ESSAYS ON INEQUALITY, POLARIZATION AND CONTESTS

BHARGAV PRATIM BHATTACHARYA

Thesis submitted to the **Indian Statistical Institute**
in partial fulfillment of the requirements for the award of the degree of
Doctor of Philosophy

Indian Statistical Institute, Kolkata
May, 2015

Acknowledgements

At this joyous moment of submission of my thesis, parts of which have already been published in various journals, I gratefully acknowledge the co-operation received from various quarters.

First of all, I express my sincere gratitude to Prof. Satya R. Chakravarty, my supervisor for his constant direction and positive, enthusiastic response at each and every step forward. Having come from an alien background, it would have been impossible for me to work in Quantitative Economics without his tireless support.

Distinguished faculty members of Economic Research Unit, ISI, Kolkata have responded very cordially on a number of occasions. Let me mention gratefully the esteemed names of Prof. Nityananda Sarkar, Prof. Abhirup Sarkar, Prof. Manoranjan Pal, Prof. Manipushpak Mitra, Prof. Tarun Kabiraj, Prof. Samarjit Das, Prof. Priyodarshi Banerjee and Prof. Broti Sankar Chakraborty for the co-operation received at different points of time.

It is my pleasure to convey my heartfelt thanks to Dr. Coonan Mukherjee, Sri Rajit Biswas, Sri Parikshit De, Sri Sandip Sarkar, Sri Tanmoy Das, Sri Mahamitra Das and other research scholars who assisted me in every possible way: in academic as well as in non-academic spheres.

I had very fruitful and enlightening discussions with Prof. Nachiketa Chattopadhyay of SOSU, ISI, Kolkata, especially in the initial phase of my work. Prof. Sankar Dihidar (PSU) and Prof. Tapas Samanta (ASU) of this institute lent their helping hands whenever they were sought after. Dr. Biswajit Das (ex-SRF, SMU, ISI, Kolkata) partook some of his brilliant ideas with me. I pay my deep respects to all of them. Assistance received from the office staff of Economic Research Unit also deserves grateful acknowledgement.

During this sojourn, I had opportune moments of coming in contact with luminaries like Prof. Debraj Ray and Prof. Joan Esteban whose valuable suggestions have been very helpful.

I was fortunate to have taken part and presented papers in the 11th congress of the Society for Social Choice and Welfare (SSCW) held at Delhi on August, 2012 and in the 9th and 10th Annual Conferences on Economic Growth and Development (ACEGD) hosted by ISI, Delhi on December, 2013 and December, 2014 respectively. A number of questions were raised by the participants of these congregations which helped me in improving the quality of my thesis.

And there are innumerable friends, colleagues and students who encouraged and assisted me all through the journey. My thanks and gratitude to all of them!

Finally, the monastic members of Ramakrishna Mission Vidyamandira deserve special mention. It was at their directive that I undertook the project, it was with them that I shared my experiences of light and darkness, of hope and despair, of acceleration and retardation and I always received their blessings and constant inspiration for pursuing the work initiated. I dedicate this piece of work to this monastic family!

Belurmath, Howrah — Bhargav Pratim Bhattacharya

28th May, 2015

Contents

List of figures

Chapter 1

GENERAL INTRODUCTION

There are two objectives of this chapter. First, to review the existing literature on inequality, polarization and contests, which have drawn considerable attention of researchers over several decades of yore. Second, to delineate the plan of the thesis.

Given the distribution of income in a population, an index of income inequality measures interpersonal differences in income. In case the population comprises a number of subgroups, a polarization index captures the extent of identification among individuals belonging to the same subgroup and alienation between individuals in different subgroups. It has been observed that income inequality plays a determinant role in the studies of development, poverty, social outcomes and public finance while polarization throws light on the evolution of the distribution of income, economic growth and social conflicts.

A contest is a non-cooperative game with at least two participants contending for a prize. The theory of contests successfully analyzes a variety of phenomena like rent seeking, electoral candidacy, sporting tournament and provision of public goods.

The chapter has been broadly classified into the following sections: (i) Properties of inequality indices: an overview (with specific thrust on subgroup decomposability), (ii) Research on polarization: the initial phase, (iii) Income bipolarization: indices and quasi-orderings, (iv) Income polarization: indices of multi-polar polarization, (v) Alternative measures of income polarization, (vi) Measurement of social polarization, (vii) Polarization in case of ordinal data and (viii) Contest success functions: a synoptic account.

1.1 Subgroup decomposable inequality indices

1.1.1 Properties of inequality indices : an overview

In Chapters 2 and 3, we deal with subgroup-decomposable inequality indices. To begin with, here we summarise some important properties (including 'subgroup-decomposability') of inequality indices from existing literature.

Consider a population of size n. Let x_i denote the income of the i^{th} individual, assumed to be drawn from the non-degenerate interval $[\upsilon, \infty)$ in the positive part R^1_{++} of the real line R^1. The vector $x = (x_1, x_2, .., x_n)$ represents the distribution of income. For any i, $x_i \in [\upsilon, \infty)$ and so, $x \in D^n = [\upsilon, \infty)^n$, the n-fold Cartesian product of $[\upsilon, \infty)$. To allow variability of population size, we consider $D = \bigcup_{n \in N} D^n$, N being the set of natural numbers. For all $n \in N$, for all $x = (x_1, x_2, ..., x_n) \in D^n$, $\sum_{i=1}^{n} (x_i / n)$, the mean of x, is denoted by $\lambda(x)$ (or simply by λ). For all $n \in N$, 1^n denotes the n-coordinated vector of ones. The non-negative (resp. positive) orthant of the n-dimensional Euclidean space R^n is denoted by R^n_+ (resp. R^n_{++}). Unless specified otherwise, we take the domain of an inequality index to be D. In other words, an inequality index is a function $I : D \to R^1_+$.

Mentioned below are some standard properties that an Inequality index is expected to obey.

Symmetry (SYM): For an arbitrary $n \in N$, if $x \in D^n$, then $I(x) = I(y)$, for any permutation y of x.

Principle of Transfers (POT): For an arbitrary $n \in N$ and $x \in D^n$, suppose that y is obtained from x by the following transformation

$$y_i = x_i + c \le y_j,$$

$$y_j = x_j - c,$$

$$\text{and} \qquad y_k = x_k \text{ for all } k \ne i, j, \tag{1.1}$$

where $c > 0$. Then $I(y) < I(x)$.

This principle is also known as the Pigou-Dalton transfers principle, named after Pigou (1912) and Dalton (1920).

Principle of Population (POP): For all $n \in N$, $x \in D^n$, $I(x) = I(y)$, where $y = (x^1, x^2, ..., x^l)$, each $x^i = x$ and $l \ge 2$ is arbitrary.

Normalization (NOM): For all $n \in N$, $I\left(c1^n\right) = 0$ for all $c > 0$.

Non-negativity (NON): For all $n \in N$, $x \in D^n$, $I(x) = 0$ if and only if $x = c1^n$ for some $c > 0$.

According to **SYM**, a condition of anonymity, I remains invariant under reordering of all incomes. Thus, **SYM** implies that any characteristic other than income has no relevance in the measurement of inequality. **POT** (popularly known as the Pigou-Dalton principle; also referred to as strict Schur-concavity) says that a transfer of income from a rich person *j* to a poor person i that does not change their relative positions reduces inequality while the reverse happens in case of a transfer from a poor to a rich. According to **POP,** inequality remains unalteredunder replications of the population. Thus, **POP** plays significant role in cross population comparisons of inequality. **NOM** stipulates that inequality vanishes if there is perfect equality in the underlying distribution. Finally, **NON** imposes a typical restriction on **NOM**. It demands that the inequality index can never vanish unless there is perfect equality.

An inequality index can satisfy invariance of scale or of location. A scale invariant index is one which remains constant if all the incomes are multiplied by the same positive scalar quantity. Similarly, a translation invariant index remains unaltered if all the incomes are increased or decreased by the same amount.

To be precise, an inequality index $I_R : D \to R_+^1$ is a relative or scale invariant index if proportional changes in all incomes do not change inequality, that is, for all $n \in N$, $x \in D^n$,

$$I_R(cx) = I_R(x), \tag{1.2}$$

where $c > 0$ is any scalar. Similarly, an inequality index $I_A : D \to R_+^1$ is absolute or translation invariant if for all $n \in N$, $x \in D^n$,

$$I_A\left(x + c1^n\right) = I_A(x), \tag{1.3}$$

where c is a scalar such that $x + c1^n \in D^n$.

1.1.2 Subgroup decomposability of an inequality index: standard notions:

In the measurement of income inequality, one may be dealing with a population partitioned into a number of subgroups (determined by age, sex, ethnicity, geographical location etc.). Then a pertinent question is on the contribution of different subgroups. The subgroup-decomposability of an inequality index captures this point. Formally, an inequality index $I : D \to R$ is subgroup decomposable if for all $k \geq 2$ and for all $x^1, x^2, \ldots, x^k \in D$ we have,

$$I(x) = \sum_{i=1}^{k} \omega_i(\underline{n}, \underline{\lambda}) \; I\left(x^i\right) + I\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots\ldots, \lambda_k 1^{n_k}\right), \tag{1.4}$$

where n_i is the population size associated with the distribution x^i, $n = \sum_{i=1}^{k} n_i$, $\lambda_i = \lambda(x^i)$ =mean of the distribution x^i, $\underline{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, $\underline{n} = (n_1, n_2, \ldots, n_k)$, $\omega_i(\underline{n}, \underline{\lambda})$ is the positive weight attached to inequality in x^i, assumed to depend on the vectors $\underline{n}$ and $\underline{\lambda}$ and $x = \left(x^1, x^2, \ldots, x^k\right)$. The first term, which is a weighted average of subgroup inequalities, is called the within group inequality (WI) while the second term is called the between group inequality (BI).

The notion of subgroup-decomposability of an inequality index came to the fore in the eighties of the last century. Shorrocks (1980) identified the class of all symmetric, scale-invariant, twice continuously differentiable and subgroup decomposable inequality indices satisfying **POP** and **NON** as follows:

$$I_c(x) = \begin{cases} \dfrac{1}{nc(c-1)} \sum_{i=1}^{n} \left[\left(\dfrac{x_i}{\lambda} \right)^c - 1 \right], c \neq 0,1, \\ \dfrac{1}{n} \sum_{i=1}^{n} \log \dfrac{\lambda}{x_i}, c = 0, \\ \dfrac{1}{n} \sum_{i=1}^{n} \dfrac{x_i}{\lambda} \log \dfrac{x_i}{\lambda}, c = 1. \end{cases} \tag{1.5}$$

The family I_c is popularly known as the generalized entropy family. If $c = 0$, I_c coincides with the Theil (1972) mean logarithmic deviation I_{ML} defined by

$$I_{ML}(x) = \frac{1}{n}\sum_{i=1}^{n} \ln \frac{\lambda}{x_i} \tag{1.6}$$

For $c = 1$, I_c becomes the Theil (1967) entropy index of inequality I_{TE} given by

$$I_{TE}(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{\lambda} \ln \frac{x_i}{\lambda} \tag{1.7}$$

For $c = 2$, I_c becomes half the squared coefficient of variation.

Shorrocks (1984) introduced the notion of '(aggregative) decomposability' of an inequality measure. According to this definition, an inequality measure $I : D \to R$ is decomposable if there exists a function A such that for all $x, y \in D$,

$$I(x, y) = A(I(x), I(y), \lambda(x), \lambda(y), n(x), n(y)) \quad , \tag{1.8}$$

where A is continuous and strictly increasing in its first two arguments.

Hence weakening the assumption of continuous differentiability by continuity, Shorrocks (1984) established that an inequality measure $I : D \to R$ is scale invariant and decomposable if and only if there exists a continuous strictly increasing function $G_1 : D \to R_1^+$ such that $G_1(0) = 0$ and $G_1(I) = I_c$ for some member I_c of the family of generalized entropy indices.

Replacing the arithmetic mean by the 'generalized' or the 'q^{th} order mean' (in (1.4)) as the representative income of a population group, Foster and Sheneyrov (1999) introduced the notion of the 'general additive decomposability'. The q^{th} order mean is defined by

$$\lambda_q(x) = \begin{cases} \left(\sum_{i=1}^{n} \frac{1}{n} x_i^q\right)^{\frac{1}{q}}, q \neq 0 \\ \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}, q = 0 \end{cases} . \tag{1.9}$$

Hence, it has been shown that a relative inequality measure satisfies the general additive decomposability if and only if it is a positive multiple of $I_{c,q}$, where

$$I_{c,q}(x) = \begin{cases} \frac{1}{c(c-q)}\left[\left(\frac{\lambda_c(x)}{\lambda_q(x)}\right)^c - 1\right], & c \neq 0, q \neq c \\ \frac{1}{q}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\lambda_q(x)}\right)^q \log\frac{x_i}{\lambda_q(x)}, & c \neq 0, q = c \\ \frac{1}{q}\ln\frac{\lambda_q(x)}{g(x)}, & c = 0, q \neq c \\ \frac{1}{2}V_L(x), & c = 0, q = c \end{cases} \tag{1.10}$$

where $V_L(x)$ stands for the variance of logarithms of x_i's and $g(x) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$ is the geometric mean of *x*. The class of generalized entropy measures is obtained by fixing $q = 1$ among which Theil mean logarithmic index and the Theil entropy index are obtained by substituting $c = 0$ and $c = 1$ respectively. $I_{c,q}$ satisfies Pigou-Dalton transfers principle if and only if $c \geq 1, q \leq 1$ or $c \leq 1, q \geq 1$. It is also worth noting here that the axiomatic derivation of $I_{c,q}$ does not require any regularity assumption on the functional form of the inequality measure. Instead, it makes use of a typical form of a transfer principle over two-person distributions.

Blackorby, Donaldson and Auersperg (1981) described a welfare-based approach for the measurement of intragroup and intergroup inequality. They considered ethical indices of inequality (viz. the Atkinson-Kolm-Sen (AKS) index defined by

$$I^k(x) = 1 - \frac{x_e}{\lambda(x)} \tag{1.11}$$

x_e being the "equally distributed equivalent" income (EDEI) of distribution x defined by

$$W(x) = W(x_e 1^n) \tag{1.12}$$

where $W : D \to R$ is any continuous, increasing and strictly S-concave social welfare function and $\lambda(x)$ the mean of *x*. The intergroup inequality has been defined as the inequality of the distribution where everyone enjoys EDEI of one's subgroup. The intragroup inequality is given by the percentage saved in moving from the original distribution to the distribution generated by EDEI's. The overall inequality can then be decomposed as

$$I(y) = I_A(y) + I_R(y) - I_A(y) I_R(y), \tag{1.13}$$

where $I_A(y)$ and $I_R(y)$ denote the intragroup and the intergroup inequality respectively.

Adopting a normative approach for measurement of inequality, Blackorby, Bossert and Donaldson (1999) demonstrated a means of decomposition of the AKS index into the between-group and within-group components.

Chakravarty and Tyagarupananda (1998) observed that the absolute counterpart of the family I_c, that is, the class of subgroup decomposable indices that remains invariant under equal translation of all incomes comprises

$$I_\theta(x) = \frac{1}{n} \sum_{i=1}^{n} \left[e^{\theta(x_i - \lambda)} - 1 \right],\ \theta \neq 0, \tag{1.14}$$

and

$$I_V(x) = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \lambda^2. \tag{1.15}$$

Bossert and Pfingsten (1990) argued that a natural generalization of the absolute and relative invariance of inequality indices is the intermediate invariance. To be specific, an inequality index $I_\mu : D \to R$ is intermediate μ-invariant $(0 \leq \mu \leq 1)$ if

$$I_\mu\left(x + c(\mu x + (1-\mu))1^n\right) = I_\mu(x) \tag{1.16}$$

for and all $x \in D^n$ and $n \in N$.

Chakravarty and Tyagarupananda (2009) classified all symmetric, intermediate μ-invariant, twice continuously differentiable and subgroup decomposable inequality indices satisfying Population Principle and normalization. The identified class is given by:

$$
I_{r,\mu}(x) = \begin{cases} \dfrac{1}{n(r,\mu)((r,\mu)-1)} \displaystyle\sum_{i=1}^{n} \left[\left(\dfrac{x_i + t}{\lambda + t} \right)^{(r,\mu)} \right], (r,\mu) \neq (0,1) \\ \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} log \dfrac{\lambda + t}{x_i + t}, (r,\mu) = 0 \\ \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \dfrac{x_i + t}{\lambda + t} log \dfrac{x_i + t}{\lambda + t}, (r,\mu) = 1 \end{cases}, \tag{1.17}
$$

where $(r,\mu) = r(1+\mu)/\mu$ and $t = \mu/(1-\mu)$, $0 < \mu < 1$. Clearly, this class contains the generalized entropy family as a special case.

Chakravarty (2001) pointed out that variance is the only proper subgroup-decomposable index satisfying absolute invariance, symmetry, twice continuous differentiability, population replication invariance and normalization.

1.1.3 Alternative notions of subgroup decomposability:

Bosmans and Cowell (2010) adopted the definition of (aggregative) decomposability (also called ‘absolute decomposability’) of an inequality index (introduced in Shorrocks (1984)) and demonstrated that an inequality measure $I : D \to R$ satisfies **SYM, POT, POP,** absolute decomposability and translation invariance if and only if a continuous and strictly increasing transform (vanishing at 0) of I is I_V or I_θ (as defined in Chakravarty and Tyagarupananda (1998)) for some θ.

Subramanian (2011) defined proper subgroup-decomposability of an inequality index by demanding that the weight $\omega_i(\underline{n}, \underline{\lambda})$ in definition (1.2) depends only on the subgroup population-share n_i/n. Introducing the notion of ‘level sensitivity’, the author showed that (a) there exists no properly decomposable inequality index satisfying ‘level sensitivity’ and (b) there exists a relative inequality index satisfying both subgroup decomposability and ‘level sensitivity’.

In a remarkable contribution, Zheng (2007) contended that the notion of invariance (scale/ translation/ intermediate) depends on the value judgment of the policy maker. To get rid of this problem, he suggested the use of ‘unit consistency’ condition (a natural generalization of 'scale invariance') which demands that for any two distributions $x, y \in D$ if $I(x) < I(y)$, then $I(\kappa x) < I(\kappa y)$ for all $\kappa \in R^1_{++}$.

To illustrate the notion of unit consistency, suppose there are two income distributions and an inequality measure demands that the first distribution has greater inequality than the second when all the incomes are measured in Indian Rupees. Now, unit consistency of the inequality index demands that relative ranking of the two distributions should remain unaltered even when the money unit is changed to Dollars. Clearly, 'unit consistency' is an ordinal property.

Zheng (2007) then characterized all unit-consistent, symmetric, differentiable, population replication invariant, strict Schur concave subgroup-decomposable inequality measures that satisfy normalization. The resulting index is a positive multiple of

$$I_{\alpha,\beta}(x) = \begin{cases} \dfrac{1}{n\alpha(\alpha-1)(\lambda(x))^{\beta}} \sum_{i=1}^{n} \left[x_i^{\alpha} - (\lambda(x))^{\alpha} \right], \alpha \neq 0,1, \\ \dfrac{1}{n(\lambda(x))^{\beta-1}} \sum_{i=1}^{n} \dfrac{x_i}{\lambda(x)} \log \dfrac{x_i}{\lambda(x)}, (\alpha = 0), \\ \dfrac{1}{n(\lambda(x))^{\beta}} \sum_{i=1}^{n} \log \dfrac{\lambda(x)}{x_i}, (\alpha = 1).. \end{cases} \tag{1.18}$$

for $\alpha, \beta \in R$.

Clearly, this family is a two parameter extension of the generalized entropy family (that is, if $\alpha = \beta = c$, then $I_{\alpha,\beta}$ coincides with I_c defined in (1.5)). The extended family includes the following generalization of the intermediate measure suggested in Krtscha (1994):

$$I(x) = \frac{1}{n(\lambda(x))^{\beta}} \sum_{i=1}^{n} \left[x_i^2 - (\lambda(x))^2 \right], \tag{1.19}$$

where $0 < \beta < 2$. However, it does not include the decomposable intermediate inequality measures characterized in Chakravarty and Tyagarupananda (2009).

Zheng (2005) further observed that while characterizing (1.18), one can replace the assumption of differentiablity of the inequality index by a weaker hypothesis viz. continuity.

Ebert (2010) put forward the notion of ‘Weak decomposability’ of inequality indices.He defined an inequality index to be weakly decomposable if for all $\boldsymbol{n} = (n^1, n^2)$ (where $n^1, n^2 \in N$) there exist strictly positive weighting functions $\alpha^1(\boldsymbol{n}), \alpha^2(\boldsymbol{n})$ and $\beta(\boldsymbol{n})$ such that

$$I(x^1, x^2) = \alpha^1(\boldsymbol{n}) I(x^1) + \alpha^2(\boldsymbol{n}) I(x^2) + \beta(\boldsymbol{n}) \sum_{i=1}^{n^1} \sum_{j=1}^{n^2} I(x_i^1, x_j^2) \tag{1.20}$$

for all $x^1 = (x_1^1, x_2^1, \dots, x_{n^1}^1) \epsilon R_{++}^{n^1}$ and $x^2 = (x_1^2, x_2^2, \dots, x_{n^2}^2) \epsilon R_{++}^{n^2}$.

The sum of the first two quantities on the right hand side may be regarded as the within group inequality while the third term, dependent on the sum of inequality between different pairs, refers to the between group term. Thus, for calculating ‘between group inequality’ according to this definition of subgroup-decomposability we don’t require to consider the smoothed distribution comprising of subgroup means.

Ebert (2010) next classified all weakly decomposable inequality indices satisfying population replication invariance and a normalization condition. The resulting expression for I satisfies

$$I(x) = \frac{2}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} I(x_i, x_j) \tag{1.21}$$

for all $x \in R_{++}^n$ and $n \geq 2$.

1.2 Research on polarization: the initial phase:

Last two decades have witnessed a tremendous surge of interest in the study of polarization. Research on inequality has been supplemented by the theory of measurementand orderings in terms of polarization. Loosely speaking, polarization refers to clustering of individuals/attributes around local poles or subgroups in a distribution, It has been observed that polarization can reasonably explain a number of social phenomena such as income distribution evolution, economic growthand social conflicts.

The initial phase of the study of polarization dates back to the eighties of the last century with the pursuit on the 'disappearing middle class'. Lester Thurow (1984) and Blackburn and Bloom (1985) observed that during the period 1967-1983, the income distribution in the United States was getting polarized in the two extremes in the sense that the percentage of income in the 'middle' income-range was coming down. The question was pertinent because it was generally believed that the existence of a thriving middle class helps in the growth of 'a healthy political democracy' and an opulent market for domestic goods and services. By the 'middle' income-group they meant those with incomes between (a) 75% and 125% and (b) 60% and 225% respectively of the median income. However, pursuing a similar query, Levy (1987a, 1987b) defined the 'middle' as the middlemost 3/5 of the population and concluded that the proportion of the 'middle' remained more or less constant. Wolfson (1994) cited a typical example to show that for two income distributions it is quite probable that the income share of the middle third of the first distribution is lower than that of the second while the ranking is opposite if we consider the middle two-thirds of the two distributions. Thus the relative ranking of two income distributions on the basis of the 'disappearing middle' may very well depend on how the 'middle' is defined.

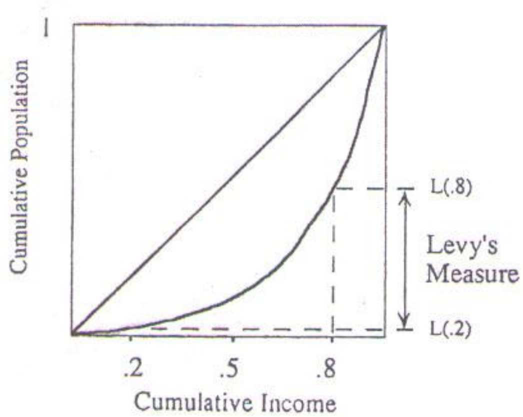

Fig. 1.1(Levy's bipolarization index)[1]

Levy (1987a, 1987b) defined the 'middle' in the 'people space' instead of the 'income space'. His definition of the 'middle class' consists of all the p-percentiles, where $20 \leq p \leq 80$. Thus, Levy's measure of the 'middle' is: $L(0.2) - L(0.8)$, where $L(p)$ denotes the Lorenz-ordinate at cumulative proportion p (which equals the share of the total income possessed by the cumulative p proportion of the population). However, this index has been criticized by Foster and Wolfson (1992, 2010) for its inability to measure the 'spread' on each side of the 'middle'.

[1] Figure No. 1.1-1.3 are borrowed from Foster and Wolfson (2010).

For example, Levy's index remains constant for all symmetric distributions; the dispersion away from the median is immaterial.

1.3 Income bipolarization: indices and quasi-orderings:

Addressing the above shortcoming of Levy's measure, Foster and Wolfson (1992, 2010) suggested a methodology which allows us to vary the cut-off points of the 'middle' income range. They introduced the notions of 'increased spread' (**IS**) and 'increased bipolarity' (**IB**) on the basis of which they described how to draw polarization curves that determine whether one distribution is unambiguously more or less polarized than another. The first property (**IS**) requires that polarization should increase via a median$\big(m(x)\big)$-preserving transfer that reduces income(s) below $m(x)$ and increases income(s) above $m(x)$ while the second (**IB**) demands that polarization should go up via a number of progressive transfers taking place on either side of the median.

To construct an index of polarization, Foster and Wolfson (1992, 2010) and Wolfson (1994) defined two types of polarization curves. The first degree polarization curve $S_F(q)$ of a distribution F is defined by the distance between the median and the income of the person at the q^{th} percentile. Thus, one distribution has an unequivocally smaller middle class if and only if the first degree polarization curve of the former lies everywhere higher than that of the latter.

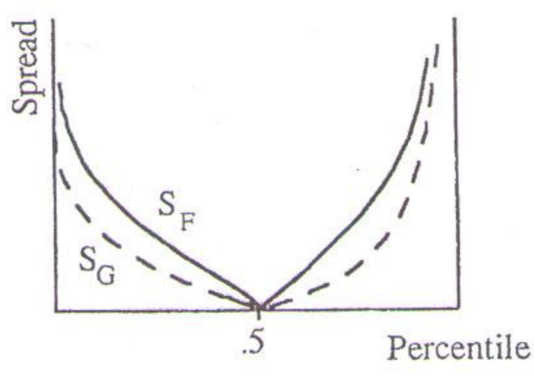

Fig. 1.2 (1st degree polarization curve)

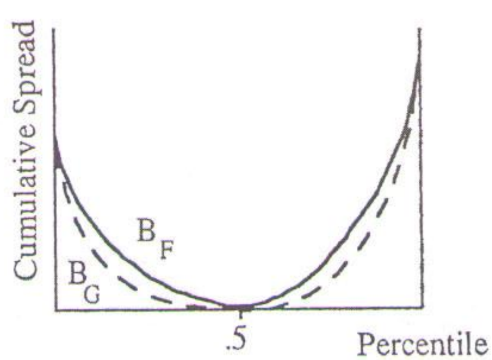

Fig. 1.3 (2nd degree polarization curve)

The second degree polarization curve $B_F(q)$ of a distribution F is given by the area under the first degree polarization curve between 0.5 and q.

Wolfson (1994) suggested three distinct indices for measuring bipolarization. The first one is the median share (*mshare*, for short) given by the income share of the bottom half of the population. The second statistic is m/λ, the ratio of the median to the mean income, which is also defined as the 'median tangent' (*mtan*). The third one defined by

$$P = 2(T - G) / mtan, \tag{1.22}$$

where T =twice the area of the trapezoid defined by the 45-degree line and the median-tangent (= the vertical distance between the Lorenz curve and the 45-degree line at the 50$^{\text{th}}$ percentile) and G is the Gini index of inequality. It has been shown that

$$T = \left(\lambda^{U} - \lambda^{L}\right) / \lambda \tag{1.23}$$

with λ^U= mean of those above the median, λ^L= mean of those below the median and λ the overall mean.

Wolfson (1994) measures are relative bipolarization indices in the sense that they remain invariant under equal relative changes in all incomes,

The third measure is identical with the one suggested by Foster and Wolfson (1992, 2010)

$$P_{FW} = 2\int_0^1 B_F(q)\, dq, \tag{1.24}$$

which is simply twice the area below the second degree polarization curve. It has been shown that the measure P_{FW} satisfies both 'increased spread' and 'increased bipolarity'.

