Chapter 4
The Measurement of Income Polarization

4.1 Introduction

Over the last 15 years or so, the study of polarization has become quite important for several reasons. Some of the major reasons are its role in analyzing the income distribution evolution, social conflict, and economic growth. Broadly speaking, polarization is concerned with appearance (or disappearance) of groups in a distribution. In politics, it is regarded as a process that leads to division of public opinion and movement of the divided opinion to the extremes. Likewise, one notion of income polarization, which we refer to as bipolarization, is concerned with the decline of the middle class. In this case, the relative frequency of observations associated with the central value of the distribution is low compared to those in the extremes. Polarization in this case is measured by the dispersion of the distribution from the central value toward the extreme points. The principal reason for looking at polarization this way is that a large and wealthy middle class contributes to economic growth in many ways and hence is important to every society. The middle class occupies the intermediate position between the poor and the rich. A person with low income may not be able to become highly rich but may have the expectation of achieving the position enjoyed by a middle-class person. Thus, such a person is likely to work hard to fulfill his expectation and unlikely to revolt against the society. Therefore, a society with thriving middle class contributes significantly to social and political stability as well. In contrast, a society with high degree of polarization may generate social conflicts, rebellions, and tensions (*see* Pressman, 2001). Therefore, in order to avoid or reduce such possible risks, it is necessary to monitor the situation in the society using indices that look at the spread of the distribution from its center. Bipolarization indices have been investigated in details by Foster and Wolfson (1992), Wolfson (1994, 1997), Wang and Tsui (2000), Chakravarty and Majumder (2001), Rodriguez and Salas (2003), Duclos and Echevin (2005), Amiel et al. (2007), Chakravarty et al. (2007), Silber et al. (2007), and others.

Esteban and Ray (1994) developed a more general notion of polarization. They assumed that the society is divided into groups or poles, where the individuals

S.R. Chakravarty, *Inequality, Polarization and Poverty: Advances in Distributional Analysis,* Economic Studies in Inequality, Social Exclusion and Well-Being 6,
DOI: 10.1007/978-0-387-79253-8_4,

belonging to the same group have a feeling of identification and there is a feeling of alienation against individuals in a different group. In other words, individuals in a group share similar characteristics with the other members of the group but in terms of the same characteristics they are different from the members of the other groups. The Esteban and Ray (1994) index regards the concept of polarization as conflict among groups (*see also* Esteban and Ray, 1999). Clearly, high degree of polarization, in terms of presence of conflicting groups, can give rise to instability in a society. Alternatives and variations of the Esteban and Ray (1994) index have been suggested, among others, by Gradin (2000), D'Ambrosio (2001), Zhang and Kanbur (2001), Duclos et al. (2004), and Esteban et al. (2007).

The objective of this chapter is to discuss the two views of polarization, the underlying axioms and the indices rigorously. We also characterize a compromise relative index of bipolarization. A relative index remains invariant under equiproportionate variations in all incomes and is said to possess the compromise property if, when multiplied by the median, becomes an absolute index that does not alter under equal absolute translation of incomes. Clearly, a particular index of bipolarization will rank alternative distributions of income in a complete manner. However, if we use more than one index, there may be different rankings of the distributions. Given the diversity of indices, it will be worthwhile to identify the class of indices that produces a similar ordering of different distributions. Finally, we look at this issue in this chapter.

4.2 Polarization: Two Views, Axioms and Indices

This section begins with a discussion on the postulates for an index of polarization rigorously. We follow Esteban and Ray (1994), Wang and Tsui (2000), Chakravarty and Majumder (2001), and Chakravarty et al. (2007) and present them using uniform notation. For a population of size n, a typical income distribution is given by a pair (p,x), where $x=(x_1,x_2,..x_k)$ and $p=(p_1,p_2,\ldots,p_k)$. Here x_i values indicate different income levels, p_i is the number of individuals with income exactly equal to x_i and $n=\sum_{i=1}^{k} p_i$. Clearly, $p=(p_1,p_2,\ldots,p_k)\in R_+^k$, the nonnegative orthant of the k dimensional Euclidean space R^k. Each x_i is assumed to be drawn from $[\mu,\gamma]$, a nondegenerate interval in the nonnegative part R_+^1 of the real line R^1. The set of income distributions for this population is denoted by S. Thus, we characterize an income distribution as a vector of population masses located on the steps in an income ladder. For any $x_i\in[\mu,\gamma]$, $x\in[\mu,\gamma]^k$, the k-fold Cartesian product of $[\mu,\gamma]$. For the sake of simplicity and convenience, the lower bound of the interval $[\mu,\gamma]$ has been taken to be nonnegative, which in turn implies nonnegativity of all the incomes. Extension of our results to the situation where some of the incomes are negative is quite straightforward.

For any $(p,x)\in S$, the mean and median of (p,x) are denoted, respectively, by $\lambda(p,x)$ and $m(p,x)$ (or, simply by λ and m). If n is odd, the median income is given by

$$m = \left\{ x_j : \sum_{i=1}^{j} p_i = \frac{1}{2}\left(\sum_{i=1}^{k} p_i + 1\right)\right\}, \tag{4.1}$$

where x_is are illfare-ranked, that is, ordered nondecreasingly and p_is are rearranged accordingly. But if n is even, the arithmetic mean of the $(n/2)$th and the $(n/2+1)$th values is taken as the median (given that the incomes are illfare-ranked and p_i's are permuted accordingly). We will assume throughout the chapter that the mean and the median are positive. For example, let $x = (2,4,10,1)$ and $p = (4,3,9,2)$. The illfare-ranked permutation of x is $(1,2,4,10)$ and the corresponding rearrangement of p is $(2,4,3,9)$. Since $n = 18$ is even here, the median m is the average of the ninth and tenth values, that is, $m = (4+10)/2 = 7$.

Since in the measurement of bipolarization, all incomes are compared with the median, persons with incomes below the median can be regarded as "deprived," where the source of deprivation is the shortfall of their incomes from the median. Likewise, all persons with incomes not below the median can be referred to as "satisfied" (*see* Runciman, 1966).

Some more preliminaries are necessary for the purpose at hand. Assuming that x_i's are illfare-ranked, we denote the vectors of such x_i's that are below the median and of those that are not below the median by x^- and x^+, respectively. The corresponding partition of p, under proper rearrangement, is (p^-, p^+). For the example taken above, $x^- = (1,2,4)$, $x^+ = (10), p^- = (2,4,3)$, and $p^+ = (9)$. The k-coordinated vector of ones is denoted by 1^k. For all $x, y \in [\mu, \gamma]^k$, we write xV_jy to represent the situation that x has been obtained from y by a simple increment in y_j, that is, $x_j > y_j$ for some j and $x_i = y_i$ for all $i \neq j$. Recall from our discussion in Chap. 1 that if income distributions are ordered, the transformation V allows only rank-preserving increments. For $x, y \in [\mu, \gamma]^k$, we write $xT_{\{i,j\}}y$ to denote that x has been obtained from y by a progressive transfer of income from the rich person j to the poor person i. Recall that the transfer does not alter the relative positions of the donor j and the recipient i and for ordered distributions, only rank-preserving transfers are allowed.

