On a class of bi-polarization variance-based measurement: Evidence from Cameroonian data.

Chameni Nembua, Célestin

Université de Yaoundé II, Cameroon

10 July 2010

On a class of bi-polarization variance-based measurement: Evidence from Cameroonian data.

C. CHAMENI NEMBUA

Chameni@yahoo.com

Université de Yaoundé II
Faculté des Sciences Economiques et de Gestion
Département des Techniques Quantitatives.
Yaoundé, Cameroun.

Abstract: This paper revisits one of the bipolarization indices of the large class of Duclos-Estebane-Ray polarization measures. The relationship between polarization, inequality and poverty is analyzed via the index. First, polarization measure for the median is defined and related to the subtraction of the between group and the within-groups components of the coefficient of variation squared. Second, the generalized bipolarization measure is defined and related to poverty via the headcount ratio, the income gap ratio and the overabundance gap ratio. In particular, it is shown that polarization is high when the headcount ratio is around 0.5 and polarization is little when the headcount ratio is far from this value. Third, the preceding results are applied to analyze the Cameroonian household's consumption distribution.

Keywords: Polarization, inequality, poverty, Cameroon.

Jel Classification: D39, D63,H30.

1- Introduction

Over the past recent years, there has been an increasing interest in the notion and measurement of polarization in the literature. The concept plays a pertinent role in the analysis of the evolution of income distribution of the consequences of economic growth and social conflict.

It is well known that Polarization is different from inequality. Polarization concentrates the income distribution on several focal and polar modes whereas inequality relates to the overall dispersion of the distribution. Wolfson (1994) noted that a more bi-polarized income distribution is one that is more spread out from the middle, so there are fewer individuals or households with mid-level incomes. It is well known that an income distribution may be facing a decrease (increase) in inequality while at the same time running into an increase (decrease) in polarization. In this context, to analyze and compare income distributions, not only inequality, poverty and welfare are considered but, nowadays, polarization is also taken into account to shed more light on the income distribution behavior.

Many polarization indices have been defined in the literature. Broadly speaking, these indices may be split in two categories of measures. The first category tries to capture the formation of any arbitrary number of the groups. This has been particularly studied by Esteban and Ray (1994)(for discrete distributions), Zhang and Kanbur (2001), Duclos ,Esteban and Ray (2004) (for continuous distributions). The second family of indices (referred as measures of bi-polarization) is elaborated in the context where polarization is apprehended as a process by which a distribution becomes bi-polar. Measures of bi-polarization initiate with the work of Foster and Wolfson (1992), Wolfson (1994,1997), and Tsui and Wang (2000).

The present paper may be situated at this area of research which it attempts to extend. A measurement of bi-polarization based on the variance is revisited. Then following Rodriguez and Salas approach (2003), polarization is linked to inequality and poverty (Rodriguez, 2004).

First, polarization measure for the median is defined. It is taken to be the quarter of the square of the relative gap between the income means of the group of individuals (or households) with their income less than the median and the opposite group. Then a direct relationship between polarization and inequality is established. The result found is in perfect agreement with the Rodriguez and Salas (2003) result concerning the Wolfson (1994) index and the Gini inequality ratio. In particular, it is shown that polarization and inequality can be respectively obtained with subtraction and addition of the within-groups dispersion of the coefficient of variation squared. The subgroups decomposition of the squared coefficient of variation used is different from its classical form. The approach is borrowed from Chameni Nembua.C. (2006, 2007) and it is similar to that used by C. Dagum (1997) when decomposing the Gini ratio. Also the relation between the bi-polarization measure and the Lorenz curve is established. Therefore the relationship between Wolfson bi-polarization index and the Gini ratio can be linked to the bi-polarization measure.

Second, the generalized polarization measure is defined. The bi-polarization in term of within-groups and between groups component of coefficient of variation squared for income groups separated by any z income value is proposed. In addition, it is shown that bi-polarization and poverty measures are related when the z income value used to separate the two income groups coincides with the poverty line. In particular, the proposed generalized bi-polarization measure is expressed as a function of the headcount ratio, the income gap ratio and the overabundance gap ratio. However, as in the case of the Wolfson index, the bi-polarization measure is not an increasing function of these three poverty measures. Moreover, it is proved that there exists a threshold from which an increase in the proportion of poor assures greater polarization.

Third, the preceding results are applied to analyze the Cameroonian household's consumption. Using micro ECAM data on expenditure, it is obtained that, results on polarization measured by the Wolfson bi-polarization index and the new polarization index may sometime differ. However, the two approaches unambiguously show that, Polarization measures for the median and evaluated on the total equivalent personal consumption, reduces a lot during the 1996-2001 period and stagnant from 2001 to 2007, as did the poverty.

The outline of the article is organized into three sections in addition to the present introduction. In section II, the theoretical formulation of the index is introduced and the main results of the papers are established. As in section III, the preceding results are implemented to analyze Cameroonian household's consumption. Finally, the paper is concluded in section IV.

