A NOTE ON THE MULTIDIMENSIONAL DECOMPOSITION OF SEN'S INDEX[1]

Stéphane Mussard

LAMETA, Université de Montpellier I
UFR Sciences Économiques

Kuan Xu

Department of Economics
Dalhousie University

[1] Corresponding author: Kuan Xu, Department of Economics, Dalhousie University, Halifax, NS - Canada B3H 3J5; E-mail: kuan.xu@dal.ca; Tel: 902-494-6995; Fax: 902-494-6917.

ABSTRACT

Given the multiplicative decomposition of the Sen index into three commonly used poverty statistics – the poverty rate (poverty incidence), poverty gap (poverty depth) and 1 plus the Gini index of poverty gap ratios of the poor (inequality of poverty) – the index becomes much easier to use and interpret for economists, policy analysts and decision makers in addition to its desirable axiomatic properties. Based on the recent findings on simultaneous subgroup and source decomposition of the Gini index, we further decompose the Sen index and its components. This extension would be of value from both theoretical and empirical perspectives.

Key words: Gini index, Sen index, Source decomposition, Subgroup decomposition.

JEL Classification: I32, D63, D31.

I. INTRODUCTION

Finding appropriate measures for inequality and poverty is of great significance to economists, policy analysts, and decision makers.[2] On one hand, these measures provide the true state of an economy or a society where inequality and poverty issues must be addressed. On the other hand, these measures must be assessable and understandable by economists, policy analysts and decision makers.

Over the last few decades, the literature on inequality and poverty measures has evolved significantly. One of the important developments was the proposal of a poverty measure made by Amartya Sen [22], now called the Sen index. This index is attractive because it measures not only the commonly used poverty rate (or headcount), but also the average poverty gap (or average relative income shortfall of the poor below the poverty line), and the inequality of the poor. Using this index would not lead to the undesirable poverty-reduction actions such as providing a small amount of additional income assistance to the richest of the poor, which would lead to the reduction of the poverty rate but not the poverty gap and inequality of the poor.

More specifically, as noted by Xu and Osberg [28, 29], the Sen index is easy to use and to understand precisely because of its decomposability into three measures of poverty: incidence (the poverty rate), depth (poverty gap), and inequality (1 plus the Gini index of poverty gaps). Naturally, economists and policy analysts would like to know whether it is possible to further decompose the Sen index components according to social groups (subgroup decomposition: age group, educational, regional, etc.) or consumption/expenditure/income sources (consumption source decomposition: food and non-food consumptions; expenditure source decomposition: consumption bundles; income source decomposition: labour income,

---

[2] For example, Ghiglino and Sorger [11] note that poverty trap has a lot to do with the wealth distribution. In this context, how to best measure poverty becomes an important task.

investment income, social assistance, child-support benefits, etc.). This note will address this question with some new findings.

In particular, consumption and expenditure source decomposition is a common practice in poverty research in developing countries. This is because absolute poverty is a major problem in developing countries while relative poverty is more focused on in developed countries. Typically, the minimum consumption bundle is used to set up the poverty line that represents minimum food and non-food consumption items to satisfy the minimum calories (around 2000-2400) and some minimum non-food necessities. If one individual or a household commands more than what the poverty line represents, this economic unit is considered non-poor; otherwise, it is considered poor. In this context, the source decomposition is indeed very important. There are two major practices in this regard – one is the direct method which decomposes the poverty line into food and non-food consumption items; the other is the indirect method which use the Engle's coefficient to allocate the minimum expenditure into the food versus non-food sources (say, food 1 versus non-food 0.2-0.7 depending on the situation of a specific country involved).[3]

The advantage for an inequality/poverty measure to possess subgroup and source decomposability is that this allows researchers to measure and, therefore, to understand how each contributing components affect the overall inequality/poverty. However, decomposition must be tractable in the sense that these components must have clear interpretations and logical relations to the aggregate inequality/poverty measure. Furthermore, decomposition can facilitate the communication between researchers and policy markers. From the technical point of view, economists sometimes require the decomposed parts to be additive for the whole of these parts. This requirement rules out many widely used inequality/poverty

---

[3] The two approaches are reported by participating countries at the 2004 International Conference on Official Poverty Statistics in Manila, Philippines, Oct. 2-6, 2004.

measures such as the Gini index and Sen index that are not viewed as additive decomposable or subgroup consistent (see Shorrocks [23], Foster and Shorrocks [10]).

However, Lambert and Aronson [15], among many other contributors, summarized and re-interpreted the subgroup decomposition of the Gini index while Fei, Ranis and Kuo [9] discussed the source decomposition. Rao [20] considered both subgroup and source decomposition. Silber [25] also used a matrix algebra approach to deal with subgroup or source decomposition. Recently Mussard [17, 18] proposed a new way to decompose the Gini index from two different but related perspectives; that is, the decomposition permits a transparent link between subgroup decomposition and source decomposition. It is well known that the Sen index has three decomposable components: the poverty rate, average poverty gap ratio and one plus the Gini index of inequality of poverty gap ratios. As pointed in Xu [26] and Irvine and Xu [13] that each of these components is subgroup decomposable. But these papers do not discuss source decomposition and how source decomposition is related to subgroup decomposition in a transparent way.

The aim of this paper is therefore to analyze how the Sen index and its components can be decomposed *jointly* by source and subgroup. This is because such a discussion may shed additional light to the measurement of poverty in terms of identifying sources of income shortfalls and most affected social subgroups. On the technical level, this note will use the results in Xu and Osberg [28, 29] and Mussard [17, 18] to obtain the theoretical findings on the multidimensional decomposition of the Sen index.

The remainder of the paper is organized as follows. In section II, notation and identification issues are introduced and discussed. Section III introduces the available relevant results that are applicable to the rest of the analysis in the paper. In section IV, the multidimensional decomposition of the Sen index is examined. The properties of the multi-decomposition are then evaluated in section V. Finally, in section VI, a brief conclusion is provided.