A polarization index L is a real valued function defined on S, that is, $L : S \to R^1$. For all $(p,x) \in S$, the functional value $L(p,x)$ indicates the level of polarization associated with the distribution (p,x).

Esteban and Ray (1994) have suggested the following axioms for an index of polarization. All of them are based on an income distribution constituted by three distinct values $x_1 = 0, x_2$, and x_3, and the corresponding population masses p_1, p_2, and p_3, where $x_1 < x_2 < x_3$.

Axiom 1. Let $p_1 > p_2 = p_3 > 0$. Fix $p_1 > 0$ and $x_2 > 0$. There exists $\tilde{c}_1 > 0$ and $\tilde{c}_2 > 0$ (possibly depending on p_1 and x_2) such that if $|x_2 - x_3| < \tilde{c}_1$ and $p_2 < \tilde{c}_2 p_1$, then joining of the masses p_2 and p_3 at their mid-point, $(x_2 + x_3)/2$, increases polarization.

Axiom 2. Let $p_1, p_2, p_3 > 0$; $p_1 > p_3$, and $|x_2 - x_3| < x_2$. There exists $\tilde{c}_3 > 0$ such that if p_2 is moved to the right, toward p_3 by an amount not exceeding $\tilde{c}_3$, polarization increases.

Axiom 3. Let $p_1, p_2, p_3 > 0$; $p_1 = p_3$; and $x_2 = x_3 - x_2 = \tilde{c}_4$. Any new distribution formed by shifting population mass from the central mass p_2 equally to the two lateral masses p_1 and p_3, each $\tilde{c}_4$ units of distance away, must increase polarization.

Before we proceed to state the axioms for a bipolarization index, let us explain the ones proposed by Esteban and Ray (1994). Axiom 1 underlines the idea that lower dispersion inside the groups and higher homogeneity of group's size should augment polarization. The next axiom argues that polarization should go up with heterogeneity among the groups. Finally, according to axiom 3, polarization should increase under a movement of the middle class into higher and lower categories.

The following axioms, for an index of bipolarization, have been suggested by Wang and Tsui (2000), Chakravarty and Majumder (2001), and Chakravarty et al. (2007) or are in common use elsewhere.

Axiom 4. (Normalization): If $(p,x) \in S$ is such that $x = c1^k$, where $c > 0$ is any scalar, then $L(p,x) = 0$.

Axiom 5. (Scale Invariance): For all $(p,x) \in S$ and all scalars $c > 0$, $L(p,x) = L(p,cx)$.

Axiom 6. (Translation Invariance): For all $(p,x) \in S$ and all scalars c such that $x + c1^k \in [\mu, \gamma]^k$, $L(p,x) = L(p, x + c1^k)$.

Axiom 7. (Symmetry): For all $(p,x) \in S$, $L(p,x) = L(p\Pi, x\Pi)$, where Π is any $k \times k$ permutation matrix.

Axiom 8. (Population Principle): For all $(p,x) \in S$, $L(p,x) = L(cp,x)$, where $c > 0$ is any scalar.

Axiom 9. (Increased Spread): If (p,x) and $(p,y) \in S$, where $m(p,x) = m(p,y)$, are related through anyone of the following cases,

(i) yV_jx and $y_j < m(p,y)$, (ii) xV_iy and $y_i > m(p,y)$, and (iii) both (i) and (ii), then $L(p,x) > L(p,y)$.

Axiom 10. (Increased Bipolarity): If (p,x) and $(p,y) \in S$, where $m(p,x) = m(p,y)$, are related through anyone of the following cases,

(i) $xT_{\{i,j\}}y$ and $y_j < m(p,y)$, (ii) $xT_{\{\hat{l},l\}}y$ and $y_{\hat{l}} > m(p,y)$, and (iii) both (i) and (ii), then $L(p,x) > L(p,y)$.

Axiom 11. (Continuity): L is continuous in its income arguments.

Axioms 4–8 and 11 are the bipolarization counterparts to the corresponding inequality axioms. As in the case of inequality indices, only a constant function can fulfill axioms 5 and 6 simultaneously. Note that Symmetry requires the same reordering of incomes and the corresponding population masses. Under the Population Principle, the population masses are changed by a fixed proportion but the incomes are kept unchanged.

Axiom 9, Increased Spread, is a monotonicity condition and close to axiom 3 of Esteban and Ray (1994). Since increments (reductions) in incomes above (below) the median widen the distribution, polarization should go up. That is, greater distancing between the groups below and above the median should make the distribution more polarized. Increased Bipolarity is a bunching or a clustering principle. Since an egalitarian transfer between two individuals on the same side of the median brings the individuals closer to each other, bipolarization should increase. As an egalitarian transfer demands decreasingness of inequality, this axiom explicitly establishes that inequality and polarization are two nonidentical concepts.[1] Thus, bipolarization involves both an inequality-like component, the greater distancing criterion, which increases both inequality and polarization, and an equality-like component, the clustering or bunching principle, which increases polarization, while lowering any inequality index that fulfills the Pigou-Dalton Transfers Principle. This shows that although there is a nice complementarity between the two concepts, there are differences as well.[2]

Using specific subsets of the axioms considered above, we may be able to characterize specific classes of polarization indices. For instance, Esteban and Ray (1994) assumed the quasi-additive structure $\sum_{i=1}^{k}\sum_{j=1}^{k} p_i p_j \tilde{H}[\tilde{g}(p_i), \tilde{A}(|x_i - x_j|)]$; where the continuous identification function $\tilde{g} : R_+^1 \to R_+^1$, satisfying the restriction that $\tilde{g}(p_i) > 0$ whenever $p_i > 0$, gives a sense of identification of an individual with other persons of the same group. The continuous and nondecreasing alienation function $\tilde{A} : R_+^1 \to R_+^1$, with $\tilde{A}(0) = 0$, gives alienation of individual i with individual j, and the continuous effective antagonism function $\tilde{H}$, which is a measure of the extent of antagonism felt by person i toward person j, fulfills the restrictions that $\tilde{H}(\tilde{g}(p_i), 0) = 0$ and increasingness in alienation on the strictly positive part of the domain. They invoked the axioms 1–3 and a population homotheticity axiom which demands that ranking of two distributions by a polarization index remains invariant with respect to the size of the population, to derive the index

$$L_{\mathrm{ER}}(p,x) = \tilde{\Xi} \sum_{i=1}^{k} \sum_{j=1}^{k} p_i^{\tilde{\alpha}+1} p_j |x_i - x_j|, \tag{4.2}$$

where $\tilde{\Xi} > 0$ is a constant and $\tilde{\alpha} \in (0, 1.6]$ (*see* Theorem 1 of the authors). If another axiom [axiom 4 of Esteban and Ray, 1994] which demands that a migration from a very small population mass at a low income to a higher income of moderate size increases polarization is imposed, then $\tilde{\alpha}$ must take on the value in the interval [1,1.6] (*see* Theorem 3 of the authors). The multiplicative constant $\tilde{\Xi}$ is used for population normalization. If the parameter $\tilde{\alpha}$ takes on the value zero, the Esteban-Ray index L_{ER} would correspond to the (absolute) Gini index. The positive value of $\tilde{\alpha}$, and hence the identification function $p_i^{\tilde{\alpha}}$, plays an important role to underline

[1] See Levy and Murname (1992), Collier and Hoeffler (2001), Knack and Keefer (2001), Garcia-Montalvo and Reynal-Querol (2002, 2005), Reynal-Querol (2002), and Bossert and Schworm (2006).