Foster and Wolfson (1992, 2010) established that

$$P_{FW} = \left(BI^{G} - WI^{G}\right)\frac{\lambda}{m}, \tag{1.25}$$

where BI^G and WI^G are the between-group and within-group inequalities of the Gini coefficient, where the population has been assumed to be divided into two subgroups: one below the median and the other above it. This clearly underscores the essential difference between polarization and inequality.

Foster and Wolfson (1992, 2010), were pioneers in developing a rigorous definition of bipolarization ordering. They introduced the notion of the relative bipolarization curve (RBC). For any $x \in D^n$, let $x^0 = \left(x_{(1)}, x_{(2)}, \ldots, x_{(n)}\right)$ be the non-decreasingly ordered permutation of x. We assume for simplicity that n is odd. Then $m(x) = x_{(\bar{n})}$, where $\bar{n} = \left(\frac{n+1}{2}\right)$. Let x_-^0 $\left(x_+^0\right)$ be the subvector of x^0 such that $x_{(i)} < m(x)$ ($x_{(i)} > m(x)$). The normalized aggregate shortfall $RB\left(x, \frac{k}{n}\right) = \frac{1}{nm(x)} \sum_{k \le i < \bar{n}} \left(m(x) - x_{(i)}\right)$ is the deviation of the total income of the population proportion k/n from the corresponding total that it would possess under the hypothetical case

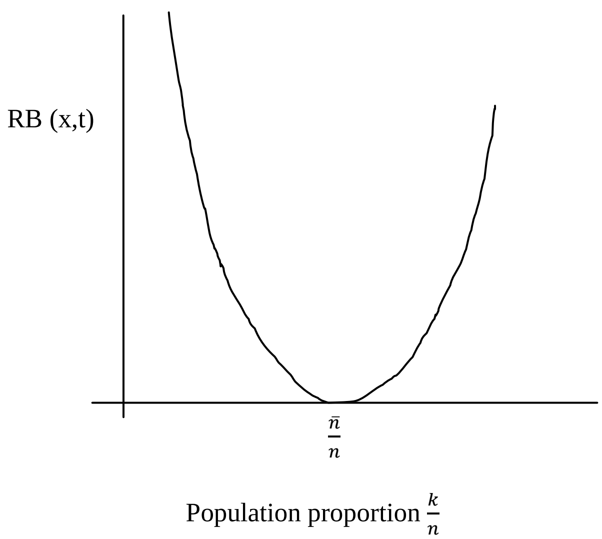

Fig. 1.4 (Relative Bipolarization Curve)[2]

where everybody enjoys the median income, as a fraction of the factor nm, where $1 \le k < \bar{n}$. This is the ordinate of the relative bipolarization curve (RBC) of x, corresponding to the population proportion k/n, where $1 \le k < \bar{n}$. For incomes not less than the median, the corresponding ordinate is $\frac{1}{nm(x)} \sum_{\bar{n} \le i \le k} \left(x_{(i)} - m(x)\right)$. A similar construction of the curve runs when the population size is even. (See Wolfson, 1997, 1999, Wang and Tsui, 2000, Chakravarty, 2009). It is shown

[2] This figure is borrowed from Chakravarty (2015).

that of two distributions $x, y \in D^n$, the RBC of y dominates that of x, that is, the RBC of y is nowhere below that of x and at some places strictly inside if and only if y is more polarized than x by all relative, symmetric bipolarization indices that satisfy **IS** and **IB** (see Wolfson, 1997, 1999, Chakravarty et al., 2007 , Chakravarty, 2009 and Foster and Wolfson , 2010).

Extending the works of Foster and Wolfson (1992, 2010), Wang and Tsui (2000) characterized a generalized family of bipolarization indices. They also established various equivalent forms of a slightly modified version of the partial ordering suggested by Foster and Wolfson (1992, 2010). In particular, the modified version of Foster-Wolfson (partial) ordering is defined as follows: for any two vectors *x* and *y* having the same dimension and a common median (*m*), we say that *y* is more polarized than *x* if the sum of the absolute deviations from the median of all the observations below the median as well as those above the median are not greater than the corresponding quantities for the vector *y* and if strict inequality holds in at least one case. Formally,

$$\sum_{k \le i < \bar{n}} \left| x_{(i)} - m \right| \le \sum_{k \le i < \bar{n}} \left| y_{(i)} - m \right|, 1 \le k < \bar{n} \tag{1.26}$$

and

$$\sum_{\bar{n} \le i < k} \left| x_{(i)} - m \right| \le \sum_{\bar{n} \le i < k} \left| y_{(i)} - m \right|, \ \bar{n} < k \le n \ . \tag{1.27}$$

Wang and Tsui (2000) showed that (i) given two such vectors *x* and *y*, *y* is not less polarized than *x* if and only if there exists a vector *z* with the same dimension and the same median such that *y* has a greater spread than *z* and *z* has a greater bipolarity than *x* and that (ii) for a given median, this is equivalent to the condition that $y_-^0 \le x_-^0 B$ and $x_+^0 C \le y_+^0$ for all bistochastic matrices B, C of appropriate orders, and $y_-^0 \ne x_-^0$ and/or $y_+^0 \ne x_+^0$. (For any two n-coordinated vectors p and q, $p \le q$ means that $p_i \le q_i$ for all $1 \le i \le n$. An $n \times n$ nonnegative matrix is called a bistochastic matrix of order n if each of its rows and columns sums to 1.)

For determining a generalized class of Foster-Wolfson family of bipolarization indices, Wang and Tsui (2000) assumed two particular additive structures at the outset and then imposed the **IS** and **IB** axioms. Finally, formulating a normalization axiom, they characterized an index which is a weighted sum of the percentage deviations from the median income:

$$P^n(x) = \sum_{i=1}^{n} b_i^N \frac{m(x) - x_{(i)}}{m(x)} \quad (1.28)$$

where $0 < b_{i-1}^n < b_i^n$ for all $i < \bar{n}$; $b_i^n > b_{i+1}^n > 0$ for all $i < \bar{n}$; $b_{\bar{n}}^n = \sum_{i=1}^{\bar{n}-1} b_i^n - \sum_{i=\bar{n}-1}^{n} b_i^n$ if n is odd and $m(x)$ is the median of x.

Another significant contribution of Wang and Tsui (2000) is a new family of polarization indices which are expressible as weighted sums of the r^{th} powers of the absolute deviations from the median:

$$P_1^n(x) = \frac{\theta_1}{n} \sum_{i=1}^{n} \left| x_{(i)} - m(x) \right|^r, \quad (1.29)$$

and $$P_2^n(x) = \frac{\theta_2}{n} \sum_{i=1}^{n} \left| \frac{x_{(i)} - m(x)}{m(x)} \right|^r. \quad (1.30)$$

where r is a positive fraction and θ_1, θ_2 are positive constants. The characterization-exercise involves a postulate on scale compatibility.

Rodriguez and Salas (2003) observed that the Foster-Wolfson index of bipolarization, computed for two groups separated by the mean (instead of the median) transforms to

$$P_\lambda = \left(BI^G - WI^G \right). \quad (1.31)$$

The authors then defined the extended Wolfson bipolarization measure with inequality-aversion parameter v as

$$P_{RS}(v) = BI^G(v) - WI^G(v), \quad (1.32)$$

where, as before, BI^G and WI^G are respectively the between-groupand the within group component associated with Donaldson and Weymark (1980, 1983) welfare-ranked S-Gini inequality index(with parameter v) given by

$$G(v) = v(v-1)\int_0^1 (1-q)^{\upsilon-2}\left[q - L(q)\right]dq \tag{1.33}$$

(For $v = 2$, (1.33) yields the well-known Gini index of inequality.) It has then been shown that this bi-polarization measure is consistent with the second degree polarization curve (that is, it satisfies increased spread and increased bipolarity) if $v \in [2,3]$.

Adopting welfare-based approach, Chakravarty and Majumder (2001) suggested a family of bipolarization indices. The first one is based on the Atkinson-Kolm-Sen (AKS) ethical inequality index and is given by

$$Q^n(x) = \frac{\lambda(x_+)\left(1 - I^k(x_+)\right) + 2\lambda(x_+)}{2m(x)} + \frac{\lambda(x_-)\left(1 - I^k(x_-)\right) - B\left(m(x)\right)\lambda(x_-)}{2m(x)}, \tag{1.34}$$

where $m(x)$, as before, is the median of x, x_+ (resp. x_-) comprises all x_i 's where $i < m(x)$ (resp. $i > m(x)$), $k = n/2$, if n is even and $k = (n-1)/2$, if n be odd; $B(m(x)) = (m(x)/\theta)^{1-r}$ and $H(m(x)) = \frac{1}{2}(m(x)/\theta)^{1-r} - 2$ and $I^k(x)$ denotes the Atkinson-Kolm-Sen (AKS) index of inequality defined in (1.10).

Dealing with absolute polarization indices (which remain invariant under equal absolute changes in all incomes), Chakravarty et al. (2007) introduced the absolute polarization curve (APC). This curve is obtained from the Foster-Wolfson bipolarization curve on multiplication by the median. Thus, the APC measures for any cumulative population proportion the absolute difference of its total income, expressed as a fraction of the total population size, from the corresponding income that it would enjoy under the hypothetical distribution where everyone has the median income.

Formally,

$$FW_A(x;k) = \begin{cases} \frac{1}{n}\sum_{k \le i \le \overline{n}} \left(m(x) - x_i\right) if \;\; 1 \le k \le \overline{n} \\ \frac{1}{n}\sum_{\overline{n} \le i \le k} \left(x_i - m(x)\right) if \;\; \overline{n} \le k \le n \;. \end{cases} \tag{1.35}$$

The area under the APC is an absolute measure of polarization. Chakravarty et al. (2007) demonstrated that for two income distributions *x* and *y*, the APC of *x* lies below that of *y* if and only if polarization of *x* is not less than that of *y* for all absolute, symmetric, population-replication invariant polarization indices satisfying **IS** and **IB**.

Undoubtedly, **IS** and **IB** are the two cardinal principles of bipolarization. (Amiel et. al. (2010) studied the acceptance of these two axioms among the researchers through questionnaires.) Bossert and Schworm (2008) presented a complete characterization of the class of all quasi-orderings satisfying **IS** and **IB** and anonymity (which demands that polarization should remain invariant under any permutation of the arguments).

To be specific, a quasi-ordering $\succcurlyeq_P$ satisfies the three axioms if and only if $\succcurlyeq_P$ includes $\succcurlyeq_0$, where two income distributions $x, y \in R^n_{++}$ (where *n* is assumed to be even) satisfy $x \succcurlyeq_0 y$ if and only if they have the same median and $-y^L$ and y^H generalized Lorenz dominates $-x^L$ and x^H respectively. (Here $x^L = \left(x_{[1]}, \ldots .., x_{[n/2]}\right)$ and $x^H = \left(x_{[n/2+1]}, \ldots .., x_{[n]}\right)$ respectively denote the rank-ordered subvector of *x*consisting of low-income and high-income individuals.)

As a corollary it follows that the quasi-ordering $\succcurlyeq_0$ is indeed the quasi-ordering defined by Foster and Wolfson (1992, 2010).

To identify a family of polarization measures, Bossert and Schworm (2008) stated the following 'Independence' axiom:

For all $x, y \in R^n_{++}$ with the same median,

$$P\left(y^L, y^H\right) \geq P\left(x^L, y^H\right) \Leftrightarrow P\left(y^L, x^H\right) \geq P\left(x^L, x^H\right) \tag{1.36}$$

and

$$P\left(y^L, y^H\right) \geq P\left(y^L, x^H\right) \Leftrightarrow P\left(x^L, y^H\right) \geq P\left(x^L, x^H\right). \tag{1.37}$$

The axiom demands that for any fixed value of the median, x^L and x^H are strictly separable from each other.

The relevant characterization theorem states that a polarization measure P satisfies Anonymity, **IS**, **IB** and Independence if and only if for all $M \in R^{1}_{++}$ and for all $x \in R^{n}_{++}$ with $m(x) = M$ we have,

$$P_M\left(x^L, x^H\right) = \Phi_M\left(\phi_M^L\left(x^L\right), \phi_M^H\left(x^H\right)\right) \tag{1.38}$$

where $\Phi_M : R^2 \to R$ is increasing, $\phi_M^L : R^{n/2}_{++} \to R$ is non-increasing and S-concave and $\phi_M^H : R^{n/2}_{++} \to R$ is non-decreasing and S-concave.

The paper concludes with the observation that there is no straightforward generalization of the partial ordering in a multi-group context.

Deustch et al. (2007) defined an index of flatness of an income distribution as follows:

$$P_G = \left(G^B - G^W\right) / G \ . \tag{1.39}$$

Then they showed that P_G can as well be regarded as an index of bipolarization since it satisfies **IS** and **IB**. This in turn establishes a clear link between an index of kurtosis and one of bipolarization.

Duclos and Echevin (2005) introduced the first order bipolarization dominance. A bi-polarization index of an income distribution x, according to their definition, is a function of $d_x(i)$'s, where $d_x(i)$ is the 'proprtional spread' of individual i's income from median income, that is,

$$d_x(i) = \left|x_i - m_x\right| / m_x \ . \tag{1.40}$$

Then looking at the increasingly ordered transformation $d_x^*(i)$ of $d_x(i)$'s, they considered four axioms (homogeneity, population replication invariance, monotonicity and symmetry) based on which they define the first order bipolarization dominance. Given two distributions x and y, the dominance criterion demands that $d_x$is more polarized than d_y if and only if $d_x^*(i) \geq d_y^*(i)$ for all i. Further, defining

$$Q_x(\vartheta) = \frac{1}{n}\sum_{i=1}^{n} I\left(d_y(i) \geq \vartheta\right) \tag{1.41}$$

(where I denotes the indicator function), Duclos and Echevin (2005) define a second bipolarization dominance as follows: x is more bipolarized than y if and only if $Q_x(\vartheta) \geq Q_y(\vartheta)$ for all $\vartheta > 0$.

Chakravarty and D'Ambrosio (2010) developed the notion of intermediate invariance of bipolarization measures, based on which they define the intermediate polarization curve (IPC). This requires that the distributions x and $x + c\left(\mu x + (1-\mu)1^n\right)$ should have the same level of polarization for any positive constant$c$ and any $\mu \epsilon (0,1)$. The authors then identified two distinct classes of bipolarization indices. The first one is:

$$F_\mu(x) = \frac{2\left(\bar{x}_+ - \bar{x}_- - A_G(x)\right)}{\mu m(x) + 1 - \mu}, \tag{1.42}$$

where $A_G(x)$, the absolute Gini index of inequality for the distribution x is defined by $A_G(x) = I^G(x)\lambda(x)$.

The other one is:

$$C_\mu(x) = \frac{\left(n^{-1}\sum_{1\leq i\leq n}\left|x_i - m(x)\right|^r\right)^{\frac{1}{r}}}{\mu m(x) + 1 - \mu}, \tag{1.43}$$

where $0 < r \leq 1$.

Chakravarty and D'Ambrosio (2010) further established that for two income distributions x and y, the IPC of x lies below that of y if and only if polarization of x is not less than that of y for all intermediate, symmetric, population-replication invariant polarization indices satisfying **IS** and **IB**.

Lasso de la Vega et al. (2010) generalized the notion of 'intermediate invariance'. Introducing the notion of 'Unit Consistency' of bipolarization measures, they showed that the only unit consistent indices are related to the Krtscha-type notion of intermediateness (see Krtscha 1994).

1.4 Income polarization: indices of multipolar polarization:

In the previous section it was assumed a priori that the population under consideration consists of precisely two groups. Widening the framework of bipolarization, Esteban and Ray (1994) (*ER* (1994), henceforth) introduced polarization in a multi-group setting. The population has been assumed to be fragmented exogenously into different subgroups (with respect to religion, ethnicity, geographical region or some such social characteristic). The individuals belonging to the same subgroup possess a feeling of identification among them and a feeling of alienation against individuals in other subgroups. Polarization, thus, has two components viz. ‘identification’ and ‘alienation’, both being related increasingly to the former. By considering numerous examples, it has been shown that in such a situation, polarization is conceptually distinct from inequality. In particular, the notion of inequality is dependent on the ‘local’ properties such as Pigou-Dalton transfers principle whereas the idea of polarization relies on the shape of the entire income distribution and thus behaves ‘globally’.

To pin down an index of polarization, *ER* (1994) confined attention to the following quasi-additive structure:

$$P\left(\underline{\pi},\underline{y}\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}\pi_i\pi_j T\left(J\left(\pi_i\right),a\left(\delta\left(y_i,y_j\right)\right)\right), \qquad (1.44)$$

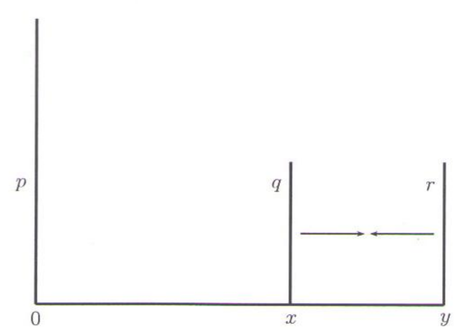

Fig. 1.5 (the 1st Axiom in ER (1994))[3]

where π_i and y_i respectively denote the population proportion and the representative income of the i^{th} subgroup, J is the identification function, a the alienation function and $T\left(J\left(p\right),a\left(\delta\left(y,y'\right)\right)\right)$ gives the effective antagonism felt by y towards y'. T has been assumed to be continuous in both the arguments and strictly increasing in the second argument.

[3] Figure No. 1.5 and 1.6 are borrowed from Esteban and Ray (1994).

For characterization of polarization indices, a number of axioms have then been invoked. The first one says that merger of two sufficiently small masses separated at sufficiently small distance at their midpoint augments polarization. The underlying idea is that lower is the dispersion inside the groups and higher is the homogeneity of groups' sizes, the greater is the polarization.

The next axiom translates the intuition that polarization increases with an increase in the heterogeneity of group-sizes. We begin with three masses *p*, *q* and *r* where $p > q$, $p > r$ and the intermediate point mass *q* is at least as close to *r* as it is to *p*. Then the axiom demands that a small shift of mass *q* that brings it closer to *r* augments polarization. The third axiom relies on the fact that the disappearance of a middle class into 'rich' and 'poor' categories results in an increment in polarization.

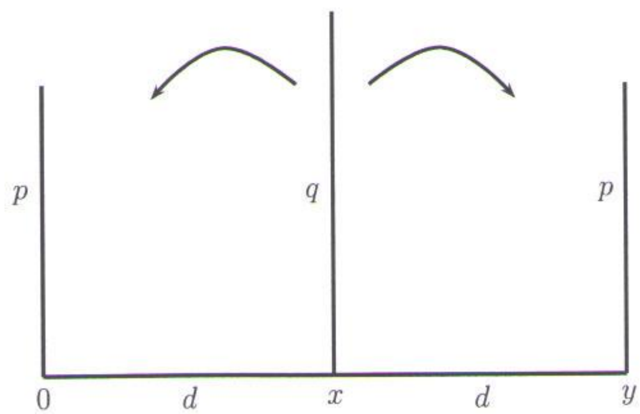

Fig. 1.6 (the 2nd Axiom in ER (1994))

The final postulate (Axiom 4) demands that given three masses *p*, *q* and *r* with $q > r$, *p* sufficiently small, polarization increases by a shift of population mass from *p* to *r*. The intuition is quite clear. The shift brings the sizes of the two dominant groups closer to one another, thereby enhancing polarization.

Employment of the first three axioms along with a condition of homotheticity (which requires invariance of polarization ordering under scalar multiplication of population-sizes) on a polarization index (of the form (1.41)) reduces it to a positive multiple of

$$ER = \sum_{i,j=1}^{n} \pi_i^{1+\alpha} \pi_j \left| y_i - y_j \right|, \tag{1.45}$$

where $\alpha > 0$ is a parameter (to be regarded as a polarization sensitivity parameter), $\alpha \grave{o} \left(0, \alpha^*\right]$ with $\alpha^* \approx 1.6$. In fact, the value of α can be restricted within a smaller domain $\left[1, \alpha^*\right]$ if we additionally impose Axiom 4. The gretater is the value of α, the greater is the divergence of *ER* from inequality. The value $\alpha = 0$ corresponds to the Gini index of inequality. The bimodal distribution (that is, a distribution with two modes) is a maximal element for the partial ordering generated by the class of polarization measures identified above in the sense that it is more polarized than any other income distribution.

To capture the intergroup distance in *ER* (1994), D'Ambrosio (2001) replaced the Euclidean distance by the Kolmogorov metric defined by

$$D(i,j) = \frac{1}{2}\int_0^\infty \left|f_i(v) - f_j(v)\right| dv \tag{1.46}$$

where f_i (resp. f_j) is the density of the i^{th} (resp. j^{th}) income group.

Notice that *ER* (1994) measures polarization in a situation where the underlying income distribution is discrete. However, in many practical problems, the income distribution may be continuous. Extending the *ER* (1994) measure to such cases, Duclos, Esteban and Ray (*DER*, henceforth) (2004) suggested an index that can be applied to distributions with density functions. The authors began with the following form of the polarization index:

$$P(F) = \iint T\left(f(x), |x-y|\right) f(x) f(y)\, dx dy\ . \tag{1.47}$$

To specify the functional form of *T*, *DER* (2004) stated a set of axioms. For this, the notion of *basic densities* has been introduced. An unnormalized, symmetric, unimodal density having a compact support is called a *basic density*. A *basic density* with mean 1 and support [0,2] is called a *root*. A basic density can undergo two types of transformations viz. the *slide* and the *squeeze*. A slide to the right (resp. left) by *x* is just a new density *g* defined by $g(y) = f(y-x)$ (resp. $g(y) = f(y+x)$). A λ-squeeze of f is a transformation f^λ defined by $f^\lambda(x) = \frac{1}{\lambda} f\left(\frac{x-[1-\lambda]\mu}{\lambda}\right)$. A squeeze, thus, implies a global compression of a density.

The statements of the axioms can be summarized as follows.

(i) Given a distribution comprising a single basic density, a squeeze of the density cannot increase polarization.

A squeeze of a basic density brings the density closer to its mean, thus making it more homogeneous. Consequently, polarization is expected to go down.

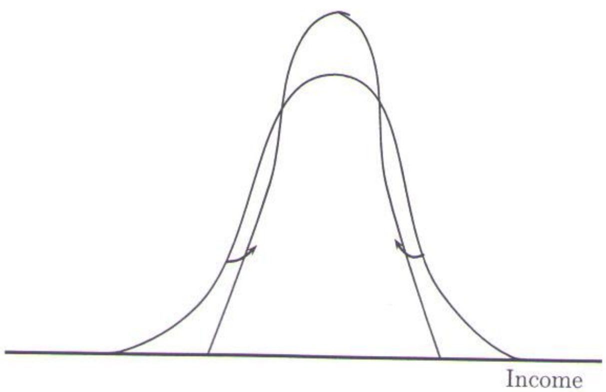

Fig. 1.6 (a 'squeeze' of a basic density)[4]

(ii) If in a symmetric distribution there are three basic densities with the same root and mutually disjoint supports, then a "double squeeze", that is, a symmetric squeeze of the two side densities cannot lower polarization.

In other words, a 'local' squeeze contributes to increment in identification and hence cannot reduce polarization.

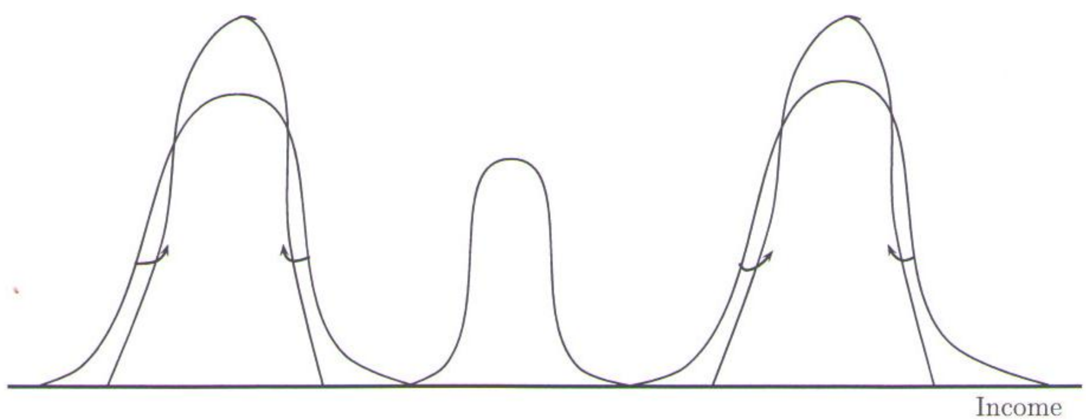

Fig. 1.7 (a 'double squeeze')

(iii) Consider a symmetric distribution comprising of four basic densities having the same

[4] Figure No. 1.6-1.8 are borrowed from Duclos, Esteban and Ray (2004).

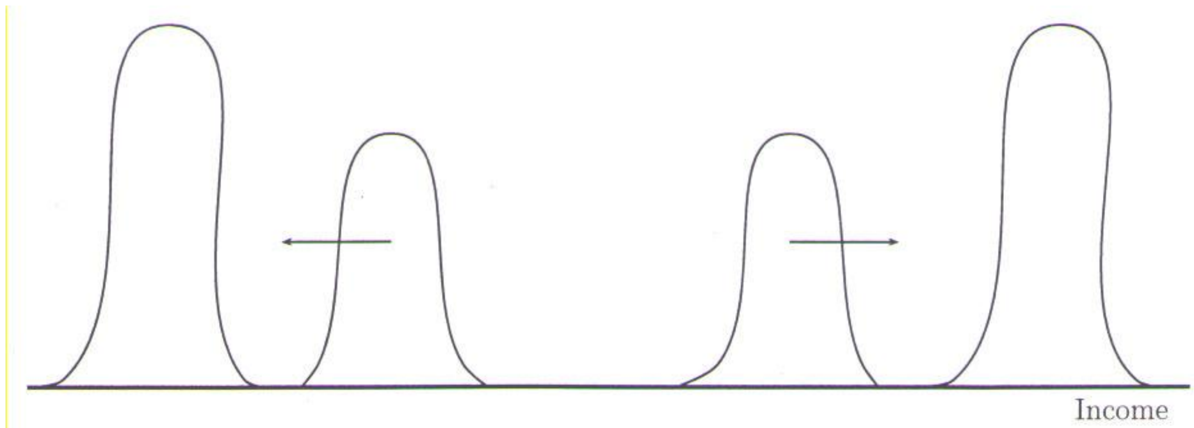

Fig. 1.8 (a 'slide of the two middle densities')

root and mutually disjoint supports. A 'slide' of the two middle densities must augment polarization.

The reason is evident: a twin 'slide' of the two inner densities brings the distribution closer to the bipolar case, thereby increasing the alienation component.

(iv) The final axiom is a homotheticity principle: the same ordering of polarization must be maintained if populations are multiplied by a scalar constant. This is the same as the Axiom 4 in *ER* (1994).

DER (2004) characterization theorem states that a polarization measure *P* (of the form (1.44)) satisfies the four axioms stated above if and only if it is a positive multiple of

$$P_{\alpha}(F) = \iint f(x)^{1+\tilde{\alpha}} f(y)\left|y - x\right| dy\,dx \tag{1.48}$$

where $\tilde{\alpha} \in [0.25, 1]$.

The parameter $\tilde{\alpha}$ may be regarded as a polarization sensitivity parameter. If $\tilde{\alpha} = 0$, then $P_{\tilde{\alpha}}$ equals the Gini coefficient of inequality. A greater value of $\tilde{\alpha}$ means a greater amount of identification and hence a greater polarization.

Next comes the problem of estimation and statistical inference. The suggested estimator for $P_{\tilde{\alpha}}(F)$ is:

$$P_{\tilde{\alpha}}\left(\hat{F}\right) = n^{-1}\sum_{i=1}^{n} \hat{f}\left(y_i\right)^{\tilde{\alpha}} \hat{a}\left(y_i\right) \tag{1.49}$$

where $\hat{a}\left(y_i\right) = \hat{\mu} + y_i\left\{n^{-1}\left(2i-1\right)-1\right\} - n^{-1}\left(2\sum_{j=1}^{i-1} y_j + y_i\right)$, $\hat{\mu}$ being the sample mean and $\hat{f}\left(y_i\right)^{\tilde{\alpha}}$ is estimated non-parametrically.

The asymptotic normality of $n^{0.5}\left(P_{\tilde{\alpha}}\left(\hat{F}\right) - P_{\tilde{\alpha}}\left(F\right)\right)$ has then been established.