II. NOTATION AND IDENTIFICATION

Researchers engaging in poverty studies must make two fundamental decisions: (1) identifying the poor, which involves establishing a suitably selected poverty line or threshold, and (2) finding a good poverty measure, which is often debated and justified on the basis of a set of commonly agreed rules called axioms. For the first decision, it is largely a responsibility of the social planner and statistics administrator. For the second decision, before the central finding of the Sen index in 1976, the commonly used measure of poverty has been the poverty rate or headcount. Now, it has been recognized that the poverty rate is a misleading measure of poverty because, unlike the Sen index, it only measures the incidence of the poverty but ignores the depth and inequality of the poor.

To facilitate the further discussion, now introduce some notation. Let the number of consumption/expenditure/income units, say individuals, in a population be n and the number of the poor individual whose consumption/expenditure/income below the poverty line z in consumption/expenditure/income be q. In this population there are K distinct subgroups. In subgroup k there are total n_k individuals and q_k poor individuals. The overall poverty rate is

$H = \dfrac{q}{n}$ and the poverty rate for subgroup k is $H_k = \dfrac{q_k}{n_k}$ with $H = \sum\limits_{k=1}^{K} \dfrac{n_k}{n} H_k$.

Given consumption/expenditure/income of the poor individual y_i and the poverty line z, one can define the poverty gap (sometime called relative poverty gap or poverty gap ratio) as:

$$x_i = \begin{cases} \dfrac{z - y_i}{z}, & \forall z > y_i \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

for all q poor individuals. Then, the vector of poverty gaps of the poor is given by:

$x_p = (x_1, ..., x_i, ..., x_q)$.

Now, consider a partition of each y_i into M sources $y_i = \left(y_i^1, ..., y_i^m, ..., y_i^M \right)$ such that:

$$\sum_{m=1}^{M} y_i^m = y_i. \tag{2}$$

In other words, the total consumption/expenditure/income of an individual is the sum of all consumption/expenditure/income components received by this individual. For simplicity of our discussion in this paper, we focus only on income sources (factor component) not on consumption or expenditure although consumption or expenditure is often used in practice.

The identification of the poor is based on to whether or not the income of an individual income y_i falls below the poverty line z. Obviously, this criterion is also applicable to any subgroups. However, when we attempt to analyse the contributions of shortfalls in income components to shortfalls in overall income, we have to consider and accept a condition, that is, the poverty line can be suitably decomposed according to different sources:

$$\sum_{m=1}^{M} z^m = z. \tag{3}$$

How to determine the values of z^m's depends on the norm on how much money an average individual is supposed to receive from source m as minimum levels (z^m's). Of course, the normal is unlikely to be unique given various social and economic conditions of individuals. However, for poverty analysis in developing countries, the minimum food and non-food consumption (and, hence, their components) can be appropriately determined. The former is based on the basic calories and protein required while the latter is settled by the normal minimum need of the local standard.

A numerical example can be used to explain the above point. Let the poverty line z be 5 in a society where there are two poor individuals. They have two consumption/expenditure/income sources: 1 (food items) and 2 (non-food items). The total consumption/expenditure/incomes for the first and the second poor individuals are respectively:

$$y_1 = y_1^1 + y_1^2 = 3 + 1 = 4,$$

$$y_2 = y_2^1 + y_2^2 = 2 + 1 = 3.$$

In order to make source decomposition operational, introduce the following additional condition for this numerical example:

$$x_i = \frac{z - y_i}{z} = \frac{\left(z^1 - y_i^1\right) + \left(z^2 - y_i^2\right)}{z}. \tag{4}$$

The question is: what will be suitable z^1 and z^2? Because all of the poor people have income less than z, their average total income must be less than z and the sum of the averages of all components must be less than z. Hence, z^1 and z^2 can be set either by the experts' view as to food versus non-food needs or proportionally to the averages of all components of the reference poor population (say 20 or 40% of the low income group of the total population). The first choice is simple to understand. The second choice differs from the first in that decomposition of z is made according to the average income source structure of the reference poor population. More specifically,

$$z^1 = z\left(\frac{y_1^1 + y_2^1}{y_1 + y_2}\right) = 5\left(\frac{3+2}{4+3}\right) = 5\left(\frac{5}{7}\right)$$
$$z^2 = z\left(\frac{y_1^2 + y_2^2}{y_1 + y_2}\right) = 5\left(\frac{1+1}{4+3}\right) = 5\left(\frac{2}{7}\right). \tag{5}$$

As a result, this leads to: $z = z^1 + z^2 = 5$. Then, the general configuration of the poverty gap involving the source decomposition is:

$$x_i^m = \frac{z^m - y_i^m}{z}, \tag{6}$$

where x_i^m is the poverty gap in source m of individual i such that $\sum_{m=1}^{M} x_i^m = x_i$. Such a decomposition of z has an important feature. That is, while x_i is nonnegative, its components x_i^m's can be positive, or zero, or negative implying that the income component y_i^m is less than, or equal to, or greater than the chosen benchmark z^m.