[2] Amiel et al. (2007) investigated whether people's perception about polarization is consistent with different axioms.

the difference between inequality and polarization. As the value of $\tilde{\alpha}$ increases, the greater is the divergence from inequality and consequently, $\tilde{\alpha}$ may be interpreted as a polarization sensitivity parameter. Given income classes and total population, the index achieves its maximum value when half the population is concentrated in the lowest income class and the remainder is in the highest income class. On the other hand, it attains its minimum value if the entire population mass is concentrated at one value, which coincides with the mean and the median.

Duclos et al. (2004) developed an axiomatic characterization of the index

$$L_{\mathrm{DER}}(f) = \int_0^\infty \int_0^\infty (f(v'))^{1+\tilde{\alpha}} f(v)|v - v'|\mathrm{d}v\mathrm{d}v', \tag{4.3}$$

for income distributions defined in the continuum with a normalized mean of unity, where f is the income density function and $\tilde{\alpha} \in [.25, 1]$. The Duclos et al. index $L_{\mathrm{DER}}(f)$ can be regarded as the continuous analogue to the index L_{ER} in (4.2). It overcomes the limitation of the original index L_{ER} that requires a population to be bunched into relevant groups. They also constructed estimators for their index to use in the case of disaggregated data.

Clustering of the population into groups such that individuals feel identified inside a group and alienated outside it looses important information about income disparity within each group. Esteban et al. (2007) proposed an index that corrects L_{ER} in (4.2) from this perspective. Consider an income distribution with density f and the mean normalized at unity. For an income distribution with J income classes, let $\pi_i = \int_{x_{i-1}}^{x_i} f(v)\mathrm{d}v$ and $\lambda_i = 1/\pi_i \int_{x_{i-1}}^{x_i} vf(v)\mathrm{d}v$, respectively, be the population frequency and mean income of the income class $[x_{i-1}, x_i]$, $i = 1, 2, \ldots, J$. The corresponding vectors are given by $\underline{\pi}$ and $\underline{\lambda}$. Then the index L_{ER} applied to the discrete grouping considered here, with a correction for within-group inequality, is given by

$$L_{\mathrm{EGR}}(\underline{\pi}, \underline{\lambda}) = \sum_{i=1}^{J} \sum_{j=1}^{J} \pi_i^{1+\tilde{\alpha}} \pi_j |\lambda_i - \lambda_j| - c_5 \mathrm{er}(\underline{\pi}, \underline{\lambda}), \tag{4.4}$$

where the error $\mathrm{er}(\underline{\pi}, \underline{\lambda})$ corresponds "to the implicit fuzziness of group identification" (Esteban et al., 2007, p. 5) and $c_5 \geq 0$ is the weight assigned to the error. The presence of the error term ensures that the Esteban et al. index $L_{\mathrm{EGR}}(\underline{\pi}, \underline{\lambda})$ is decreasing in within-group and increasing in between-group disparities. They also considered the problem of grouping the population such that the error function, which has been chosen as the average of income distances within all groups, is minimized. The modified index can be applied to all kinds of income distributions.[3] If in (4.4), the term under double summation is multiplied by $(1 - I_{\mathrm{G}}^i)^{\tilde{\mu}}$ and the second term $-c_5\mathrm{er}(\underline{\pi}, \underline{\lambda})$ is dropped, then the resulting index becomes the variant of $L_{\mathrm{EGR}}(\underline{\pi}, \underline{\lambda})$ suggested by Lasso de la Vega and Urrutia (2006), where I_{G}^i is the Gini index of group i and $\tilde{\mu} \geq 0$ is a constant. Formally, their index is given by

[3] Gradin (2000) extended the index in (4.3) to the case when groups are defined according to attributes other than income, for example, education level, health.

$L_{\mathrm{LU}}(\underline{\pi},\underline{\lambda}) = \sum_{i=1}^{J}\sum_{j=1}^{J}\pi_i^{1+\tilde{\alpha}}\pi_j(1-I_{\mathrm{G}}^i)^{\tilde{\mu}}|\lambda_i-\lambda_j|$. The constant $\tilde{\mu} \geq 0$ represents the degree of sensitivity toward group cohesion. The sense of identification of each member of group i is now given by $\pi_i^{\tilde{\alpha}}(1-I_{\mathrm{G}}^i)^{\tilde{\mu}}$. Multiplication by the increasing function $(1-I_{\mathrm{G}}^i)^{\tilde{\mu}}$ of the Gini index of equity $(1-I_{\mathrm{G}}^i)$ makes the polarization index L_{LU} a decreasing function of within-group dispersion. It is evidently increasing in between-group inequality. It should be clear that Esteban et al. index and its variant L_{LU} have several common properties.

D'Ambrosio (2001) proposed a modification of L_{ER} in (4.2) using the Kolomogorov measure of distance as the alienation function instead of the simple distance function $|x_i - x_j|$ for taking into account the intergroup measure of disparity. The D'Ambrosio index is then given by

$$L_{\mathrm{D}}(\underline{\pi},\underline{f}) = \frac{1}{2}\sum_{i=1}^{J}\sum_{j=1}^{J}\pi_i^{\tilde{\alpha}+1}\pi_j\int_0^{\infty}|f_i(v)-f_j(v)|\mathrm{d}v, \tag{4.5}$$

where π_is are population frequencies; $1/2\int_0^{\infty}|f_i(v)-f_j(v)|\mathrm{d}v$ is the Kolomogorov measure of distance between groups i and j; f_i and f_j, are respectively, the densities of income distributions corresponding to these groups and $\underline{f} = (f_1, f_2, \ldots, f_J)$. An advantage of the use of this alternative alienation function is that the disparity between groups are now compared using their income distributions, not by their means, as is done in (4.2).

Milanovic (2000) suggested an index with the objective that it (1) achieves the minimal value zero for a distribution with population mass concentrated at a single point; (2) takes on the maximal value one for a society subdivided into two extreme groups of equal size, where all the incomes in the first group are zero and all the incomes in the other group are twice the mean; (3) increases if the difference between the incomes of the two groups increases, keeping the population masses in the groups fixed; and (4) satisfies scale invariance but decreases under equal absolute augmentation in all incomes. Postulates (1) and (2) are similar to the corresponding properties of L_{ER} in (4.2). Postulate (3) is analogous to the alienation function of Esteban and Ray (1994). The Milanovic index, which incorporates the idea of alienation in its formulation, has a Gini-type structure and measures the divergence of incomes from the situation of minimum polarization.