Using Luxemberg Income Study (LIS) data, the authors finally show that the behaviors of inequality and polarization are entirely different.

Considering the Shapley decomposition procedure (introduced by Chantreuil and Trannoy (1999)) meant originally for inequality indices, Deustch and Silber (2010) showed how this could be applied for measuring the marginal impact of an income source on the overall polarization.

Other Modifications of ER (1994):

A considerable amount of information may be lost in clubbing individuals into different subgroups. Keeping this in mind, Esteban, Gradin and Ray (2007) suggested the following modification of the *ER*- index:

$$EGR\left(\underline{\pi},\underline{\lambda}\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\left|\lambda_i - \lambda_j\right|\pi_i^{1+\alpha}\pi_j - \kappa\, er\left(\underline{\pi},\underline{\lambda}\right), \tag{1.50}$$

where $\underline{\pi} = \left(\pi_1, \pi_2, \ldots, \pi_n\right)$ and $\underline{\lambda} = \left(\lambda_1, \lambda_2, \ldots, \lambda_n\right)$, π_i and λ_i being the population frequency and the mean of the income class $[x_{i-1}, x_i]$ for $i = 1, \ldots, n$ and the error term $er\left(\underline{\pi},\underline{\lambda}\right)$ corresponds to 'the implicit fuzziness of group identification' and $\kappa \geq 0$ is a constant. It is fairly interesting to note that when $n = 2$ and $\alpha = \kappa = 1$ we have,

$$EGR\left(\underline{\pi},\underline{\lambda}\right) = \frac{m}{2} P_{FW} \tag{1.51}$$

Maintaining the same spirit, Lasso de la Vega and Urrutia (2006) suggested the use of the following index:

$$LU(\underline{\pi},\underline{\lambda}),=\sum_{i=1}^{n}\sum_{j=1}^{n}\left|\lambda_i-\lambda_j\right|\pi_i^{1+\tilde{\alpha}}\pi_j\left(1-I_i^G\right)^{\tau}, \tag{1.52}$$

τ being a nonnegative constant representing the degree of sensitivity towards group cohesion and I_i^G the Gini index of group i. It is clear that the LU index is increasing in between-group inequality and decreasing in within group inequality.

1.5 Alternative measures of income polarization:

Alesina and Spolaore (1997) suggested an alternative measure of income polarization. This is given by the median distance from the median. Let F be the distribution with median m. Then the proposed measure P^{AS} is given implicitly by the formula

$$F\left(m+P^{AS}\right)-F\left(m-P^{AS}\right)=\frac{1}{2}. \tag{1.53}$$

In course of their empirical study on the regional variation of expenditure in rural and urban China, Kanbur and Zhang (2001) proposed an index of polarization, based on a subgroup-decomposable inequality index. Their index is given by the ratio between the between-group (BI) and within-group (WI) components of inequality (I):

$$P_{ZK}=BI/WI. \tag{1.54}$$

Here the between-group term can be taken as an indicator of alienation and the within-group component is inversely related to identification.

Employing Gini index in (1.54) we get

$$P_{ZK}^G=BI^G/WI^G. \tag{1.55}$$

The increasing transformation

$$P_{SDH}=\frac{P_{ZK}^G-1}{P_{ZK}^G+1} \tag{1.56}$$

yields the index of bipolarization suggested by Silber et al (2007). Obviously P_{SDH} and P_{ZK} have the same properties as they are increasingly related. It has been mentioned that $1 - P_{SDH}$ is a measure of the kurtosis (that is, the degree of peakedness) of the distribution.

1.6 Measurement of social polarization:

Besides the study on income polarization, the mid-eighties of the last century saw the beginning of another dimension of social sciences viz. the study on conflict. The initial empirical research presented income inequality (and distribution of landownership) as the major cause of conflict. However, perceptions began to change with the conclusion in Lichbach (1989) that theempirical results were not of much significance. Tracing the history of the last century, a need was being felt for introduction of ethnic/religious components into the scope of empirical study. Following Easterly and Levine (1997), researchers pointed to the ethnic and religious social divides as a cause of conflict and low collective action. Some of the pioneering contributions include Alesina and La Ferrara (2000), Alesina and La Ferrara (2005), Collier and Hoeffler (2004), Desmet et al (2008, 2009), Fearon (2003), Fearon and Laitin (2000), Miguel et al. (2004) and Montalvo and Reynal-Querol (2005).

In the beginning, the focus was on the ethnic diversity, of which the most popular index is *FRAC*, the (ethno-linguistic) index of fractionalization. Considering a population with subgroup-proportions $\pi_1, \pi_2, \ldots, \pi_n$ the fractionalizationindex is defined by

$$FRAC = \sum_{i=1}^{n} \pi_i (1 - \pi_i) . \tag{1.57}$$

Clearly, *FRAC* measures the probability that two randomly selected individuals belong to two different groups.

A number of papers accepted this as an efficient predictor of ethnic violence and civil war. Moreover, they opined that in an ethnically diverse society, the chance of empirical growth is low and the level of corruption is likely to be high. Collier and Hoeffler (2004), Fearon and Laitin (2003) and Miguel et al. (2004) accepted *FRAC* as a regressor in their regression analysis on conflict. Vigdor (2002) considered a model of differential altruism and used *FRAC* to show that estimated fragmentation effects can be regarded as a weighted average of within-group affinity in the population.

A generalization of the fractionalization index can be found in Bossert et al. (2011). Building the notion of ‘similarity matrices’ containing similarity values s_{ij} between individuals *i* and *j* that takes values on [0,1] (0 being the case of perfect dissimilarity and 1 the case of perfect similarity), they defined the Generalized index of Ethno-linguistic Fractionalization (*GELF*). For characterization of *GELF*, the following axioms have been employed: (1) *Normalization* that demands the proposed measure to vanish in a situation of maximal dissimilarity in a society and to take on value 1 in case of maximal similarity; (2) *Symmetry* that demands the index to remain invariant under permutations of the arguments; (3) *Additivity*, which is a typical decomposition rule and (4) *Replication Invariance* that demands the measure to remain constant if the population gets multiplied. It has been shown that a diversity measure (that is, a map $: S \rightarrow R^{1}_{+}$, where $= \bigcup_n S_n$, S_n being the set of all n-dimensional similarity matrices) satisfies the four axioms mentioned above if and only if it is a positive multiple of $G(S) = 1 - \frac{1}{n^2}\sum_{i,j} s_{ij}$, *n* being the population size.

In spite of having its intuitive appeal, *FRAC* was unable to explain a number of issues related to the occurrence of civil war. Collier and Hoeffler (2004) claimed that contested dominance of one large group increases the chance of civil conflict rather than a high level of fractionalization.

Esteban and Ray (1999) developed a behavioral model that links the onset and intensity of social conflict to the society-wide distribution of individual characteristics and studied the problems of existence and uniqueness of conflict-equilibrium. Conflict has been viewed as a situation in which in the absence of a collective decision rule, social groups with contradictory interests make losses in order to increase the chance of obtaining their preferred outcomes. Further, it has been assumed that individuals with the same preferred outcome agree in their valuations of the remaining outcomes. Conflict is presented by the equilibrium sum of resources that are expended for achieving the preferred outcomes.

A striking conclusion of the paper is that the probability of conflict attains its maximum in a symmetric bipolar situation.This establishes a clear link between conflict and polarization.

Drawing inspiration from these findings, Montalvo and Reynal-Querol (*MR*, henceforth) (2005) put forward the notion of 'ethnic polarization' in place of 'ethno-linguistic fractionalization'. Citing the seminal paper by Horowitz (1985), they argued that chance of ethnic clash is small in a perfectly homogeneous/heterogeneous society. This probability is high in case the ethnic majority faces a large ethnic minority.

Reynal Querol (2002) replaced *FRAC* by a measure of ethnic polarization defined by

$$RQ = 4\sum_{i=1}^{N} \pi_i^{2}\left(1-\pi_i\right) \tag{1.58}$$

(which essentially measures the probability that out of three randomly selected individuals, two belong to the same group).

The *RQ* index equals twice the Ethno-linguistic Fractionalization Index in case there are only two ethnic groups. However, the equality relationship breaks up in case there are more than two groups. The two indices differ in another respect. In a multi-group setting, the proportional contribution of the largest group in case of *FRAC* is smaller than its relative size whereas the opposite is the case with *RQ*.

MR (2005) contended that *RQ* is a strong indicator of civil wars. This has been supplemented by the regression analysis in the empirical section.

Fitting a logit regression model for the incidence of civil wars as a function of *RQ* and *FRAC*, they found that the effect of *FRAC* on the onset of conflict is not significant whereas the impact of *RQ* is very much significant.

The theoretical justification behind introducing the *RQ* index has been furnished in Montalvo and Reynal-Querol (2008) (*MR* (2008), henceforth). The basic structure of a polarization index, as assumed by Esteban and Ray (1994) has been kept unaltered. However, given the categorical nature of the data, the *ER* index is required to be modified suitably. As the distance across groups (which, in case of *ER* measure is given by the Euclidean distance) in this situation is to be measured by a 'discrete metric (1-0)', *MR* (2008) suggested the family of Discrete Polarization ($DP(\alpha, \mu)$) measures defined by

$$DP\,(\alpha\,,\mu_1) = \mu_1 \sum_{i=1}^{n} \pi_i^{1+\alpha}\,(1-\pi_i) \tag{1.59}$$

By subsequent imposition of some desirable properties borrowed and redefined from the corresponding criteria cited in Esteban and Ray (1994), the authors characterized the *RQ*- index.

MR (2008) also studied the causal factors behind genocides. Replacing *FRAC* by *RQ* as a measure of ethnic heterogeneity, they demonstrated that the latter can successfully recognize the dependence of occurrence of genocides on ethnic heterogeneity, as was proposed by Horowitz .

MR (2005) further concluded that *RQ* has a highly positive correlation with *FRAC*.

Thus, there are two candidates for a plausible indicator of ethnic conflict (viz. *FRAC* and *RQ*). Considering conflict as a game, Esteban and Ray (2008b) developed a behavioral model to compare these two measures as indicators of (a) the intensity and (b) the onset of conflict. Their findings are as follows: (i) The intensity and the likelihood of conflict are two different notions. They usually run along opposite directions. (ii) The relationship between polarization or fractionalization and conflict is not monotonic. (iii) The occurrence of conflict is monotonically related to the degree of fractionalization while (iv) the intensity of conflict has a positive association with the degree of polarization. Finally, the incidence of conflict depends not only on the shape of the distribution, but also on the existing political system.

In a similar vein, Esteban and Ray (2011) used a second behavioral model of conflict to justify the use of *FRAC*, *RQ* and the Gini index as measures of polarization. A monotone transform of the equilibrium level of resources expended in conflict has been shown to be approximately equal to a linear combination of the three indices. In particular, for a large population, per-capita conflict is proportional to a convex combination of only *FRAC* and *RQ*. Further, the higher is the 'altruism' (α), the more pertinent are *FRAC* and *RQ* in explaining conflict. Here $\alpha \epsilon [0,1]$ is given by

$$U_i(k) = (1-\alpha)\,\pi_i(k) + \alpha \sum_{l \neq i} \pi_i(l)\,, \tag{1.60}$$

$U_i(k)$ and $\pi_i(k)$ being respectively the extended utility and the expected pay-off of a group- i member k.

In an empirical follow up, Esteban et al. (2012) verified that the three distributional indices suggested above are significant correlates of conflict. Performing a number of robustness tests, they concluded that *RQ* has a highly significant and positive contribution to it, the effect of *FRAC* is significantly positive (to a comparatively lesser extent, though) while the Greenberg-Gini index (see Greenberg, 1956) defined by

$$G = \sum_{i,j=1}^{m} \pi_i \pi_i d_{ij} , \tag{1.61}$$

with π_i= population share of the group i and d_{ij}= "intergroup distance" between groups i and j, affects the chance of conflict negatively. To compute ethno-linguistic distances across groups, the authors considered the cardinality of intervening nodes on the language tree.

As a generalization of *FRAC* and *RQ*, Esteban and Ray (2008b) introduced the following index:

$$R = \sum_{i} \frac{\pi_i}{p_i} \pi_i^2 (1 - \pi_i) b_i , \tag{1.62}$$

whereb_i is the alienation felt by the i^{th} group from other groups and p_i is the probability of success for the i^{th} group in the contest game. Then the effect of a transfer of mass (from one group to another), effect of a split (of a group), maximality etc. have been studied.

One of the basic assumptions in *MR* (2005) was that the distance between two ethnic groups is to be measured by a discrete metric. However, Fearon (2003) argued that the intergroup-distance plays a key role in measuring ethnic heterogeneity. This view has been seconded in Desmet et al (2009, 2010). Incorporating intergroup distances, they found that the resulting index outperforms the *RQ* index. In their empirical study, Esteban and Mayoral (2011) strongly favored the use of interpersonal distances "driven by the intensity of the ethnic and religious attitudes." (This *intensity of feelings* is obtained by summing up the responses to a number of relevant queries.) Then the interpersonal distance between two persons belonging to two different groups has been computed by adding up the intensities while for two persons belonging to the same group, this is given by the absolute distance of their feelings. Subsequently, performing a series of robustness tests, they established that the distance-based indices of both religious and ethnic polarization (that take into account the intensity of feelings) are significant in explaining the occurrence of conflict.

Some indices of social polarization:

Ethnicity is a social characteristic. Therefore, ethnic polarization can be regarded as a particular case of social polarization. Replacing the discrete metric to measure the distance between two groups in case of categorical data, Permaneyer (2012) introduced the notion of *radicalism degree*, which is the intensity with which one feels identified with the group one belongs to. The alienation between two groups has been posited to be the sum of their radicalism degrees.[5] Next, stating a number of assumptions which bear similarity with the postulates of *ER* (1994), the author characterized the following polarization measure:

$$P_{N,\alpha}^{b}(\mathbf{f}) = \sum_{i=1}^{N} \sum_{j\neq i} \pi_i^{1+\alpha} \pi_j \left(\lambda_i + \lambda_j\right), \tag{1.63}$$

where $\alpha \in (0,1]$ and μ_i is the mean of the radicalism distribution f_i (of the i^{th} group). It is easily seen that the discrete polarization indices form a subfamily of $P_{N,\alpha}^{b}$ (in case all the λ_i's are equal).

Permaneyer (2012) then stated an additional axiom demanding that when other things are kept fixed, the greater is the number of groups, the less polarized the distribution is. Using this postulate, he characterized a second class of polarization indices. Moreover, a lower bound on α has been shown to be: $\alpha^* \approx 0.71$.

A second model has then been suggested taking into account the within-group alienation. The alienation felt by two persons with radicalism degrees x and y belonging to the same group is assumed to be $|x - y|$. The author then employed three additional axioms. The first two axioms are based on the principle that if different groups are made more homogeneous, polarization should go up. The third one is a direct translation of the idea that polarization attains its maximum in case of a symmetric bipolar distribution.

[5]This approach is similar to the one adopted by Esteban and Mayoral (2011).

Finally, application of this new set of axioms and an earlier one together forces that the polarization index has to take the form

$$P_{n,\alpha}(\mathbf{f}) = \sum_{i=1}^{n} \iint f_i^{1+\alpha}(x)\, f_i(y)\left|x-y\right| dy dx + \sum_{i=1}^{n}\sum_{j\neq i} \iint f_i^{1+\alpha}(x)\, f_i(y)(x+y)\, dy dx\,, \qquad (1.60)$$

where f_i is the density of the i^{th}group, $\mathbf{f} = (f_1, f_2, \ldots, f_n)$ and $\alpha \in \left[\frac{1}{3n-2}, 1\right]$, n being the number of social groups.

1.7 Polarization in case of ordinal data:

An ordinal variable is similar to a categorical variable like gender, ethnicity and religion that has one or more categories or types. However, for an ordinal variable there is a well-defined ordering rule. For instance, consider self-reported health data of a population. The six health categories ‘very poor’, ‘poor’, ‘fair’, ‘good’, ‘very good’ and ‘excellent’ can be provided with positive integral values in an increasing order. This assignment of integral values is arbitrary; the only restriction is that to preserve the ordering a higher number should be assigned to a better category. Thus, in any assignment of numbers to the categories ‘very good’ should get a higher number than ‘good’ (see Allison and Foster, 2004). A second example can be ordering of educational achievement levels of individuals in a society starting from illiteracy to university education by assigning numbers in an increasing way (see Chakravarty and Zoli, 2012). Thus, in all these cases we need to select an ordinal scale which assigns a numerical value to each category of the characteristic under consideration so that ranking of categories is maintained. This approach is robust to changes in the scale in the sense that if instead of assigning numbers in an increasing strictly convex manner, the assignment is done using an affine transformation, the ordering remains preserved.

Using the ‘self-reported health status’ (SRHS) data, Allison and Foster (2003) demonstrated an ‘ordinal’ method for calculating overall health inequality. Their approach is median-based and the inequality is viewed as spread away from the median. . It illustrates the role of first order stochastic dominance as an unambiguous indicator of population health.

Chakravarty and Zoli (2012) demonstrated usefulness of second order stochastic dominance for integer variables in this context (see also Chakravarty and D'Ambrosio, 2006).

For two health distributions x and y with the same median category m, x is said to have a

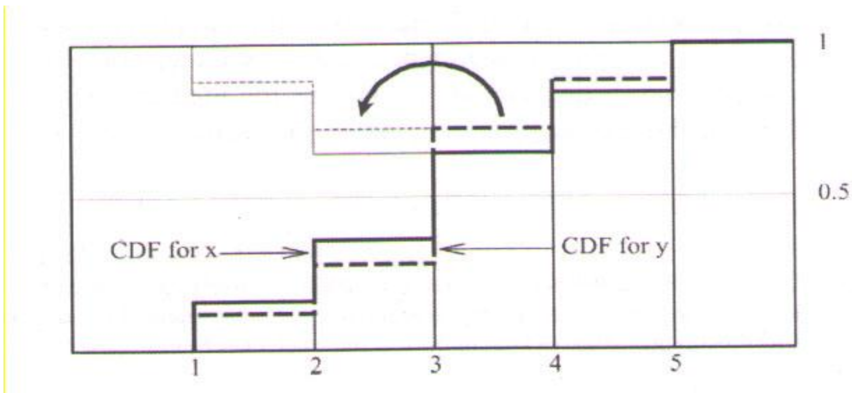

Fig. 1.9 (Construction of *S*-curves)[6]

greater spread than y (what we denote by xSy) if the cumulative population share of the bottom k categories of x is greater than the corresponding quantity of y for $k < m$ while the reverse happens for all $k \geq m$.

If F_x^i denotes the sum of the population shares of the distribution x of all the categories upto category i, then xSy if and only if

$$F_x^i > F_y^i \text{ for all } i < m \text{ and } F_x^i \leq F_y^i \text{ for all } i \geq m . \tag{1.64}$$

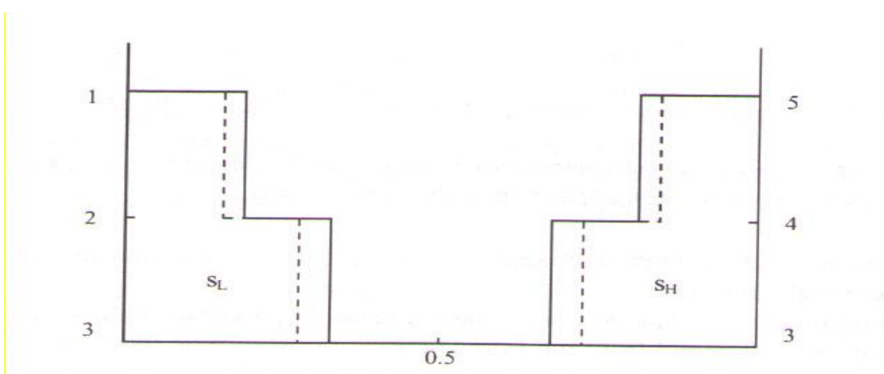

Fig. 1.10 (The partial ordering *S*)

[6] Figures 1.9 and 1.10 are borrowed from Allison and Foster (2004).

Denoting by $S_L(x)$ (resp. $S_R(x)$)twice the area below the S-curve of x to the left (resp. right) of 0.5, it has been shown that (1.64) is equivalent to $S_L(x) > S_L(y)$ and $S_R(x) > S_R(y)$. The partial ordering S is clearly reflexive and transitive but incomplete.

Moving along the same direction and beginning with self-assessed health data, Apouey (2007) developed the notion of polarization in an ordinal context. Taking the median as the reference point, Apouey (2007) proposed a number of axioms borrowed from Wang and Tsui (2000) and modified suitably. The first two axioms are: Increased spread (**IS**) and Increased bipolarity (**IB**).

The **IS** axiom demands that polarization should not decrease if there is a spread in the distribution away from the median. In other words, greater distancing between the categories below and not below the median should not make the distribution less polarized.

Before stating the **IB** axiom, Apouey (2007) introduces the notion of 'transfer', which essentially means bunching or clustering of mass in categories lying on the same side of the median category. The outcome of a 'transfer' is enhancement of homogeneity among the individuals within the categories, which in turn raises polarization. (For discussion on **IS** and **IB** in greater details, see Chapter 5.)

The third axiom, similar to the corresponding one in social polarization, states that the maximum value of the polarization index is attained by the symmetric bipolar distribution.

The final postulate is a compatibility assumption.

Then assuming an additive structure of the polarization index whose components are continuous transforms of the deviation from the symmetric bipolar distribution, Apouey (2007) characterized the following polarization indices:

$$P_{A_1}(F) = 1 - \frac{2^{\varsigma}}{n-1}\sum_{c=1}^{n-1}\left|F_c - \frac{1}{2}\right|^{\varsigma} \tag{1.65}$$

and

$$P_{A_2}(F) = \varpi\left[\left(\frac{N_n}{2}\right)^{\varsigma} - \frac{1}{n-1}\sum_{c=1}^{n-1}\left|\frac{N_c}{N_n} - \frac{1}{2}\right|^{\varsigma}\right], \tag{1.66}$$

where $\varpi > 0$ is a constant, $\varsigma \in (0,1)$ and N_c (resp. F_c) is the cumulative frequency (resp. proportion) of the c^{th} category, the categories being arranged in the ascending order of some ordinal characteristic.

Using the notion of 'concentration curve' for health, obtained by plotting the cumulative shares of health levels ranked by increasing incomes, Apouey (2010) suggested two 'social health bipolarization indices'.

Based on Reardon's (2009) method of cardinal measurement of the degree of inequality of an ordered variable, Fusco and Silber (2014) formulated a number of axioms (eg. 'swap' of individuals between unordered population subgroups/ ordered categories) on polarization indices in an ordinal context. Considering an I by J matrix (where I = number of population subgroups and J = number of ordered categories), several indices have been proposed, some of which are closely related to the measures used in information theory, the theory of diversity and the notion of dissimilarity.

1.8 Multidimensional polarization measures:

Multidimensional extensions of the measures of polarization remain largely unexplored till now. Esteban and Ray (2012) maintain that the index suggested by Zhang and Kanbur (1.54) may serve well in some cases as an appropriate measure of multidimensional polarization. The reason is that the groups are exogenously given using some social characteristic such as religion, ethnicity or geographical location, whereas the 'identification' and 'alienation' components are measured in terms of income differences. In a recent survey paper, Permaneyaer (2015) terms it a 'hybrid polarization measure'.

Mogues (2008) suggested a two-dimensional extension (in which one dimension is a variable and the other is an attribute) of the (1-dimensional) *ER*-measure using four basic postulates viz. Shrinking of the Middle Class, Population Concentration around Poles, Separation of the Poles and Strong Relationship among Individuals' Attributes.

Recently Merz and Scherg (2013) have proposed extended multidimensional polarization indices based on a CES-type well-being function and have presented a new measure to multidimensional polarization.

Arcagni and Fattore (2014) have showed the relevance of partial order theory in measuring multidimensional polarization in case of ordinal data.

1.9 Contest success functions:

In a discussion on a new behavioral model, Esteban and Ray (1999) present conflict as a contest game, thus establishing the link between polarization and contests.

A contest refers to a non-cooperative game in which two or more participants contend for a prize. Models of contest have been employed extensively to analyse a variety of phenomena like rent seeking (Tullock 1980, Nitzan 1991, Corchon 2000, Baye and Hoppe 2003, Amegashie 2006), conflict (Hirshleifer 1991, Skaperdas 1992), polarization (Esteban and Ray 2011, Chakravarty 2015), electoral candidacy (Snyder 1989, Skaperdas and Grofman 1995), sporting tournament (Szymanzki 2003), provision of public goods (Kolmar and Wagener 2011) and reward structure in firms (Rosen 1986)[7]. In a contest, agents make irretrievable investments, which depending on the situation; can be money, effort or any other valuable resource.

Hirshleifer (1989) formally introduced the notion of Contest Success Function. According to his definition, a contest success function (CSF) determines each player's probability of winning for a given level of efforts. Skaperdas (1996) characterized all CSFs satisfying a number of reasonable axioms. Axioms employed include homogeneity and translation invariance of the CSF. Deduced functional forms of the CSF are power and logit respectively. Extending these results, Clark and Riis (1998) characterized an asymmetric form of the 'power' success function. Generalizing further, Rai and Sarin (2009) considered CSFs in a multidimensional framework where each contestant is allowed to invest in multiple areas. Munster (2009) considered contest between groups. Using axioms which look quite similar to those used in Skaperdas (1996), the author explored the impact of homogeneity in the extended framework. Difference form contest success function, introduced by Hirshleifer (1989), was studied in greater detail by Baik (1998), Che and Gale (2000), Gersbach and Haller (2009) and Corchon and Dahm (2011).

[7]The literature has been surveyed by Nitzan (1994), Corchon (2007), Konrad (2009) and Skaperdas and Garfinkel (2012). See also Dixit (1987) for a general discussion.

In a contest, an increase in each contestant's outlay increases his chances of winning and reduces his opponents' chances. In a highly interesting contribution, Skaperdas (1996) characterized this probability for any contestant as the ratio between the level of effective investment made by the contestant and the sum of effective investments across all the contestants. The effective investment of a contestant can be interpreted as the output determined by his effort, which may be regarded as his input in the contest. It is assumed to be an increasing and positive valued function of effort. This is the basic structure of Skaperdas (1996).

Using his basic structure, Skaperdas (1996) also developed axiomatic characterizations of the Tullock (1980)-Hirschleifer (1989) functional forms of CSFs. One of the axioms employed by Skaperdas (1996) is an anonymity principle which demands that a contestant's probability of success depends only on his outlays. Thus, the agents are not distinguished by any characteristic other than their outlays. Clark and Riis (1998) broadened the Skaperdas (1996) framework by allowing the contestants to differ with respect to their contest-related personal characteristics. Rai and Sarin (2009) generalized the characterizations of Skaperdas (1996) to the situation where agents can have investments that are of multiple types in nature. Münster (2009) extended the Skaperdas (1996) and Clark and Riis (1998) characterizations to contests between groups. Arbatskaya and Mialon (2010) developed a model for a multi-armed contest and characterized the CSF axiomatically in this context.

1.10 Plan of the thesis:

There are five chapters in this treatise. Chapter 1 gives a general introduction to the existing literature on inequality, polarization and the theory of contests. The bulk of the thesis (chapters 3, 4 and 5) deals with the theory of (income as well as social) polarization while inequality and contests have been discussed in chapters 2 and 6 respectively.

Chapter 2 of this thesis suggests a two-parameter extension of the family of subgroup decomposable absolute inequality indices identified in Chakravarty and Tyagarupananda (1998) and in Bosmans and Cowell (2010). Maintaining similarity with Zheng (2007), we replace the notion of 'translation invariance' by 'translation consistency' and hence characterize the relevant class of subgroup decomposable inequality indices.

Chapter 3 is on a generalization of the polarization indices suggested by Zhang and Kanbur (2001) and Rodriguez and Salas (2003). Beginning with the 'identification-alienation' framework suggested by Esteban and Ray (1994), we define a reduced-form polarization index in the following way. We consider a population subgroup decomposable inequality index. A reduced form polarization index is an increasing function of the between-group term and a decreasing function of the within-group term of the inequality index. The between-group term represents the 'alienation' component of polarization and the within-group term can be regarded as an inverse indicator of its 'identification' component. A quasi-ordering for ranking alternative distributions of income using such polarization indices has been developed. Several polarization indices of the said type have been characterized using intuitively reasonable axioms. Finally, we consider the dual problem of retrieving the inequality index from the specified form of a polarization index.