2- Definitions and main results

Consider a population $P = \{1,2,\dots,i,\dots n\}$ of size n and an income variable X from which we have a distribution vector $x = (x_1, x_2, \dots x_n)$ where x_i is the income of person i . Each x_i is assumed to be drawn from $\mathbb{R}^+$ and X is a continuous variable. The vector $x = (x_1, x_2, \dots x_n)$ is ranked, that is $x_1 \leq x_2 \leq \cdots \leq x_n$ and we suppose that $x_1 \neq x_n$. The mean, the variance and the median of the vector x are denoted respectively by $\mu(x), \sigma^2(x)$ and $m(x)$. If n is odd $m(x)$ is the $\left(\frac{n+1}{2}\right)^{th}$ observation in x and if n is even $m(X) = \frac{x_{\frac{n}{2}}+x_{\frac{n}{2}+1}}{2}$. We assume that n is sufficiently high such that $m(x)$ separates the vector x into two equal size groups , that is $\frac{n+1}{n} \approx 1$.

For any real $z \epsilon \mathbb{R}^+$, we write x_z^+ and x_z^- for the subvector of x that includes x_i such that $x_i \geq z$ and x_i such that $x_i < z$.

Firstly, we need to recall the following classical result concerning the variance:

Lemma 1:

$$\sigma^2(\mathrm{x}) = \frac{1}{2\mathrm{n}^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x_j)^2 \qquad (1)$$

Secondly, For any real value z, $x_1 < z < x_n$ we can subdivide P into two non overlapping groups:

$P_z^- = \{i | x_i < z\}$ and $P_z^+ = \{i | x_i \geq z\}$ and we denote by n_z^- and n_z^+ their respective size. x_z^+ and x_z^- are the restriction of x in $P_z^+$and P_z^- respectively. Note that if $z = m(x)$, the assumption on the size of n leads to accept that $n_z^- = n_z^+ = \frac{n}{2}$.

We can then split the squared coefficient of variation as follow:

$$CV^2(X) = \frac{\sigma^2(X)}{\mu^2(x)} = \frac{1}{2\mathrm{n}^2\mu^2(\mathrm{x})} \sum_{\mathrm{i}\epsilon\mathrm{P}_\mathrm{z}^-} \sum_{\mathrm{j}\epsilon\mathrm{P}_\mathrm{z}^-} (\mathrm{x_i} - \mathrm{x_j})^2 + \frac{1}{2\mathrm{n}^2\mu^2(\mathrm{x})} \sum_{\mathrm{i}\epsilon\mathrm{P}_\mathrm{z}^+} \sum_{\mathrm{j}\epsilon\mathrm{P}_\mathrm{z}^+} (\mathrm{x_i} - \mathrm{x_j})^2 + \frac{1}{\mathrm{n}^2\mu^2(\mathrm{x})} \sum_{\mathrm{i}\epsilon\mathrm{P}_\mathrm{z}^-} \sum_{\mathrm{j}\epsilon\mathrm{P}_\mathrm{z}^+} (\mathrm{x_i} - \mathrm{x_j})^2$$

It comes that,

$$CV^2(x) = \left(\frac{\mu_{z^-}}{\mu(x)}\right)^2 \left(\frac{n_z^-}{n}\right)^2 CV^2(x_z^-) + \left(\frac{\mu_{z^+}}{\mu(x)}\right)^2 \left(\frac{n_z^+}{n}\right)^2 CV^2(x_z^+) + \frac{1}{\mathrm{n}^2\mu^2(\mathrm{x})} \sum_{\mathrm{i}\epsilon\mathrm{P}_\mathrm{z}^-} \sum_{\mathrm{j}\epsilon\mathrm{P}_\mathrm{z}^+} (\mathrm{x_i} - \mathrm{x_j})^2$$

Hence, $CV^2(x) = CV_z^{2W} + CV_z^{2B}$ (2)

where $CV_z^{2W} = \left(\frac{\mu_{z^-}}{\mu(x)}\right)^2 \left(\frac{n_z^-}{n}\right)^2 CV^2(x_z^-) + \left(\frac{\mu_{z^+}}{\mu(x)}\right)^2 \left(\frac{n_z^+}{n}\right)^2 CV^2(x_z^+)$

and $CV_z^{2B} = \frac{1}{\mathrm{n}^2\mu^2(\mathrm{x})} \sum_{\mathrm{i}\epsilon\mathrm{P}_\mathrm{z}^-} \sum_{\mathrm{j}\epsilon\mathrm{P}_\mathrm{z}^+} (\mathrm{x_i} - \mathrm{x_j})^2 = \frac{n_z^-}{n} \frac{n_z^+}{n} \left[\left(\frac{\mu_{z^-}}{\mu(x)}\right)^2 + \left(\frac{\mu_{z^+}}{\mu(x)}\right)^2\right] \frac{\sum_{\mathrm{i}\epsilon\mathrm{P}_\mathrm{z}^-} \sum_{\mathrm{j}\epsilon\mathrm{P}_\mathrm{z}^+} (\mathrm{x_i} - \mathrm{x_j})^2}{n_z^- n_z^+ \left(\mu_{z^-}^2 + \mu_{z^+}^2\right)}$

We can now state the second lemma:

Lemma 2:

For any reel value z*, such that* $x_1 < z < x_n$*, the coefficient of variation squared can be split into a within groups and a between groups component:* $CV^2(x) = CV_z^{2W} + CV_z^{2B}$ *(3)*

$$With\ \ CV_z^{2W} = \left(\frac{\mu_{z^-}}{\mu(x)}\right)^2 \left(\frac{n_z^-}{n}\right)^2 CV^2(x_z^-) + \left(\frac{\mu_{z^+}}{\mu(x)}\right)^2 \left(\frac{n_z^+}{n}\right)^2 CV^2(x_z^+)\ \ and$$

$$CV_z^{2B} = \frac{n_z^-}{n}\frac{n_z^+}{n}\left[\left(\frac{\mu_{z^-}}{\mu(x)}\right)^2 + \left(\frac{\mu_{z^+}}{\mu(x)}\right)^2\right]\frac{\sum_{i\epsilon P_z^-}\sum_{j\epsilon P_z^+}(x_i - x_j)^2}{n_z^- n_z^+\left(\mu_{z^-}^2 + \mu_{z^+}^2\right)}$$

It is interesting to note that this decomposition method yields a between group component that is different from the classical well known component of the coefficient of variation squared. The latter represents a difference in means of the two subgroups whereas the former seems to have a better specification. It is based on the inequalities between the subgroups. The approach here is similar to that used by C. Dagum (1997) when decomposing the Gini coefficient. For more details on the method, see C. Chameni Nembua (2006, 2007).

Having introduced some concept that will be used, let us now define our bi-polarization measure. Following Wolfson (1994) we consider that the population is subdivided into two groups via the median $m(x)$ (z= m) of the income distribution $= (x_1, x_2, \dots x_n)$.

Definition 1 :

For any income distribution vector x *with median* $m(x)$*, the polarization is measured by :*

$$P_m(x) = \frac{1}{4}\left(\frac{\mu_{m_+} - \mu_{m_-}}{m(x)}\right)^2 \qquad (4)$$

Where μ_{m_+} *and* μ_{m_-} *are the mean of* x *in* P_m^+ *and* P_m^- *respectively.*

The idea in formula (4) is simple and clear, the polarization is captured by the square of the relative gap between the income means of the two groups, the group of individuals (or households) with their income less than the median and the opposite group.

On the other hand, it is interesting to note that $P_m(x)$ belongs to the large class of Duclos-Estebane-Ray polarization measures $I(x) = C\sum_{i=1}^{k}\sum_{j=1}^{k}\pi_i^{\alpha+1}\pi_j T\left(|m_i - m_j|\right)$ but with $T\left(|m_i - m_j|\right) = |m_i - m_j|^2$ instead of $T\left(|m_i - m_j|\right) = |m_i - m_j|$ as it is often the case in the literature.

In order to motivate the pertinence of the proposed measure, we have to study its properties. There are particularly three properties that literature seems to consider to be indispensable to a measure of polarization:

(1) Polarization must be little when there is only one group.
(2) Polarization rises when within group inequality is reduced.
(3) Polarization rises when between group inequality increases.

The following proposition shows that the polarization measure defined in (4) satisfies the mentioned properties.

Proposition 1:

$$P_m(x) = \frac{\mu^2(x)}{m^2(x)}\left(CV_m^{2B} - CV_m^{2W}\right) \qquad (5)$$

Proof: $CV_m^{2B} - CV_m^{2W} = CV^2(x) - 2CV_m^{2W}$

$$= \frac{\mu(x^2)-\mu^2(x)}{\mu^2(x)} - 2\left[\left(\frac{\mu_{m^-}}{\mu(x)}\right)^2\left(\frac{n_m^-}{n}\right)^2 CV^2(x_m^-) + \left(\frac{\mu_{m^+}}{\mu(x)}\right)^2\left(\frac{n_m^+}{n}\right)^2 CV^2(x_m^+)\right]$$

$$= \frac{\mu(x^2)-\mu^2(x)}{\mu^2(x)} - \frac{2}{\mu^2(x)}\left[\left(\frac{n_m^-}{n}\right)^2 \sigma^2(x_m^-) + \left(\frac{n_m^+}{n}\right)^2 \sigma^2(x_m^+)\right]$$

$$= \frac{1}{\mu^2(x)}\left[\mu(x^2) - \mu^2(x) - \frac{1}{2}(\mu(x_m^-)^2 - \mu^2(x_m^-) + \mu(x_m^+)^2 - \mu^2(x_m^+))\right]$$

Consider that

$$\mu^2(x) = \left(\frac{\mu(x_m^-) \;\; +\mu(x_m^+)}{2}\right)^2 = \frac{\mu^2(x_m^-)+\mu^2(x_m^+)+2\mu(x_m^-) \;\; \mu(x_m^+)}{4} \quad \text{and} \quad \mu(x^2) = \frac{\mu(x_m^-)^2+\mu(x_m^+)^2}{2}$$

We obtain,

$$CV_m^{2B} - CV_m^{2W} = \frac{\mu^2(x_m^-)+\mu^2(x_m^+)-2\mu(x_m^-) \;\; \mu(x_m^+)}{4\mu^2(x)} = \left(\frac{\mu(x_m^-) \;\; -\mu(x_m^+)}{2\mu(x)}\right)^2 = \frac{m^2(x)}{\mu^2(x)} P_m(x). \blacksquare$$

The assertion in the proposition 1 is similar to the Rodriguez and Salas (2003) result about the Wolfson index. The authors obtained a reformulation of the index in term of the difference of the between and the within group component of the Gini coefficient. Such a reformulation has at least two advantages. First, a link is directly established between polarization and inequality. It is immediate that polarization rises when between groups inequality increases or when within groups inequality reduces. Second, the reformulation permit to join the polarization model of Esteban and Ray (1994) and Duclos et al. (2004) where polarization is relied on identification –alienation concept. The alienation corresponds to the accentuation of the polarization through the heterogeneity of the groups (that is a high between groups component) while identification relates to the accentuation of the polarization via intragroups homogeneity (that is a modest within groups component).

