\right);\mu_*=\frac{\theta^2+3\theta}{2(1+\theta)(2+\theta)},$$

$$Q(u)=\mu\left[1(1-u)^\theta+u\theta(1-u)^{\theta-1}\right],$$

no closed form distribution function.

(vi) [24]:

$$L(u)=\frac{e^{ku}-1}{e^k-1},k>0,$$

$$B(u)=\mu\left[1-\frac{e^{k(1-u)}}{e^k-1}\right];B^*(u)=\frac{e^{k(1-u)-1}}{e^k-1};\mu_*=\frac{1}{e^k-1},$$

$$Q(u)=\frac{ke^{ku}}{e^k-1},$$

$$F(x)=\frac{1}{k}\log\frac{e^k-1}{k}x,\frac{k}{e^k-1}\le x\le\frac{ke^k}{e^k-1}.$$

(vii) [25]:

$$L(u)=\frac{u(\beta-1)}{\theta-u},\beta>1,$$

$$B(u)=\mu\frac{\beta u}{\beta+u-1};B^{*}(u)=\frac{\beta u}{\beta-1+u};\mu_{*}=(\beta-1)\left(\beta\log\frac{\beta}{\beta-1}\right),$$

$$Q(u)=\frac{\beta(\beta-1)}{(\beta-u)^{2}},$$

$$F(x)=\beta-\left(\frac{\beta(\beta-1)}{x}\right)^{\frac{1}{2}},\frac{\beta-1}{\beta}\leq x\leq\frac{\beta}{\beta-1}.$$

(viii) [26]:

$$L(u)=\frac{e^{-r(1-u)}-1}{e^{-r(1-e)}-1},r>0,$$

$$B^{*}(u)=\frac{e^{-r(1-e)}-e^{-r(1-e^{1-u})}}{e^{-r(1-e)}-1};\mu_{*}=\frac{r-1-re^{1-r(1-e)}}{e^{-r(1-e)}-1},$$

$$Q(u)=re^{-r(1-u)+u},$$

$F(x)$ has no closed form.

(ix) Pareto:

$$L(u)=1-(1-u)^{\frac{1}{\alpha}},$$

$$B(u)=\mu u^{1/\alpha};B^{*}(u)=u^{1/\alpha};\mu_{*}(u)=\frac{1}{\alpha+1},$$

and

$$\overline{\overline{F}}(x)=\left(\frac{x}{\sigma}\right)^{-\beta},\sigma=\frac{\mu}{\alpha},\beta=\frac{\alpha}{\alpha-1},x>\sigma>0.$$

Remark 4.1.

1. In the above DP's we have considered only one-parameter Lorenz curves. The construction is quite similar in multiparameter cases.
2. In most cases, the quantile function of X has simple form, while the distribution function have neither closed form nor comparatively simple analytic form.
3. The income distributions are generally finite range models.

5. Applications to income analysis

The usefulness of the theoretical results derived in the previous sections will now be illustrated for real income data.

5.1. Computation of SST

We make use of the formula (2.5)

$$P_1(z,y) = 1 - \frac{1}{z}\int_0^z \overline{F}^2(y)dy \tag{5.1}$$

to calculate the SST, after finding the income distribution that had generated the given data set. For this purpose we consider the data on the household income arising from the Socio Economic Survey conducted by the National Statistical Office of Thailand for the year 1975-76. The data on 3941 households that comprises the lower-tail from a total sample of 12189 observations on all incomes, is used for our demonstration. See Table 1 for the data obtained from [27].

Household: Class	**0 - 249**	**250 - 374**	**375 - 499**	**500 - 624**	**625 - 749**	**750 - 874**	**875 - 999**
Income (y): in Baht	211	309	439	561	692	808	921