We next look at polarization in case of ethnic data. The main goal of Chapter 3 is to characterize the *RQ* index, suggested in Reynal-Querol (2002). Further, we mention two distinct ethnic polarization quasi-orderings that can rank ethnic distributions unambiguously in terms of all ethnic polarization indices satisfying certain intuitively reasonable postulates. These two quasi-orderings are proven to be independent of one another. In the process, we characterize a generalized form of the *RQ* index.

The next chapter is meant for measuring polarization for an ordinal data. A family of generalized Gini indices of polarization has been introduced. These indices can be applied to dimensions of human well-being with ordinal significance such as self-assessed health data and literacy. We investigate several properties of this general index and characterize it axiomatically. We also look at a quasi-ordering induced by the generalized Gini indices for ranking alternative distributions of an ordinally measurable dimension. In addition, implications of some of the axioms have been explored.

The final chapter is on the study of some structural properties of contest success functions (CSFs), which stipulate the winning probabilities of the contestants. Two major axioms (viz. the scale invariance and translation invariance) used in Skaperdas (1996) have been relaxed to yield two ordinal postulates (viz. scale consistency and translation consistency). Further, an intermediate invariance axiom, a convex mixture of the two invariance axioms has been formulated and the corresponding class of CSFs has been identified. This family contains the

Tullock and Hirschleifer CSFs as special cases. We also explore the possibility of existence of Nash equilibrium of the corresponding contest game. Next, we look at the two consistency conditions and characterize the respective classes of CSFs. It has been demonstrated that if the number of contestants is at least three, then scale consistency and translation consistency, in the presence of other axioms, characterize the same functional forms identified by scale and translation invariances respectively. Finally, we define an intermediate consistency condition and classify all CSFs satisfying the same.

Chapter 2

TRANSLATION CONSISTENT SUBGROUP DECOMPOSABLE INEQUALITY INDICES

2.1 Introduction

As mentioned in Chapter 1, ‘subgroup-decomposability’ of an inequality index, is an important property for various reasons. In this chapter, we take up a problem related to the ‘subgroup-decomposable’ inequality indices satisfying certain standard assumptions (discussed in the previous chapter).

Following the argument put forward by Zheng (2007) and Zheng (2005), one may wonder if it is possible to find an ordinal counterpart of ‘translation invariance’ and axiomatize the corresponding class of decomposable inequality indices. This paper makes an attempt to answer this question. We first define a ‘translation consistent’ inequality index. To illustrate the notion, consider two income distributions D_1 and D_2 and let the inequality index I rank D_1 higher than D_2. Now, if all the incomes in both the distributions aree increased/decreased by a constant amount, then ‘translation consistency’ demands that I should rank the former higher than the latter. Evidently, ‘translation consistency’ is an ordinal property whereas ‘translation invariance’ is a cardinal one. It is equally clear that a translation invariant inequality index is ‘translation consistent’, but the converse is not true. Thus, the class of all ‘translation consistent’ inequality indices includes I_θ and the variance. In other words, the class characterized in this chapter can be viewed as a generalization of the family mentioned in Chakravarty and Tyagarupananda (1998).

Since the basic framework has already been discussed in the previous chapter, we avoid a repetition of that. The characterization theorems have been presented in the following section and the subsequent one is a formal conclusion.

2.2 **The Characterization Theorem**

In Chapter 1, we have defined scale and translation invariance of inequality indices and have explicitly mentioned the structure of the 'unit consistent' family of subgroup-decomposable inequality indices characterized by Zheng (2007) and Zheng (2005).

We begin this section with the formal definition of translation consistency of an inequality measure.

Definition 2.1: An inequality index $I : D \to R^1_+$ is said to be translation consistent if for all $x, y \in D^n$, $I(x) \le I(y)$ implies $I\left(x + c1^n\right) \le I\left(y + c1^n\right)$ for all scalar c such that $x + c1^n,\ y + c1^n \in D^n$.[8]

As already mentioned in the previous section, both variance and the Kolm measure, being translation invariant, are translation consistent as well. However, it can be demonstrated that no member of the generalized entropy class satisfies translation consistency. For example, if we consider I_2, that is, half the squared coefficient of variation, then with $x = (1,3,8)$, $y = (2,3,10)$ and $c = 10$ we have, $I_2(x) = 1.625$ and $I_2(y) = 1.52$ while $I_2\left(x + c1^3\right) = 0.1326$ and $I_2\left(y + c1^3\right) = 0.1689$. Thus, $I_2(x) > I_2(y)$ but $I_2\left(x + c1^3\right) < I_2\left(y + c1^3\right)$. In other words, I_2 fails to satisfy translation consistency.

The first result of this section, whose proof is similar to that of Proposition 1 in Zheng (2007), is on the necessary and sufficient condition of translation consistency.

Proposition 2.1: An inequality index $I : D \to R^1_+$ is translation consistent if and only if for all $x \in D$ and for all $c > 0$, there exists a continuous function $f : R^1_{++} \times R^1_+ \to R^1_+$, which is nondecreasing in the second argument such that

$$I\left(x + c1^n\right) = f\left(c, I(x)\right) \tag{2.1}$$

We next mention a result borrowed from Shorrocks (1980).

---

[8] It may be mentioned here that Zheng (2007) defined 'unit consistency' in terms of strict inequality. But in this chapter we have adopted a more general definition using weak inequality.

Proposition 2.2: A differentiable inequality index I satisfies **SYM, POP, SUD** and **NOM** if and only if there exist functions $\psi : R^1_+ \to R^1_{++}$ and $\phi : R^1_+ \to R^1_+$ such that for any $x \in R^n_+$,

$$I(x) = \frac{1}{n\,\psi(\lambda(x))} \sum_{i=1}^{n} [\phi(x_i) - \phi(\lambda(x))] \tag{2.2}$$

where ψ is differentiable; ϕ is strictly convex and continuously differentiable.

Our first result is on an immediate implication of translation consistency.

Proposition 2.3: If an inequality index I satisfies **SYM, NOM, SUD, POT** and translation consistency, then

$$I\left(x + c1^n\right) = \tau^c I(x) \tag{2.3}$$

for all $c > 0$ and some constant $\tau > 0$.

Proof: See Appendix.

We now state the main result of the chapter.

Theorem 2.1: An inequality index $I : D \to R^1_+$ satisfies **SYM, NOM, POP, SUD, POT,** continuous differentiability and translation consistency if and only if it is a positive multiple of the form

$$I_{\gamma,\delta}(x) = \begin{cases} \dfrac{1}{\delta^{\lambda(x)}} I_V(x), \gamma = 0, \delta > 0 \\ \dfrac{1}{n(x)\,\delta^{\lambda(x)}} \displaystyle\sum_{i=1}^{n} \left[e^{\gamma x_i} - e^{\gamma\lambda(x)}\right], \gamma > 0, \delta > 0 \end{cases}, \tag{2.4}$$

Proof of the theorem uses the following lemma.

Lemma 2.1: Whenever I satisfies (2.3) we have,

$$\sum_{i=1}^{n} I_i(x) = (\ln \tau) I(x), \tag{2.5}$$

for all $x \in R^n_{++}$.

Proofs of Lemma 2.1 and Theorem 2.1: See Appendix.

Remark 2.1: The substitution $\delta = e^{\gamma}$ in (2.4) transforms $I_{\gamma,\delta}$ to I_{θ} (with $\theta = \gamma$). Thus, $I_{\gamma,\delta}$ can be viewed as a 2-parameter extension of the family (I_{θ}, I_{V}) characterized in Chakravarty and Tyagarupananda (1998).

Remark 2.2: It can be easily verified that none of the measures $I_{\gamma,\delta}$ is scale invariant. However, the variance is the only member of this family which is unit consistent. Thus, there is a subgroup decomposable inequality index which is both unit consistent and translation consistent.

Zheng (2007) talks of extreme rightist and extreme leftist views of inequality measurement. An extreme rightist measure *I* is one which is reduced when all the incomes are increased by the same proportion, that is, if $I(ax) < I(x)$ for all $a > 1$. One can easily see that none of the $I_{\gamma,\delta}$ indices agrees with the extreme rightist view. Similarly, *I* is an extreme leftist measure if it is increased when all the incomes are augmented by the same amount, that is, if $I\left(x + c1^{n}\right) > I(x)$ for all $c > 0$. A simple calculation shows that $I_{\gamma,\delta}$ conforms to the extreme leftist view if $\delta < e^{\gamma}$.

2.3. Conclusion

Following Zheng (2007), we have characterized in this chapter a generalization of the class of subgroup-decomposable absolute indices of inequality. It is fairly interesting because we have replaced a cardinal property (viz, translation invariance) of a subgroup-decomposable inequality index by an ordinal one (viz.translation consistency) and the identified class extends the family characterized earlier.

2.3 Appendix

Proof of Proposition 2.3: Proceeding as in Proposition 3 of Zheng (2007) and maintaining the same set of notations we arrive at

$$f(c,k) = a(c)k \tag{2.6}$$

whenever $c > 0$, for some constant $a(c) \neq 0$.

Equation (2.1), along with equation (2.6) implies that for any $x \in D$,

$$I(x + c1^n) = a(c)I(x) \tag{2.7}$$

from which it follows that for arbitrary $c_1, c_2 > 0$,

$$I\{x + (c_1 + c_2)1^n\} = a(c_1 + c_2)I(x) \tag{2.8}$$

and
$$\begin{aligned} I\{x + (c_1 + c_2)1^n\} &= a(c_2)I(x + c_1 1^n) \\ &= a(c_1)a(c_2)I(x) \end{aligned} \tag{2.9}$$

Equations (2.8) and (2.9) together yield:

$$a(c_1 + c_2) = a(c_1)a(c_2) \tag{2.10}$$

for all $c_1, c_2 > 0$.

The only continuous solution to this equation is given by

$$a(c) = \tau^c \tag{2.11}$$

where $c > 0$ (see Aczel, 1966, p. 84). Positivity of τ is a consequence of non-negativity of I. This completes the proof of the proposition. □

Proof of Lemma 2.1: Fix $x \in R^n_{++}$ and define $g : R^1_+ \to R^1_+$ by

$$g(c) = I(x + c1^n) \tag{2.12}$$

By differentiability of I it follows that

$$g'(c) = \nabla I(x + c1^n).1^n \tag{2.13}$$

where ∇v is the gradient of v.

By continuous differentiability of I we have,

$$g'(0) = \nabla I(x).1^n = \sum_{i=1}^{n} I_i(x) \tag{2.14}$$

But

$$g(c) = \tau^c I(x). \tag{2.15}$$

Differentiating both sides of (2.15) we get,

$$g'(c) = \tau^c \ln \tau I(x) \tag{2.16}$$

From (2.16) it readily follows that

$$g'(0) = \ln \tau I(x) \tag{2.17}$$

Comparison of equations (2.14) and (2.17) yields the desired result. □

Proof of Theorem 2.1: Let $x \in R^n_{++}$. Differentiating (2.2) partially w.r.t. x_i we get,

$$n\, I_i(x) = \frac{\psi(\lambda)\left[\phi'(x_i) - \phi'(\lambda)\dfrac{1}{n}\right] - \sum_{i=1}^{n}\left[\phi(x_i) - \phi(\lambda)\right]\psi'(\lambda)\dfrac{1}{n}}{\left(\psi(\lambda)\right)^2}. \tag{2.18}$$

Next, taking sum over all i,

$$\sum_{i=1}^{n} n\, I_i(x) = \frac{\psi(\lambda)\left[\phi'(x_i) - \phi'(\lambda)\right] - \psi'(\lambda)\sum_{i=1}^{n}\left[\phi(x_i) - \phi(\lambda)\right]}{\left(\psi(\lambda)\right)^2}. \tag{2.19}$$

Combining (2.2), (2.5) and (2.19) we get,

$$\begin{aligned}\psi(\lambda)\left[\phi'(x_i) - \phi'(\lambda)\right] - \psi'(\lambda)\sum_{i=1}^{n}\left[\phi(x_i) - \phi(\lambda)\right]&\\ = (\ln\tau)\,\psi(\lambda)\sum_{i=1}^{n}\left[\phi(x_i) - \phi(\lambda)\right].&\end{aligned} \tag{2.20}$$

A rearrangement of (2.20) entails

$$\psi(\lambda)\sum_{i=1}^{\mu}\left[\phi'(x_{i'}) - \phi'(\lambda)\right] - \left\{\psi'(\lambda) + (\ln\tau)\,\psi(\lambda)\right\}\sum_{i=1}^{n}\left\{\phi(x_{i'}) - \phi(\lambda)\right\} = 0\,. \tag{2.21}$$

Differentiating (2.21) partially w.r.t. x_i we obtain,

$$\begin{aligned}&\psi'(\lambda)\left[\phi''(x_i) - \phi''(\lambda)\frac{1}{n}\right] + \sum_{i=1}^{n}\left[\phi'(x_i) - \phi'(\lambda)\psi'(\lambda)\frac{1}{n}\right]\\ &= \left\{\psi'(\lambda) + (\ln\tau)\,\tau(\lambda)\right\}\left[\phi'(x_i) - \phi'(\lambda)\frac{1}{n}\right]\\ &\qquad + \left[\psi''(\lambda)\frac{1}{n} + (\ln\tau)\,\psi'(\lambda)\frac{1}{n}\right]\sum_{i=1}^{n}\left\{\phi(x_i) - \phi(\lambda)\right\}.\end{aligned} \tag{2.22}$$

Replacing x_i by x_j in (2.22) and taking difference of both sides we get,

$$\psi(\lambda)\left\{\phi''(x_i) - \phi''(x_j)\right\} = \left\{\psi'(\lambda) + (\ln\tau)\psi(\lambda)\right\}\left[\phi'(x_i) - \phi'(x_j)\right]. \tag{2.23}$$

Since this holds for all $x_i, x_j \in R^1_{++}$, it follows that

$$\frac{\psi'(\lambda)}{\psi(\lambda)} = \delta, \tag{2.24}$$

where δ is a constant.

Solution to (2.24) is given by:

$$\psi(\lambda) = K\delta^{\lambda} \tag{2.25}$$

for some constant K. Since ψ is positive-valued, it follows that $K > 0$.

Substituting (2.25) in (2.23) we get,

$$\phi''(x_i) - \phi''(x_j) = (\delta + \ln\tau)\left[\phi'(x_i) - \phi'(x_j)\right], \tag{2.26}$$

that is,

$$\phi''(x_i) - (\delta + \ln\tau)\phi'(x_i) = \phi''(x_j) - (\delta + \ln\tau)\phi'(x_j) \tag{2.27}$$

It follows that

$$\phi''(x_i) - (\delta + \ln\tau)\phi'(x_i) = b\,, \tag{2.28}$$

for some constant *b*.

Thus, $y = \phi(x)$ satisfies the differential equation

$$\left(D^2 - \gamma D\right)y = b \tag{2.29}$$

where $Dy = \frac{dy}{dx}, D^2 y = \frac{d^2 y}{dx^2}$ and $\gamma = (\delta + \ln\tau)$.

The complete solution to (2.29) is given by

$$y = \phi(x) = \begin{cases} c_0 + c_1 x + \dfrac{b}{2}x^2\,;\; b, c_0, c, \in R\;(\gamma = 0) \\ c_1 - \dfrac{b}{\gamma}x + c_2 e^{\gamma x};\; c_1, c_2, b \in R\;(\gamma \neq 0) \end{cases}. \tag{2.30}$$

Using strict convexity of ϕ it further follows that $b > 0$ in the first case and $c_2 > 0$ in the second.

Simplified forms of I corresponding to the solutions in (2.30) can be described as follows. □

Case I: $\gamma = 0$.

$$I(x) = \frac{1}{nK\delta^{\lambda(x)}}\sum_{i=1}^{n}\left\{c_0 + c_1 x_i + \frac{b}{2}x_i^2 - c_0 - c_1\lambda(x) - \frac{b}{2}\lambda^2\right\} = \frac{b\;\operatorname{var}(x)}{2K\delta^{\lambda(x)}} \tag{2.31}$$

Case II: $\gamma \neq 0$.

$$I(x) = \frac{1}{Kn\delta^{\lambda(x)}} \sum_{i=1}^{n} \left[c_1 - \frac{b}{\alpha} x_i + c_2 e^{\gamma x_i} - c_1 + \frac{b}{\alpha} \lambda(x) - c_2 e^{\gamma \lambda(x)} \right] = \frac{c_2}{Kn\delta^{\lambda(x)}} \left\{ e^{\gamma x_i} - c_2 e^{\gamma \lambda(x)} \right\} \tag{2.32}$$

Equations (2.31) and (2.32) jointly imply the desired result. □

Chapter 3

REDUCED FORM INDICES OF INCOME POLARIZATION[9]

3.1 Introduction

In Chapters 1 and 2 we have studied the properties of subgroup-decomposable inequality indices and have subsequently classified a family of such indices satisfying some desirable properties. In this chapter we try to connect the notion of subgroup-decomposability of inequality indices with the notion of polarization.

To recall from Chapter 1, Esteban and Ray (1994) defined an index of polarization using identification-alienation framework (where both the components are increasingly related to the former). The authors then developed an axiomatic characterization assuming a quasi-additive structure of the index.

It can also be recalled that beginning with a subgroup-decomposable index of inequality, Zhang and Kanbur (2001) suggested the ratio of the between-group and within-group components of inequality as an indicator of polarization. Clearly, the 'between group' inequality incorporates the intuition behind the 'alienation' factor while an inverse measure of 'identification' is given by the 'within group' inequality term.

A similar approach was adopted by Rodriguez and Salas (2003), who considered bi-partitioning of the population using the median and defined a bipolarization index as the difference between the between-group and within-group terms of the Donaldson-Weymark (1980) S-Gini index of inequality.

A common feature of the Zhang-Kanbur and the Rodriguez-Salas index is that both are 'reduced-form' or 'abbreviated' indices that can be used to characterize the trade-off between the alienation and identification components of polarization.

As Esteban and Ray (2005, p.27) noted, the Zhang-Kanbur formulation is a 'direct translation of the intuition behind' the postulates that polarization is increasing in between-group inequality and decreasing in within-group inequality. Since the Zhang-Kanbur -Rodriguez-Salas approach enables us to understand the two main components of polarization, identification and alienation, in an intuitive way, this chapter makes some analytical and rigorous investigation

[9] A major portion of this chapter has been published in Chakravarty and Maharaj (2011b). The content of this chapter is also related to Chakravarty, Chattopadhyay and Maharaj (2010).

using the idea that polarization is related to between-group inequality and within-group inequality in increasing and decreasing ways respectively.

Now, polarization indices can give quite different results. Evidently, a particular index will rank income distributions in a complete manner. However, two different indices may rank two alternative income distributions in opposite directions. In view of this, it becomes worthwhile to develop necessary and sufficient conditions that make one distribution more or less polarized than another unambiguously. This is one objective of this chapter. We can then say whether one income distribution has higher or lower polarization than another by all abbreviated polarization indices that satisfy certain conditions. In such a case it does not become necessary to calculate the values of the polarization indices to check polarization ranking of distributions. If the population is bi-partitioned using the median, then this notion of polarization quasi-ordering becomes close to the Wolfson (1994, 1997) concept of bipolarization quasi-ordering.

Next, given the diversity of numerical indices it will be a worthwhile exercise to characterize alternative indices axiomatically for understanding which index becomes more appropriate in which situation. An axiomatic characterization gives us insight of the underlying index in a specific way through the axioms employed in the characterization exercise. This is the second objective of this chapter. We characterize several polarization indices, including a generalization of the Rodriguez-Salas form. The structure of a normalized ratio form index parallels that of the Zhang-Kanbur index. We then show that the different sets of intuitively reasonable axioms considered in the characterization exercises are independent, that is, each set is minimal in the sense that none of its proper subset can characterize the index.

Finally, we show that it is also possible to start with a functional form of a polarization index and determine the inequality index which would generate the given polarization index. Specifically, we wish to determine a set of sufficient conditions on the form of a polarization index to guarantee that there exists an inequality index, which would produce the polarization index. This may be regarded as the dual of the characterization results for polarization indices.

In the next section of the chapter we make a requisite discussion pertaining to a specific property of subgroup-decomposable inequality indices. The polarization quasi-ordering is discussed and analyzed in the following section. The characterization theorems and a duality theorem are presented in Section 3.4. Section 3.5 concludes the chapter. Proofs of all the theorems are relegated to an Appendix (Section 3.6).

3.2 The Background

Consider a population of size *n*. Let x_i denote the income of the i^{th} individual, assumed to be drawn from the non-degenerate interval $[\upsilon, \infty)$ in the positive part R^1_{++} of the real line R^1. Since for inequality and **SUD** to be well defined, we need $n, k \in \Gamma$ and $n_i \in \Gamma$ for all $1 \le i \le k$, we assume throughout this chapter that $n \ge 4$, where $\Gamma = N \setminus \{1\}$.

Maintaining the same set of notations as in Section 1.1 of Chapter 1, the weight attached to the inequality of subgroup *i* in the decomposition of the generalized entropy family I_c (see equation 1.5) is given by $\omega_i(\underline{n}, \underline{\lambda}) = (n_i/n)\big/(\lambda/\lambda_i)^c$. The corresponding weights in the decomposition of I_θ and I_V (see equations 1.11 and 1.12) are given by $\omega_i(\underline{n}, \underline{\lambda}) = \left(n_i e^{\theta\lambda_i}\right)\big/\left(n e^{\theta\lambda}\right)$ and n_i/n respectively. Evidently, the sum of these weights across subgroups becomes unity only for the two Theil indices and the variance.

If there is a progressive transfer of income between two persons in a subgroup then inequality within the subgroup decreases without affecting between-group inequality. But polarization increases because of higher homogeneity/identification of individuals within a subgroup. Of two subgroups, a proportionate (an absolute) reduction in all incomes of the one with lower mean keeps the subgroup relative (absolute) inequality unchanged but reduces its mean income further. Likewise, a proportionate (an absolute) increase in the incomes of the other subgroup increases its mean but keeps relative (absolute) inequality unaltered. This in turn implies that BI increases. In other words, a greater distancing between subgroup means, keeping within-group inequality unchanged, increases between-group inequality making the subgroups more heterogeneous. A sufficient condition that ensures fulfilment of this requirement is that the decomposition coefficient $\omega_i(\underline{n}, \underline{\lambda})$ depends only on n_i/n. The only subgroup decomposable indices for which this condition holds are the Theil mean logarithmic deviation index I_{ML}, which corresponds to $c = 0$ in (1.5)[10], and the variance. We denote the set $\{I_{ML}, I_V\}$ of these two

[10] Buourguignon (1979) developed a characterization of I_{ML} using $\omega_i(\underline{n}, \underline{\lambda}) = n_i/n$.

indices by SD . For further analysis, we restrict our attention to the set SD . Note that the members of SD are onto functions and they vary continuously over the entire non-negative part of the real line. (It may be mentioned here that the Esteban-Ray (2005) discussion on the Kanbur-Zhang index is based on the functional form I_{ML} .) We also assume throughout the chapter that the number of subgroups (k) is exogenously given.

3.3 The Polarization Quasi-ordering

Following our discussion in Section 3.1, we define a polarization index P as a real valued function of income distributions of arbitrary number of subgroups of a population, partitioned with respect to some homogeneous characteristic. Formally,

Definition 3.1: By a polarization index we mean a continuous function $P:\Omega \rightarrow R^1$, where

$$\Omega = \bigcup_{k\in\Gamma}\left(\prod_{n_i\in\Gamma,1\le i\le k} D^{n_i}\right).$$

For any $x=(x^1,x^2,...,x^k)\in\Omega$, $k\in\Gamma$,the real number $P(x)$ indicates the level of polarization associated with x .

Often economic indicators abbreviate the entire income distribution in terms of two or more characteristics of the distribution. For instance, a 'reduced-form' welfare function expresses social welfare as an increasing function of efficiency (mean income) and a decreasing function of inequality (see Ebert, 1987; Amiel and Cowell, 2003 and Chakravarty, 2009, 2009a). Likewise, we have

Definition 3.2: A polarization index P is called abbreviated or reduced-form if for all $x=(x^1,x^2,...,x^k)\in\Omega$, $k\in\Gamma$, $P(x)$ can be expressed as $P(x)=f(BI(x),WI(x))$, where $I\in SD$ is arbitrary and the real valued function f defined on R_+^2 is continuous.

We refer to the function f considered above as a characteristic function. Clearly, the polarization index defined above will be a relative or an absolute index depending on whether we choose I_{ML} or I_V as the inequality index.

Since the characteristics 'identification' and 'alienation' are regarded as being intrinsic to the concept of polarization, in order to take them into account correctly we assume that the function f is monotonic, that is, it is increasing in BI and decreasing in WI. Such polarization indices are called feasible. Formally,

Definition 3.3: A reduced–form polarization index $P(x) = f(BI(x), WI(x))$, where $I \in SD$, $x = (x^1, x^2, ..., x^k) \in \Omega$, $k \in \Gamma$ are arbitrary and the real valued function f defined on R_+^2 is continuous, is called feasible if f is increasing in BI and decreasing in WI.

It will now be worthwhile to compare the index presented in Definition 3.3 with the Esteban-Ray (1994) index, which is given by

$$ER(p, z) = A \sum_{i=1}^{k} \sum_{j=1}^{k} p_i^{1+\kappa} p_j \left| z_i - z_j \right|,$$

where z_i is the representative income, defined in an unambiguous way, of subgroup *i*, p_i is its population size, z is the vector of z_i's, $A > 0$ is a constant and $\kappa \in (0, 1.6]$ (see (1.45)). A positive value of κ, and hence the identification function p_i^{κ}, underlines the differences between inequality and polarization. On the other hand, the distance function $\left| z_i - z_j \right|$ is an indicator of the alienation component. Clearly, both identification and alienation are directly related to ER. Thus, while in our case, identification is formulated in terms of inverse within-group inequality, in the case of ER, it is a function of population proportions. In contrast, in both cases, the alienation component is based on income distances. While ER directly incorporates the subgroup-sizes, in the reduced-form index the subgroup-sizes are taken into account in the within-group component of inequality. (See the definition of the family SD in Section 3.2.) Thus, for the latter, identification is formulated involving both subgroup-sizes and subgroup inequality levels.

In the Esteban-Ray framework, the postulates are formulated in terms of population shift and minimum polarization arises when there is perfect homogeneity in the sense that the entire population is concentrated in a subgroup, that is, identification is maximum. In the reduced-form set up the notion of polarization is based on inequality indices and therefore, the postulates involve, among other conditions, scaling/ translation of incomes and redistribution of incomes. The minimum polarization arises in this case when both alienation and identification are minimum, that is, when $BI = 0$ and WI is maximum. In the ER-case, polarization is maximized when the population is equally split into two subgroups and the remaining subgroups have zero population-size, whereas in our case, maximum polarization arises if identification is maximized $(WI = 0)$ and alienation (BI) is also maximum. Thus, while for the ER -case, these extreme situations are specified in terms of population concentration, in the present case, they are consequences of income concentration. These differences arise because of different basic formulations.

Note that as the number of subgroups increases and k ends up in n, each individual constitutes a subgroup. Since for the concept of subgroup inequality to be defined, there should be at least two persons in a subgroup, within-group inquality is undefined. That is, now there is only one subgroup, the entire population. Consequently, inequality is represented only by the between-group term, a direct indicator of polarization. Thus, in this polar case in the absence of identification component inequality and polarization are directly related. In fact, Esteban and Ray (1994) also did not 'claim that the notion of polarization always conflicts with that of inequality (op. cit., p.825)'.

There are some more differences between our approach and *ER*-approach. For instance, in the *ER*-approach, the impact of merger of two equally-sized groups at the midpoint will depend on the shape of the entire distribution. However, in the Zhang-Kanbur set up, this will lead to reduction of inequality as well as polarization. This difference arises because while the latter looks at polarization simply in terms of identification and alientation with a fixed number of groups, the former allows variability of groups as well as shifts of populations across groups. While our objective is definitely not to supplant the *ER* -index, we see a clear merit in the Zhang-Kanbur approach given that the number of groups as well as group sizes are fixed, because it takes into account the alienation and identification factors in a very easy and intuitive way. Since polarization is a multifaceted phenomenon, our attempt to look at polarization from a different perspective appears to be quite sensible. Intersting on this issue is the remark by Nissanov et al. (2011): "The *ZK*-measure and the *ER* measure seem to be complementary measures of polarization since the former is able to capture the effects of within-group inequality (that the *ER* measure leaves out of the analysis) while the latter performs better when no changes in within-group inequality are observed."