Then, it is possible to define the total income shortfall of individual i who is in subgroup k:

$$x_{ik} = \frac{z - x_{ik}}{z} \ . \tag{7}$$

Now, the poverty gap in source m of individual i in group k is:

$$x_{ik}^m = \frac{z^m - x_{ik}^m}{z} \ . \tag{8}$$

Then, the average poverty gap for the population and that for subgroup k are

$$\bar{x}_p = \sum_{i=1}^{q} \frac{x_i}{q} \quad \text{and} \quad \bar{x}_{p(k)} = \sum_{i=1}^{q_k} \frac{x_{ik}}{q_k}, \tag{9}$$

respectively. These two average gaps are related as follows:

$$\bar{x}_p = \sum_{k=1}^{K} \frac{q_k}{q} \bar{x}_{p(k)} . \tag{10}$$

III. SOME EXISTING RESULTS FOR MULTIDIMENSIONAL DECOMPOSITION OF THE SEN INDEX

The findings of this note will be based on some earlier findings in Xu and Osberg [28, 29] and Irvine and Xu [13]. It is noted that multiplicative decomposition of the Sen index and a further subgroup decomposition of each component are possible:

$$S = H \bar{x}_p (1 + G)$$

$$\Leftrightarrow \quad S = \sum_{k=1}^{K} \frac{n_k}{n} H_k \times \sum_{k=1}^{K} \frac{q_k}{q} \bar{x}_{p(k)} \times \left(1 + \left(G_w + G_b + G_t\right)\right), \tag{11}$$

where G is the Gini index of poverty gap ratios of the poor, G_w is the contribution of the inequalities within K subgroups, G_b is the contribution of the inequalities among K subgroups excluding the overlap between the distributions of these groups, and G_t is the inequalities between K subgroups limited to the overlap between the conditional distributions or the intensity of transvariation (see Gini [12], Dagum [2, 3, 4, 5, 6]). Note that here G_w, G_b, and G_t are not Gini indices in their original sense. Rather, these are quantities for inequality of different kinds Dagum names G_w 'the Gini inequality within subpopulations'; G_b 'the net

extended Gini inequality between subpopulations' and G_t 'the intensity of transvariation between subpopulations'. However, Xu and Osberg [28, 29] and Irvine and Xu [13] do not suggest any source decomposition for the Sen index.

Mussard [17, 18] addresses the multi-decomposition issue of the Gini index and finds that the Gini index is simultaneously decomposable by income source and by subgroup:

$$G = \sum_{m=1}^{M}\left(G_w^m + G_b^m + G_t^m\right) = \sum_{m=1}^{M}\left(G_w^m + G_{gb}^m\right), \tag{12}$$

where G_w^m, G_b^m, G_t^m, and G_{gb}^m are respectively the contributions of the mth income source to G_w, G_b, G_t and G_{gb}. The latter [G_{gb}] stands for the gross between-group inequality. Indeed, following Dagum's methodology [5], one can combine the net between-group inequality and intensity of transvariation as follows:

$$G_{gb} = G_b + G_t. \tag{13}$$

The gross between-group Gini index, G_{gb}, measures inequality between each and every pair of the overall population in a more complete sense than the standard net between-group Gini index, G_b, that only measures inequality existing among the mean incomes of all subgroups. This simultaneous method enables one to compute the contribution of each source to the quantities G_w, G_b, G_t and G_{gb}. Of course, the Gin index to be composed in the Sen index is for poverty gap ratios of the poor not incomes of the population.

The Gini index of poverty gap ratios of the poor can be written as:

$$G = \frac{\sum_{i=1}^{q}\sum_{r=1}^{q}\left|x_i - x_r\right|}{2\bar{x}_p q^2}, \tag{14}$$

where $x_i = \max\left\{0, \dfrac{z - y_i}{z}\right\}$, ($i = 1, \ldots, q$), is the poverty gap ratio of individual i. Note that poverty gap ratios for the non-poor are zeros. The Gini index of poverty gap ratios of the poor can then be expressed as:

$$G = \frac{\sum\limits_{i=1}^{q}\sum\limits_{r=1}^{q}(x_i+x_r-2\min\{x_i, x_r\})}{2\bar{x}_p q^2},$$ (15)

since

$$|x_i-x_r| = (x_i+x_r-2\min\{x_i, x_r\}).$$ (16)

For $x_i = x_r$, $|x_i-x_r| = 0$; that is, we do not decompose two x's when $x_i = x_r$. This condition will make our decomposition consistent with the aggregate Gini index. However, if $x_i \neq x_r$, we can decompose the term $2\min\{x_i, x_r\}$ by factor components:

$$\sum\limits_{m=1}^{M} 2x_{ir}^{*m} = 2\min\{x_i, x_r\},$$ (17)

where x_{ir}^{*m} is an operator which selects the minimum of x_i and x_r and then decompose the minimum poverty gap (either x_i or x_r) by M sources. For example, if $x_i = \min\{x_i, x_r\}$ and there are only two sources or $M = 2$, then $x_i = x_i^1 + x_i^2$ and $\sum\limits_{m=1}^{M} 2x_{ir}^{*m} = 2(x_i^1 + x_i^2)$. Consequently, the Gini index of poverty gaps of the poor can be expressed as:

$$G = \sum\limits_{m=1}^{M}\left(\frac{\sum\limits_{i=1}^{q}\sum\limits_{r=1}^{q}(x_i^m+x_r^m-2x_{ir}^{*m})}{2\bar{x}_p q^2}\right).$$ (18)

The Gini index [Eq. (18)] is broken down into M contributions representing the weight of each factor component to the global inequality of poverty gaps of the poor.