Another remark that can be made at this stage concerns the role of the ratio $\frac{\mu(x)}{m(x)}$ which appears in its squared form in the expression (5) of $P_m(x)$. Indeed, the presence of the ratio is useful to capture the gap between the mean and the median of x. But the question is that: Is there any particular reason to prefer $\left(\frac{\mu(x)}{m(x)}\right)^2$ rather than $\frac{\mu(x)}{m(x)}$ (as it is the case in the Wolfson index)?

Consider that, the only reason is the sensitivity of $\left(\frac{\mu(x)}{m(x)}\right)^2$ toward the small and great values; one could also suggest an alternative formulation of the polarization index as:

$$P_m^*(x) = \frac{m(x)}{\mu(x)} P_m(x) = \frac{\mu(x)}{m(x)} \left(CV_m^{2B} - CV_m^{2W} \right) \qquad (5^*)$$

Note that, $P_m(x)$ may be greater or less than $P_m^*(x)$ according to the position of $\frac{\mu(x)}{m(x)}$ relatively to 1. However, in developing countries, one will often have $P_m(x)$ largely greater than $P_m^*(x)$ as income (or consumption/ expenditure) distribution x will often be largely right skewed.

Now we generalize our bipolarization index by considering that the population is separated into two groups by any income value z.

Definition 2 :
For any income distribution vector x with mean $\mu(x)$, if the population is separated into two groups by the income value z, the polarization is measured by :

$$P_z(x) = \frac{n_z^+}{n}\frac{n_z^-}{n}\left(\frac{\mu_{z+}-\mu_{z-}}{z}\right)^2 = \frac{n_z^+}{n}\frac{n_z^-}{n}\left(\frac{\mu(x)}{z}\right)^2\left(\frac{\mu_{z+}-\mu_{z-}}{\mu(x)}\right)^2 \qquad (6)$$

Or

$$P_z^*(x) = \frac{n_z^+}{n}\frac{n_z^-}{n}\frac{z}{\mu(x)}\left(\frac{\mu_{z+}-\mu_{z-}}{z}\right)^2 = \frac{n_z^+}{n}\frac{n_z^-}{n}\frac{\mu(x)}{z}\left(\frac{\mu_{z+}-\mu_{z-}}{\mu(x)}\right)^2 \qquad (6^*)$$

Where μ_{z_+} and μ_{z_-} are the mean of x in P_z^+ and P_z^- respectively.

The significance of the generalized polarization measure in expression (6) is clear. When the income distribution (or the population) is separated into two groups by the income value z, the polarization is captured by the weighted square of the relative difference between the total incomes in the two groups. Nothing guaranties that $P_z(x)$ is a normalized index in the sense that its values lie between 0 and

1. $P_z(x)$ may be greater than 1 even if z equal to the mean $\mu(x)$, or to the median $m(x)$ of x. This is not surprising because $P_z(x)$ is based on the coefficient of variation squared which is itself a non normalized inequality index. If x is symmetric, such that $\mu(x)= m(x) = \frac{\mu_{m_+}+\mu_{m_-}}{2}$, it is immediate that: $P_m(x) = \left(1 - \frac{\mu_{m_-}}{\mu(x)}\right)^2$ and hence $P_m(x)$ is less than 1.

Note that, the polarization measure for the median in the expression (4) is a special case of the generalized polarization measure. This expression is related to the Lorenz curve at the median population percentile. More generally, the following results are straightforward.

Corollary 2

1) $$P_z(x) = \frac{\mu^2(x)}{z^2}\left(1 - 2L(q_z)\right)^2 + \frac{1}{z^2}\left(1 - 2\frac{n_z^-}{n}\right)\left(\frac{n_z^-}{n}\mu_{z_-}^2 - \frac{n_z^+}{n}\mu_{z_+}^2\right) \quad (7)$$

2) $$P_z(x) = \frac{\mu^2(x)}{z^2}\frac{[L(q_z)-q_z]^2}{q_z(1-q_z)} \quad (7^*)$$

Where q_z is the population percentile at the value z and $L(q_z)$ is the value of the Lorenz curve evaluated at q_z .