3.3.1. The Quasi-ordering

In order to develop a polarization quasi-ordering of the income distributions, consider the distributions $x = (x^1, x^2,..., x^k), y = (y^1, y^2,..., y^k) \in \prod_{i=1}^{k} D^{n_i}$, where $k \geq 2$, $n_i \geq 2$, $1 \leq i \leq k$ are arbitrary. Then we say that x is more polarized than y, what we write $x \succ_P y$, if $P(x) > P(y)$ for all feasible polarization indices $P : \prod_{i=1}^{k} D^{n_i} \to R^1$. Our definition of $\succ_P$ is general in the sense that we do not assume equality of the total income of the distributions.

As we have noted in the previous section, given $y = (y^1, y^2,..., y^k) \in \prod_{i=1}^{k} D^{n_i}$, we can generate $x = (x^1, x^2,..., x^k) \in \prod_{i=1}^{k} D^{n_i}$, which is more polarized than y, by one of the following three polarization increasing transformations: (i) decreasing WI (keeping BI unchanged), (ii) increasing BI (keeping WI unchanged), and (iii) decreasing WI and increasing BI. We can write these three conditions more compactly as $BI(x) \geq BI(y)$ and $WI(x) \leq WI(y)$ with strict inequality in at least one case. The following theorem demonstrates equivalence of this with $x \succ_P y$.

Theorem 3.1: Let $x = (x^1, x^2,..., x^k), y = (y^1, y^2,..., y^k) \in \prod_{i=1}^{k} D^{n_i}$, where $k \geq 2$, $n_i \geq 2$, $1 \leq i \leq k$, are arbitrary. Then the following conditions are equivalent:

(i) $x \succ_P y$.

(ii) $BI(x) \geq BI(y)$ and $WI(x) \leq WI(y)$ for any inequality index I in SD, with strict inequality in at least one case.

Proof: See Appendix.

What Theorem 3.1 says is the following: if condition (ii) holds then we can unambiguously say that distribution *x* is regarded as more polarized than distribution *y* by all reduced-form polarization indices that are increasing in BI and decreasing in WI. Note that we do not require equality of the mean incomes of the distributions for this result to hold. Clearly, condion (ii) in the theorem can be verified easily.

3.3.2. Discussion

The polarization quasi-ordering defined in the above theorem is a quasi-ordering – it is transitive but not complete. To see this, consider the bi-partitioned distributions $x = ((1,3,5),(2,6))$ and $y = ((1,3,5),(2,4))$. Let us choose I_V as the index of inequality and denote its between and within-group components by BI_V and WI_V respectively. Then $BI_V(x) = (6/25), BI_V(y) = 0$. Also $WI_V(x) = (16/5)$, $WI_V(y) = 2$. Thus, we have $BI_V(x) > BI_V(y)$ and $WI_V(x) > WI_V(y)$. This shows that the distributions x and y are not comparable with respect to $\succ_P$ and hence $\succ_P$ is not a complete ordering. Next, suppose that for three distributions x, y and z, partitioned with respect to the same characteristic into equal number of subgroups, we have $x \succ_P y$ and $y \succ_P z$. Then it is easy to check that $x \succ_P z$ holds, which demonstrates transitivity of $\succ_P$.

Now, to see that inequality quasi-ordering of income distributions is different from polarization quasi-ordering, consider the bi-partitioned distributions $y = ((a,c),(b,d))$ and $x = ((a, c-\varepsilon),(b+\varepsilon, d))$, where $a < b < c < d$ and $0 < \varepsilon < (c-b)/2$. Then it is easy to see that $BI_V(y) < BI_V(x)$ but $WI_V(y) > WI_V(x)$. Hence for all feasible polarization indices P, we have $P(y) < P(x)$. But by the Pigou-Dalton transfers principle, $I_V(y) > I_V(x)$. Next, let us consider the income distribution $x = (x^1, x^2, \ldots, x^k) \in \prod_{i=1}^{k} D^{n_i}$ and generate the distribution $y = (y^1, y^2, \ldots, y^k)$ from x by the following transformation: $y^i = x^i$ for all $i \neq j$ and y^j is obtained from x^j by a progressive transfer of income between two persons in subgroup j. By construction, $BI(x) = BI(y)$ and $WI(y) < WI(x)$, where $I \in SD$. This in turn implies that for any orderis $I(x) > I(y)$. Thus, in these two cases polarization and inequality rank the distributions in completely opposite ways. The intuitive reasoning behind this is that while each of the two components BI and WI is related to inequality in an increasing manner, for polarization the former has an increasing relationship but for the latter the relationship is a decreasing one. It should be evident that polarization quasi-ordering will depend on the way partitioning of the population is done. For instance, with ethnic group partitioning, one population may be regarded as more polarized than another while for geographic location partitioning the reverse situation may arise. This is natural because the identification of the subgroups depends on the characteristic on the basis of which the partitioning is done.

3.3.3. A Comparison with the Bi-polarization quasi-ordering

To relate $\succ_P$ with the bi-polarization quasi-ordering, which relies on the increased spread and increased bipolarity axioms, suppose that the distributions are partitioned into two subgroups with incomes below and above the median. The increased spread axiom says that polarization should go up under increments (reductions) in incomes above (below) the median. The increased bipolarity axiom, which requires bi-polarization to increase under a progressive transfer of income on the either side of the median, is a bunching or clustering principle.

Note, on the other hand, that in the same set up, alienation refers to increase in the distance between the subgroups below and above the median and this can be achieved by increasing (decreasing) incomes proportionately above (below) the median. Hence, alienation is similar in spirit to the increased spread axiom. Now, a progressive transfer of incomes between two individuals on the same side of the median increases identification. Thus, the increased bipolarity axiom possesses the same flavor as the identification criterion. Hence the two notions of polarization ordering are essentially the same when the two population subgroups are formed using the median[11].

3.4. The Characterization Theorems

A polarization quasi-ordering often may not be able to rank two distributions conclusively. Then in order to look at the directional rankings of the distributions in terms of polarization, it becomes necessary to calculate values of one or more polarization indices. Use of a particular index involves a set of implicit value judgements. We know that a characterization exercise gives us a set of necessary and sufficient conditions for identifying an index uniquely. These conditions, which are referred to as axioms, become helpful in understanding the underlying polarization index in an intuitive way. In other words, characterization of an index

[11] In a recent contribution, Bossert and Schworm (2008) showed that the two-group approach can be interpreted in terms of treating polarization as an aggregate of inverse welfare measures of the two groups under consideration.See also Duclos and Echevin (2005) and Chakravarty et al. (2007) for a related discussion.

enables us to get insight of the implicit value judgements in an explicit manner. These axioms seem to be appropriate for a polarization index in a particular framework.

All the polarization indices considered in this section are assumed to be feasible (as defined in Definition 3.3).

We can very well conceive of a 'threshold level'/ 'tolerance limit' of polarization exceeding which a society becomes turbulent [12]. In this case, a small increment in alienation/identification is likely to escalate tension to a degree, which may generate conflict, as characterized by higher polarization. This is strengthened further by an argument of Esteban and Ray (1994, p. 844) which says that "...when the population is already largely bunched at the two extreme points, further bunching will serve to accentuate polarization." It is likely that the net increment in polarization will not be lower for a society characterized by a higher level of conflict/ polarization. Now, the tolerance limit is likely to vary from society to society, particularly, for a highly peaceful society it is expected to be quite low. This, therefore, permits us to assume that the change in polarization is non-decreasingly related to alienation and identification over the entire domain. As we have said, while in the Esteban-Ray set up the axioms are based on population concentration, in our case the notion of polarization is based on income concentration between and within-groups. Consequently, for the latter polarization change should be related to inequality change.

The following two axioms can now be stated:

$(A1)$ For all $x = \left(x^1, x^2, ..., x^k\right) \in \Omega$, $k \in \Gamma$ and for any non-negative α, $f(BI(x) + \alpha, WI(x)) - f(BI(x), WI(x)) = \psi(BI(x), WI(x))g(\alpha)$ for some continuous functions $\psi : R_+^2 \to R_+^1$ and $g : R_+^1 \to R_+^1$, where ψ is non-decreasing in its first argument, g is increasing, $g(0) = 0$ and $I \in SD$.

$(A2)$ For all $x = \left(x^1, x^2, ..., x^k\right) \in \Omega$, $k \in \Gamma$ and for any non-negative β, $f(BI(x), WI(x)) - f(BI(x), WI(x) + \beta) = \varphi(BI(x), WI(x))h(\beta)$ for some continuous functions φ :

[12] The term 'tolerance limit' is borrowed from the theory of Statistical Quality Control.

$R_+^2 \to R_+^1$ and $h : R_+^1 \to R_+^1$, where φ is non-increasing in its second argument, h is increasing, $h(0) = 0$ and $I \in SD$.

Clearly, these two axioms specify the rate of increase in BI and that of decrease in WI respectively in a specific but very simple way. Axiom $(A1)$ says that increment in polarization resulting from an increase in *BI* by the amount α is proportional to an increasing transform of α . More precisely, it stipulates that the increment can be decomposed into two continuous factors, one a non-negative function of α alone and the other a non-negative valued function of *BI* and *WI*, which is non-decreasing in BI . In other words, given differentiability of the function f , the polarization index becomes convex in *BI*. Increasingness of the function g reflects the view that polarization is increasing in *BI*. The assumption $g(0) = 0$ ensures that if there is no change in *BI*, there will be no change in the value of the polarization index (assuming that *WI* remains unaltered). Given other things, with a higher value of α , there will be more increment in alienation. Axiom $(A2)$ can be explained similarly. The functions g and h may be interpreted respectively as alienation and identification sensitivity functions.

It may be worthwhile to note that decompositions of the type specified in axioms $(A1)$ and $(A2)$ can as well be satisfied by some bipolarization indices. To see this, consider the distribution $x = (x_1, x_2, x_3 = m, x_4, x_5)$, where x_i 's are non-decreasingly ordered and m is the median. Now, consider the bipolarization index $Q(x) = 1 - \left(\frac{1}{5}\right) \exp\left\{-\sum_{i=1}^{5} \left|x_i - m\right|\right\}$. This absolute, symmetric index of bipolarization satisfies the increased spread and increased bipolarity axioms. It takes on the value 0 when the income distribution is perfectly equal. Next, suppose that the distribution y is obtained from the distribution x by increasing the highest income x_5 by an amount $c > 0$, that is, $y_i = x_i$, for $1 \le i \le 4$ and $y_5 = x_5 + c$. Then the change $Q(y) - Q(x)$ can be expressed as the product $\frac{1}{5} \exp\left\{-\sum_{i=1}^{5} \left|x_i - m\right|\right\}\{1 - \exp(-c)\}$. That is, the change has been decomposed into two components, one depends on the original distribution x and other on the increment c .

Often we may need to assume that a polarization index is normalized, that is, for a perfectly equal distribution the value of the polarization index is zero. Formally,

$(A3)$ For arbitrary $k \in \Gamma$, if $x = \left(x^1, x^2, ..., x^k\right) \in \Omega$ is of the form $x^i = c1^{n_i}$, where $n_i \in \Gamma$ for all $1 \le i \le k$ and $c > 0$ is a scalar, then for any $I \in SD$, $f\left(BI(x), WI(x)\right) = 0$.

Since for a perfectly equal distribution x, $BI(x) = WI(x) = 0$, we may restate axiom $(A3)$ as $f(0,0) = 0$.

The following theorem can now be stated.

Theorem 3.2: Assume that the characteristic function is continuously differentiable. Assume also that the right partial derivative of the characteristic function at zero with respect to each argument exists and is positive for the first argument and negative for the second argument. Then a feasible polarization index $P : \Omega \to R^1$ with such a characteristic function satisfies axioms $(A1)$, $(A2)$ and $(A3)$ if and only if it is of one of the following forms for some arbitrary positive constants c_1 and c_2:

$(i)\ P_1(x) = c_1 BI(x) - c_2 WI(x)$,

$(ii)\ P_2(x) = \dfrac{c_1}{\log a}\left(a^{BI(x)} - 1\right) - c_2 WI(x),\ a > 1$,

$(iii)\ P_3(x) = \left(a^{BI(x)} - 1\right)\left(\dfrac{c_1}{\log a} + \rho WI(x)\right) - c_2 WI(x),\ 0 < a < 1, -c_2 \le \rho \le 0$,

$(iv)\ P_4(x) = c_1 BI(x) - \dfrac{c_2}{\log b}\left(b^{WI(x)} - 1\right),\ b > 1$,

$(v)\ P_5(x) = c_1 BI(x) - \left(b^{WI(x)} - 1\right)\left(\dfrac{c_2}{\log b} + \sigma BI(x)\right), 0 < b < 1, -c_1 \le \sigma \le 0$,

$(vi)\ P_6(x) = \dfrac{c_1}{\log a}\left(a^{BI(x)} - 1\right) - \dfrac{c_2}{\log b}\left(b^{WI(x)} - 1\right),\ a > 1, b > 1$,

$(vii)\ P_7(x) = \dfrac{c_1}{\log a}\left(a^{BI(x)} - 1\right) - \dfrac{c_2}{\log b}\left(b^{WI(x)} - 1\right) + \eta\left(a^{BI(x)} - 1\right)\left(b^{WI(x)} - 1\right),\ a > 1, 0 < b < 1$, $0 \le \eta \log a \le c_1$,

$(viii)\ P_8(x) = \dfrac{c_1}{\log a}\left(a^{BI(x)} - 1\right) - \dfrac{c_2}{\log b}\left(b^{WI(x)} - 1\right) + \eta\left(a^{BI(x)} - 1\right)\left(b^{WI(x)} - 1\right),\ 0 < a < 1, b > 1$, $-c_2 \le \eta \log b \le 0$,

(ix) $\quad P_9(x) = \frac{c_1}{\log a}\left(a^{BI(x)} - 1\right) - \frac{c_2}{\log b}\left(b^{WI(x)} - 1\right) + \eta\left(a^{BI(x)} - 1\right)\left(b^{WI(x)} - 1\right)$, $\quad 0 < a, b < 1,$

$\frac{c_1}{\log a} \leq \eta \leq -\frac{c_2}{\log b}$,

where $x = \left(x^1, x^2, ..., x^k\right) \in \Omega$, $k \in \Gamma$ and $I \in SD$ are arbitrary.

Proof: See Appendix.

In Theorem 3.2 the only assumptions we make about f are its continuous differentiability and existence of partial derivatives at the end point 0. Many economic indicators satisfy these assumptions. It is known that if the partial derivatives exist at the end point 0, then they are right partial derivatives (Rudin, 1987, p.104).

The constants c_1 and c_2 reflect the importance of alienation and identification in the aggregation. They can be interpreted as scale parameters in the sense that, given other things, an increase in c_1 increases polarization. Likewise, ceteris paribus, if c_2 decreases then polarization increases. The other parameters can be interpreted similarly. For $c_1 = c_2 = 1$, P_1 becomes the Rodriguez-Salas index of polarization, if we subdivide the population into two non-overlapping groups using the median and use the Donaldson-Weymark S-Gini index $I_{\hat{\varepsilon}}(x) = 1 - \sum_{i=1}^{n}\left(i^{\hat{\varepsilon}} - (i-1)^{\hat{\varepsilon}}\right)\hat{x}_i \Big/ \lambda n^{\hat{\varepsilon}}$ as the index of inequality , where $\hat{\varepsilon} > 1$ is an inequality sensitivity parameter and $\hat{x} = (\hat{x}_1, \hat{x}_2, ..., \hat{x}_n)$ is that permutation of x such that $\hat{x}_1 \geq \hat{x}_2 \geq .... \geq \hat{x}_n$. For $\hat{\varepsilon} = 2$, $I_{\hat{\varepsilon}}$ becomes the Gini index. In the Rodriguez-Salas case for P_1 to increase under a progressive transfer on the same side of the median, it is necessary that $2 \leq \hat{\varepsilon} \leq 3$.

However, Rodriguez-Salas index regards all income distributions that have equal between-group and within-group components of inequality as equally polarized. Thus, a distribution x with $BI(x) = WI(x) = .3$ becomes equally polarized as the equal distribution y with $BI(y) = WI(y) = 0$. Therefore, in situations of the type where $BI = WI$, P_1 can avoid this problem if we make different choices of c_1 and c_2. The same remark applies to the choices of a_1 and a_2 in the normalized ratio form index $P_{a_1, a_2}(x) = \left(\frac{a_1^{BI(x)}}{a_2^{WI(x)}} - 1\right)$, which is obtained as a particular case of P_7 as follows. If in P_7 we set $\frac{c_1}{\log a} = -\frac{c_2}{\log b} = \eta = 1$, then on simplification we get $P_7(x) = a^{BI(x)} b^{WI(x)} - 1$, which we can rewrite as $P_7(x) = \left(\frac{a_1^{BI(x)}}{a_2^{WI(x)}} - 1 \right)$, where $a_1 = a > 1$ and $1/b = a_2 > 1$. Therefore for suitable choices of the parameters we get the normalized ratio form index $\left(\frac{a_1^{BI(x)}}{a_2^{WI(x)}} - 1 \right)$ as a special case of P_7.

In order to discuss eventual differences among the indices $P_1 - P_9$, we look at the following properties.

Property 1: *P* is strictly convex in *BI.*

Property 2: *P* is strictly concave in *WI.*

These propertie underline the choice of the policy-maker in fixing up the rate of increase in identification and alienation factors. It is readily seen that P_1 satisfies none of these properties (and hence can be seen as a rather 'weak' indicator); P_2 and P_3 satisfy the first property, but not the second one; P_4 and P_5 obey Property 2, but not Property 1 while each one of the indices $P_6 - P_9$ meets both the properties (and so, they can be considered as 'strong' indicators). Indices $P_7 - P_9$ are identical; they vary only in terms of the restrictions on the parameters.

However, if two distributions x and y can be ranked unambiguously by the quasi-ordering discussed in Section 3.3, then from quasi-ordering perspective essentially no difference arises among the indices characterized in Theorem 3.2.

In order to demonstrate independence of the three axioms, we need to construct indicators of polarization that will fulfill any two of the three axioms but not the remaining one. The feasible characteristic function $f_1(s,t) = (s - t^2)$ satisfies axioms $(A1)$ and $(A3)$ but not axiom $(A2)$. Likewise, the feasible characteristic function $f_2(s,t) = (s^2 - t)$ fulfills axioms $(A2)$ and

$(A3)$ but not axiom $(A1)$. Finally, the feasible characteristic function $f_3(s,t)=(s-t-1)$ is a violator of axiom $(A3)$ but not of axioms $(A1)$ and $(A2)$. We can therefore state the following:

Remark 3.1: Axioms $(A1)$, $(A2)$ and $(A3)$ are independent.

For the index given by (i) the ratio c_2/c_1 is the marginal rate of substitution of alienation for identification along an iso-polarization contour. This ratio shows how WI can be traded off for BI along the contour. In fact, we can take this trade-off into account in a more general way through some changes in the original distribution. Suppose all the incomes in the subgroup with the minimum subgroup mean are proportionately scaled down or reduced by the same absolute amount. Because of increased differences in subgroup means BI, that is, alienation increases, by some amount δ, say. The resulting increase in polarization can be compensated by a decrease in identification through a sequence of regressive transfers within one or more subgroups. Since the corresponding reduction in identfication depends on the size of δ, we denote it by $g_1(\delta)$. That is, because of an increase in BI by δ, for keeping the level of polarization unaltered it becomes necessary to increase WI by some amount $g_1(\delta)$. By a similar argument, if WI increases by δ then a corresponding positive change in BI by $g_2(\delta)$, say, will be necessary to keep level of polarization constant (see also Esteban and Ray, 1994, p.828, pp.845-6 and Chakravarty and D'Ambrosio, 2010, for a related discussion). Formally,

$(A4)$ For all $x=(x^1,x^2,\ldots,x^k)\in\Omega$, $k\in\Gamma$ and for any non-negative δ, $f(BI(x),WI(x))=$ $f(BI(x)+\delta,WI(x)+g_1(\delta))$ $=$ $f(BI(x)+g_2(\delta),WI(x)+\delta)$ for some continuous functions $g_1,g_2:R^1_+\to R^1_+$.

Using axiom $(A4)$ we can develop a joint characterization of the normalized ratio form index P_{a_1,a_2} and the difference form index P_1. This is shown below.

Theorem 3.3: Assume that the characteristic function is continuously differentiable. Assume also that the right partial derivative of the characteristic function at zero with respect to the first argument exists and is positive. Then a feasible polarization index $P:\Omega\to R^1$ with such a characteristic function satisfies axioms $(A1)$ (or $(A2)$), $(A3)$ and $(A4)$ if and only if it is of one of the following forms:

$(i)\, P_{c_1,c_2}(x) = c_1 BI(x) - c_2 WI(x)$ for some arbitrary constants $c_1, c_2 > 0$,

$(ii)\, P_{a_1,a_2}(x) = c\left(\dfrac{a_1^{BI(x)}}{a_2^{WI(x)}} - 1\right)$ for some arbitrary constants $c > 0$, $a_1, a_2 > 1$,

where $x = (x^1, x^2, ..., x^k) \in \Omega$, $k \in \Gamma$ and $I \in SD$ are arbitrary.

Proof: See Appendix.

Since the constants c_1 and c_2 in the above theorem are arbitrary, we can choose them to be equal to the corresponding constants in Theorem 3.2 and therefore use the same notation. The same remark applies for the constants a_1 and a_2.

To check independence of axioms $(A1)$, $(A3)$ and $(A4)$, consider the characteristic functions f_1, f_3 (as defined earlier) and $f_4(s,t) = \left(2^{s-t} + s - t - 1\right)$. Then f_1 satisfies axioms $(A1)$ and $(A3)$ but not axiom $(A4)$, f_3 is a violator of axiom $(A3)$ but not of the other two, while f_4 fulfills all the axioms except $(A1)$. We therefore have

Remark 3.2: Axioms $(A1)$, $(A3)$ and $(A4)$ are independent.

Again, the characteristic function f_2 meets axioms $(A2)$ and $(A3)$ but not $(A4)$. On the other hand f_3 violates axiom $(A3)$ but not the remaining two. Finally, f_4 fulfills all the axioms except $(A2)$. This enables us to state the following:

Remark 3.3: Axioms $(A2)$, $(A3)$ and $(A4)$ are independent.

The transformed ratio form index $\left(1 + P_{a_1,a_2}\right)$ has a structure similar to the Zhang-Kanbur index $P_{ZK}(x) = BI(x)/WI(x)$. However, one minor problem with P_{ZK} is its discontinuity if $WI(x) = 0$. The transformed index and hence P_{a_1,a_2} do not suffer from this shortcoming. However, the alienation and identification components of polarization are incorporated correctly in the formulation of P_{ZK}.

In the literature on income-inequality measurement, it is a common practice to relate an inequality index with a welfare function in a negative monotonic way and vice -versa. For instance, we may define the welfare function U associated with any inequality index I defined on D as $U(x) = \lambda(x) e^{-I(x)}$. When efficiency considerations are absent, that is, the mean income

$\lambda(x)$ is fixed, an increase in inequality is equivalent to a reduction in welfare and vice-versa. A proportionate or an absolute increase in all incomes will increase U depending on whether I is a relative or an absolute index (see Shorrocks, 1988 and Chakravarty, 2009). Note also that given a functional form of U, we can generate the form of the inequality index I. In a similar attempt, Chakravarty et al. (1985) determined the functional form of the underlying social welfare function from the knowledge of the ethical income mobility index suggested by them.

Likewise, a similar problem can be the issue of generating an inequality index from a specific polarization index. More precisely, for a polarization index with a particular structure, we identify one possible corresponding subgroup decomposable inequality index. In other words, given the polarization index, we determine the functional form of the underlying subgroup decomposable inequality index by constructing an appropriate algorithm. Thus, we may regard the problem as the dual of generating polarization indices from inequality indices. For this purpose we assume at the outset that for fixed $k \in \Gamma$ and $(n_1, n_2, \ldots, n_k) \in \Gamma^k$, the polarization index $P : \prod_{i=1}^{k} D^{n_i} \to R^1$ satisfies the following axiom:

$(A5)$: For all $x = \left(x^1, x^2, \ldots, x^k\right) \in \prod_{i=1}^{k} D^{n_i}$, $P(y) - P(x) = v_i(\underline{n}, \underline{\lambda}) g\left(x^i\right)$, where $y = \left(y^1, y^2, \ldots, y^k\right)$ with $y^i = \lambda\left(x^i\right) 1^{n_i}$ and $y^j = x^j$ for $j \neq i$; v_i is a positive real number, assumed to depend on the vector $(\underline{n}, \underline{\lambda})$ and g is a non-negative valued function defined on $\bigcup_{i=1}^{k} D^{n_i}$.

Note that we are not assuming here that the polarization index is feasible. However, it will be demonstrated that feasibility drops out as an implication of our structure. The transformation that takes us from x to y makes the distribution y^i in subgroup i perfectly equal and leaves distributions in all other subgroups unchanged. Given positivity of v_i, axiom $(A5)$ states that the resulting change in polarization, as indicated by $P(y) - P(x)$, is non-negative (since g is non-negative). This is quite sensible. Assuming that x^i is unequal, a movement towards perfect equality makes the subgroup more homogeneous and because of closer identification of the individuals in the subgroup, polarization should not decrease. Since the transformation does not affect the distributions in all subgroups other than subgroup i, we are assuming that the change does not depend on unaffected subgroups' distributions. However, it is assumed to depend on x^i, the original distribution in subgroup i, and the vectors of population sizes of the subgroups and their mean incomes.

Theorem 3.4: If the continuous polarization index $P : \prod_{i=1}^{k} D^{n_i} \to R^1$ satisfies axiom $(A5)$, then there exists a corresponding subgroup decomposable continuous inequality index $I : \left(\prod_{i=1}^{k} D^{n_i} \right) \cup \left(\bigcup_{i=1}^{k} D^{n_i} \right) \to R_+^1$ of the type $I\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots\ldots, \lambda_k 1^{n_k}\right) + \sum_{i=1}^{k} \omega_i(\underline{n}, \underline{\lambda}) I\left(x^i\right)$ which takes on the value zero for the perfectly equal distribution on $\bigcup_{i=1}^{k} D^{n_i}$.

Proof: See Appendix.

Note that Axiom $(A5)$ does not say anything about the identification and alienation factors of P. However, using Theorem 3.4, we can clearly extract them since the retrieved index is subgroup decomposable.

Remark 3.4: From equation (3.44) in the appendix we observe that P can be expressed as $(c_1 BI - c_2 WI)$ for some subgroup decomposable inequality index I that becomes zero for the perfectly equal distribution on $\bigcup_{i=1}^{k} D^{n_i}$, where $c_1, c_2 > 0$ are arbitrary constants. Therefore, it is a feasible index of polarization for the inequality index defined in equation (3.43) in the appendix.

Remark 3.5: Since Theorem 3.4 is concerned with the existence of a subgroup decomposable inequality index, we have considered an inequality index that can be generated by an algorithm from the polarization index satisfying Axiom $(A5)$ and which satisfies subgroup-decomposability. If we assume that $\omega_i(\underline{n}, \underline{\lambda}) = v_i(\underline{n}, \underline{\lambda})/c_2$ depends only on n_i/n, then given the domain, this inequality index is a member of SD. Furthermore, I will be symmetric whenever P and g are. Finally, if g takes on positive values for all distributions which are not perfectly equal, then I will satisfy **NON** also.

Remark 3.6: Since $\left(\prod_{i=1}^{k} D^{n_i}\right) \cup \left(\bigcup_{i=1}^{k} D^{n_i}\right)$ is a closed subset of D and I is continuous, I can be continuously extended to D (Rudin, 1987, p.99). (Here we assume that D can be identified with $\bigcup_{\substack{m_j \in \Gamma, 1 \le j \le l, \\ l \in \Gamma}} \left(\prod_{j=1}^{l} D^{m_j}\right) \cup \left(\bigcup_{j=1}^{l} D^{m_j}\right)$.)

3.5 Conclusion

Polarization is concerned with clustering of incomes in subgroups of a population, where the partitioning of the population into subgroups is done in an unambiguous way. A reduced-form polarization index is one which abbreviates an income distribution in terms of 'alienation' and 'identification' components of polarization. The between-group term of a subgroup decomposable inequality index is taken as an indicator of alienation, whereas within-group inequality is regarded as an inverse indicator of identification. A criterion for ranking different income distributions by all reduced-form indices is developed under certain mild conditions. Some polarization indices have been characterized using alternative sets of independent axioms. Finally, the dual problem of generating an index of inequality from a given form of polarization index is investigated.