According to Dagum [5, 6], it is possible to decompose the Gini index of poverty gap ratios for K subgroups:

$$G = \frac{\sum\limits_{k=1}^{K}\left(\sum\limits_{i=1}^{q_k}\sum\limits_{r=1}^{q_k}|x_{ik}-x_{rk}|\right)}{2\bar{x}_p q^2} + \frac{2\sum\limits_{k=2}^{K}\sum\limits_{h=1}^{k-1}\left(\sum\limits_{i=1}^{q_k}\sum\limits_{r=1}^{q_h}|x_{ik}-x_{rh}|\right)}{2\bar{x}_p q^2},$$ (19)

where x_{ik} is the poverty gap ratios of individual i who is in subgroup k. Therefore, the Gini index defined in Eq. (19) can be expressed as the sum of two components:

$$G = G_w + G_{gb}, \tag{20}$$

where (*i*) G_w is the contribution of the inequality of poverty gap ratios within K subgroups (*ii*) G_{gb} is the gross contribution of the inequality of poverty gap ratios among K subgroups. Now combine the previous two findings by substituting the source decomposition of the Gini index in Eq. (18) into the subgroup decomposition in Eq. (19) giving

$$G = \sum_{m=1}^{M} \left(\frac{\sum\limits_{k=1}^{k} \left(\sum\limits_{i=1}^{q_k} \sum\limits_{r=1}^{q_k} \left(x_{ik}^m + x_{rk}^m - 2x_{irk}^{*m} \right) \right)}{2\bar{x}_p q^2} \right) + \sum_{m=1}^{M} \left(\frac{2\sum\limits_{k=2}^{K} \sum\limits_{h=1}^{k-1} \left(\sum\limits_{i=1}^{q_k} \sum\limits_{r=1}^{q_k} \left(x_{ik}^m + x_{rh}^m - 2x_{irkh}^{*m} \right) \right)}{2\bar{x}_p q^2} \right), \tag{21}$$

$$= \sum_{m=1}^{M} \left(G_w^m + G_{gb}^m \right),$$

where x_{irkh}^{*m} is the *m*th source of the minimum value of x_{ik} and x_{rh} (that is x_{ik}^m or x_{rh}^m), and x_{irk}^{*m} is the *m*th source of the minimum value of x_{ik} and x_{rk} (that is x_{ik}^m or x_{rk}^m). The within-group and the gross between-group elements are decomposed by income source.

The decomposition in Eq. (21) highlights the possibility of decomposing the Gini index simultaneously according to income source and subgroup (see Mussard [17, 18]). However, it is also important to note that G_w^m and G_{gb}^m are not Gini indices in their original sense. In fact, these are quantities proportional to the relative Gini differences (see Xu [27, p. 15]).

It is well-known that the Gini index is not subgroup consistent. The concept of subgroup consistency of an inequality measure (SCIM) can be explained as follows. Let p_k be the proportion of population belonging to subgroup k and s_k the income share of subgroup k ($\forall k = 1,\ldots,K$). A measure of inequality I satisfies the subgroup consistency property if:

$$I = f\left(I_1,\ldots, I_k,\ldots, I_K \, ; \, p_1,\ldots, p_k,\ldots, p_K \, ; \, s_1,\ldots, s_k,\ldots, s_K \right), \tag{SCIM}$$

where f is increasing in its first K arguments. In other words, consider a situation where group k has a change in incomes, *ceteris paribus*, such that the mean income and the number of individuals remain constant. The measure of inequality is said to be subgroup consistent if an

increase (or a decrease) in group k's inequality leads to an increase (or a decrease) in the overall inequality. In this clearly defined way, the Gini index is not subgroup consistent (see Shorrocks [24]). Neither is the Sen index.

Now let us review the concept of subgroup consistency of poverty measures (SCPM) due to Foster and Shorrocks [10]. Let A and B be two distributions within an aggregate distribution of the two sub-distributions (A, B), $P(. ; z)$ an index of poverty, and $n(A)$ the number of individuals in distribution A. For simplicity, two distributions A and B may change to distributions A' and B'. Also assume the number of observations n in these distributions maintains the following properties: $n(A) = n(A')$ and $n(B) = n(B')$. If poverty increases when distribution A' changes to distribution A so that $P(A ; z) > P(A' ; z)$ while poverty in distribution B remains the same, then the overall poverty level increases :

$$P(A, B ; z) > P(A', B ; z). \tag{SCPM}$$

While subgroup consistency is a useful concept for constructing inequality (poverty) measures, it can be also argued that some other principles may be just as important. Pyatt [19] notes that the Gini index enables the interpersonal comparison between each and every pairs of individual's incomes. Behind this insight, quite consistent with the sociological findings, is that the relative position of an individual in an income distribution has substantial impact on the sense of inequality (poverty). Sen [21] also argues that " [I]n any pair-wise comparison the man with the lower income can be thought to be suffering from some depression on finding his income to be lower. Let this depression be proportional to the difference in income. The sum total of all such depressions in all possible pair-wise comparisons takes us to the Gini coefficient."

The above views highlight an interesting and powerful point. When inequality (or poverty) decreases in one subgroup, say group 1, and remains constant in other subgroups, the overall inequality (poverty) can increase if the individuals of other subgroups feel that their relative

13

rankings are not improving at the same rate from the change in inequality (poverty) of group 1. This intuitive argument indeed violates the often-cited subgroup consistency. However, it is accommodated by the Gini index of inequality, and hence the Sen index of poverty intensity, since these two indices employ the Gini social welfare function implicitly (see Xu and Osberg [29]) and both involve all the interpersonal comparisons (see also Dagum [7] and the related pair-based measures of income inequality of Kolm [14]).

With the same token, the interpretation of the Gini, and hence Sen, index is sociological consistent but not necessarily subgroup consistent. An interesting case about the third subgroup component of the Gini index G_t, it is not subgroup consistent but it measures the intensity of transvariation and still makes sense in the subgroup decomposition context − a positive value of this term can be interpreted as a factor contributing to the overall inequality. Of course, this notion is not new at all. It was also introduced by Gini himself in 1916 and extended by Dagum [2, 3, 4] later.

Because of the above observations, many authors still consider the decomposition of the Gini index, and hence the Sen index, represents a valuable exercise from theoretical and practical point of view (Battacharya and Mahalanobis [1], Rao [20], Pyatt [19], Silber [25], Lerman and Yitzhaki [16], Lambert and Aronson [15], Dagum [6, 7], Deutsch and Silber [8], for example). It becomes apparent that subgroup consistency is a useful concept but it should not limit our use of many inequality (poverty) measures that are attractive in other ways.