Proof:

1) Let us set $A_z(x) = \frac{\mu^2(x)}{z^2}\left(\frac{n_z^+\mu_{z_+}-n_z^-\mu_{z_-}}{n\mu(x)}\right)^2 = \frac{\mu^2(x)}{z^2}[1 - 2L(q_z)]^2$

Hence, $P_z(x) - A_z(x) = \frac{n_z^+}{n}\frac{n_z^-}{n}\left(\frac{\mu(x)}{z}\right)^2\left(\frac{\mu_{z_+}-\mu_{z_-}}{\mu(x)}\right)^2 - \frac{\mu^2(x)}{z^2}\left(\frac{n_z^+\mu_{z_+}-n_z^-\mu_{z_-}}{n\mu(x)}\right)^2$

$$= \left(\frac{\mu(x)}{z}\right)^2\left[\frac{n_z^+}{n}\frac{n_z^-}{n}\frac{\mu_{z_+}^2+\mu_{z_-}^2}{\mu^2(x)} - \frac{\left(n_z^+\mu_{z_+}\right)^2+\left(n_z^-\mu_{z_-}\right)^2}{n^2\mu^2(x)}\right]=\left(\frac{\mu(x)}{z}\right)^2(n_z^- - n_z^+)\frac{(n_z^+\mu_{z_+}^2-n_z^-\mu_{z_-}^2)}{n^2\mu^2(x)}$$

$$= \frac{1}{z^2}\left(1 - 2\frac{n_z^-}{n}\right)\left(\frac{n_z^-}{n}\mu_{z_-}^2 - \frac{n_z^+}{n}\mu_{z_+}^2\right)$$

It comes that, $P_z(x) = \frac{\mu^2(x)}{z^2}\left(1 - 2L(q_z)\right)^2 + \frac{1}{z^2}\left(1 - 2\frac{n_z^-}{n}\right)\left(\frac{n_z^-}{n}\mu_{z_-}^2 - \frac{n_z^+}{n}\mu_{z_+}^2\right)$

2) $\mu_{z_-} = \mu(x)\frac{L(q_z)}{q_z}$ and $\mu_{z_+} = \mu(x)\frac{1-L(q_z)}{1-q_z} \Longrightarrow P_z(x) = \frac{1}{z^2}\left(\mu(x)\frac{L(q_z)}{q_z} - \mu(x)\frac{1-L(q_z)}{1-q_z}\right)^2$

$$= \frac{\mu^2(x)}{z^2}\frac{[L(q_z)-q_z]^2}{q_z(1-q_z)}$$

■

The assertions in corollary 2 suggest three specific values of z with a specific computation of $P_z(x)$:

1) if $z = z_1$ is such that $\frac{n_z^+}{n_z^-} = 1 \Leftrightarrow q_z = \frac{1}{2} \Leftrightarrow$z is the median of x, then

$P_z(x) = \frac{\mu^2(x)}{z^2}\left(1 - 2L(0.5)\right)^2$.

2) if $z = z_2$ is such that $\frac{n_z^+}{n_z^-} = \frac{\mu_{z-}}{\mu_{z+}} \Leftrightarrow L(q_z) = 0.5 \Leftrightarrow$ z is the medial of x, then

$P_z(x) = \frac{\mu_{z-}^2}{z^2}\frac{n_z^-}{n}\left(1 - \frac{n_z^-}{n_z^+}\right)\left(1 - 2\frac{n_z^-}{n}\right) = \frac{\mu^2(x)}{z^2}\frac{[0.5-q_z]^2}{q_z(1-q_z)}$

3) if $z = z_3$ is such that $\frac{n_z^+}{n_z^-} = \left(\frac{\mu_{z-}}{\mu_{z+}}\right)^2$, then $P_z(x) = \frac{\mu^2(x)}{z^2}\left(1 - 2L(q_z)\right)^2$.

Note that, $z_1$and z_2 are the two flipping points of the Lorenz curve of x ; $z_1 \leq z_2 \leq z_3$ and two of the three points coincide if and only if x is constant, what means that $P_z(x) = 0$ for any z.

Another suggestion coming from corollary 2 is that, there exists a link between Wolfson index and our index. We tend toward the issue in the following corollary.

Corollary 3

If $W(x)$, $W_z(x)$ and G(x) are respectively the Wolfson index , the generalized Wolfson index and the Gini ratio of x then:

1) $$P_m(x) = \left(\frac{1}{2}W(x) + \frac{\mu(x)}{m(x)}G(x)\right)^2 \quad (8)$$

2) $$P_z(x) = \frac{\left(\frac{1}{4}W_z(x) + \frac{\mu(x)}{2z}G(x)\right)^2}{q_z(1-q_z)} \quad (8^*)$$

Proof: We prove the equation (8*).

By definition $W_z(x) = 4\frac{\mu(x)}{z}\left(q_z - L(q_z) - \frac{1}{2}G(x)\right)$

Therefore, $1 - L(q_z) = \frac{1}{4}\frac{z}{\mu(x)}W_z(x) + \frac{1}{2}G(x) \Rightarrow \left(1 - L(q_z)\right)^2 = \left(\frac{1}{4}\frac{z}{\mu(x)}W_z(x) + \frac{1}{2}G(x)\right)^2$

According to corollary 2,

$$P_z(x) = \frac{\mu^2(x)}{z^2}\frac{[L(q_z)-q_z]^2}{q_z(1-q_z)} = \frac{\mu^2(x)}{m^2(x)}\frac{1}{q_z(1-q_z)}\left(\frac{1}{4}\frac{z}{\mu(x)}W_z(x) + \frac{1}{2}G(x)\right)^2 = \frac{\left(\frac{1}{4}W(x) + \frac{\mu(x)}{2z}G(x)\right)^2}{q_z(1-q_z)} \blacksquare$$

From equation (8*) it is easy to obtain that $P_z(x)$ is greater than $W_z(x)$ whenever $\frac{\mu(x)}{m(x)}G(x) > \frac{1}{2}$. This condition will often arrive in developing countries where inequality is accentuated and the income distribution is most of the time right skewed so that $\frac{\mu(x)}{m(x)} > 1$.