3.6 Appendix

Proof of Theorem 3.1: Suppose $x \succ_P y$ holds. Consider the polarization index $P_\varepsilon(x) = BI(x) - \varepsilon WI(x)$, where $\varepsilon > 0$ is arbitrary. By definition, $P_\varepsilon(x)$ is a feasible index. Now, $P_\varepsilon(x) > P_\varepsilon(y)$ implies that $BI(x) - BI(y) > \varepsilon(WI(x) - WI(y))$. Since $\varepsilon > 0$ is arbitrary, letting $\varepsilon \to 0$, we get $BI(x) \ge BI(y)$.

Next, consider the feasible index $P'_\varepsilon(x) = \varepsilon BI(x) - WI(x)$, where $\varepsilon > 0$ is arbitrary. Then $P'_\varepsilon(x) > P'_\varepsilon(y)$ implies that $WI(x) - WI(y) < \varepsilon(BI(x) - BI(y))$. Again because of arbitrariness of $\varepsilon > 0$, we let $\varepsilon \to 0$ and find that $WI(x) \le WI(y)$.

Now, at least one of the inequalities $BI(x) \ge BI(y)$ and $WI(x) \le WI(y)$ has to be strict. This is because if $BI(x) = BI(y)$ and $WI(x) = WI(y)$, then $P(x) = f(BI(x), WI(x)) =$ $f(BI(y), WI(y))$, that is, $P(x) = P(y)$, which contradicts the assumption $x \succ_P y$.

The proof of the converse follows from the defining condition of the feasible polarization index, that is, increasingness in the first argument and decreasingness in the second argument.□

Proof of Theorem 3.2: Since the components of the two inequality indices considered are onto functions, we can restate axioms $(A1)$ and $(A2)$ as follows:

$$f(s+\alpha,t)-f(s,t)=\psi(s,t)g(\alpha), \tag{3.1}$$

$$f(s,t)-f(s,t+\beta)=\varphi(s,t)h(\beta), \tag{3.2}$$

where $s,t,\alpha,\beta \geq 0$ are arbitrary. Putting $s=0$ in (3.1) and assuming positivity of α we get

$$f(\alpha,t)-f(0,t)=\psi(0,t)g(\alpha). \tag{3.3}$$

For a fixed $t \in R^1_+$, define $f_t : R^1_+ \to R^1$ by $f_t(s)=f(s,t)$, where $s \geq 0$. Then continuous differentiability of f implies that f_t is also continuously differentiable and moreover, it is increasing. Further, by assumption, $f_t'(0)>0$ which implies that $f(\alpha,t)>f(0,t)$ for all $\alpha>0$. Also, by increasingness of g we have, $g(\alpha)>g(0)=0$. This, along with (3.3) yields: $\psi(0,t)>0$ for all $t \in R^1_+$. Hence, for all $s,t \in R^1_+$ we have, $\psi(s,t) \geq \psi(0,t) > 0$.

From (3.1) and (3.3) it then follows that

$$\frac{f(s+\alpha,t)-f(s,t)}{f(\alpha,t)-f(0,t)}=\frac{\psi(s,t)}{\psi(0,t)}, \tag{3.4}$$

for all $s,t \geq 0$.

We rewrite (3.4) in terms of f_t as follows:

$$\frac{f_t(s+\alpha)-f_t(s)}{f_t(\alpha)-f_t(0)}=\frac{\psi(s,t)}{\psi(0,t)}. \tag{3.5}$$

Note that the right hand side of (3.5) is independent of α . So we can divide the denominator and numerator of the left hand side of (3.5) by α and take the limit of the resulting expressions as $\alpha \to 0$. Then (3.5) becomes

$$\frac{f_t'(s)}{f_t'(0)}=\frac{\psi(s,t)}{\psi(0,t)}, \tag{3.6}$$

where f_t' stands for the derivative of f_t. By assumption the right hand side of (3.6) is positive. This along with positivity of $f_t'(0)$ (by assumption) implies that $f_t'(s) > 0$ for all $s \geq 0$. From this it follows that $\frac{\partial f(s,t)}{\partial s} > 0$ for all $s,t \geq 0$.

Because of independence of the right hand side of (3.5) of α, the derivative of the left hand side of (3.5) with respect to α is zero. This gives $(f_t(\alpha) - f_t(0)) f_t'(s+\alpha) = (f_t(s+\alpha) - f_t(s)) f_t'(\alpha)$, from which it follows that

$$\frac{f_t(s+\alpha) - f_t(s)}{f_t(\alpha) - f_t(0)} = \frac{f_t'(s+\alpha)}{f_t'(\alpha)}. \tag{3.7}$$

Equations (3.5), (3.6) and (3.7) jointly imply that $\frac{f_t'(s+\alpha)}{f_t'(\alpha)} = \frac{f_t'(s)}{f_t'(0)}$, which gives $f_t'(s+\alpha) = (f_t'(s) f_t'(\alpha)) / f_t'(0)$. Define the function $\mu_t : R^1_+ \to R^1$ by $\mu_t(s) = f_t'(s) / f_t'(0)$. Then the previous equation becomes

$$\mu_t(s+\alpha) = \mu_t(s)\mu_t(\alpha) \tag{3.8}$$

for all $s,\alpha \geq 0$. Since f is continuously differentiable, μ_t is continuous. The general nontrivial solution to the functional equation (3.8) is given by $\mu_t(s) = (a(t))^s$ for some continuous function $a : R^1_+ \to R^1_{++}$, where $s \geq 0$ is arbitrary (Aczel, 1966, p.41). Letting $f_t'(0) = w(t)$, we can now write f_t' as $f_t'(s) = (a(t))^s w(t)$ for some continuously differentiable maps $a, w : R^1_+ \to R^1_{++}$. Integrating f_t' we get

$$f_t(s) = \begin{cases} \frac{(a(t))^s w(t)}{\log\ a(t)} + w_1(t),\ a(t) \neq 1, \\ sw(t) + w_1(t),\ a(t) = 1, \end{cases} \tag{3.9}$$

where $s \geq 0$ is arbitrary and $w_1 : R^1_+ \to R^1$ is continuously differentiable. We rewrite (3.9) more explicitly as

$$f(s,t) = \begin{cases} \frac{(a(t))^s w(t)}{\log\ a(t)} + w_1(t),\ a(t) \neq 1, \\ sw(t) + w_1(t),\ a(t) = 1. \end{cases} \tag{3.10}$$

where $s,t \geq 0$ are arbitrary.

We now show that $a(t)$ is a constant for all $t \geq 0$. First, note that there is nothing to prove if $a(t) = 1$ for all $t \geq 0$. If $a(t) \neq 1$ for some $t \geq 0$, then consider the set $B = \{t \geq 0 : a(t) \neq 1\}$, which is assumed to be non-empty. Now, (3.4) along with the first equation in (3.10) implies that for all $t \in B$ and for all $s \geq 0$,

$$\frac{(a(t))^{s+\alpha}\, w(t)}{\log\ a(t)} - \frac{(a(t))^{s}\, w(t)}{\log\ a(t)} = \psi(s,t)g(\alpha). \tag{3.11}$$

Putting $s = 0$ in (3.11) we get $\frac{\left((a(t))^{\alpha} - 1\right)w(t)}{\log\ a(t)} = \psi(0,t)g(\alpha)$, which gives

$$\left((a(t))^{\alpha} - 1\right) = \phi(t)g(\alpha), \tag{3.12}$$

where $\phi(t) = (\psi(0,t)\log\ a(t))/w(t)$ and $t \in B$ is arbitrary. Since by assumption $a(t) \neq 1$ for all $t \in B$, the right hand side of (3.12) is non-zero for all $\alpha > 0$. Substituting $\alpha = 1$ and 2 in (3.12) we get $((a(t)) - 1) = \phi(t)g(1)$ and $\left((a(t))^{2} - 1\right) = \phi(t)g(2)$ respectively. Dividing the right (left) hand side of the second equation by the corresponding side of the first equation, we get $((a(t)) + 1) = g(2)/g(1)$, which implies that for all $t \in B$, $a(t) = -1 + g(2)/g(1) = c$, a positive constant. But $a(t) = 1$ for all nonnegative $t \in B^{c}$, the complement of B. Since $a(t)$ is a continuous map on its domain and B is a non-empty set, B^{c} must be empty. Thus, $a(t) = c$, a positive constant not equal to one, for all $t \geq 0$. Hence in either case $a(t)$ is a constant. In the sequel we will write a in place of $a(t)$.

Therefore, equation (3.10) now can be written as

$$f(s,t) = \begin{cases} \dfrac{a^{s} w(t)}{\log\ a} + w_1(t), 0 < a \neq 1, \\ sw(t) + w_1(t), a = 1, \end{cases} \tag{3.13}$$

where $s, t \geq 0$ are arbitrary, w, w_1 are continuously differentiable and w is positive valued.

Proceeding in a similar manner and making use of axiom $(A2)$ we get

$$f(s,t) = \begin{cases} \dfrac{b^{t} \gamma(s)}{\log\ b} + \gamma_1(s), 0 < b \neq 1, \\ t\gamma(s) + \gamma_1(s), b = 1, \end{cases} \tag{3.14}$$

for some continuously differentiable maps $\gamma, \gamma_1 : R_+^1 \to R$, γ being negative valued. We can also show that $\frac{\partial f(s,t)}{\partial t} < 0$ for all $s, t \geq 0$.

Now, for comparing (3.13) and (3.14) we need to consider various cases.

Case I: $f(s,t) = sw(t) + w_1(t) = t\gamma(s) + \gamma_1(s)$. (3.15)

By axiom $(A3)$, $w_1(0) = \gamma_1(0) = 0$. Putting $s = 0$ in (3.15), we get $w_1(t) = t\gamma(0)$. Likewise, for $t = 0$, we have $sw(0) = \gamma_1(s)$. Substituting these expressions for w_1 and γ_1 in (3.15), we get $sw(t) + t\gamma(0) = t\gamma(s) + sw(0)$, from which it follows that $s(w(t) - w(0)) = t(\gamma(s) - \gamma(0))$. Since this holds for all $s, t \geq 0$, there exists a constant θ such that $w(t) = w(0) + \theta t$ and $\gamma(s) = \gamma(0) + \theta s$. Hence $f(s,t) = s(w(0) + \theta t) + t\gamma(0)$. Differentiating this form of f partially with respect to s and t, we get $\frac{\partial f(s,t)}{\partial s} = (w(0) + \theta t) > 0$ and $\frac{\partial f(s,t)}{\partial t} = (\gamma(0) + \theta s) < 0$. Now, if $\theta > 0$, then negativity of $\frac{\partial f(s,t)}{\partial t}$ cannot hold for all $s \geq 0$. On the other hand, if $\theta < 0$, then positivity of $\frac{\partial f(s,t)}{\partial s}$ cannot hold for all sufficiently large positive t. Hence the only possibility is that $\theta = 0$. Consequently, $f(s,t) = sw(0) + t\gamma(0) = c_1 s - c_2 t$, where $c_1 = w(0) > 0$ and $c_2 = -\gamma(0) > 0$ (by positivity and negativity of partial derivatives of f with respect to s and t respectively, as shown earlier).

Case II: $f(s,t) = \frac{a^s w(t)}{\log a} + w_1(t) = t\gamma(s) + \gamma_1(s), 0 < a \neq 1$. (3.16)

By axiom $(A3)$,

$$\frac{w(0)}{\log a} + w_1(0) = \gamma_1(0) = 0. \tag{3.17}$$

Putting $s = 0$ in (3.16) and using the information $\gamma_1(0) = 0$ from (3.17) in the resulting expression we get $f(0,t) = \frac{w(t)}{\log a} + w_1(t) = t\gamma(0)$. Substituting the expression for $w_1(t)$ obtained from this equation into (3.16) we have

$$f(s,t) = \frac{(a^s - 1)w(t)}{\log a} + t\gamma(0). \tag{3.18}$$

Similarly, putting $t = 0$ in (3.16) we find $\frac{a^s w(0)}{\log a} + w_1(0) = \gamma_1(s)$, which, in view of $w_1(0) = -w(0)/\log a$ (obtained from (3.17)) gives $\gamma_1(s) = \frac{(a^s - 1)w(0)}{\log a}$. Substituting this value of $\gamma_1(s)$ into (3.16) we get

$$f(s,t) = \frac{(a^s - 1)w(0)}{\log a} + t\gamma(s). \tag{3.19}$$

Equating the functional forms of f given by (3.18) and (3.19) we then have $\frac{(a^s - 1)(w(t) - w(0))}{\log a} = t(\gamma(s) - \gamma(0))$, from which it follows that for all $s, t > 0$, $\left(\frac{\gamma(s) - \gamma(0)}{(a^s - 1)/\log a}\right) = \frac{(w(t) - w(0))}{t}$ =constant= θ (say). This gives $\gamma(s) = \gamma(0) + \theta\frac{(a^s - 1)}{\log a}$ for all $s, t \geq 0$, and $w(t) = w(0) + \theta t$. Substitution of the functional form of $\gamma(s)$ into (3.19) yields

$$f(s,t) = \frac{(a^s - 1)(w(0) + \theta t)}{\log a} + t\gamma(0) \ . \tag{3.20}$$

Now, $\frac{\partial f(s,t)}{\partial s} = a^s(w(0) + \theta t) > 0$ for all $s, t \geq 0$. For $s = 0$ this implies that

$$(w(0) + \theta t) > 0 \tag{3.21}$$

holds for all $t \geq 0$. Hence $\theta \geq 0$, otherwise for a sufficiently high value of t, $(w(0) + \theta t)$ will be negative.

Also

$$\frac{\partial f(s,t)}{\partial t} = \theta\frac{(a^s - 1)}{\log a} + \gamma(0) < 0 \ , \tag{3.22}$$

for all $s, t \geq 0$.

Sub-case I: $a > 1$. Then $\frac{(a^s - 1)}{\log a}$ is increasing and unbounded in $s \geq 0$. So if $\theta > 0$, then choosing $s > 0$ sufficiently large, we can make the left hand side of the inequality in (3.22) positive, which is a contradiction. So the only possibility is that $\theta = 0$. Plugging $\theta = 0$ into

(3.20) we get, $f(s,t)=\frac{(w(0))\left(a^{s}-1\right)}{\log a}+t\gamma(0)$, which, in view of our earlier notation, can be rewritten as $f(s,t)=\frac{c_1\left(a^{s}-1\right)}{\log a}-c_2 t$ with $c_1=w(0)>0$ and $c_2=-\gamma(0)>0$.

Sub-case II: $0<a<1$. In this case also (3.21) holds so that $\theta\geq 0$. We rewrite the inequality in (3.22) as $\theta<\frac{\gamma(0)\log a}{\left(1-a^{s}\right)}$ for all $s>0$, which implies that $\theta\leq\gamma(0)\log a$. Using our earlier notation, we have $f(s,t)=\left(a^{s}-1\right)\left(\frac{c_1}{\log a}+\rho t\right)-c_2 t$, where, $c_1=w(0)>0$, $c_2=-\gamma(0)>0$ and $\rho=\theta/\log a$. Also $0\geq\rho=\theta/\log a\geq\gamma(0)=-c_2$.

Case III: $f(s,t)=sw(t)+w_1(t)=\frac{b^{t}\gamma(s)}{\log b}+\gamma_1(s), 0<b\neq 1.$

Solution in this case is similar to that of Case II and (by symmetry) is given by

$$f(s,t)=\begin{cases} c_1 s-c_2\frac{\left(b^{t}-1\right)}{\log b}, b>1, \\ c_1 s-\left(b^{t}-1\right)\left(\frac{c_2}{\log b}+\sigma t\right), 0<b<1, \end{cases}$$

where $c_1, c_2>0$ are same as before and $\sigma\,(-c_1\leq\sigma\leq 0)$ is a constant.

Case IV: $f(s,t)=\frac{a^{s}w(t)}{\log a}+w_1(t)=\frac{b^{t}\gamma(s)}{\log b}+\gamma_1(s), 0<a,b\neq 1,$ (3.23)

for all $s,t\geq 0$.

Applying axiom $(A3)$ to (3.23) we get

$$\frac{w(0)}{\log a}+w_1(0)=0 \text{ and } \frac{\gamma(0)}{\log b}+\gamma_1(0)=0. \tag{3.24}$$

Putting $s=0$ in (3.23) we get $\frac{w(t)}{\log a}+w_1(t)=\frac{b^{t}\gamma(0)}{\log b}+\gamma_1(0)$, which in view of the second equation in (3.24) can be rewritten as $\frac{w(t)}{\log a}+w_1(t)=\frac{\left(b^{t}-1\right)\gamma(0)}{\log b}$. Substituting the value of $w_1(t)$ obtained from this equation into the first expression for $f(s,t)$ in (3.23) we have

$$f(s,t)=\frac{(a^s-1)w(t)}{\log a}+\frac{(b^t-1)\gamma(0)}{\log b}. \tag{3.25}$$

Next, put $t=0$ in (3.23) to get $\frac{a^s w(0)}{\log a}+w_1(0)=\frac{\gamma(s)}{\log b}+\gamma_1(s)$. We solve these two equations to get $\gamma_1(s)=\frac{a^s w(0)}{\log a}+w_1(0)-\frac{\gamma(s)}{\log b}$, which in view of $w_1(0)=-\frac{w(0)}{\log a}$ (from the first equation in (3.24)) gives $\gamma_1(s)=\frac{(a^s-1)w(0)}{\log a}-\frac{\gamma(s)}{\log b}$. Substitution of this form of $\gamma_1(s)$ into the second expression for $f(s,t)$ in (3.23) yields

$$f(s,t)=\frac{(a^s-1)w(0)}{\log a}+\frac{(b^t-1)\gamma(s)}{\log b}. \tag{3.26}$$

Equating (3.25) and (3.26) and simplifying we get

$$\frac{(a^s-1)(w(t)-w(0))}{\log a}=\frac{(b^t-1)(\gamma(s)-\gamma(0))}{\log b}, \tag{3.27}$$

for all $s,t\geq 0$. As in the earlier cases $\gamma(s)=\gamma(0)+\theta\frac{(a^s-1)}{\log a}$ and $w(t)=w(0)+\theta\frac{(b^t-1)}{\log b}$ for some constant θ. Substituting this form of $\gamma(s)$ into (3.26) we get

$$f(s,t)=\frac{(a^s-1)w(0)}{\log a}+\frac{(b^t-1)}{\log b}\left(\gamma(0)+\theta\left(\frac{a^s-1}{\log a}\right)\right). \tag{3.28}$$

Now, $\frac{\partial f(s,t)}{\partial s}>0$ implies that

$$\frac{\theta(b^t-1)}{\log b}+w(0)>0 \tag{3.29}$$

for all $t\geq 0$. On the other hand, $\frac{\partial f(s,t)}{\partial t}<0$ implies that

$$\frac{\theta(a^s-1)}{\log a}+\gamma(0)<0, \tag{3.30}$$

for all $s\geq 0$.

Again various sub-cases come under consideration.

Sub-case I: $a>1, b>1$. Applying the same logic as in the case II, we get $\theta=0$. So the general solution in this case is $f(s,t)=\frac{c_1}{\log a}\left(a^s-1\right)-\frac{c_2}{\log b}\left(b^t-1\right)$, where $c_1=w(0), c_2=-\gamma(0)>0$ are the same as in Case I.

Sub-case II: $a>1, 0<b<1$. Considering (3.30) and noting that $\frac{\left(a^s-1\right)}{\log a}$ is positive and unbounded above we conclude that $\theta \le 0$. From (3.29) we get $\theta>\frac{w(0)\log b}{\left(1-b^t\right)}$ for all $t>0$, which implies that $\theta \ge w(0)\log b$. Thus, the general solution given by (3.28) becomes $f(s,t)=\frac{c_1}{\log a}\left(a^s-1\right)-\frac{c_2}{\log b}\left(b^t-1\right)+\eta\left(a^s-1\right)\left(b^t-1\right)$, where $c_1=w(0)>0$, $c_2=-\gamma(0)>0$ and $\eta=\frac{\theta}{\log a \log b}$, with $0 \le \eta \log a \le c_1$.

Sub-case III: $0<a<1, b>1$. Here using (3.29) we conclude that $\theta \ge 0$. Moreover, from (3.30), $\theta<\frac{\gamma(0)\log a}{\left(1-a^s\right)}$ for all $s>0$, which implies that $\theta \le \gamma(0)\log a$. Thus, $0 \le \theta \le \gamma(0)\log a$. Consequently, $f(s,t)=\frac{c_1}{\log a}\left(a^s-1\right)-\frac{c_2}{\log b}\left(b^t-1\right)+\eta\left(a^s-1\right)\left(b^t-1\right)$, where $c_1=w(0)$ and $c_2=-\gamma(0)$ are positive and $-c_2 \le \eta \log b \le 0$ with $\eta=\frac{\theta}{\log a \log b}$.

Sub-case IV: $0<a<1, 0<b<1$. Applying the same logic as before we get $f(s,t)=\frac{c_1}{\log a}\left(a^s-1\right)-\frac{c_2}{\log b}\left(b^t-1\right)+\eta\left(a^s-1\right)\left(b^t-1\right)$, where $w(0)\log b \le \theta \le \gamma(0)\log a$, which implies that $\frac{c_1}{\log a} \le \eta \le -\frac{c_2}{\log b}$, with $\eta=\frac{\theta}{\log a \log b}$. This completes the necessity part of the proof. The sufficiency is easy to check.□

Proof of Theorem 3.3: We will prove the Theorem for axioms $(A1)$, $(A3)$ and $(A4)$. A similar proof will run if axiom $(A1)$ is replaced by axiom $(A2)$. From the proof of Theorem 3.2

we know that axioms $(A1)$ and $(A3)$ force f to take one of the two forms given by (3.13). Now, suppose f is given by the second form in (3.13). Applying axiom $(A4)$ to this case we have

$$sw(t)+w_1(t)=(s+g_2(\delta))w(t+\delta)+w_1(t+\delta), \tag{3.31}$$

for all $s,t,\delta \geq 0$. Putting $s=0$ in (3.31) we get $w_1(t)-w_1(t+\delta)=g_2(\delta)w(t+\delta)$, which when subtracted from (3.31), on simplification, gives $s(w(t+\delta)-w(t))=0$, from which we get $w(t+\delta)=w(t)$ for all $t,\delta \geq 0$. Thus, $w(t)=$ a constant$=c_1$, say. Substituting this value of $w(t)$ in the equation $w_1(t)-w_1(t+\delta)=g_2(\delta)w(t+\delta)$, we get $w_1(t)-w_1(t+\delta)=g_3(\delta)$ for all $t,\delta \geq 0$, where $g_3(\delta)=c_1 g_2(\delta)$. Note that by axiom $(A3)$, $w_1(0)=0$. So, $g_3(\delta)=-w_1(\delta)$, which implies that $w_1(t+\delta)=w_1(t)+w_1(\delta)$ for all $t,\delta \geq 0$. The only continuous solution to this functional equation is $w_1(t)=q't$ for some $q' \in R^1$ (see Aczel, 1966, p.34). Hence in this case f is given by $f(s,t)=c_1 s+q't$. By increasingness of f in s, $c_1>0$. Note also that $q'=f(0,1)<f(0,0)=0$ (by axiom $(A3)$). So we rewrite the general solution as $f(s,t)=c_1 s-c_2 t$, where $c_1,c_2>0$.

Next, we take up the first form in (3.13). By axiom $(A4)$,

$$\frac{a^s w(t)}{\log a}+w_1(t)=\frac{a^{s+g_2(\delta)}w(t+\delta)}{\log a}+w_1(t+\delta) \tag{3.32}$$

for all $s,t,\delta \geq 0$. Putting $s=0$ in both sides of (3.32) we have

$$\frac{w(t)}{\log a}+w_1(t)=\frac{a^{g_2(\delta)}w(t+\delta)}{\log a}+w_1(t+\delta). \tag{3.33}$$

Subtracting the left (right) hand side of (3.33) from the corresponding side of (3.32) and then rearranging the resulting expression we get

$$\frac{(a^s-1)}{\log a}\left(a^{g_2(\delta)}w(t+\delta)-w(t)\right)=0. \tag{3.34}$$

But $\frac{(a^s-1)}{\log a}>0$ for all $s>0$. This shows that

$$\left(a^{g_2(\delta)}w(t+\delta)-w(t)\right)=0 \tag{3.35}$$

for all $t,\delta \geq 0$.

Now, recall from (3.13) that $w(t)>0$ for all $t \geq 0$. Therefore, from (3.35) we get

$$\frac{w(t+\delta)}{w(t)} = a^{-g_2(\delta)} \tag{3.36}$$

for all $t, \delta \geq 0$. Putting $t = 0$ in (3.36) we have

$$\frac{w(\delta)}{w(0)} = a^{-g_2(\delta)}. \tag{3.37}$$

From (3.36) and (3.37) it follows that

$$\frac{w(t+\delta)}{w(t)} = \frac{w(\delta)}{w(0)} \tag{3.38}$$

for all $t, \delta \geq 0$. As we have noted in the proof of Theorem 3.2, the general solution to this equation is given by $w(t) = c'\varsigma^t$ for some constants $c', \varsigma > 0$. A comparison of (3.33) and (3.35) gives $w_1(t) = w_1(t+\delta)$ for all $t, \delta \geq 0$, so that $w_1(t) =$ constant$= \xi$, say. Hence the complete solution in this case is $f(s,t) = \dfrac{a^s c'\varsigma^t}{\log a} + \xi$. By axiom $(A3)$, $\xi = -\dfrac{c'}{\log a}$. Consequently, $f(s,t) = \dfrac{c'}{\log a}\left(a^s\varsigma^t - 1\right)$. Increasingness and decreasingness of f in its first and second arguments respectively require that $a > 1$ and $\varsigma < 1$. So the solution can be written as $f(s,t) = c\left(\dfrac{a_1^s}{a_2^t} - 1\right)$, where $c > 0$ and $a_1, a_2 > 1$ are constants. This completes the necessity part of the proof. The sufficiency is easy to check. □

Proof of Theorem 3.4: Given $x = \left(x^1, x^2, \ldots, x^k\right) \in \prod_{i=1}^{k} D^{n_i}$ and $\lambda_i = \lambda\left(x^i\right)$, define a sequence $\{y(i)\}$ as follows:

$y(0) = x$,

$y(1) = \left(\lambda_1 1^{n_1}, x^2, \ldots x^k\right)$,

$y^j(2) = y^j(1)$ for $j \neq 2$, $y^2(2) = \lambda_2 1^{n_2}$,

$y^j(3) = y^j(2)$ for $j \neq 3$, $y^3(3) = \lambda_3 1^{n_3}$, and so on. Finally,

$y^j(k) = y^j(k-1)$ for $j \neq k$ and $y^k(k) = \lambda_k 1^{n_k}$.

Thus, for any $i, 1 \le i \le k$, we have $y(i) = \left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_i 1^{n_i}, x^{i+1}, \ldots, x^k\right)$. Note that for all i and j, $\lambda\left(y^j(i)\right) = \lambda\left(x^j\right), \lambda(y(i)) = \lambda(x)$ and $y(k) = \left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right)$.