In this note, it is noted that the Sen index is not subgroup consistent. However it should not pose any problem for us to explore decomposability as many earlier authors did. As may be illustrated later in the paper, the proposed decomposition is indeed useful.

IV. AN EXTENSION: THE MULTI-DECOMPOSITION OF THE SEN INDEX AND ITS

COMPONENTS

Rewriting the decomposition in Eq. (11) using Eq. (13) yields:

$$S=\sum_{k=1}^{K}\frac{n_k}{n}H_k\times\sum_{k=1}^{K}\frac{q_k}{q}\,\overline{x}_{p(k)}\times\left(1+\left(G_w+G_{gb}\right)\right). \tag{22}$$

Remark 1.

The Gini index measures the inequality of poverty gap ratios, instead of that of incomes, of the poor. Further, the component of the Sen index, $1+(G_w+G_{gb}) = 1+G$, is an extended Gini index of inequality of poverty gap ratios that takes value in $[1, 2]$.[4]

Proposition 1.

If there exists a partition in factor components of both incomes and the poverty lines, then the Gini index of poverty gap ratios of the poor is multi-decomposable and the multiplicative decomposition of the Sen index is permitted.

Proof.

Given (6), (12), (13) and (22), the Sen index can be written as:

$$S=\sum_{k=1}^{K}\frac{n_k}{n}H_k\times\sum_{k=1}^{K}\frac{q_k}{q}\,\overline{x}_{p(k)}\times\left(1+\left(\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)\right). \quad\blacksquare \tag{23}$$

This mixture decomposition is potentially useful for applied researchers because it permits source and subgroup decomposition of the determinants of the overall poverty. Let G_e be the extended Gini index of inequality[5]: $G_e = 1+G$. If G_e is multi-decomposable as in Eq. (21), then source and subgroup contributions to G_e can be obtained.

---

[4] This term can be interpreted as a measure for inequality aversion, which is studied thoroughly by Zheng [30].
[5] In this case the extended Gini index of inequality G_e belongs to the close interval $[1,2]$.

Proposition 2.

The extended Gini index of inequality is multi-decomposable.

Proof.

The extended Gini index of inequality is defined as:

$$G_e = 1+G = \frac{2\bar{x}_p q^2 + \sum_{i=1}^{q}\sum_{r=1}^{q}|x_i-x_r|}{2\bar{x}_p q^2}. \tag{24}$$

The denominator can be rewritten such as:

$$2\bar{x}_p q^2 = \sum_{i=1}^{q}\sum_{r=1}^{q}(x_i+x_r). \tag{25}$$

Then, the extended Gini index of inequality becomes:

$$G_e = 1+G = \frac{\sum_{i=1}^{q}\sum_{r=1}^{q}(x_i+x_r) + \sum_{i=1}^{q}\sum_{r=1}^{q}|x_i-x_r|}{2\bar{x}_p q^2}. \tag{26}$$

Note that term A is:

$$A = 1 = \frac{\sum_{i=1}^{q}\sum_{r=1}^{q}(x_i+x_r)}{2\bar{x}_p q^2}, \tag{27}$$

and term B is:

$$B = \frac{\sum_{i=1}^{q}\sum_{r=1}^{q}|x_i-x_r|}{2\bar{x}_p q^2}. \tag{28}$$

To reach an exact multi-decomposition without redundant term, A and B must be decomposable both by subgroup and income source. As we have pointed out before, a good way to decompose inequality indexes by subgroup is to use the interpersonal comparisons. Gathering these comparisons within subgroups and between subgroups (between each and pairs of subgroups), we have:

$$A = \frac{\sum\limits_{i=1}^{q}\sum\limits_{r=1}^{q}(x_i + x_r)}{2\bar{x}_p q^2} = \frac{\sum\limits_{k=1}^{K}\left(\sum\limits_{i=1}^{q_k}\sum\limits_{r=1}^{q_k}(x_{ik}+x_{rk})\right)}{2\bar{x}_p q^2} + \frac{2\sum\limits_{k=2}^{K}\sum\limits_{h=1}^{k-1}\left(\sum\limits_{i=1}^{q_k}\sum\limits_{r=1}^{q_h}(x_{ik}+x_{rh})\right)}{2\bar{x}_p q^2}. \tag{29}$$

Moreover, term A is multi-decomposable by both subgroup and income source into K subgroups and M income sources. By applying the source decomposition over the subgroup decomposition we get:

$$A = \frac{\sum\limits_{m=1}^{M}\left(\sum\limits_{k=1}^{K}\left(\sum\limits_{i=1}^{q_k}\sum\limits_{r=1}^{q_k}(x_{ik}^m+x_{rk}^m)\right)\right)}{2\bar{x}_p q^2} + \frac{\sum\limits_{m=1}^{M}\left(2\sum\limits_{k=2}^{K}\sum\limits_{h=1}^{k-1}\left(\sum\limits_{i=1}^{q_k}\sum\limits_{r=1}^{q_h}(x_{ik}^m+x_{rh}^m)\right)\right)}{2\bar{x}_p q^2}. \tag{30}$$

$$= \sum\limits_{m=1}^{M} A_w^m + \sum\limits_{m=1}^{M} A_{gb}^m.$$

Given the multi-decomposition of term B [Eq. (21)], the multi-decomposition of G_e is:

$$G_e = \overbrace{\sum\limits_{m=1}^{M} A_w^m + \sum\limits_{m=1}^{M} G_w^m}^{\sum\limits_{m=1}^{M} G_{ew}^m} + \overbrace{\sum\limits_{m=1}^{M} A_{gb}^m + \sum\limits_{m=1}^{M} G_{gb}^m}^{\sum\limits_{m=1}^{M} G_{egb}^m} \tag{31}$$

$$= \quad G_{ew} \quad + \quad G_{egb}.$$

Therefore, the extended Gini index of inequality is decomposable by both source and subgroup. This implies that both the contribution of source m to within-group equality G_{ew} and the contribution of source m to between-group equality G_{egb} can be computed. ∎

Corollary 1.