Table 1. Household income of Thailand

Just as the Pareto distribution is a reasonable hypothesis of incomes in the upper tail in many cases, we propose the power distribution $F(x) = \left(\frac{x}{\alpha}\right)^{\beta},\ 0 \leq x \leq \alpha$, as the initial choice of income distibution in the lower tail. Making use of the data in the second and sixth classes, the method of percentiles is applied to get the estimates of α and β as $\alpha = 0.994$ and $\beta = 2.0712$. The observed and expected frequencies in the various classes shown in Table 2 (see Figure 1 for the closeness), gave a chi square value of 5.2711 for 4 degrees of freedom, which does not reject the hypothesis of power distribution with the estimated parameter values. By direct substitution of $\overline{F}(y)$ in (5.1), we have functional form of the SST indices as

$$P_1(y,z) = 0.6512\Big(\frac{y}{0.994}\Big)^{2.0712} - 0.1945\Big(\frac{y}{0.994}\Big)^{4.1424}.$$

Class	0 - 249	250 - 374	375 - 499	500 - 624	625 - 749	750 - 874	875 - 999
Observed	211	309	439	561	692	808	921
Expected	224	289	408	582	690	826	922

Table 2. Observed and expected frequencies

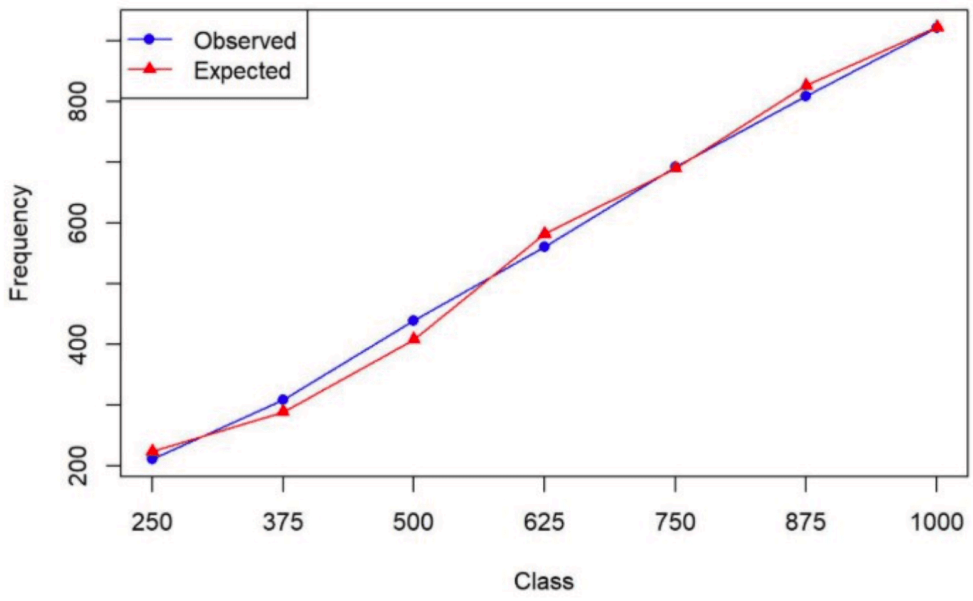


Figure 1. Frequency curves of the observed and expected distributions

The values of the Sen-Shorrocks index at the 5% through 40% points of the income distribution, taken as the poverty lines are given in Table 3.

Poverty percentile	5	10	15	20	25	30	35	40
$p_1(u)$	0.0325	0.0639	0.0945	0.1239	0.1523	0.1801	0.2063	0.2318

Table 3. Sen-Shorrocks index for Thailand data

doi.org/10.32388/5364ED

5.2. Recovering income distribution from SST

The input in this case is the values of $P_1(y,z)$ or $p_1(u)$ at various values of $Z(u)$ computed from the data. For explaining the methodology and easy verification of its validity, we assume the data in Table 3, without any other information. The first problem is to fix a functional form for $p_1(u)$ from values of $u = 0.05, 1.0, \cdots, 0.35, 0.4$. In general, this is done by using the standard curve fitting technique or by appealing to a suitable parametric form as explained in Section 4. For the present data the forward differences are computed, which shows that the second order differeences are approximately constant around 0.0007. This suggests a quadratic function of $p_1(u)$ in the form $p_1(u) = Au + Bu^2$, without involving a constant term as $p_1(u) = 0$ as u tends to zero. Employing the method of least squares to fit the model to the data, we find the estimates of A and B as $\hat{A} = 0.6577$ and $\hat{B} = -0.1992$. The mean squared error of estimation is