It is given that for any i, $1 \le i \le k$, $P(y(i)) - P(y(i-1)) = v_i(\underline{n}, \underline{\lambda}) g\left(x^i\right)$. Summing over all i, we get $P(y(k)) - P(y(0)) = \sum_{i=1}^{k} v_i(\underline{n}, \underline{\lambda}) g\left(x^i\right)$. That is,

$$P\left(\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right)\right) - P(x) = \sum_{i=1}^{k} v_i(\underline{n}, \underline{\lambda}) g\left(x^i\right). \tag{3.39}$$

Now define $I : \left(\prod_{i=1}^{k} D^{n_i}\right) \cup \left(\bigcup_{i=1}^{k} D^{n_i}\right) \to R^1_+$ by the following relation:

$$I(x) = \begin{cases} \left(\dfrac{1}{c_1} + \dfrac{1}{c_2}\right) P\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right) - \dfrac{1}{c_2} P(x) \text{ for } x = \left(x^1, x^2, \ldots, x^k\right) \in \prod_{i=1}^{k} D^{n_i}, \\ g(x) \text{ if } x \in \bigcup_{i=1}^{k} D^{n_i}, \end{cases} \tag{3.40}$$

where $c_1, c_2 > 0$ are arbitrary constants. Clearly, there is no ambiguity in the definition of I. By continuity of P, I is continuous. From the above definition it follows that $P\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right) = c_1 I\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right)$ and $g\left(x^i\right) = I\left(x^i\right)$, $1 \le i \le k$. Substituting this into (3.39) we get

$$P(x) = c_1 I\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right) - c_2 \sum_{i=1}^{k} \omega_i(\underline{n}, \underline{\lambda}) g\left(x^i\right), \tag{3.41}$$

where $\omega_i(\underline{n}, \underline{\lambda}) = v_i(\underline{n}, \underline{\lambda})/c_2$. This in turn gives:

$$\begin{aligned} I(x) &= \frac{1}{c_1} P\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right) + \frac{1}{c_2}\left\{P\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right) - P(x)\right\} \\ &= I\left(\lambda_1 1^{n_1}, \lambda_2 1^{n_2}, \ldots, \lambda_k 1^{n_k}\right) + \sum_{i=1}^{k} \omega_i(\underline{n}, \underline{\lambda}) I\left(x^i\right). \end{aligned}$$

Thus, I is subgroup decomposable. To show that I takes on the value zero for the perfectly equal distribution on $\bigcup_{i=1}^{k} D^{n_i}$; observe that $I\left(x^i\right) = (P(y) - P(x))/v_i(\underline{n}, \underline{\lambda})$, which implies that $I\left(c 1^{n_i}\right) = 0$ for all $i, 1 \le i \le k$ and for all $c > 0$. □

Chapter 4

ETHNIC POLARIZATION QUASI-ORDERINGS AND INDICES[13]

4.1. Introduction

The previous chapter was devoted to a particular problem on income polarization. Several indices were characterized and a polarization quasi-ordering was developed. It goes without saying that the same treatment applies to an arbitrary continuous variable.

However, in many important situations there may not be information on a continuous attribute to measure distance across groups or individuals. For example, in the case of ethnic classification the only information we may have is whether an individual belongs to a particular ethnic group or not. Such cases pertain to the discussion on ‘social polarization’ in Chapter 1, of which ‘ethnic polarization’ is a particular case. As already mentioned thereat, the *RQ* index suggested by Montalvo and Reynal-Querol (2005) is the most popular index of ethnic polarization.

Several properties of ethnic polarization indices have been investigated by Montalvo and Reynal-Querol (2005, 2008), Esteban and Ray (1999, 2008b) and Chakravarty and Maharaj (2011). Taking cue from such investigations, we develop some new reasonable axioms for ethnic polarization indices. The first goal of this chapter is to characterize the *RQ* index using these axioms and some additional ones borrowed from Montalvo and Reynal-Querol (2005, 2008) and Esteban and Ray (1999, 2008b). These characterizations enable us to understand the *RQ* index from alternative perspectives in greater detail. None of our characterization results begins with any specific assumption like additivity. From this perspective these results are quite general. More precisely, our characterization reveals how within a general structure we can isolate a set of necessary and sufficient conditions for identifying the *RQ* index uniquely. In the process we characterize a generalization of the *RQ* index, which we refer to as the *‘Generalized RQ-Index of order* θ *’*.

[13] A major portion of this chapter has been published in Chakravarty and Maharaj (2012).

It is true that a particular index of ethnic polarization will order different ethnic groups in a complete manner. However, two alternative indices of ethnic polarization may not rank the ethnic groups in the same way. It is, therefore, natural to develop an ethnic polarization quasi-ordering that will rank two different ethnic groups in an identical manner. Such a quasi-ordering determines the necessary and sufficient conditions for one ethnic group to be regarded as more or less polarized than another by all ethnic polarization indices that satisfy certain desirable criteria. It may be worthwhile to mention here that no such ethnic polarization quasi-ordering has been suggested in the existing literature.

One of the major objectives of this chapter is to develop two quasi-orderings of this type without any specific structural assumption e.g., additivity. One of these quasi-orderings can be easily verified using the population concentration curve. The population concentration curve of an ethnic distribution is the graph of the cumulative population shares against the cumulative number of groups, with groups ranked from the largest to the smallest.

The chapter is organized as follows. After discussing the background material in Section 4.2, we present the characterization theorems in Section 4.3. The two quasi-orderings are discussed in Section 4.4. Finally, Section 4.5 concludes.

4.2 The Background

For a population consisting of k ethnic groups $E_1, E_2, ..., E_k$, where $k \in \Gamma = N \setminus \{1\}$, N being the set of positive integers, let π_i denote the proportion of individuals in E_i. Therefore, $0 \le \pi_i \le 1$, $1 \le i \le k$ and $\sum_{i=1}^{k} \pi_i = 1$, $k \in \Gamma$ being arbitrary. This generates a probability distribution $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_k)$, which we will refer to as 'ethnic distribution'.

The *RQ* index suggested in Montalvo and Reynal-Querol (2005, 2008) is based on the *ER* index of income polarization (defined in (1.45)).

In order to identify ethnicity of an individual, it becomes necessary to verify if he ‘belongs to’ or ‘does not belong to’ a particular ethnic group. Montalvo and Reynal-Querol

(2005, 2008) argued that in such a situation it is essential to replace the Euclidean metric in (1.45) by the discrete metric[14] and hence deal with the class of discrete polarization indices $DP(\alpha, \mu_1)$ defined in (1.59).

For relating $DP(\alpha, \mu_1)$ with the *RQ* index, Montalvo and Reynal-Querol (2008) considered three properties, which are taken directly from Esteban and Ray (1994) and redefined in terms of group's size only. We will also invoke these properties for our characterizations. The first property is:

Property 1: If there are three groups of sizes *p*, *q* and *r*, and $p > q$ and $q \geq r$, then if we merge the two smallest groups into a new group, $\tilde{q}$, the new distribution is not less polarized than the original one. That is, $POL(p, q, r) \leq POL(p, \tilde{q})$, with $\tilde{q} = (q + r)$.

According to this property for three groups with relative frequencies *p*, *q* and *r*, where $p > q$ and $q \geq r$, polarization should not decrease under merger of the two smallest groups [15]. It corresponds to the Esteban-Ray (1994) Axioms 1 and 2.

Montalvo and Reynal-Querol (2008) showed that a necessary and sufficient condition for Property 1 to hold for $DP(\alpha, \mu_1)$ is that $\alpha \geq 1$. We obtain the same boundary restriction on the value of α if Property 1 is replaced by Property 1b, whose formulation does not need the assumption that the number of groups is three.

Property 1b: Suppose that there are two groups with sizes π_1 and π_2. Take any one group, say π_2 and split it into $m \geq 2$ groups in such a way that $\pi_1 = \tilde{\pi}_1 \geq \tilde{\pi}_i$ for all $i = 2, \ldots, (m+1)$, where

[14]It may be added here that an ethnic polarization index need not take the distance between two ethnic groups as unity. For instance, in Fearon, (2003) and Desmet et al. (2008, 2009) computation of intergroup distance is based on linguistic distance, whereas Spolaore and Wacziarg (2009) used genetic distance. However, following Reynal-Querol (2005, 2008) we assume that the intergroup distance is unity.

[15] In Montalvo and Reynal-Querol (2008) the three properties have been stated using strict inequality. However, we use weaker versions of these properties and none of our results changes if we replace weak inequality by strict inequality. It may be mentioned that in Esteban and Ray (1994) the axioms have been stated using weak inequality.

$\tilde{\pi}$ is the new vector of population shares, and clearly $\sum_{i=2}^{m+1} \tilde{\pi}_i = \pi_2$. Then the polarization under $\tilde{\pi}$ is not higher than that under π.

The following property, which may be regarded as a particular version of Property 1b, has been established and analyzed by Esteban and Ray (1999).

Property 1c: Let $G \geq 3$. Consider a distribution of the population across *G* groups. Suppose that $\pi_i \neq \pi_j$ for some *i* and *j*. Then polarization is increased by a merger of any $G-1$ smallest groups into one. However, if the initial distribution of population is uniform, then polarization is unchanged.

The next property, which is based on Axiom 3 of Esteban and Ray (1994), says that if there are three groups, two of which have equal population share, then polarization should not decrease under shift of population mass from the group with unequal size equally to the other two groups. Formally,

Property 2: Assume that there are three groups of sizes *p*, *q* and *p*. Then if we shift mass from the *q* group equally to the other two groups, polarization does not decrease. That is, $POL(p,q,p) \leq POL(p+x, q-2x, p+x)$, where $0 < x < q/2$.

Montalvo and Reynal-Querol (2008) also demonstrated that $DP(\alpha, \mu_1)$ satisfies Property 2 for any distribution if and only if $\alpha = 1$. Given $\alpha = 1$, if we choose $\mu_1 = 4$, then the resulting index $DP(1,4)$ becomes the *RQ* index defined in (1.58), which can also be expressed as

$$RQ = 1 - \sum_{i=1}^{k} \left(\frac{0.5 - \pi_i}{0.5} \right)^2 \pi_i . \tag{4.1}$$

This expression for *RQ* incorporates a weighted sum of population fractions, where the weights are squared deviations of population shares of different groups from the maximum population share 0.5 as a proportion of 0.5. It is worth noting here that the index *RQ* is actually a positive multiple of the probability that out of three randomly selected persons in the population, two will belong to the same group.[16]

Dealing with a general model of conflict, Esteban and Ray (1999, 2008b) considered a particular situation in which each group feel equally alien towards all other groups and suggested

$$IC = \sum_{i=1}^{k} \frac{\pi_i^3}{p_i}(1 - \pi_i)b_i \,, \tag{4.2}$$

as an indicator of intensity of conflict, where b_i = equal distance of the i^{th} group from any other group, p_i = probability of success of the i^{th} group in a conflict game, defined in an unambiguous way. If $p_i = \pi_i, b_i = 1$ for all *i*, then IC equals $RQ/4$.

4.3 Axioms, Discussion and Characterizations

We begin this section by presenting some axioms existing in the literature. We then develop two new axioms as implications of some observations made by Esteban and Ray (1994, 1999, 2008b) and Montalvo and Reynal-Querol (2005, 2008). All these axioms are used to characterize the *RQ* index and study the partial orderings. While the characterization theorems are presented in this section, the partial orderings will be analyzed in the next section.

We write Δ_k for the set of all discrete probability distributions of dimension k on the real line *R* and **Δ** for the set of all probability distributions on *R*. Obviously, $\Delta = \bigcup_{k=2}^{\infty} \Delta_k$.

We start with a general definition of a polarization measure in case of an ethnic data.

Definition 4.1: An 'Ethnic Polarization Index' (EPI) is a continuous real-valued function defined on Δ , that is, $P : \Delta \to R$, which is symmetric in its arguments (that is, for all $k \in \Gamma$, $\underline{\pi} \in \Delta_k$, $P(\underline{\pi}) = P\left(\sigma(\underline{\pi})\right)$, where $\sigma(\underline{\pi})$ is an arbitrary permutation of $\underline{\pi}$) and which satisfies zero-

[16] We are grateful to Joan Esteban for drawing our attention to this fact in course of a personal correspondence.

frequency independence (that is, for all $k \in \Gamma$, $\underline{\pi} \in \Delta_k$, we have, $P(\pi_1, \pi_2, ..., \pi_k) = P(\pi_1, \pi_2, ..., \pi_k, 0)$).

For any $\underline{\pi} \in \Delta$, an EPI simply aggregates components of $\underline{\pi}$ in an unambiguous way. Given $\underline{\pi} \in \Delta$, the real number $P(\underline{\pi})$ indicates the level of ethnic polarization associated with $\underline{\pi} \in \Delta$.

Continuity of an EPI ensures that minor changes in π_i's will generate only minor changes in P. Anonymity or symmetry demands that P remains invariant under any reordering of π_i's. Next, given the ethnic groups $E_1, E_2, ..., E_k$ and their relative frequencies , if a new ethnic group is created with zero frequency, then this does not have any impact on ethnic conflict and hence the level of polarization should remain unaffected. In other words, we say that an EPI satisfies zero-frequency independence[17]. It may be noted that these defining minimal conditions on an EPI are satisfied by *RQ*.

The following axioms, which have been discussed by Montalvo and Reynal-Querol (2005, 2008) and are satisfied by *RQ*, will be necessary for our characterizations (see also Esteban and Ray, 1994).

Axiom 1: For all $k \in \Gamma$, $\underline{\pi} \in \Delta_k, 0 \le P(\underline{\pi}) \le 1$.

Axiom 2: For all $k \in \Gamma$, $P(\underline{\pi}) = 0$ if $\underline{\pi} \in \Delta_k$ is some permutation of (1, 0,…, 0).

Axiom 3: For all $k \in \Gamma$, $P(\underline{\pi}) = 1$ if $\underline{\pi} \in \Delta_k$ is some permutation of $\left(\frac{1}{2}, \frac{1}{2}, 0, ..., 0\right)$.

Axiom 1 is a boundedness principle. The next axiom says that the EPI achieves its minimum value, zero, if there is complete homogeneity in the sense that all the individuals

[17]We may mention here that a continuous, symmetric function need not satisfy zero-frequency independence. For example, the continuous, symmetric function $P(\pi_1, \pi_2, ..., \pi_k) = \min_{1 \le i \le k} \pi_i$ does not satisfy zero-frequency independence.

belong to a particular ethnic group. Finally, according to Axiom 3 the EPI is maximized if there is an equal splitting of the entire population into two groups, that is, in the bipolar situation. Given the existence of a large ethnic group, if the ethnic minority is not divided into many groups and is large as well, then chances of ethnic conflicts increase (Horowitz, 1985). Since ethnic conflicts are likely to increase with a bipolar ethnic distribution, it is sensible to assume that ethnic polarization is maximized in the case of a bipolar ethnic distribution (see also Esteban and Ray, 1999, 2008b)[18]. Thus, an EPI is an indicator of divergence of the actual ethnic distribution from the extreme distribution $\left(1/2 , 1/2 , 0 , \ldots , 0 \right)$.

While for our characterizations we will employ Properties 1, 1b, 1c, 2 and Axioms 1-3, one additional postulate we wish to use is a multiplicative decomposability condition. Let $\pi^{(i)} = (\pi_1, \pi_2, \ldots, \pi_i, 0, \ldots, 0) \in \Delta_k$, where $1 \le i < k$. For $0 \le \varepsilon \le \pi_i$, define $\underline{\pi}_\varepsilon^{(i)} = (\pi_1, \pi_2, \ldots, \pi_i - \varepsilon, \varepsilon, 0, \ldots, 0) \in \Delta_k$. Thus, $\underline{\pi}_\varepsilon^{(i)}$ is obtained from $\underline{\pi}^{(i)}$ by splitting the population coming from E_i into two subpopulations with respective masses $\pi_i - \varepsilon$ and ε, that is, by shifting a mass ε from E_i to E_{i+1}. A question that arises naturally in this context is: how does polarization change due to such a split? Answers to this query have been provided by Esteban and Ray (1994) (see Lemma 1, p. 838) and Montalvo and Reynal-Querol (2008) (see Properties 1, 1b and 2, pp. 7-8) in terms of 'merger of groups' and 'shift of mass'. However, they did not specify any functional form for the resulting change in polarization.

To be precise, in the above set-up we are interested in looking at the polarization difference $P\left(\underline{\pi}_\varepsilon^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right)$. Since we are keeping $E_1, E_2, \ldots, E_{i-1}$ constant, it seems reasonable to enquire if the reduction/increase in the level of conflict due to the population shift can be attributed to the internal conflict between E_i and E_{i+1}. In other words, does this difference have any direct relation with $P\left(1 - \frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right)$, the level of polarization of the population consisting of

[18] Esteban and Ray (1994) proved that a polarization index should achieve its maximum value for the distribution $\left(1/2 , 1/2 , 0 , \ldots , 0 \right)$.

the ethnic groups E_i and E_{i+1}? Given the population masses of $E_1, E_2, ..., E_{i-1}$; Axiom 4 provides a very simple way of relating the difference with $P\left(1-\frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right)$.

For a concrete formulation of the axiom, we use arguments put forward in Esteban and Ray (2008b). In their discussion on IC these authors noted that "a split of a group with population π into two groups with π' and π'', $\pi' + \pi'' = \pi$, increases IC if and only if the group-size is sufficiently large. If π is small, the split will decrease IC." (They also studied the effect of such a split on *RQ* and *FRAC*.) Taking cue from this, it is apparent that there should a threshold level, say η_0, of the size (π) of the splintered group, such that whenever $\pi \le \eta_0$, a split results in a decrease of polarization, whereas the reverse happens if $\pi > \eta_0$.

Observation 4.1: A necessary condition for Property 1 to hold is that $\eta_0 \ge 2/3$.

Proof: See Appendix.

Now, consider $\underline{\pi}_{\varepsilon}^{(i)} = (\pi_1, \pi_2, ..., \pi_i - \varepsilon, \varepsilon, 0, ..., 0)$, which is obtained from $\underline{\pi}^{(i)} = (\pi_1, \pi_2, ..., \pi_i, 0, ..., 0)$ by a 'rank-preserving equalizing transfer' from E_i to E_{i+1}. For small π_i, Esteban and Ray (2008b, p. 175) contends that 'the equalization of population brings polarization down'. So, $P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) < P\left(\underline{\pi}^{(i)}\right)$ whenever $0 < \varepsilon \le \frac{1}{2}\pi_i$. Symmetry of P ensures that the same result holds for $\frac{1}{2}\pi_i \le \varepsilon < \pi_i$.

Next, let $0 < \varepsilon_1 < \varepsilon_2 < \pi_i/2$. Clearly, $\underline{\pi}_{\varepsilon_1}^{(i)}$ is obtained from $\pi^{(i)}$ by an 'equalizing progressive transfer' and $\underline{\pi}_{\varepsilon_2}^{(i)}$ is derived from $\underline{\pi}_{\varepsilon_1}^{(i)}$ by another such transfer. Thus, $P\left(\underline{\pi}_{\varepsilon_2}^{(i)}\right) < P\left(\underline{\pi}_{\varepsilon_1}^{(i)}\right) < P\left(\underline{\pi}^{(i)}\right)$ which implies that

$$P\left(\underline{\pi}_{\varepsilon_2}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) < P\left(\underline{\pi}_{\varepsilon_1}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) < 0, \tag{4.3}$$

whenever $0 < \varepsilon_1 < \varepsilon_2 < \pi_i/2$.

Now, observe that for fixed π_i, $P\left(1-\frac{\varepsilon}{\pi_i},\frac{\varepsilon}{\pi_i}\right)$ is an increasing function of ε for $0<\varepsilon<\pi_i/2$ (since an increment in ε in this range brings the distribution $\left(1-\frac{\varepsilon}{\pi_i},\frac{\varepsilon}{\pi_i}\right)$ closer to the bipolar situation $\left(\frac{1}{2},\frac{1}{2}\right)$, see Montalvo and Reynal-Querol, 2005). This along with the non-negativity of the polarization index implies that

$$0\le P\left(1-\frac{\varepsilon_1}{\pi_i},\frac{\varepsilon_1}{\pi_i}\right)<P\left(1-\frac{\varepsilon_2}{\pi_i},\frac{\varepsilon_2}{\pi_i}\right) \tag{4.4}$$

whenever $0<\varepsilon_1<\varepsilon_2<\pi_i/2$. Thus, from (4.3) and (4.4) it follows that in the range $0<\varepsilon_1<\varepsilon_2<\pi_i/2$, the polarization-difference $P\left(\underline{\pi}_\varepsilon^{(i)}\right)-P\left(\underline{\pi}^{(i)}\right)$ and $P\left(1-\frac{\varepsilon}{\pi_i},\frac{\varepsilon}{\pi_i}\right)$ are related in a negative monotonic way. We can arrive at the same conclusion for $\pi_i/2\le\varepsilon_1<\varepsilon_2\le\pi_i$ and $0\le\varepsilon_1\le\pi_i/2<\varepsilon_2\le\pi_i$. The above implication of the Esteban-Ray observation suggests that the polarization difference $P\left(\underline{\pi}_\varepsilon^{(i)}\right)-P\left(\underline{\pi}^{(i)}\right)$ can be taken to be negatively proportional to $P\left(1-\frac{\varepsilon}{\pi_i},\frac{\varepsilon}{\pi_i}\right)$. Given that this two-group polarization has no relationship with population masses of the unaffected groups, we can decompose the change $P\left(\underline{\pi}_\varepsilon^{(i)}\right)-P\left(\underline{\pi}^{(i)}\right)$ into two components, one depending on the population masses of the unaffected groups and other is $P\left(1-\frac{\varepsilon}{\pi_i},\frac{\varepsilon}{\pi_i}\right)$. This motivates us to state the following axiom:

Axiom 4: For all $k\ge 3$, let $\underline{\pi}^{(i)}$, $\underline{\pi}_\varepsilon^{(i)}$, $1\le i<k$, η_0 be as above and $\pi_i\le\eta_0$. Then

$$P\left(\underline{\pi}_\varepsilon^{(i)}\right)-P\left(\underline{\pi}^{(i)}\right)=f\left(\pi_1,\pi_2,\ldots\ldots,\pi_{i-1}\right)P\left(1-\frac{\varepsilon}{\pi_i},\frac{\varepsilon}{\pi_i}\right), \tag{4.5}$$

where $f:[0,1]^{(i-1)}\to R$ is continuous and negative-valued.

In Axiom 4 it is assumed that the population proportion of E_i is not of highly significant size. According to the axiom, the marginal polarization, the decrement in ethnic polarization, which can be attributed to the inner conflict in E_i resulting from the split, is the product of a negative-valued continuous function of the population proportions of the unaffected groups and the polarization of the two groups resulting from the split. Clearly, alternative formulations of marginal polarization are possible, for instance, the additive decomposability $f(\pi_1, \pi_2, \cdots, \pi_{i-1}) +$ $P\left(1-\frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right)$ is also a very simple possibility. Since $f(\pi_1, \pi_2, \cdots, \pi_{i-1})$ is arbitrarily negative and by Axiom 1, $0 \le P\left(1-\frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right) \le 1$, additive decomposability in the presence of Axiom 1 does not ensure that $P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right)$ is non-positive. Axiom 4 avoids this problem and the formulation specified in the axiom is simple; easy to understand, and it recognizes the dependence of the change on the population sizes of the unaffected groups in a general way. We will use this general form of the axiom for characterizations, but for ordering of ethnic profiles, we will assume that *f* is additive (see Axiom $4'$ below).

To understand the essential message of Axiom 4, let's consider $\underline{\pi} = \left(\frac{2}{3}, \frac{1}{3}, 0\right)$ and $\underline{\pi}^{(1/6)} = \left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right)$, $\underline{\pi}^{(1/9)} = \left(\frac{2}{3}, \frac{2}{9}, \frac{1}{9}\right)$ obtained from $\underline{\pi}$ by transfers of masses $\frac{1}{6}$ and $\frac{1}{9}$ respectively from the ethnic group E_2 to E_3. Now, let us consider a property viz. Property A, which states that for an ethnic distribution $\underline{\pi} = (\pi_1, \pi_2, \dots, \pi_k) \in \Delta_k$, a shift of mass from a larger group E_i to a smaller group E_j (where $j > i \ge 2$) reduces polarization. (Ths property is discussed more elaborately in the next section.)

By Property A, (i) each one of $\underline{\pi}^{(1/9)}$ and $\underline{\pi}^{(1/6)}$ is less polarized than $\underline{\pi}$ and (ii) $\underline{\pi}^{(1/9)}$ is more polarized than $\underline{\pi}^{(1/6)}$. This clearly implies that the polarization difference $P\left(\underline{\pi}^{(1/9)}\right) - P(\underline{\pi})$ is greater than $P\left(\underline{\pi}^{(1/6)}\right) - P(\underline{\pi})$. On the other hand, as discussed earlier, the level of polarization of the subpopulation comprising ethnic groups E_2 and E_3 is $P\left(\frac{1/6}{1/3}, \frac{1/6}{1/3}\right) = P\left(\frac{1}{2}, \frac{1}{2}\right)$ in the first case and $P\left(\frac{2/9}{1/3}, \frac{1/9}{1/3}\right) = P\left(\frac{2}{3}, \frac{1}{3}\right)$ in the second case so that the internal conflict of the former is larger than that of the latter. Thus, the decrement in polarization and the measure of the internal conflict of the distribution formed by the splintered groups move along opposite directions.

This is precisely the intuition behind Axiom 4.

The next axiom is concerned with polarization in 2-group situations. For $0 \le \pi \le 1$, consider the map $\varphi(\pi) = P(\pi, 1-\pi)$. Montalvo and Reynal Querol (2005) argued that polarization is maximized at $(1/2, 1/2)$ and the greater is the distance between $(\pi, 1-\pi)$ and $(1/2, 1/2)$, the lower is the level of polarization in the 2-group case.

To motivate the axiom from another angle, let us begin with $\underline{\pi}^{(1)} = \left(\frac{1}{3}, \frac{2}{3}\right)$ and $\underline{\pi}^{(2)} = \left(\frac{1}{4}, \frac{3}{4}\right)$. For comparing polarizations between $\underline{\pi}^{(1)}$ and $\underline{\pi}^{(2)}$, we note that $\underline{\pi}^{(2)}$ is more homogeneous than $\underline{\pi}^{(1)}$ (that is, closer to the perfectly homogeneous (1,0) distribution). Alternatively, we can apply Property A (put forward in the subsequent section) and observe that $\underline{\pi}^{(2)}$ is obtained from $\underline{\pi}^{(1)}$ by a rank preserving progressive transfer. Using either argument, a polarization index should give the former a higher value than the latter. However, it is interesting to note that $\underline{\pi}^{(1)}$ is closer to the extremal $\left(\frac{1}{2}, \frac{1}{2}\right)$ distribution, whereat the polarization measure takes its maximum value (1). So, it is reasonable to demand that the polarization measure in the 2-dimensional case is related to the distance from $\left(\frac{1}{2}, \frac{1}{2}\right)$ in a negative monotonic way.

Therefore, $\varphi(\pi)$ can be taken as a decreasing function of $\delta(\pi) = d^2((1/2, 1/2), (\pi, 1-\pi))$, where d denotes the Euclidean distance in R_+^2. In other words, $\varphi(\pi)$ can be regarded as a decreasing function of $\delta(\pi)$.

Note that any function meeting this requirement can be taken as a candidate for $\varphi(\pi)$. For simplicity of exposition, we take a decreasing linear transform of $\delta(\pi)$ and suggest the following axiom, which will be used for both characterization and the partial ordering:

Axiom 5: For $0 \le \pi \le 1$, $P(\pi, 1-\pi)$ is a decreasing linear transform of $d^2((1/2, 1/2),(\pi, 1-\pi))$.

We are now in a position to state and prove our characterization theorems.

Theorem 4.1: The only EPI $P : \Delta \to R$ that satisfies Axioms 2-5 and Property 1c is *RQ*.

Proof of Theorem 4.1 relies on the following lemma.

Lemma 4.1: An EPI $P : \Delta \to R$ satisfies Axioms 2-5 and Property 1 only if it is of the form

$$RQ_{(\theta_3, \theta_4, \ldots, \theta_k)}(\underline{\pi}) = 4\sum_{i=1}^{k} \pi_i^2 (1 - \pi_i) + \sum_{i=3}^{k} \theta_i S_i(\underline{\pi}), \tag{4.6}$$

where $S_l(\underline{\pi}) = \sum_{1 \le i_1 < i_2 < i_3 < \ldots < i_l \le k} \pi_{i_1} \pi_{i_2} \pi_{i_3} \ldots . \pi_{i_l}$, $\theta_l \in R$ are arbitrary constants, $3 \le l \le k$ and $\underline{\pi} = (\pi_1, \pi_2, \ldots, \pi_k) \in \Delta_k$ are arbitrary.

Remark 4.1: Clearly, $S_l(\underline{\pi})$ is a positive multiple of the probability that l persons chosen at random from the population belong to l different groups. If $\theta_3 = \theta_4 = \ldots . = \theta_k = 0$, $RQ_{(\theta_3, \theta_4, \ldots, \theta_k)}$ coincides with *RQ*. Therefore, $RQ_{(\theta_3, \theta_4, \ldots, \theta_k)}$ may be regarded as a *'Generalization of the RQ-Index.'*

Proofs of Lemma 4.1 & Theorem 4.1: See Appendix.

The function *f*, introduced in Axiom 4, is general in the sense of its dependence on the population share vector of the unaffected groups. However, as a special case, one can consider a situation where *f* is dependent on the merged entity $\sum_{j=1}^{i-1} \pi_j$ only. In this case, (4.6) becomes:

$$P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) = f_1(\pi_1 + \pi_2 + \ldots .. + \pi_{i-1}) P\left(1 - \frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right), \tag{4.7}$$

where $f_1 : [0,1] \to R$ is continuous and negative-valued.