The multiplicative decomposition of the Sen index involves the multi-decomposition of the extended Gini index of inequality.

Proof.

It is straightforward:

$$S = \sum\limits_{k=1}^{K} \frac{n_k}{n} H_k \times \sum\limits_{k=1}^{K} \frac{q_k}{q} \bar{x}_{p(k)} \times \left(\sum\limits_{m=1}^{M}\left(G_{ew}^m + G_{egb}^m\right)\right). \ \blacksquare \tag{32}$$

Consequently, it is possible to measure the contribution of poverty gap ratios in source *m* in terms of within-group equality or between-group equality to the overall Sen index.

V. PROPERTIES OF THE MULTI-DECOMPOSITION OF THE SEN INDEX

Proposition 3.

The Sen index permits further source decomposition of the average poverty gap ratio:

$$
\begin{aligned}
S &= \sum_{k=1}^{K} \frac{n_k}{n} H_k \times \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{q_k}{q} \, \overline{x}_{p(k)}^{m} \times \left(1 + \sum_{m=1}^{M} \left(G_w^m + G_{gb}^m \right) \right) \\
&= \sum_{k=1}^{K} \frac{n_k}{n} H_k \times \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{q_k}{q} \, \overline{x}_{p(k)}^{m} \times \left(\sum_{m=1}^{M} \left(G_{ew}^m + G_{egb}^m \right) \right).
\end{aligned}
\tag{33}
$$

Proof.

Given that $\overline{x}_p = \sum_{k=1}^{K} \frac{q_k}{q} \overline{x}_p^{(k)}$, and $\tag{34}$

$$
\overline{x}_{p(k)} = \sum_{i=1}^{q_k} \frac{q_k \, x_i}{q_k} = \sum_{i=1}^{q_k} \frac{q_k \, \sum_{m=1}^{M} x_i^m}{q_k}
$$

$$
\Leftrightarrow \quad \overline{x}_{p(k)} = \sum_{m=1}^{M} \left(\frac{\sum_{i=1}^{q_k} x_i^m}{q_k} \right) = \sum_{m=1}^{M} \overline{x}_{p(k)}^{m} ,
\tag{35}
$$

the average poverty gap ratio in source *m* for subgroup *k* is

$$
\overline{x}_{p(k)}^{m} = \frac{\sum_{i=1}^{q_k} x_i^m}{q_k} \quad .
\tag{36}
$$

Then,

$$
\overline{x}_p = \sum_{k=1}^{K} \frac{q_k}{q} \, \overline{x}_{p(k)}
\tag{37}
$$

$$
\Leftrightarrow \quad \overline{x}_p = \sum_{k=1}^{K} \frac{q_k}{q} \sum_{m=1}^{M} \overline{x}_{p(k)}^{m}
$$

$$
\Leftrightarrow \quad \overline{x}_p = \sum_{m=1}^{M} \sum_{k=1}^{K} \frac{q_k}{q} \, \overline{x}_{p(k)}^{m} . \quad \blacksquare
\tag{38}
$$

18

Proposition 3 shows that the Sen index is totally decomposed by both income source and subgroup.

Proposition 4.

The multiplicative decomposition of the Sen index given in (22) and (32) permits the measurement of time variations in terms of the components of the Sen index.

Proof.

Given the operator $\Delta x = \mathrm{Ln} x_t - \mathrm{Ln} x_{t-1}$, one can show:

$$\Delta S = \Delta H + \Delta \bar{x}_p + \Delta\left(1 + G(x_p)\right) . \tag{39}$$

Then,

$$\Delta S = \Delta \sum_{k=1}^{K} \frac{n_k}{n} H_k + \Delta \sum_{m=1}^{M}\sum_{k=1}^{K} \frac{q_k}{q}\, \bar{x}_{p(k)}^m + \Delta\left(1 + \sum_{m=1}^{M}\left(G_w^m + G_{gb}^m\right)\right)$$

$$= \Delta \sum_{k=1}^{K} \frac{n_k}{n} H_k + \Delta \sum_{m=1}^{M}\sum_{k=1}^{K} \frac{q_k}{q}\, \bar{x}_{p(k)}^m + \Delta\left(\sum_{m=1}^{M}\left(G_{ew}^m + G_{egb}^m\right)\right). \ \blacksquare \tag{40}$$

This decomposition indicates that the poverty rate, the average poverty gap ratio and the extended Gini index of poverty gap ratios of the poor yield its contribution to the variation of Sen index between two periods.

Proposition 5.

According to the first-order Taylor approximation, the additive decomposition of the contributions of components of the Sen index to the overall Sen index is permitted.

Proof.

$$\Delta S = \mathrm{Ln}\left(\sum_{k=1}^{K}\frac{n_k}{n}H_k\right)_t - \mathrm{Ln}\left(\sum_{k=1}^{K}\frac{n_k}{n}H_k\right)_{t-1} + \mathrm{Ln}\left(\sum_{m=1}^{M}\sum_{k=1}^{K}\frac{q_k}{q}\,\bar{x}_{p(k)}^m\right)_t - \mathrm{Ln}\left(\sum_{m=1}^{M}\sum_{k=1}^{K}\frac{q_k}{q}\,\bar{x}_{p(k)}^m\right)_{t-1}$$
$$+ \mathrm{Ln}\left(1 + \sum_{m=1}^{M}\left(G_w^m + G_{gb}^m\right)\right)_t - \mathrm{Ln}\left(1 + \sum_{m=1}^{M}\left(G_w^m + G_{gb}^m\right)\right)_{t-1} . \tag{41}$$