To carry on with the reasoning, we show that generalized bi-polarization measure is related to the between groups and the within groups coefficient of variation squared.

Proposition 4:

$$P_z(x) = \frac{\mu^2(x)}{z^2}\left(\mathrm{CV}_z^{2B} - \mathrm{CV}_z^{2W}\right) + \frac{1}{z^2}\left(2\frac{n_z^-}{n} - 1\right)\left(\frac{n_z^-}{n}\sigma^2(x_z^-) - \frac{n_z^+}{n}\sigma^2(x_z^+)\right) \qquad (9)$$

Proof: $\mathrm{CV}_z^{2B} - \mathrm{CV}_z^{2W} = CV^2(x) - 2CV_z^{2W}$

$$= \frac{\mu(x^2)-\mu^2(x)}{\mu^2(x)} - 2\left[\left(\frac{\mu_{z^-}}{\mu(x)}\right)^2\left(\frac{n_z^-}{n}\right)^2 CV^2(x_z^-) + \left(\frac{\mu_{z^+}}{\mu(x)}\right)^2\left(\frac{n_z^+}{n}\right)^2 CV^2(x_z^+)\right]$$

$$= \frac{\mu(x^2)-\mu^2(x)}{\mu^2(x)} - \frac{2}{\mu^2(x)}\left[\left(\frac{n_z^-}{n}\right)^2\sigma^2(x_z^-) + \left(\frac{n_z^+}{n}\right)^2\sigma^2(x_z^+)\right]$$

$$= \frac{1}{\mu^2(x)}\left[\mu(x^2) - \mu^2(x) - 2\left(\frac{n_z^-}{n}\right)^2\left(\mu(x_z^-)^2 - \mu^2(x_z^-)\right) - 2\left(\frac{n_z^+}{n}\right)^2\left(\mu(x_z^+)^2 - \mu^2(x_z^+)\right)\right]$$

Since $\mu^2(x) = \left(\frac{n_z^-}{n}\mu(x_z^-) + \frac{n_z^+}{n}\mu(x_z^+)\right)^2 = \left(\frac{n_z^-}{n}\right)^2\mu^2(x_z^-) + \left(\frac{n_z^+}{n}\right)^2\mu^2(x_z^+) + 2\frac{n_z^-}{n}\frac{n_z^+}{n}\mu(x_z^-)\mu(x_z^+)$

$$\mathrm{CV}_z^{2B} - \mathrm{CV}_z^{2W} = \frac{1}{\mu^2(x)}\left[\mu(x^2) - 2\left(\frac{n_z^-}{n}\right)^2\mu(x_z^-)^2 - 2\left(\frac{n_z^+}{n}\right)^2\mu(x_z^+)^2 + \left(\frac{n_z^-}{n}\mu(x_z^-) - \frac{n_z^+}{n}\mu(x_z^+)\right)^2\right]$$

As $\mu(x^2) = \frac{n_z^-}{n}\mu(x_z^-)^2 + \frac{n_z^+}{n}\mu(x_z^+)^2$ and $n_z^- + n_z^+ = n$,

We have,

$$\mathrm{CV}_z^{2B} - \mathrm{CV}_z^{2W} = \frac{1}{\mu^2(x)}\left[\left(\frac{n_z^-}{n}\mu(x_z^-) - \frac{n_z^+}{n}\mu(x_z^+)\right)^2 + \frac{n_z^-}{n}\mu(x_z^-)^2\left(1 - 2\frac{n_z^-}{n}\right) + \frac{n_z^+}{n}\mu(x_z^+)^2\left(1 - 2\frac{n_z^+}{n}\right)\right]$$

$$= \frac{1}{\mu^2(x)}\left[\left(\frac{n_z^-}{n}\mu(x_z^-) - \frac{n_z^+}{n}\mu(x_z^+)\right)^2 + \left(1 - 2\frac{n_z^-}{n}\right)\left(\frac{n_z^-}{n}\mu(x_z^-)^2 - \frac{n_z^+}{n}\mu(x_z^+)^2\right)\right]$$

Hence,

$$\left(\frac{n_z^-}{n}\mu(x_z^-)-\frac{n_z^+}{n}\mu(x_z^+)\right)^2=\mu^2(x)\left(\mathrm{CV}_z^{2B}-\mathrm{CV}_z^{2W}\right)+\left(2\frac{n_z^-}{n}-1\right)\left(\frac{n_z^-}{n}\mu(x_z^-)^2-\frac{n_z^+}{n}\mu(x_z^+)^2\right)$$

Considering that, $\left(\frac{n_z^-}{n}\mu(x_z^-)-\frac{n_z^+}{n}\mu(x_z^+)\right)^2=z^2P_z(x)+\left(2\frac{n_z^-}{n}-1\right)\left(\frac{n_z^-}{n}\mu_{z_-}^2-\frac{n_z^+}{n}\mu_{z_+}^2\right)$