Zhang and Kanbur (2001) employed the ratio between the between-group and within-group components of the Shorrocks (1980) weighted generalized entropy index of inequality for measuring polarization. Formally, the Zhang-Kanbur index is defined as

$$L_{\mathrm{ZK}}(p,x) = \frac{I(p;\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots\ldots, \lambda_J 1^{n_J})}{\sum_{i=1}^{J} w_i(\underline{\lambda},\underline{n})I(p^i,x^i)}, \tag{4.6}$$

where for any partitioning of the population into J groups with respect to some homogeneous characteristic (say, age, sex, region, etc.), n^i is the population size of group i whose income distribution and mean income are respectively (p^i, x^i) and $\lambda_i, \underline{\lambda} = (\lambda^1, \lambda^2, \ldots\ldots, \lambda^J)$, $\underline{n} = (n^1, n^2, \ldots\ldots, n^J)$, $w_i(\underline{\lambda},\underline{n})$ is the positive weight attached to inequality in (p^i, x^i), assumed to depend on the vectors $\underline{n}$ and $\underline{\lambda}$, and

$p = (p^1, p^2, \ldots.., p^J)$, $x = (x^1, x^2, \ldots.., x^J)$. Although the Zhang-Kanbur approach is different from that of Esteban and Ray (1994), there is a similarity in interpretation. The within-group term may be interpreted as an inverse indicator of feelings of identification between similar individuals – L_{ZK} increases if the groups become more concentrated, that is, if within-group inequality reduces. Also, the further apart are the means, the greater is the degree of polarization. Thus, the between-group term is an indicator of feelings of alienation between dissimilar individuals. The weighted generalized entropy family, which forms the basis of the index L_{ZK} and can be expressed as the sum of the between-group and within-group components considered in (4.6), is defined as

$$I_{\mathrm{S}}(p,x) = \begin{cases} \frac{1}{n\bar{c}(\bar{c}-1)} \sum_{i=1}^{k} p_i \left[\left(\frac{x_i}{\lambda}\right)^{\bar{c}} - 1 \right], \bar{c} \neq 0, 1, \\ \frac{1}{n} \sum_{i=1}^{k} p_i \left[\log\left(\frac{\lambda}{x_i}\right) \right], \bar{c} = 0, \\ \frac{1}{n} \sum_{i=1}^{k} p_i \left[\left(\frac{x_i}{\lambda}\right) \log\left(\frac{x_i}{\lambda}\right) \right], \bar{c} = 1, \end{cases} \tag{4.7}$$

where positivity of the lower bound μ of the income interval $[\mu, \gamma]$ is necessary for the index I_{S} to be well-defined in all cases. The real number $\bar{c}$ is a transfer sensitivity parameter – a transfer of income from a person to anyone who has a lower income decreases I_{S} by a larger amount, the lower is the value of $\bar{c}$ (*see* Shorrocks, 1980).

In the remainder of this section, we will discuss polarization indices that are concerned with the decline of the middle class, more precisely, with bipolarization. Several attempts considered some income interval around the median and defined the decline of the middle class in terms of reduction in population/income share corresponding to the interval. (*See*, e.g., Beach, 1989; Beach et al., 1997; Blackburn and Bloom, 1985; Horrigan and Haugen, 1988; Ilg and Haugen, 2000; McMahon and Tsechetter, 1986; Rosenthal, 1985; Wolfson, 1997).

Rigorous attempts to study the decline of the middle class have first been made by Foster and Wolfson (1992) and Wolfson (1994, 1997). Given any income distribution, they defined its (relative) bipolarization curve that shows for any population proportion, how far a normalized value of the share of the total income enjoyed by that proportion is from the corresponding share that it would receive under the hypothetical situation where everybody enjoys the median income (*see* Sect. 4.4). The area under the curve, which is popularly known as the Wolfson polarization index, is given by

$$L_{\mathrm{W}}(p,x) = \frac{2\lambda(2Q - I_{\mathrm{G}}(p,x))}{m}, \tag{4.8}$$

where $Q = (\lambda(p^+, x^+) - \lambda(p^-, x^-))/2\lambda$ and $I_{\mathrm{G}}(p,x) = 1/2n^2\lambda \sum_{i=1}^{k} \sum_{j=1}^{k} p_i p_j |x_i - x_j|$ is the (relative) Gini index of the income distribution (p,x). L_{W} fulfills all the postulates for a bipolarization index. For bipartitioning of the population into deprived and satisfied groups, L_{W} can be rewritten as

$$L_{\mathrm{W}}(p,x) = \frac{2\lambda\,(I_{\mathrm{G}}^{\mathrm{BI}}(p,x) - I_{\mathrm{G}}^{\mathrm{WI}}(p,x))}{m}, \tag{4.9}$$

where $I_{\mathrm{G}}^{\mathrm{BI}}(I_{\mathrm{G}}^{\mathrm{WI}})$ is the corresponding between-group(within-group) component of the Gini index (*see* Rodriguez and Salas, 2003). Under ceteris paribus assumptions, L_{W} is increasing in $I_{\mathrm{G}}^{\mathrm{BI}}$, the alienation component, and decreasing in $I_{\mathrm{G}}^{\mathrm{WI}}$, an inverse indicator of identification.

Rodriguez and Salas (2003) also suggested the use of the difference

$$L_{\mathrm{RS}}(F) = I_{\mathrm{DWW}}^{\mathrm{BI}}(F) - I_{\mathrm{DWW}}^{\mathrm{WI}}(F) \tag{4.10}$$

as a bipolarization index and referred to this as the extended Wolfson index, where F is the income distribution function. $I_{\mathrm{DWW}}^{\mathrm{BI}}(F)(I_{\mathrm{DWW}}^{\mathrm{WI}}(F))$ is the between-group (within-group) component associated with the Donaldson and Weymark (1980, 1983) welfare ranked S-Gini inequality index, given that the population is bi-partitioned using the median as the reference point. The boundedness condition $2 \leq \delta \leq 3$ is necessary for the extended index to satisfy the postulate Increased Bipolarity. The higher is the value of δ, the higher is the weight assigned by the Rodriguez-Salas index to the identification and alienation terms.

If we employ the Gini index in (4.6) under the bipartitioning of the population using the median, then L_{ZK} will be $L_{\mathrm{ZK}}^{I_{\mathrm{G}}}(p,x) = I_{\mathrm{G}}^{\mathrm{BI}}/I_{\mathrm{G}}^{\mathrm{WI}}$. The increasing transformation

$$L_{\mathrm{SDH}}(p,x) = \frac{L_{\mathrm{ZK}}^{I_{\mathrm{G}}} - 1}{L_{\mathrm{ZK}}^{I_{\mathrm{G}}} + 1} \tag{4.11}$$

of $L_{\mathrm{ZK}}^{I_{\mathrm{G}}}(p,x)$ was suggested as an index of bipolarization by Silber et al. (2007). Since the Silber et al. index L_{SDH} is increasingly related to $L_{\mathrm{ZK}}(p,x)$, applied to the Gini index, it shares the properties of the latter. As Silber et al. (2007) noted $1 - L_{\mathrm{SDH}}$ is an indicator of kurtosis of the income distribution (Berrebi and Silber, 1989). A measure of kurtosis indicates the degree of steepness or peakedness of the distribution.