Given that $\mathrm{Ln}(1+X) \cong X$ (the first-order of Taylor's approximation) then:

$$\text{Ln}\left(1+\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)_t - \text{Ln}\left(1+\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)_{t-1}$$

$$\cong \left(\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)_t - \left(\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)_{t-1}. \tag{42}$$

Consequently,

$$\Delta S \cong \Delta \sum_{k=1}^{K}\frac{n_k}{n}H_k + \Delta \sum_{m=1}^{M}\sum_{k=1}^{K}\frac{q_k}{q}\,\overline{x}_{p(k)}^m + \left(\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)_t - \left(\sum_{m=1}^{M}\left(G_w^m+G_{gb}^m\right)\right)_{t-1}. \tag{43}$$

Eq. (43) provides an additive approximation of the time variations of the multi-decomposition of the Sen index. ∎

It is interesting to note from the proof of Proposition 5, in particular Eq. (43), a change in the proportion of the poor and/or the average of income shortfalls below the poverty line are positive related to the change of the Sen index. Furthermore, the increase of inequality between period t and $t-1$ is positively related to the Sen index's variation. This observation confirms the principle of transfer in this literature on the basis of the changes in the Sen index over time. This is also true for the term $1+G$ in the Sen index since a Pigou-Dalton transfer decreases G (at period t) and then this decreases S (at period t), *ceteris paribus*.

The variations of the Sen index [Eq. (43)] can be captured by the contribution of the inequality of poverty gap ratios in addition to the contributions from the poverty rate and average poverty gap ratio. For instance, an increase (decrease) of the Sen index, *ceteris paribus*, can be caused by an increase (decrease) in the within-group and/or the gross between-group Gini index of poverty gap ratios. Indeed, given that interpersonal comparisons are permitted, each pair of individuals (within groups or between two groups) can compare their source poverty gap ratios.

The multi-decomposition of the Sen index enables the policy makers to detect precisely which source components (various incomes, food and non-food items, or various expenditures) are mainly responsible for the increasing poverty intensity. Further, the proposed decomposition

method allows decision makers to evaluate the impact of a economic policy on the variation of poverty intensity. For example, a change of fiscal policy may lead to the variation of the income sources such as tax and transfers across different social groups and ultimately to the variation of poverty intensity.

Finally, the multi-decomposition of the Sen index can be used to evaluate the economic policies in a much tighter framework, which links that individual poverty gaps with M variables X's describing individual characteristics, social and economic conditions and social policy indicators:

$$x_i = \sum_{m=1}^{M} \alpha_m X_i^m + \varepsilon_i, \tag{44}$$

where $X_i^1 = 1$, ε_i is the random error term and i is a subscript for individual i. In model selection, the hypothesis tests, say the Student-t tests, can be implemented to single out the significant variables for explaining the variations in poverty gap ratios. These variables include, not limited to, education, work experience, health condition, region, business conditions, social group membership, social policy, social program, and so on. Then we can compute the proportion of the poverty intensity predicted by these explanatory variables in the total poverty intensity.

Proposition 6.

The proportion of poverty intensity attributable to a set of explanatory variables in the aggregate poverty intensity can be derived from the parametric multi-decomposition of the Sen index.

Proof.

Let us rewrite the explained components: $\hat{x}_i^m = \hat{\alpha}_m X_i^m$, $i=1,...,n$. We still have a linear structure of poverty gap ratios: $\sum_{m=1}^{M}\hat{x}_i^m = x_i$. Introduce it in (43), we obtain the parametric Sen multi-decomposition:

$$\Delta\hat{S} \cong \Delta\sum_{k=1}^{K}\frac{n_k}{n}H_k + \Delta\sum_{m=1}^{M}\sum_{k=1}^{K}\frac{q_k}{q}\,\hat{\bar{x}}_p^m{}_{(k)} + \left(\sum_{m=1}^{M}\left(\hat{G}_w^m+\hat{G}_{gb}^m\right)\right)_t - \left(\sum_{m=1}^{M}\left(\hat{G}_w^m+\hat{G}_{gb}^m\right)\right)_{t-1}, \tag{45}$$

where the variables with '^' are estimated with (44). This parametric multi-decomposition yields the contribution of the variations in poverty gap ratios and inequality of poverty to the overall variation of the Sen index. ∎

Corollary 2.

The parametric Sen multi-decomposition entails the isolation of the marginal impact of each explanatory variables X^m.

Proof.

From (45), we obtain:

$$\frac{\Delta\hat{S}}{\Delta S} \cong \frac{\Delta\sum_{k=1}^{K}\frac{n_k}{n}H_k + \Delta\sum_{m=1}^{M}\sum_{k=1}^{K}\frac{q_k}{q}\,\hat{\bar{x}}_p^m{}_{(k)} + \left(\sum_{m=1}^{M}\left(\hat{G}_w^m+\hat{G}_{gb}^m\right)\right)_t - \left(\sum_{m=1}^{M}\left(\hat{G}_w^m+\hat{G}_{gb}^m\right)\right)_{t-1}}{\Delta S}. \tag{46}$$

Then, we can isolate the impact of a partial/marginal change of X^m by using $\Delta\hat{S}_{z^m}/\Delta\hat{S}$, where $\Delta\hat{S}_{z^m}$ is the Sen index with predicated values of the mth poverty gap ratio component based on $\hat{x}_i^m = \alpha_m X_i^m$ and the original data of other poverty gap ratio components x_i^l for $l \neq m$. By doing so, researchers and analysts can trace the roles that individual characteristics, social and economic conditions, and social policy indicators play in reducing poverty over time. ∎

VI. CONCLUSION

The note documents some extension of the decomposition of the Sen index. First, it shows that it is possible to obtain a mixture decomposition of the index along the line of subgroups as well as income sources. Second, it shows that the mixture decomposition can be extended to the possible decompositions to different time horizons. Third, the note shows that the logarithmic transformation permits the transformation of the multiplicative decomposition into an additive one. Fourth, it also demonstrates that the Taylor series approximation transforms the multi-decomposition of the extended Gini index of inequality into that of the Gini index of inequality. Fifth, the multi-decomposition can be extended into the parametric multi-decomposition, where a wide range of individual characteristics, social and economic conditions, and social policy indicators affecting poverty gap ratios can be identified and their impacts on poverty intensity can be evaluated.