We obtain,

$$z^2P_z(x)=\mu^2(x)\left(\mathrm{CV}_z^{2B}-\mathrm{CV}_z^{2W}\right)+\left(2\frac{n_z^-}{n}-1\right)\left(\frac{n_z^-}{n}\mu(x_z^-)^2-\frac{n_z^+}{n}\mu(x_z^+)^2\right)$$
$$-\left(2\frac{n_z^-}{n}-1\right)\left(\frac{n_z^-}{n}\mu_{z_-}^2-\frac{n_z^+}{n}\mu_{z_+}^2\right)$$

And finally,

$P_z(x)=\frac{\mu^2(x)}{z^2}\left(\mathrm{CV}_z^{2B}-\mathrm{CV}_z^{2W}\right)+\frac{1}{z^2}\left(2\frac{n_z^-}{n}-1\right)\left(\frac{n_z^-}{n}\sigma^2(x_z^-)-\frac{n_z^+}{n}\sigma^2(x_z^+)\right)$ ■

Again, proposition 4 is similar to a result obtained by Rodriguez (2004) in the context of a generalized Wolfson polarization index and the between and within groups components of the Gini coefficient. Note that the second term in the right side of equation (9) is zero when z equal to $m(x)$.

To complete this section, we study the relation between our generalized bi-polarization index and the poverty. It is well kwon that polarization and poverty measures can be related when the value z that subdivided the population in two groups represents the poverty line. In this context, P_z^- coincides with the poor group and denoted G_z^P while P_z^+ is the rich group (precisely the non poor group) and denoted G_z^R; their sizes are respectively denoted n_z^P and n_z^R , the mean of x in P_z^- and P_z^+ is respectively μ_p and μ_R .

Let us rapidly recall some poverty concepts. When z is the poverty line $\mathrm{q_z}=\frac{\mathrm{n_p}}{\mathrm{n}}$ is the headcount ratio or the proportion of the population who are poor. $g_z=\frac{n_pz-n_p\mu_p}{n_pz}$= 1- $\frac{\mu_p}{z}$ is the income gap ratio while $o_z=\frac{n_R\mu_R-zn_R}{zn_R}=\frac{\mu_R}{z}-1$ is the overabundance gap ratio.

The following result is straightforward from the expression (6):

Proposition 5:

$$P_z(x) = q_z(1\text{-}q_z)(g_z + o_z)^2 \tag{10}$$

The expression in proposition 5 clearly shows that bipolarization between poor and rich is a function of poverty via the headcount ratio, the income gap ratio and the overabundance gap ratio. However, the proposed bipolarization index is not an increasing function of these three measures. The proportion of the poor can change and the polarization changes in the same or in the opposite direction according to the effect of income gap ratio and/or the overabundance gap ratio. This in particular reveals that more poor in the population does not necessary implies social conflict according to polarization concept, which is nowadays a well known result.

Suppose for instance that, the mean income of the rich and the mean income of the poor remain unchanged, the polarization will increase as the proportion of poor increases from zero to 0.5 and then decreases as the proportion of poor will increase from 0.5 to 1. The maximum value of the polarization is obtained when the proportion of poor equal to 0.5. This clearly shows that polarization is high when the sizes of the two groups are not too different and polarization is little when one group is much bigger than the other.

On the other hand, if the proportion of poor is maintained unchanged, polarization becomes an increase function of the income gap ratio and overabundance gap ratio. This clearly shows that, polarization increases as the gap between poor incomes and rich incomes increases.

3- Applications

Illustration From Cameroonian households Data

Data from the country's household survey known as ECAM ('Enquête Camerounaise auprès des ménages') is used. It is conducted every 5 years by the National Institute of Statistics in Cameroon. Due to data availability, we consider ECAM I, II, III which correspond to the years 1996, 2001 and 2007.

3.1- Polarization and inequality.

The total consumption is considered as a proxy of the household total income. Table 1 gives statistics on inequality and polarization on the households total equivalent personnel consumption. We use the Gini coefficient and the coefficient of variation squared for the inequality measure while the Foster-Wolfson bi-polarization index and the new indices are utilized to assess the level of polarization. We also consider other statistics like ratio of mean to median and the percentage of households between 85 - 130% of the median (the middle class) and between 50-200% of the median.

The statistics displayed in the table1 unambiguously imply a decline in polarization between 1996 and 2001. The Foster-Wolfson index fell about 5.41% while the P_m index was more sensitive to the decline and fell about 20% and the P_m^* index fell 13.3%. These gaps in the sensitivity on the decline may be partly explained by the different effect of the decrease in the ratio of mean to median (the ratio fell 7.68% between 1996 and 2001), on the three indices. On the other hand, the right skeweness of the consumption distribution acts to amplify the gap between the Foster-Wolfson index and P_m^* in one side and P_m in the other side.

The decline of the bipolarization implies, in particular, the increase of the middle class in the country during the period. This result is also confirmed by the growth, during the period, of 42.54% in the proportion of households between 85-130% of the median.

During the 1996-2001 period, the situation on inequality is less clear (compare to polarization) but not enough different according to the Gini ratio and the square of coefficient of variation. The former index evenly decreased from 0.421 to 0.404 while the latter increased from 1.299 to 1.535. At the same time, the value of L(0.5) remained almost unchanged. This leads to suspect stagnation rather than a substantial change in inequality during the period.