In an interesting paper, Wang and Tsui (2000) suggested the use of

$$L_{\mathrm{WT}}^{\phi}(p,x) = \frac{1}{n}\sum_{i=1}^{k} p_i \phi\left(\left|\frac{x_i - m}{m}\right|\right) \tag{4.12}$$

and

$$L_{\mathrm{WT}}^{\varphi}(p,x) = \frac{1}{n}\sum_{i=1}^{k} p_i \varphi(|x_i - m|), \tag{4.13}$$

as relative and absolute indices of bipolarization. The Wang-Tsui indices aggregate the deviations of individual incomes from the median through the continuous transformations ϕ and φ, respectively. They are easy to understand and quite reasonable intuitively. Wang and Tsui (2000) also showed that they satisfy Increased Spread and Increased Bipolarityif and only ϕ and φ are increasing and strictly concave.

Alesina and Spolaore (1997) proposed a median-based index L_{AS}, which is implicitly defined by

$$F(m+L_{\mathrm{AS}}) - F(m-L_{\mathrm{AS}}) = \frac{1}{2}. \tag{4.14}$$

Since $F(v)$ gives the cumulative proportion of the population with income less than or equal to v, we can interpret the Alesina-Spolaore index I_{AS} as follows. It is that level of income which when added to and subtracted from the median makes the difference between the resulting cumulative population proportions equal to half. Since L_{AS} identifies a symmetric income interval around the median, it has some similarity with the interval-based indices we have discussed earlier.

So far the indices we have considered are descriptive; they are derived without using any concept of welfare. Such indices contrast with ethical indices that are designed from explicit social welfare functions. Needless to say, neither type of indices is meant to supplant the other type. Chakravarty and Majumder (2001) and Chakravarty et al. (2007) suggested relative and absolute indices of bipolarization using explicit forms of social welfare function. In their framework, bipolarization is measured in terms of welfare related to the given distribution. The relative index proposed by Chakravarty and Majumder (2001) is defined as

$$\begin{aligned} L_{\mathrm{CM}}(p,x) = & \frac{\Xi(\lambda(p^+,x^+), I(p^+,x^+)) + 2\lambda(p^+,x^+)}{2m} \\ & + \frac{\Xi(\lambda(p^-,x^-), I(p^-,x^-)) - B_1(m)\lambda(p^-,x^-)}{2m} + B_2(m) \end{aligned} \tag{4.15}$$

where the reduced form social welfare function Ξ is increasing in efficiency (λ) and decreasing in relative inequality (I). The continuous normalization coefficients $B_1(m)$ and $B_2(m)$ have to be chosen such that different postulates for a bipolarization index are satisfied.

To illustrate the Chakravarty-Majumder index in (4.15), suppose that $\mu > 0$ and welfare evaluation is done with the weighted mean of order $\theta < 1$, the Atkinson (1970) abbreviated welfare function for (p,x), that is,

$$\Xi_{\mathrm{A}}(\lambda(p,x), I_{\mathrm{A}}(p,x)) = \begin{cases} \left(\frac{1}{n}\sum\limits_{i=1}^{k} p_i x_i^{\theta}\right)^{1/\theta}, \theta < 1, \theta \neq 0, \\ \prod\limits_{i=1}^{k} (x_i)^{p_i/n}, \theta = 0, \end{cases} \tag{4.16}$$

and $B_1(m) = (m/\mu)^{1-\theta}$, $B_2(m) = 1/2(m/\mu)^{1-\theta} - 2$. Then L_{CM} becomes a fairly natural translation of the Atkinson (1970) index of inequality into bipolarization measurement. Under a progressive income transfer on the either side of the median, a reduction in the value of θ increases polarization by a larger amount, the lower is the value of θ. As $\theta \to -\infty$, Ξ_{A} approaches the Rawls (1971) maximin welfare function and I_{CM} becomes the relative maximin index of bipolarization. Next, if $\Xi(p,x) = \lambda(1-I_{\mathrm{G}})$, the Gini welfare function, $B_1(m) = 4$ and $B_2(m) = 0$, then L_{CM} becomes the Wolfson index (for even values of n). Thus, given any relative inequality index (or its associated welfare function), we can generate a corresponding relative bipolarization index using (4.15).

The absolute counterpart to L_{CM} suggested by Chakravarty et al. (2007) is given by

$$L_{\mathrm{CMR}}(p,x) = \frac{\Xi(\lambda(p^+,x^+),I(p^+,x^+)) + 2\lambda(p^+,x^+)}{2} + \frac{\Xi(\lambda(p^-,x^-),I(p^-,x^-)) - B_3(m)\lambda(p^-,x^-)}{2} + B_4(m), \quad (4.17)$$

where the continuous normalization coefficients $B_3(m)$ and $B_4(m)$ serve the same purpose as $B_1(m)$ and $B_2(m)$ in (4.15), and the abbreviated welfare function Ξ retains all the assumptions, except relativity of I. In this case, we assume that I is an absolute index. For the purpose of illustration, assuming that n is even, $B_3(m) = 4$ and $B_4(m) = 0$, we can use the absolute Gini index in (4.17). The resulting Chakravarty et al. index L_{CMR} may be referred to as the absolute Gini index of bipolarization. Alternatively, we may employ the weighted Kolm (1976a)-Pollak (1971) welfare function

$$\Xi_{\mathrm{KP}}(\lambda(p,x), I_{\mathrm{KP}}(p,x)) = -\frac{1}{\beta}\log\frac{1}{n}\sum_{i=1}^{k} p_i(\exp(-\beta x_i)) \quad (4.18)$$

in (4.17) to get the corresponding form of the bipolarization index, where the free parameter $\beta > 0$ determines the curvature of the social indifference surfaces. An increase in the value of β makes the social indifference curve more convex to the origin. The normalization coefficients chosen in this case are $B_3(m) = \exp(-\beta(\mu - m))$ and $B_4(m) = m\exp(-\beta(\mu - m))/2 - 2m$. Thus, given any absolute inequality index, we have a corresponding index of bipolarization. These indices will differ in the way we make welfare evaluation.[4]

4.3 A New Compromise Bipolarization Index, its Properties, and Characterization

In bipolarization measurement, we are concerned with deviations of incomes from the median. This motivates us to construct a compromise index of polarization based on transformed values of such deviations. Formally, for any income distribution (p,x), we consider the transformed deviations $\psi(|m - x_i|)$, where ψ is continuous, increasing and $\psi(0) = 0$. A median-based deviation function ψ satisfying these