In addition, this note also shows that the Gini and Sen indices can be decomposed by subgroup but the decomposition in this context is not subgroup consistent. However, it is known that the subgroup decomposition of Gini and Sen indices are inherently based on interpersonal comparisons, which are also of great value as noted and used by Pyatt (1976) and Dagum (1998). This note shows that value of interpersonal comparisons and how it embodies in the Sen index and various decompositions associated with it.

REFERENCES

1. Bhattacharya, N. and Mahalanobis, B., "Regional disparities in household consumption in India", *J. Amer. Statistical Assoc.* 62 (1967), 143-161.

2. Dagum, C., "Transvariazione fra più di due distribuzioni", in C. Gini, eds., *Memorie di Metodologia Statistica* 2 (1959), Roma, Libreria Goliardica.

3. Dagum, C., "Teoria de la transvariacion, sus aplicaciones a la economia", *Metron* 20 (1960), 1-206.

4. Dagum, C., "Transvariacion en la hipotesis de varaibles aleatorias normales multidimensionales", *Proceedings of the International Statistical Institute* 38(4) (1961), 473-486, Tokyo.

5. Dagum, C., "A new approach to the decomposition of the Gini income inequality ratio", *Empirical Econ.* 22 (1997), 515-531.

6. Dagum, C., "Decomposition and interpretation of Gini and the generalized entropy inequality measures", *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 157th Meeting (1997), 200-205.

7. Dagum, C., "Fondements de bien-être social et décomposition des mesures d'inégalité dans la répartition du revenu", *Econ. Appl.* L1(4) (1998), 151-202 [in French].

8. Deutsch, J. and Silber, J., "Inequality Decomposition by Population Subgroups and the Analysis of Interdistributional Inequality", in Silber J. (ed.), *Handbook of Income Inequality Measurement* (1999), Kluwer Academic Publishers, 163-186.

9. Fei, J.C.H., Ranis, G. and Kuo, S.W.Y, "Growth and the family distribution of income by factor components", *Quart. J. Econ.* (1978), 17-53.

10. Foster, J. E., and Shorrocks, A. F., "Subgroup Consistent Poverty Indices", *Econometrica* 59 (1991), 687-709.

11. Ghiglino, C. and Sorger, G. "Poverty traps, indeterminacy, and the wealth distribution", *J. Econ. Theory* 105 (2002), 120-139.

12. Gini, C., "Il concetto di transvariazione e le sue prime applicazioni", (1916) in C. Gini, eds., *Giornale degli Economisti e Rivista di Statistica* (1954), 21-44.

13. Irvine, I. and Xu, K., "Crime, punishment and the measurement of poverty in the United States 1979-1997" (2003), (mimeo), Department of Economics, Dalhousie University.

14. Kolm, S-C., "Rational Foundations of Income Inequality Measurement", in Silber J. (ed.), *Handbook of Income Inequality Measurement* (1999), Kluwer Academic Publishers, 19-94.

15. Lambert, P. J. and Aronson, J. R., "Inequality decomposition analysis and the Gini coefficient revisited", *Econ. J.* 103 (1993), 1221-1227.

16. Lerman, R., and Yitzhaki, S., "Income Stratification and Income Inequality", *Rev. Income Wealth* 37(3) (1991), 313-329.

17. Mussard, S., "A new approach to the Gini decomposition by income sources and the Gini decomposition by subpopulations: a reconciliation" paper presented at the WIDER (2003), Helsinki, Finland.

18. Mussard, S., "The Bidimensional Decomposition of the Gini Ratio. A Case Study: Italy", *Appl. Econ. Letters* 11(8) (2004), 503-505.

19. Pyatt, G., "On the Interpretation and Disaggregation of Gini Coefficients", *Econ. J.* 86 (1976), 243-25.

20. Rao, V. M., "Two decompositions of concentration ratio", *J. Roy. Statistical Society* 132 (1969), 418-425.

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

22. Sen, A., "Poverty, an ordinal approach to measurement", *Econometrica* 44 (1976), 219-231.

23. Shorrocks, A. F., "The class of additively decomposable inequality measures", *Econometrica* 48 (1980), 613-625.

24. Shorrocks, A. F., "Aggregation Issues in Inequality Measurement", In Heichhorn, W. (ed.), *Measurement in Economics* (1988), New York, Physica-Verlag, 429-452.

25. Silber, J., "Factor components, population subgroups and the computation of the Gini index of inequality", *Rev. Econ. Statist.* 71 (1989), 107-115.

26. Xu, K., "Assimilation of immigrant household cohorts over time in Canada: an examination of the lower tail of income distributions" (2003), (mimeo), Department of Economics, Dalhousie University.

27. Xu, K., "How has the literature on Gini's index evolved in the past 80 years?" (2003), (mimeo), Department of Economics, Dalhousie University.

28. Xu, K. and Osberg, L., "The social welfare implications, decomposability, and geometry of the Sen family of poverty indices", *Can. J. Econ.* 35 (2001), 138-152.

29. Xu, K. and Osberg, L., "On Sen's approach to poverty measures and recent developments", paper presented at the Sixth International Meeting of the Society for Social Choice and Welfare (2002), Printed in Chinese in *China Econ. Quart.* 1, 151-170.

30. Zheng, B. "Minimum distribution-sensitivity, poverty aversion, and poverty orderings", *J. Econ. Theory* 95 (2000), 116-137.