Table 1: Polarization and inequality indicators : households total equivalent personal consumption

	1996	2001	2007
CV^2	1.299	1.535	0.889
Gini	0.421	0.404	0.389
L(0.5)	0.229	0.237	0.240
Mean	218279.642	372742.551	432899.234
Median	147200.906	272228.812	321000.812
Mean/Median	1.483	1.369	1.348
N (Size of sample)	1728	10992	11391
85-130%	18.31	26.1	24.3
50-200%	58.66	72.4	70.1
Foster-Wolfson	0.351	0.332	0.347
Our index P_m	0.646	0.517	0.476
Our index P_m^*	0.435	0.377	0.35

The 2001-2007 period case is in opposite with the 1996-2001 period. The inequality seems to decrease while the polarization is almost constant. The proportion of the middle class (85-130% of the median) lightly changes from 26.1 to 24.3%. Even though some of the results of the two periods analysis may be in contrast regarding to the different indices, there is a perfect concordance in the results obtained in the long trend analysis: The statistics in table 1 clearly show that, from 1996 to 2007, inequality and polarization have lightly decrease in the country.

3.2- Polarization and Poverty.

The statistics in table 2 shed light on the relationship between the poverty and the polarization between the poor and non poor people. It is much clear that from 1996 to 2001 poverty decreased as did polarization between poor and non poor people. This arrived for the reasons that , not only the proportion of poor moved from a value near 0.5 (exactly 0.53) to a value less and near 0.40 but also, the relative gap between the income mean of poor and the income mean of rich decreased from 2.497 to 2.371.

Table 2: Polarization and poverty indicators :Households total equivalent personal consumption

	1996	**2001**	**2007**
Poverty line (z)*	155600	232547	269400
Poor income mean (μ_p)	103541,285	158545,00	186259,968
Non poor income mean (μ_R)	349425,952	516622,3	596538,517
$\left(\frac{\mu_R - \mu_P}{z}\right)^2$	2,497	2,371	2,319
Headcount ratio (q_z)	0,533	0,402	0,399
Our index P_z	0,621	0,57	0,556
Our index P_z^*	0,443	0,355	0,346

(*) The poverty lines and the headcount ratios considered here are the official poverty indicators in Cameroon computed by the National Institute of Statistics.

From 2001 to 2007, the change in both poverty and polarization seems not significant. The headcount ratio moved from 0.402 to 0.399 while P_z and P_z^* indices show a very modest decline of polarization between poor and non poor people by decreasing from 0,57to 0,556 for the former and from 0,355 to 0,346 for the latter. Nevertheless, it is of interesting to note that, from 1996 to 2007, both polarization between poor and non poor people and poverty significantly decreased.

4- **Conclusion**

In this paper, bipolarization for the median is measured by the square of the normalized gap between the income means of the two groups, the group of individuals with their income less than the median and the opposite group. This index has been generalized in the case where the population is separated into two groups by any income value z.The link between the polarization index and the inequality has been established. The result found is similar to the Rodriguez and Salas (2003) result about the Wolfson index and its link with the inequality. It has also been shown that polarization is related to poverty when the z income value represents the poverty line. An application using Cameroonian data has been provided to support the appropriateness of the index and to contrast our results to those obtained with the Wolfson index. In particular, the empirical analysis revealed a decrease in poverty and polarization in the country during the 1996-2001 period and a stationary state of these two phenomena during the 2001-2007 period.

References

Chameni Nembua. C. (2006) A note on the decomposition of the coefficient of variation squared: Comparing entropy and Dagum's methods. Economics Bulletin, Vol 4, N°8, pp1-8.

Chameni Nembua C. (2007), The multi-decomposition of the Hirschman-Herfindahl index: Measuring households inequalities in Cameroon, 1996-2001, Applied Economics Letters, 14,27-43.

Dagum, C. (1997) A new approach to decomposition of the gini income inequality ratio. Empirical Economics 22, 515-531.

Duclos .J-Y, Esteban, J. and D. Ray (2004), *Polarization: Concepts, measurement, estimation*. Econometrica, 72(6), 1737-1772.

Esteban, J. and D. Ray (1994), *On the measurement of polarization*, Econometrica, 62,4,819-851.

Foster and Wolfson (1992), *Polarization and the decline of the middle class: Canada and the U.S.,* mimeo, Vanderbilt University.

Rodriguez J.G and Salas R. (2003), *Extended bi-polarization and inequality measures*, Research on Economic Inequality, Vol 9, 69-83.

Rodriguez J.G. (2004), *Measuring Polarization, Inequality, Welfare and Poverty,* Documentos de trabajo, Centro de Estudios Andaluces, Series 1, N° 75, 2004.

Tsui, K.Y. and Y.Q. Wang (2000), *Polarization orderings and new classes of polarization indices*, Journal of Public Economic Theory, 2(3), 349-363.

Wolfson,M.C. (1994), *When inequalities diverge,* American Economics Review, 84,353-358.

Wolfson, M.C. (1997), *Divergent inequalities theory and empirical results*, Review of Income and Wealth, 43,401-421.

Zhang and Kanbur (2001), *What difference do polarization measures make? An application to China* Journal of Development Studies, 37, 85-98.