[4] Some of these studies and several other studies have examined the extent of polarization in different countries over different periods. See, for example, Thurow (1984), Kosters and Ross (1988), Morris et al. (1994), Jenkins (1995), Kovacevic and Binder (1997), Quah (1997), Wolfson (1997), Gradin (2000, 2002), Chakravarty and Majumder (2001), Zhang and Kanbur (2001), Anderson (2004a,b), Duclos and Echevin (2005), Gigliarano (2006), Chakravarty et al. (2007), and Esteban et al. (2007).

conditions will be called regular. Given any income distribution, let d_e be the associated representative deviation, that is, d_e is that level of deviation which, if assigned to each individual, will make the resulting distribution median-based deviation indifferent to the existing distribution. Formally, given the income distribution (p,x) and a regular ψ, the corresponding d_e is implicitly defined by,

$$\sum_{i=1}^{k} p_i \psi(d_e) = \sum_{i=1}^{k} p_i \psi(|m-x_i|). \tag{4.19}$$

As an index of bipolarization, we now suggest the use of

$$L_\psi(p,x) = \frac{d_e}{m}. \tag{4.20}$$

The index L_ψ simply is an average of income deviations from the median as a fraction of the median itself.

The following theorem summarizes some properties of L_ψ.

Theorem 4.1. *Assume that ψ is regular.*

(i) Then L_ψ satisfies Normalization, Symmetry, the Population Principle, Increased Spread, and Continuity.
(ii) L_ψ satisfies Increased Bipolarity if and only if ψ is strictly concave.

Proof (i). From (4.19), we note that we can write d_e explicitly as $d_e(p,x) = \psi^{-1}[(1/n)\sum_{i=1}^{k} p_i \psi(|m-x_i|)]$. Since each x_i is drawn from the compact set $[\mu,\gamma]$ and the deviations $|m-x_i|$ are nonnegative, they will also take values in a compact set of the form $[0,\gamma']$. Thus, the domain of the function $\psi(|m-x_i|)$ is $[0,\gamma']$. Now, since ψ is increasing and the continuous image of a compact set is compact (Rudin, 1976, p. 89), $\psi(|m-x_i|)$ takes values in the compact set $[\psi(0),\psi(\gamma')]$, which, in view of the fact that $\psi(0)=0$, can be rewritten as $[0,\psi(\gamma')]$. For a given p, continuity and increasingness of the function ψ implies that the average function $(1/n)\sum_{i=1}^{k} p_i \psi(|m-x_i|)$ is continuous and takes values in $[0,\psi(\gamma')]$. Observe that increasingness of ψ ensures the existence of ψ^{-1}. Continuity and increasingness of ψ^{-1} on $[0,\psi(\gamma')]$ now follows from Theorem 4.53 of Apostol (1975, p. 95). This in turn demonstrates continuity of L_ψ.

Since $\psi(0)=0$ and ψ is increasing, $\psi^{-1}(0)=0$. This establishes that L_ψ satisfies Normalization. It is easy to check that L_ψ satisfies Symmetry and the Population Principle. Given that ψ^{-1} is increasing, the proof of satisfaction of Increased Spread by L_ψ follows from Proposition 5 of Wang and Tsui (2000).

Proof (ii). Using the fact that ψ^{-1} is increasing and Proposition 5 of Wang and Tsui (2000) again, we can show that L_ψ satisfies Increased Bipolarity if and only if ψ is strictly concave. This completes the proof of the theorem. □

Since the index in (4.20) has been expressed in a ratio form, it is reasonable to expect that it will be a relative index. However, Theorem 4.1 does not say anything about this. There can be many regular ψ functions for which the theorem holds. Examples of such functions are: $\psi_1(v) = \eta_1 v^{\varepsilon}, 0 < \varepsilon < 1, \eta_1 > 0$; $\psi_2(v) = 1 - e^{-\eta_2 v}, \eta_2 > 0$; and $\psi_3(v) = (v/1+v)$. The following theorem shows that ψ_1 is the only regular ψ function for which L_{ψ} is a relative index.

Theorem 4.2. *Assume that ψ is regular and strictly concave. Then L_{ψ} in (4.20) is a relative index if and only if $\psi(v) = \eta_1 v^{\varepsilon}$, where $0 < \varepsilon < 1$ and $\eta_1 > 0$ are constants.*

Proof. Since the denominator of (4.20) is linear homogeneous, for L_{ψ} to be a relative index, we need linear homogeneity of the numerator as well. Note that $d_e(p,x) = \psi^{-1}[(1/n)\sum_{i=1}^{k} p_i \psi(|m - x_i|)]$ is a quasi-linear mean of income deviations and, given continuity of ψ, it satisfies linear homogeneity if and only if $\psi(v) = \eta_1 v^{\varepsilon} + \tilde{\eta}_1$, where η, ε, and v_1 are constants (Aczel, 1966, p.153). Since ψ is increasing and strictly concave, we must have $0 < \varepsilon < 1$ and $\eta > 0$. Next, $\psi(0) = 0$ ensures that $\tilde{\eta}_1 = 0$. This establishes the necessity part of the theorem. The sufficiency is easy to check. □

Substitution of the form of ψ, identified in Theorem 4.2, in (4.20) yields the following form of the bipolarization index:

$$L_{\varepsilon}(p,x) = \frac{(1/n\sum_{i=1}^{k} |m - x_i|^{\varepsilon})^{1/\varepsilon}}{m}, \quad 0 < \varepsilon < 1. \tag{4.21}$$

L_{ε} in (4.21) is the ratio between the weighted mean of order ε of deviations of individual incomes from the median and the median. Given (p,x), an increase in the value of ε increases L_{ε}. A progressive transfer of income on the either side of the median increases L_{ε} by a larger amount, the lower is the value of ε. As $\varepsilon \to 1$, L_{ε} approaches the simple average of the relative deviations of individual incomes from the median. In this particular case, L_{ε} satisfies Increased Spread but not Increased Bipolarity.

The absolute version $L_{A\varepsilon}$ of L_{ε} is given by mL_{ε}, that is,

$$L_{A\varepsilon}(p,x) = \left(\frac{1}{n}\sum_{i=1}^{k} p_i |m - x_i|^{\varepsilon}\right)^{1/\varepsilon}, \quad 0 < \varepsilon < 1. \tag{4.22}$$

Conversely, we can start with the absolute index $L_{A\varepsilon}$ and translate it into its relative counterpart L_{ε} by dividing by the median. This compromise property is shared by the Wolfson index also.

4.4 Bipolarization Dominance

Evidently, different indices of bipolarization may rank alternative distributions of income in different directions. To avoid such different directional rankings, this section attempts to develop criteria for ordering income distributions in the same direction using bipolarization indices. Since median is the reference income in the measurement of bipolarization, our orderings rely on deviations of incomes from the median. For simplicity of exposition, we assume that the population mass vector is given by $p = 1^k$, that is, the frequency of each income is 1. Therefore, we now write $L(x)$ instead of $L(p,x)$ to denote the level of bipolarization of the distribution (p,x). Further, all income distributions are assumed to be illfare-ranked and let $\hat{k} = (k+1)/2$.