$$MSE = \frac{1}{n}\sum \left(Au_i + Bu_i^2 - 0.6577u + 0.1992u^2\right)^2 = 0.00182$$

with $n = 8, A = 0.6512$ and $B = 0.1945$ corresponding to Table 3. With the error being considerably small, we take

$$p_1(u) = 0.6577u - 0.1992u^2$$

that corresponds to a power distribution with, $\beta = 2.0409$ and $\alpha = 959$

$$Q(u) = 995u^{2.0409} \text{or} F(y) = \left(\frac{y}{995}\right)^{2.0409}.$$

The discrepancy with the original value $\beta = 2.0729$ and $\alpha = 994$ is due to the error in estimation and approximation of the values to four decimal places. As the error being considerably small, the closeness of the densities of the actual and recovered distributions is evident from their plots given in Figure 2.

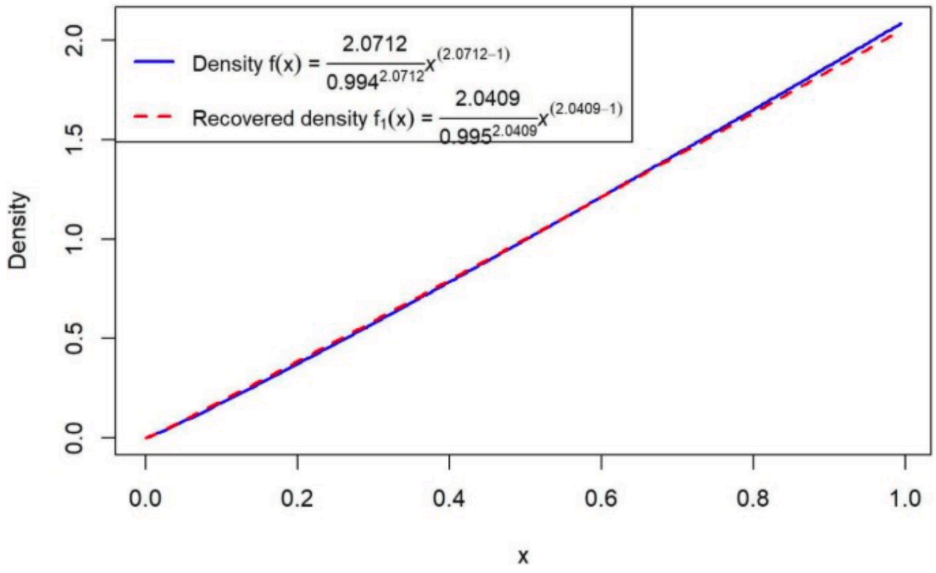


Figure 2. Density functions of the actual and recovered distributions

6. Conclusion

We have discussed some new properties of the Sen-Shorrocks poverty index and the deprivation profile arising from the income distribution. The definitions of the Sen-index, Sen-Shorrocks index and the deprivation profile are extended to the case when the income distribution is represented by quantile functions. An attempt is also made to find the income distribution from the above indices and to ascertain whether their functional forms uniquely determines the distribution. It is shown that there can be more than one income distribution corresponding to a given Sen-index. Some general forms of deprivation profile curve and their distributions are derived. The problem of determining the income model from the poverty index is illustrated by means fo a real data. The possibility of extending similar results for measures of income inequality and welfare functions and their implementation to practical problems are problems of further interest.

Statements and Declarations

Data Availability

The data that support the findings of this study are openly available in documents1.worldbank.org at https://documents1.worldbank.org/curated/zh/821871468132882698/pdf/DWE8201000Data0ribution0in0Thailand.pdf, reference number DWE - 8201.

Ethics

The data used in this study were obtained from Kansal, S. (1982). EPD income distribution project data on income distribution in Thailand. Division Working Paper No. 1982-1, Economic and Social Data Division, Economic Analysis and Projections Department, Development Policy Staff, The World Bank, and ethical approval was not required.