For any $x \in [\mu,\gamma]^k$, the normalized aggregate deviation $\mathrm{RB}(x, j/k) = 1/km \sum_{j \le i < \hat{k}} (m - x_i)$ is the shortfall of the total income of the population proportion j/k from the corresponding total that it would enjoy under the hypothetical distribution where everybody possesses the median income, as a fraction of the factor km, where $1 \le j < \hat{k}$. This is, in fact, the ordinate corresponding to the population proportion j/k of the relative bipolarization curve (RBC) of x, where $1 \le j < \hat{k}$. For incomes not below the median, the corresponding ordinate is $(1/km) \sum_{\hat{k} \le i \le j} (x_i - m), \hat{k} \le j \le k$. If k is odd, the RBC of x, $\mathrm{RB}(x,t)$, where $t \in [0,1]$, is completed by assuming $\mathrm{RB}(x,0) = 1$ and by defining

$$\mathrm{RB}\left(x, \frac{j+\tau}{k}\right) = (1-\tau)\mathrm{RB}\left(x, \frac{j}{k}\right) + \tau\mathrm{RB}\left(x, \frac{j+1}{k}\right) \tag{4.23}$$

for all $0 \le \tau \le 1$ and $1 \le j \le (k-1)$. Recall that if k is odd, then $m = x_{\hat{k}}$ is the middle most income of the distribution and the ordinate at $\hat{k}/k$ is well-defined.

If k is even, the curve is completed by setting $\mathrm{RB}(x,0) = 1$ and by defining

$$\begin{gathered}\mathrm{RB}\left(x, \frac{j+\tau}{k}\right) = (1-\tau)\mathrm{RB}\left(x, \frac{j}{k}\right) + \tau\mathrm{RB}\left(x, \frac{j+1}{k}\right), \\ \text{for all } 0 \le \tau \le 1 \text{ and } 1 \le j \le (k-1), \quad j \ne \hat{k},\end{gathered} \tag{4.24}$$

and

$$\mathrm{RB}\left(x, \frac{\hat{k}-.5+\tau}{k}\right) = (1-\tau)\mathrm{RB}\left(x, \frac{\hat{k}-.5}{k}\right) + \tau\mathrm{RB}\left(x, \frac{\hat{k}+.5}{k}\right), \text{ for all } 0 \le \tau \le 1.$$

Note that when k is even, $x_{\hat{k}}$ is not in x. However, the ordinate corresponding to the proportion $\hat{k}/k$ is defined (*see* Chakravarty et al., 2007).

If the income distribution is perfectly equal, the RBC coincides with the horizontal axis. For a typical unequal income distribution, the curve decreases until we reach the midpoint, where it coincides with the horizontal axis and then increases monotonically (Fig. 4.1).

Given any two income distributions $x, y \in [\mu,\gamma]^k$, x is said to dominate y with respect to relative bipolarization ($x \ge_{\mathrm{RB}} y$, for short) if

$$\mathrm{RB}(x,t) \ge \mathrm{RB}(y,t) \tag{4.25}$$

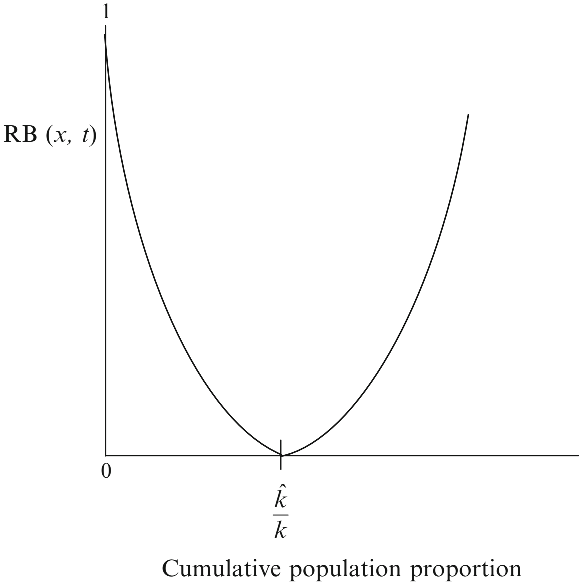

Fig. 4.1 Relative bipolarization curve

for all $t \in [0,1]$, with strict inequality for some t. That is, $x \geq_{\text{RB}} y$ means that the RBC of x lies nowhere below that of y and at some places (at least), the former lies above. Clearly, the relative bipolarization dominance relation $\geq_{\text{RB}}$ is transitive, that is, for any $x, y, u \in [\mu, \gamma]^k$ if $x \geq_{\text{RB}} y$ and $y \geq_{\text{RB}} u$ hold, then we must have $x \geq_{\text{RB}} u$. However, it is not complete, that is, we may be able to find $x, y \in [\mu, \gamma]^k$ such that neither $x \geq_{\text{RB}} y$ nor $y \geq_{\text{RB}} x$ holds. Obviously, this is a consequence of intersection of the two curves. Thus, like its Lorenz counterpart, $\geq_{\text{RB}}$ is a quasi-ordering.

To illustrate the construction of the RBC, consider the distribution $x = (1,2,5,6,10)$. Here $m = 5, \hat{k} = 3, x^- = (1,2)$, and $x^+ = (5,6,10)$. Then the ordinates of the RBC of x corresponding to the population proportions $j/5$, where $j = 1,2,\ldots,5$, are given, respectively, by $7/25, 3/25, 0, 1/25$, and $6/25$.

The following result is an implication of the dominance relation $\geq_{\text{RB}}$ for income distributions with the same population size and arbitrary medians.

Theorem 4.3. *Let $x, y \in [\mu, \gamma]^k$ be arbitrary. Then the following conditions are equivalent:*

(i) $x \geq_{RB} y$.
(ii) $L(x) > L(y)$ for all relative bipolarization indices $L : [\mu, \gamma]^k \to R^1$ that satisfy Increased Spread, Increased Bipolarity, and Symmetry.

Proof. The idea of the proof is taken from Foster and Shorrocks (1988b) and Chakravarty et al. (2007). In proving the theorem, we assume for simplicity that n is odd. A similar proof will hold when n is even.

(i) $\Rightarrow$ (ii): Define $u = m(x)/m(y)y$. Since RBC is scale invariant, we have $\mathrm{RB}(y,t) = \mathrm{RB}(u,t)$, which in turn says that $x \geq_{\mathrm{RB}} y$ is same as $x \geq_{\mathrm{RB}} u$. Observe also that $m(x) = m(u)$. Assume that the curves do not coincide for the subvectors x^+ and y^+ (hence u^+). Then $x \geq_{\mathrm{RB}} u$ along with $m(x) = m(u)$ implies

$$\sum_{i=\hat{k}}^{j} x_i \geq \sum_{i=\hat{k}}^{j} u_i, \hat{k} \leq j \leq k, \quad \text{with} > \text{ for some } j. \tag{4.26}$$

This gives rise to one of following two possibilities: (iii) $\lambda(u^+) = \lambda(x^+)$ and (iv) $\lambda(u^+) \neq \lambda(x^+)$. If the former holds then we have $x^+ \geq_{\mathrm{LC}} u^+$ and x^+ is obtained from u^+ by a finite sequence of rank-preserving progressive income transfers among persons above the median (Hardy et al., 1934). If condition (iv) holds then we note from (4.26) that $\lambda(u^+) < \lambda(x^+)$. Define $\tilde{u}_i = u_i^+$ for $\hat{k} \leq i < k$ and $\tilde{u}_k = u_k^+ + (k - \ \hat{k} - 1)(\lambda(x^+) - \lambda(u^+))$. That is, $\tilde{u}$ is obtained from u^+ by a simple increment. Then (4.26) implies either $x^+ = \tilde{u}$ or $x^+ \geq_{\mathrm{LC}} \tilde{u}$, in which case we can obtain x^+ from $\tilde{u}$ by a finite sequence of rank-preserving progressive transfers as before.

Likewise, if we assume that the two curves do not coincide for subvectors x^- and y^- (hence u^-), then x^- is obtained from u^- by reducing some incomes and/or by some equalizing transfers below the median. This means that the overall distribution x can be derived from the distribution u through the transformations specified in Increased Spread and/or Increased Bipolarity.

Since L satisfies Increased Spread and Increased Bipolarity, we have $L(x) > L(u)$. Symmetry of L follows from the fact that it has been defined on ordered distributions. As L is a relative index, $L(y) = L(u)$. This implies that $L(x) > L(y)$.

(ii) $\Rightarrow$ (i): Our demonstration above shows that the deduction of x from u by a sequence of spread increasing movements and/or egalitarian transfers on the same side of the median is equivalent to relative bipolarization dominance on that side. This in turn proves the implication (ii) $\Rightarrow$ (i) of the theorem. (See Theorem 1.4.) This completes the proof of the theorem. □

Theorem 4.3 shows that a unanimous ranking of income distributions over a given population size by all symmetric, relative bipolarization indices satisfying Increased Spread and Increased Bipolarity can be obtained if and only if relative bipolarization dominance holds. But if the two curves cross, we can get two different indices with these properties that will rank the underlying income distributions in opposite directions. Note that in the proof of the theorem, if condition (iv) holds, then x^+ second order stochastic dominates u^+. Equivalently, we can say that $x^+ \geq_{\mathrm{GL}} u^+$ holds.

Comparisons of polarization across populations generally involve different population sizes. For polarization ranking of distributions with differing population sizes, we have the following result.

Theorem 4.4. *Let $x \in [\mu, \gamma]^k, y \in [\mu, \gamma]^l$ be arbitrary. Then the following conditions are equivalent:*

(i) $x \geq_{\mathrm{RB}} y$.
(ii) $L(x) > L(y)$ *for all relative bipolarization indices* $L : \Psi \to R^1$ *that satisfy Increased Spread, Increased Bipolarity, Symmetry, and the Population Principle, where* $\Psi = \bigcup_{k \in N} [\mu, \gamma]^k$ *and N is the set of positive integers.*

Proof. (i) $\Rightarrow$ (ii): Let u^1 and u^2 be l- and k-fold replications of x and y, respectively. Since RBC is population replication invariant, we have $\mathrm{RB}(x,t) = \mathrm{RB}(u^1,t)$ and $\mathrm{RB}(y,t) = \mathrm{RB}(u^2,t)$. Therefore, $x \geq_{\mathrm{RB}} y$ is same as $u^1 \geq_{\mathrm{RB}} u^2$. As u^1 and u^2 are two distributions over the population size kl, using Theorem 4.3, we have $L(u^1) > L(u^2)$ for all relative bipolarization indices L that meet the properties stated in condition (ii) of Theorem 4.3. By the Population Principle, we have $L(u^1) = L(x)$ and $L(u^2) = L(y)$. Hence, $L(x) > L(y)$. A similar argument will demonstrate that the reverse implication is also true. □

Theorem 4.4 states that an unambiguous ranking of two arbitrary income distributions by relative bipolarization indices can be achieved through pairwise comparisons of their RBCs. Since we do not assume equality of the medians and the population sizes, this is the most general result we can have along this direction.

We can also focus our attention on the fixed median arbitrary population size case. In this case, the domain of definition of bipolarization indices is $\Psi_{\bar{m}} = \{x \in \Psi | m(x) = \bar{m}\}$. For all indices defined on $\Psi_{\bar{m}}$, we now have the following equivalence theorem, whose proof is similar to those of Theorems 4.3 and 4.4.

Theorem 4.5. *Let* $x, y \in \Psi_{\bar{m}}$ *be arbitrary. Then the following conditions are equivalent:*

(i) $x \geq_{\mathrm{RB}} y$.
(ii) $L(x) > L(y)$ *for all bipolarization indices* $L : \Psi_{\bar{m}} \to R^1$ *that fulfill Increased Spread, Increased Bipolarity, Symmetry, and the Population Principle.*

Given the median, relative bipolarization dominance becomes a sufficient condition for all relative and absolute (hence compromise) bipolarization indices, satisfying the axioms stated in condition (ii) of Theorem 4.5, to rank different income distributions in the same way.

Finally, if both mean income and median are fixed, the postulates we need for the indices to be consistent with the relation $\geq_{\mathrm{RB}}$ are Increased Bipolarity, Symmetry, and the Population Principle. There can also be situations where mean is fixed, median is different, and population size is equal/unequal. For consistency with $\geq_{\mathrm{RB}}$, while in the former case, the relative indices should be symmetric and increasing under the permissible egalitarian transfers; in the latter case, they should be population replication invariant as well.

For ranking income distributions by absolute bipolarization indices, Chakravarty et al. (2007) scaled up the RBC by the median to generate the absolute bipolarization curve (ABC). Formally, we have $\mathrm{AB}(x,t) = m\mathrm{RB}(x,t)$ for $0 \leq t \leq 1$. The area under this curve turns out to be mL_{W}, the absolute version of the Wolfson index. Thus, the Wolfson index can be converted into an absolute index by multiplying with the median.

Clearly, we can have absolute counterparts to Theorems 4.3–4.4 if we replace relative indices by absolute indices and $\geq_{\text{RB}}$ by $\geq_{\text{AB}}$, the absolute bipolarization dominance relation, defined in the same way as $\geq_{\text{RB}}$. In fact, in some cases, ambiguous comparison under $\geq_{\text{RB}}$ can be unambiguous under $\geq_{\text{AB}}$. For example, consider two distributions x and y, where $m(x) > m(y)$ and the RBC of the former lies below that of the latter up to a point t_0 below the midpoint of the horizontal axis. But after that the RBC of x does not lie below that of y. Given that we have $m(x) > m(y)$, multiplication of these RBCs by the corresponding medians may give rise to an upward shift in the ABC of x at the left of the point of intersection t_0 such that $x \geq_{\text{AB}} y$ holds. Thus, the higher median has a scaling effect for pushing the lower curve upward to guarantee absolute bipolarization dominance.