References

1. ^a, b, c, d *Sen A (1976). "Poverty: An Ordinal Approach to Measurement." Econometrica: J Econom Soc. **44**(2):219–231.*
2. ^ *Takayama N (1979). "Poverty, Income Inequality, and Their Measures: Professor Sen's Axiomatic Approach Reconsidered." Econometrica: J Econom Soc. **47**(3):747–759.*
3. ^ *Thon D (1979). "On Measuring Poverty." Rev Income Wealth. **25**(4):429–439.*
4. ^a, b *Kakwani N (1980). "On a Class of Poverty Measures." Econometrica: J Econom Soc. **48**(2):437–446.*
5. ^ *Thon D (1983). "A Poverty Measure." Indian Econ J Econom Soc. **36**:55–70.*
6. ^a, b, c, d, e *Shorrocks A (1995). "Revisiting the Sen Poverty Index." Econometrica: J Econom Soc. **63**(5):1225–1230.*
7. ^ *Silber J, Espinoza-Delgado J (2023). "A Rank-Dependent Multidimensional Deprivation Index (MDI) for Binary Data."*
8. ^ *Kulkarni V, Gaiha R (2021). "Beyond Piketty: A New Perspective on Poverty and Inequality in India." J Policy Model. **43**(2):317–336.*
9. ^ *Dizon R (2021). "Tax Incidence of Philippine Tax Reform: Poverty and Distributional Effect." J Asian Finance, Econ Bus. **8**(2):281–288.*

10. ^Igbatayo S, Yakubu M, Maji I (2021). "Socio Economic and Demographic Analysis of Poverty Among Micro and Small Scale Entrepreneurs in Nigeria: A Case Study of Jalingo, Taraba State." Afr J Bus Econ Dev. **2782**:7 658.
11. ^García-Gómez C, Perez A, Prieto-Alaiz M (2019). "A Review of Stochastic Dominance Methods for Poverty Analysis." J Econ Surv. **33**(5):1437–1462.
12. ^Zheng B (2023). "Poverty Measurement and Stochastic Dominance." Res Handb Meas Poverty Deprivatio n. :82–94.
13. a, bShorrocks A (1998). Deprivation Profiles and Deprivation Indices. Colchester: Cambridge University Pres s.
14. ^Tarsitano A (2004). "Fitting the Generalized Lambda Distribution to Income Data."
15. ^Tarsitano A, others (2006). "A New Q-Q Plot and Its Application to Income Data." Stat Appl. **4**(1):1–26.
16. ^Haritha N, Nair N, Nair K (2008). "Modelling Incomes Using Generalized Lambda Distributions." J Income Distrib. **17**(2):37–51.
17. a, bNair N, Vineshkumar B (2022). "Cumulative Entropy and Income Analysis." Stochast Qual Control. **37**(2): 165–179.
18. ^Varkey A, Haitha H (2024). "Singh-Maddala-Dagum Distribution With Applications to Income Data." J Soc Stat, Comput Appl. **22**(1):321–341.
19. ^Safari M, Masseran N, Ibrahim K, AL-Dhurafi N (2020). "The Power-Law Distribution for the Income of Po or Households." Physica A. **557**:124893.
20. ^Kakwani N, Podder N (1973). "On the Estimation of Lorenz Curves From Grouped Observations." Int Econ Rev. **14**(2):278–292.
21. ^Aggarwal V (1984). "On Optimum Aggregation of Income Distribution Data." Sankhyā B. **46**(3):343–355.
22. ^Gupta M (1984). "Functional Form for Estimating the Lorenz Curve." Econometrica. **52**(5):1313–1314.
23. ^Ortega P, Martin G, Fernandez A, Ladoux M, Garcia A (1991). "A New Functional Form for Estimating Loren z Curves." Rev Income Wealth. **37**(4):447–452.
24. ^Chotikapanich D (1993). "A Comparison of Alternative Functional Forms for the Lorenz Curve." Econ Lett. **41**(2):129–138.
25. ^Rohde N (2009). "An Alternative Functional Form for Estimating the Lorenz Curve." Econ Lett. **105**(1):61–6 3.

26. ^Paul S, Shankar S (2020). "An Alternative Single Parameter Functional Form for Lorenz Curve." Empir Econ. **59**:1393–1402.
27. ^Kansal S (1982). "EPD Income Distribution Project Data on Income Distribution in Thailand." Economic and Social Data Division, Economic Analysis and Projections Department, Development Policy Staff, The World Bank.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.