In view of the relation

$$\sum_{j=1}^{i} \pi_j = 1, \tag{4.8}$$

this is same as:

$$P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) = g_1(\pi_i) P\left(1 - \frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right) \tag{4.9}$$

Therefore, in this case, we can reformulate Axiom 4 as:

Axiom $4'$**:** For all $k \geq 3$, let $\underline{\pi}^{(i)}$ and $\underline{\pi}_{\varepsilon}^{(i)}$, $1 \leq i < k$, η_0 be as in Axiom 4 and $\pi_i \leq \eta_0$. Then (4.9) holds for some is continuous and negative-valued function $g_1 : [0,1] \to R$.[19]

Alternatively, Axiom $4'$ can also be motivated in the following way. As proposed in the discussion preceding Axiom 4, $P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right)$ can be assumed to be negatively proportional to $P\left(1 - \frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right)$. Notice that in the transformation $\underline{\pi}^{(i)} \to \underline{\pi}_{\varepsilon}^{(i)}$, sizes of the the groups $E_1, E_2, \ldots, E_{i-1}$ remains unchanged. So, in calculating the above polarization difference one can ignore the sizes of these $(i-1)$ groups. (For example, consider the additive set up suggested in Chakravarty and Maharaj (2011a), where the polarization index P is assumed to have the structure:

$$P(\underline{\pi}) = \sum_{i=1}^{k} \psi_P(\pi_i), \tag{4.10}$$

$\psi_P(\pi_i)$ being the influence function of E_i. The influence functions of $E_1, E_2, \ldots, E_{i-1}$ in the aforesaid polarization difference cancel out each other). On the other hand, in course of this transformation, the group-sizes of E_i and E_{i+1} vary, but their the sum remains unaltered (and

[19] It can be shown that this assumption leads to a violation of 'globality' mentioned in Esteban and Ray (1994). However, this is a common drawback to all EPIs that are additive across their arguments.

equals to π_i). So, we posit that the constant of proportionality in this transformation depends exclusively on π_i. This is precisely what Axiom $4'$ has to say.

We now state our second result, which replaces Axiom 4 and Property 1c in the previous Theorem by Axiom $4'$ and Property 1 respectively.

Theorem 4.2: The only EPI $P : \Delta \to R$ that satisfies Axioms 1-3, $4'$, 5 and Property 1 is *RQ*.

The following lemma will be necessary for proving Theorem 4.2.

Lemma 4.2: An EPI $P : \Delta \to R$ satisfies Axioms 1-3, $4'$ and 5 if and only if it is of the form

$$RQ_\theta(\underline{\pi}) = 4\sum_{i=1}^{k} \pi_i^2 (1 - \pi_i) + \theta \sum_{1 \le i_1 < i_2 < i_3 \le k} \pi_{i_1}\pi_{i_2}\pi_{i_3} , \tag{4.11}$$

where $\theta \in [0,3]$.

Remark 4.2: Notice that the terms $\sum_{i=1}^{k} \pi_i^2 (1 - \pi_i)$ and $\sum_{1 \le i_1 < i_2 < i_3 \le k} \pi_{i_1}\pi_{i_2}\pi_{i_3}$ appearing in (4.10) give respectively positive multiples of the probability that out of 3 randomly selected persons, 2 will belong to a single group and the third to another and that all the three belong to three different groups. A perfect homogeneity in the population demands that all the three persons belong to the same ethnic group. Thus, the right hand side of (4.11), being a linear combination of the two probabilities, measures an extent of heterogeneity of the population and hence can reasonably be taken as an index of ethnic polarization. We define RQ_θ to be the *'Generalized RQ-index of order θ '*.

Remark 4.3: It follows from the proof of Lemma 4.2 that the only EPI satisfying Axioms 2, 3, $4'$ and 5 is of the form RQ_θ where $\theta \in R$. Thus, given these axioms, Axiom 1 becomes necessary and sufficient for fixing the lower and upper bounds on θ as 0 and 3 respectively.

Proofs of Lemma 4.2 & Theorem 4.2: See Appendix.

The next characterization of *RQ* is based on Property 1b.

Theorem 4.3: The only EPI $P:\Delta \to R$ that satisfies Axioms 1-3, $4'$, 5 and Property 1b is RQ.

Proof: See Appendix.

Finally, we show how Property 2 can be employed to characterize RQ.

Theorem 4.4: The only EPI $P:\Delta \to R$ that satisfies Axioms 2, 3, $4'$, 5 and Property 2 is RQ.

Proof: See Appendix.

The next axiom, which is taken from Chakravarty and Maharaj (2011a), is specified in terms of population heterogeneity. Following arguments put forward by Horowitz (1985), Montalvo and Reynal Querol (2005, p. 797) argued that "there is less violence in …highly heterogeneous societies". Now in a society consisting of k ethnic groups $E_1, E_2, ..., E_k$, heterogeneity will be maximum if all the E_i's are of equal population share $1/k$. The index value $P(1/k, 1/k, ..., 1/k)$ should be quite small here. So, if k is allowed to vary, then for larger values of k, size of E_i will be even smaller and one can expect that polarization vanishes in the limit. This motivates us to state

Axiom 6: $P(1/k, 1/k, ..., 1/k) \to 0$ as $k \to \infty$.

The following theorem can now be demonstrated.

Theorem 4.5: The only EPI $P:\Delta \to R$ that satisfies Axioms 2, 3, $4'$, 5 and 6 is RQ.

Proof: See Appendix.

Since Theorems 4.1-4.5 axiomatically characterize the same polarization index using alternative sets of postulates, the following theorem can be stated:

Theorem 4.6: Let $P:\Delta \to R$ be an EPI. Then the following statements are equivalent:

(i) P satisfies Axioms 2-5 and Property 1c.

(ii) P satisfies Axioms 1-3, $4'$, 5 and Property 1.

(iii) P satisfies Axioms 1-3, $4'$, 5 and Property 1b.

(iv) P satisfies Axioms 2, 3, $4'$, 5 and Property 2.

(v) P satisfies Axioms 2, 3, $4'$, 5 and 6

(vi) P coincides with *RQ* given by (4.1).

The following corollary to Lemma 2 shows that in each of the Theorems 4.1-4.5 we can replace Axioms 2 and 3 by their stronger versions (that is, Axioms $2'$ and $3'$ respectively), where

Axiom $2'$: For all $k \in \Omega$, $P(\underline{\pi}) = 0$ if and only if $\underline{\pi} \in \Delta_k$ is some permutation of (1,0,…,0).

Axiom $3'$**:** For all $k \in \Omega$, $P(\underline{\pi})=1$ if and only if $\underline{\pi} \in \Delta_k$ is some permutation of $\left(\frac{1}{2}, \frac{1}{2}, 0, \ldots, 0\right)$.

Corollary 4.1: Let $P : \Delta \to R$ be an arbitrary EPI satisfying Axioms 1, $4'$ and the strong versions of Axioms 2 and 3 (Axioms $2'$ and $3'$ respectively) if and only if it is of the form

$$RQ_\theta = 4\sum_{i=1}^{k} \pi_i^2 (1-\pi_i) + \theta \sum_{1 \le i_1 < i_2 < i_3 \le k} \pi_{i_1}\pi_{i_2}\pi_{i_3},$$

where $\theta \in [0,3)$ is an arbitrary constant, $k \in \Gamma$ and $\underline{\pi} = (\pi_1, \pi_2, \ldots, \pi_k) \in \Delta_k$ are arbitrary,

Proof: See Appendix.

4.4 Ethnic Polarization Quasi-orderings

In Section 4.3 we have characterized *RQ* and its generalizations using alternative sets of axioms. However, all these indices may not rank two different ethnic distributions in the same way. It therefore becomes worthwhile to investigate whether two ethnic distributions can be ordered by a class of EPIs satisfying certain desirable properties. The first question we address along this line is: is there any possibility of uniform ranking of ethnic profiles by the family $\{RQ_\theta : 0 \le \theta \le 3\}$? That is, we examine the possibility of identical ranking of ethnic profiles by RQ_θ for all values of $\theta \in [0,3]$.

Definition 4.2: An ethnic distribution $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_k) \in \Delta_k$ is said to be at most as polarized as $\underline{\pi'} = (\pi'_1, \pi'_2, ..., \pi'_k) \in \Delta_k$, what we write $\underline{\pi} \leq_{POL} \underline{\pi'}$, if we have, $P(\underline{\pi}) \leq P(\underline{\pi'})$ for all EPI $P : \Delta \rightarrow R$ satisfying Axioms 1-3, $4'$ and 5.

The following result identifies an equivalent form of the quasi-ordering. An attractive feature of the result is that instead of looking at all values of $\theta, 0 \leq \theta \leq 3$, we can confine our attention to only two values of θ viz. $\theta = 0$ and $\theta = 3$. Thus, once specific directional inequalities involving RQ_θ for these two values of θ are satisfied, we can be sure about the partial ordering of ethnic distributions using $\leq_{POL}$.

Theorem 4.7: For arbitrary $\underline{\pi}, \underline{\pi'} \in \Delta_k$ the following conditions are equivalent:

(i) $\underline{\pi} \leq_{POL} \underline{\pi'}$.

(ii) (a) $RQ(\underline{\pi}) \leq RQ(\underline{\pi'})$ and (b) $RQ(\underline{\pi}) + 3\Sigma_3(\underline{\pi}) \leq RQ(\underline{\pi'}) + 3\Sigma_3(\underline{\pi'})$, where $\Sigma_3(\underline{\pi}) = \sum_{1 \leq i_1 < i_2 < i_3 \leq k} \pi_{i_1}\pi_{i_2}\pi_{i_3}$.

Proof: See Appendix.

Remark 4.4: The quasi-ordering $\leq_{POL}$ is clearly transitive but not complete. To see this, consider two ethnic distributions $\left(\frac{1}{2}, \frac{1}{2} - \varepsilon, \varepsilon\right)$ and $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, where $\varepsilon > 0$ is small. Then $RQ\left(\frac{1}{2}, \frac{1}{2} - \varepsilon, \varepsilon\right) > RQ\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$; but $RQ_3\left(\frac{1}{2}, \frac{1}{2} - \varepsilon, \varepsilon\right) < RQ_3\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) (= 1)$.

It is easy to see that the distribution $\underline{m} = \left(1/2, 1/2, 0, ..., 0\right)$ [0 being repeated $(k-2)$ times for any pre-fixed $k \geq 3$] is a maximal element in Δ_k with respect to $\leq_{POL}$ in the sense that there does not exist any other ethnic distribution $\underline{\pi} \in \Delta_k$ such that $\underline{m} \leq_{POL} \underline{\pi}$. Similarly, one can verify that $(1, , 0, ..., 0)$ is a minimal element with respect to $\leq_{POL}$.

Condition (ii) in Theorem 4.7 is not easily implementable. We therefore develop an alternative notion of quasi-ordering, which is easy to verify using a simple graphical device. For this purpose, in the rest of the section we assume that the given k ethnic groups $E_1, E_2, \ldots, E_k$ ($k \geq 3$) are non-increasingly ordered with respect to their population sizes. We denote the set of all non-increasingly ordered ethnic profiles for k ethnic groups by Δ^*_k. Following Esteban and Ray (1994), we also assume that there is a 'small number of significantly sized groups'.

We begin by specifying the following property of an EPI.

Property A: A rank-preserving population shift from E_i to E_j (with $j > i \geq 2$) cannot increase polarization.

The idea behind this property relies on the following observation of Esteban and Ray (2008b; p.175): "…. If the two groups are small enough, the equalization of population will" make *IC* lower. It is also seen that RQ satisfies Property A. [This follows from the simple fact that $(\pi_i + \pi_{i+1}) \leq 2/3$ as each one of π_i and π_{i+1} is at most $1/3$; for otherwise $\pi_1 > 1/3$ which implies that $(\pi_1 + \pi_i + \pi_{i+1}) > 1$, a contradiction.]

The next property is specified in terms of the population shift between the first two groups.

PropertyB: A rank-preserving population shift from E_1 to E_2 cannot decrease polarization.

The reasoning behind this postulate is simple. As is widely known in the literature, when the sizes of other groups are not significantly large, the ethnic polarization attains its maximum when sizes of the two largest groups come close to one another (Montalvo and Reynal Querol, 2005; p. 798). Since we have assumed that there is 'a small number of significantly sized groups', an equalization of the sizes of the two largest groups brings them closer so that polarization goes up. This again is supported by the Esteban and Ray (2008b, p.175) observation: "A transfer of population from a group to a smaller one increases *IC* if both groups are larger than $1/3$". It may also be noted that RQ obeys property B if $(\pi_1 + \pi_2) \geq 2/3$. In particular, *RQ* satisfies it for $k = 3$.

Definition 4.3: For two ethnic distributions $\underline{\pi} = (\pi_1, \pi_2, \ldots, \pi_k)$, $\underline{\pi}' = (\pi_1', \pi_2', \ldots, \pi_k') \in \Delta_k^*$, we say that $\underline{\pi}$ is at most as polarized as $\underline{\pi}'$, what we denote as $\underline{\pi} \leq_P \underline{\pi}'$, if for all EPI $P : \Delta_k^* \to R$ satisfying Axioms 1-3 and Properties A and B we have $P(\underline{\pi}) \leq P(\underline{\pi}')$.

An example of an EPI satisfying Axioms 1-3 and Properties A and B is: $P_0(\underline{\pi}) = 4\pi_1\pi_2$.

The following theorem gives some simple necessary conditions for $\underline{\pi} \leq_P \underline{\pi}'$ to hold.

Theorem 4.8: Let $\underline{\pi}, \underline{\pi}' \in \Delta_k^*$. A set of necessary conditions for $\underline{\pi} \leq_P \underline{\pi}'$ to hold is given by:

$$\pi_2 \leq \pi_2' \text{ and } \left(2\pi_2 + \sum_{i=3}^{l} \pi_i\right) \leq \left(2\pi_2' + \sum_{i=3}^{l} \pi_i'\right) \text{ for all } 3 \leq l \leq k. \tag{4.12}$$

Proof: See Appendix.

Remark 4.4: The above necessary conditions stated in terms of $(k-1)$ inequalities involving population proportions of the two distributions are very easy to check. The last inequality implies that $(\pi_1 - \pi_2) \geq (\pi_1' - \pi_2')$. However, these conditions are not sufficient for $\underline{\pi} \leq_P \underline{\pi}'$. To see this, let $\underline{\pi} = (0.4, 0.3, 0.3)$ and $\underline{\hat{\pi}} = (0.34, 0.34, 0.32)$. Then the system of inequalities in (4.12) holds. But observe that $\pi_1^*\pi_2^* > \hat{\pi}_1\hat{\pi}_2$ so that for the EPI $P_0(\underline{\pi}) = 4\pi_1\pi_2$, $P_0(\underline{\pi}^*) > P_0(\underline{\hat{\pi}})$ and hence $\underline{\pi}^* \leq_P \underline{\hat{\pi}}$ does not hold.

The following result gives an easily verifiable sufficient condition in terms of population concentration curve for $\leq_P$ to hold. The population concentration curve of an ethnic distribution is a plot of the cumulative population shares against the cumulative number of groups, with groups ranked from the largest to the smallest.

Theorem 4.9: For $\underline{\pi}, \underline{\pi}' \in \Delta_k^*$, a set of sufficient conditions for $\underline{\pi} \leq_P \underline{\pi}'$ is given by:

$$\pi_1 \geq \pi_1' \text{ and } \sum_{i=1}^{l} \pi_i \leq \sum_{i=1}^{l} \pi_i' \text{ for all } 2 \leq l \leq (k-1). \tag{4.13}$$

What Theorem 4.9 says is the following. If the population share of the largest ethnic group under $\underline{\pi}$ is at least as large as that under $\underline{\pi}'$ and the population concentration curve of

$(\pi_1 + \pi_2, \pi_3, \ldots\ldots, \pi_k)$ is dominated by (that is, does not lie above) that of $(\pi_1' + \pi_2', \pi_3', \ldots\ldots, \pi_k')$, then $\underline{\pi}'$ is regarded at least as polarized as $\underline{\pi}$ by all EPIs satisfying Axioms 1-3 and Properties

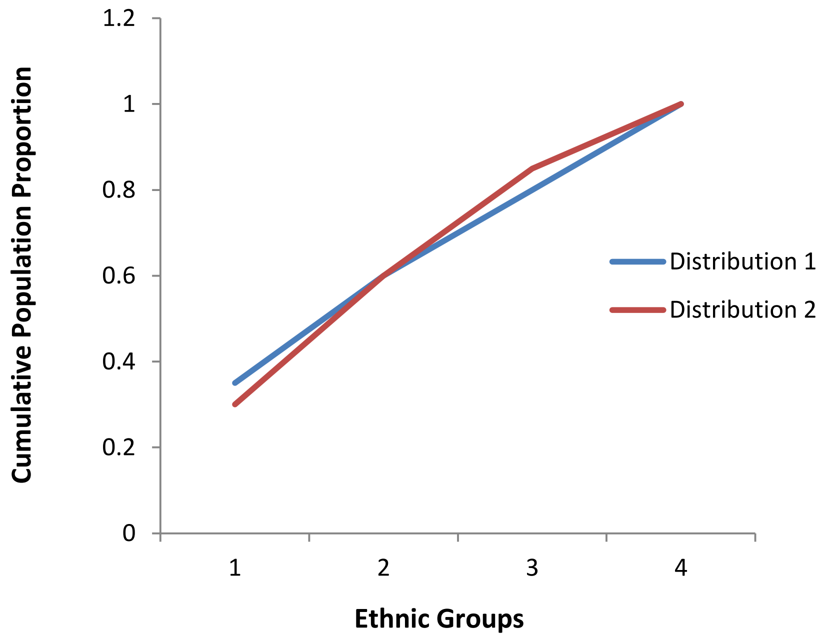

Fig. 4.1 (a sufficient condition for the odering $\leq_P$ using population concentration curves)

A and B. Since it is easy to check the inequality between the highest population groups of the distributions concerned and dominance between the required population concentration curves, condition (4.13) can be quickly implemented. Therefore, the sufficient condition for the quasi-ordering $\leq_P$ given in Definition 4.3 becomes a useful tool for ranking alternative ethnic distributions.

Proof of Theorem 4.9: See Appendix.

Remark 4.5: (a) Since $\sum_{i=1}^{k} \pi_i = 1 = \sum_{i=1}^{k} \pi_i'$, the last condition in (12) implies that $\pi_k \geq \pi_k'$.

(b) To see that the set of conditions in (4.13) is not necessary, consider $\underline{\pi}^1 = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$ and $\underline{\pi}^2 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \in \Delta_3$, for which condition (4.13) is violated but by definition, $\underline{\pi}^2 \leq_P \underline{\pi}^1$.

(c) Since a population concentration curve satisfies Zero-frequency- independence, we can extend Theorem 4.9 to the case where the number of ethnic groups is variable. The corresponding EPIs will, of course, have to satisfy Axioms 1-3 and Properties A and B.

(d) The quasi-ordering $\leq_P$ defined on Δ is obviously transitive but is incomplete, as has been observed in the previous remark. Since a necessary condition for the quasi-ordering $\underline{\pi} \leq_P \underline{\pi}'$ to hold is given by (4.12), incompleteness in the quasi-ordering arises because of violation of one or more of the $(k-1)$ conditions. Given two ethnic distributions, there is no guarantee that these two conditions will be satisfied simultaneously and hence the quasi-ordering is incomplete.

Given $\underline{\pi}, \underline{\pi}' \in \Delta_k^*$, we have thus evolved a mechanism for verifying the ranking relation $\underline{\pi} \leq_P \underline{\pi}'$. First, we have to check the conditions in (4.12). If one or more of them fails to hold, the possibility of $\underline{\pi} \leq_P \underline{\pi}'$ is ruled out. Once they are satisfied, we move to the conditions in (4.13). If they also hold, then the underlying profiles can be ranked by $\leq_P$. In case of violation of one or more of them, no specific conclusion can be drawn.

The following proposition demonstrates a relationship between the two orderings considered in this section.

Proposition 4.1: $\leq_{POL}$ does not imply $\leq_P$ and $\leq_P$ does not imply $\leq_{POL}$.

Proof: See Appendix.

4.5 Conclusion

An indicator of ethnic polarization is an indicator of the extent of ‘identification’ and ‘alienation’ of an ethnic population. Reynal-Querol (2002) introduced an index of ethnic polarization, which is referred to as the *RQ* index. Montalvo and Reynal-Querol (2005, 2008) investigated its properties in detail and explained its role empirically as an explanatory variable for incidence of civil wars. In this chapter we have developed two different ethnic polarization quasi-orderings that can rank alternative ethnic profiles unambiguously by all possible ethnic polarization indices satisfying certain axioms. One of these quasi-orderings can be easily checked using the population concentration curve. In the process, we also characterize the *RQ* index and a generalized version of it. Our quasi-orderings and characterizations rely mostly on some axioms and propositions of polarization suggested by Montalvo and Reynal-Querol (2005, 2008) and Esteban and Ray (1994, 1999, 2008b).

4.6 Appendix

Proof of Observation 4.1: If possible, let $\eta_0 = (2/3 - \rho)$, where $\rho > 0$ is small. In view of the above discussion, the distribution $(2/3 - \rho/2, 1/3 + \rho/2)$ should be less polarized than $(1/3 + \rho/2, 1/3, 1/3 - \rho/2)$), which is obtained from the former by a split of the larger group. But this contradicts Property 1. Hence, $\eta_0 \geq 2/3$.□

Proof of Lemma 4.1: Suppose *P* is an EPI satisfying Axioms 2–5 and Property 1. By Axiom 5 we have,

$$P(\pi, 1-\pi) = a_1 + a_2\, d^2((1/2, 1/2), (\pi, 1-\pi)), \tag{4.14}$$

where a_1, a_2 are constants with $a_2 < 0$. By Axiom 3, $P(1/2, 1/2) = 1$. Putting $\pi = 1/2$ in (4.14) we get, $a_1 = 1$. Also, by Axiom 2, $P(1, 0) = 0$. Putting $\pi = 0$ and using $a_1 = 1$ in (4.14) we further get $a_2 = -2$.

Substitution of $a_1 = 1$ and $a_2 = -2$ in (4.14) yield:

$$P(\pi, 1-\pi) = 1 - (1 - 2\pi)^2 = 4\pi(1-\pi). \tag{4.15}$$

Consequently, (4.5) becomes:

$$P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) = \frac{4\varepsilon(\pi_i - \varepsilon)}{\pi_i^2} f(\pi_1, \pi_2, \ldots\ldots, \pi_{i-1}), \tag{4.16}$$

whenever $\pi_i < \eta_0$.

Now, let $k = 3$ and $(\pi_1, \pi_2, \pi_3) \in \Delta$. Note that one of the sums $(\pi_1 + \pi_2)$, $(\pi_2 + \pi_3)$ and $(\pi_3 + \pi_1)$ has to be at most η_0, for, otherwise their sum, that is, $2(\pi_1 + \pi_2 + \pi_3)$ exceeds $3\eta_0 \geq 2$

(in view of Observation 1, $\eta_0 \geq 2/3$). This then gives $(\pi_1 + \pi_2 + \pi_3) > 1$, which is a contradiction. So, without loss of generality we can assume that $(\pi_2 + \pi_3) \leq \eta_0$.

For $k = 3$, equation (4.16) gives $P(\pi_1, \pi_2, \pi_3) = P(\pi_1, \pi_2 + \pi_3, 0) + \dfrac{4\pi_2\pi_3}{(\pi_2 + \pi_3)^2} f(\pi_1)$,

which in view of zero frequency independence becomes

$$P(\pi_1, \pi_2 + \pi_3) + \frac{4\pi_2(1 - \pi_1 - \pi_2)}{(1 - \pi_1)^2} f(\pi_1)$$

$$= 4\pi_1(1 - \pi_1) + \pi_2(1 - \pi_1 - \pi_2)\, g(\pi_1), \tag{4.17}$$

where $g(\pi_1) = 4f(\pi_1) \big/ (1 - \pi_1)^2$ and $\pi_1 > 0$.

By symmetry we can replace the second expression in (4.17) by $4\pi_2(1 - \pi_2) + \pi_1(1 - \pi_1 - \pi_2)\, g(\pi_2)$ and rearrange terms in the resulting expression to get

$$4\{\pi_1(1 - \pi_1) - \pi_2(1 - \pi_2)\} = (1 - \pi_1 - \pi_2)\{\pi_1 g(\pi_2) - \pi_2 g(\pi_1)\}, \tag{4.18}$$

from which it follows that

$$4(\pi_1 - \pi_2)(1 - \pi_1 - \pi_2) = (1 - \pi_1 - \pi_2)\{\pi_1 g(\pi_2) - \pi_2 g(\pi_1)\}. \tag{4.19}$$

Equation (4.19), on simplification, becomes

$\{\pi_1 g(\pi_2) - \pi_2 g(\pi_1)\} = 4(\pi_1 - \pi_2)$ (since $(\pi_1 + \pi_2) < 1$), from which we get

$$\pi_1\{g(\pi_2) - 4\} = \pi_2\{g(\pi_1) - 4\}. \tag{4.20}$$

Equation (4.20) implies that

$$\frac{g(\pi_1)-4}{\pi_1} = \frac{g(\pi_2)-4}{\pi_2}, \tag{4.21}$$

whenever $\pi_1, \pi_2 \in (0,1)$, $(\pi_1 + \pi_2) < 1$ and $\pi_2 \le \pi_1$.

Define $h:(0,1) \to R$ by $h(p) = \{g(p)-4\}/p$, where $0 < p < 1$. Then from (4.21) it follows that $h(\pi_1) = h(\pi_2)$ whenever $\pi_1, \pi_2 \in (0,1)$, $(\pi_1 + \pi_2) < 1$ and $\pi_2 \le \pi_1$. Clearly,

$$h(p) = h(1/2), \tag{4.22}$$

whenever $0 < p < 1$ and $(p + 1/2)$ <1, that is, whenever $0 < p < 1/2$. If $1/2 < p < 1$, then there exists $\delta \in (0, 1/2)$ such that $p < (1-\delta)$. So, by (4.20) we have, $h(p) = h(\delta)$ [$\because (p+\delta) < 1$]. But by (4.21), $h(\delta) = h(1/2)$. Combining (4.20) and (4.21) we have, $h(p) = h(1/2)$. Thus, in all cases, $h(p) = h(1/2) = c$, a constant, whenever $0 < p < 1$. Then $g(p) - 4 = cp$, for some constant c, which implies that

$$g(p) = (cp + 4), \tag{4.23}$$

for some constant c, where $0 < p < 1$. By continuity of g, this holds for all $p \in [0,1]$.

From (4.15) we have, $P(\pi_1, \pi_2) = 4\sum_{i=1}^{2} \pi_i^2 (1-\pi_i)$, where $(\pi_1, \pi_2) \in \Delta$. Using (4.22), it can be shown that $P(\pi_1, \pi_2, \pi_3) = 4\sum_{i=1}^{3} \pi_i^2 (1-\pi_i) + \theta_3 \pi_1 \pi_2 \pi_3$, where $\theta_3 \in R$ and $(\pi_1, \pi_2, \pi_3) \in \Delta$.

Next, let us consider the case $k = 4$. Let $(\pi_1, \pi_2, \pi_3, \pi_4) \in \Delta$. Without loss of generality we can assume that $(\pi_3 + \pi_4) \le \eta_0$. (Since given $\eta_0 \ge 2/3$, $(\pi_3 + \pi_4) > \eta_0$ implies that $(\pi_1 + \pi_2) = \{1 - (\pi_3 + \pi_4)\} < (1-\eta_0) < \eta_0$.) From (4.16) for $k = 4$, $\pi^{(i)} = (\pi_1, \pi_2, \pi_3 + \pi_4, 0)$ and $\varepsilon = \pi_4$, we then have

$P(\pi_1, \pi_2, \pi_3, \pi_4)$

$$= P\left(\pi_1, \pi_2, \pi_3 + \pi_4, 0\right) + \frac{4\pi_3\pi_4}{\left(\pi_3 + \pi_4\right)^2} f_2\left(\pi_1, \pi_2\right)$$

$$= 4\sum_{i=1}^{4} \pi_i^{\,2}\left(1 - \pi_i\right) + 4\pi_3\pi_4\left\{2 - 3\left(\pi_3 + \pi_4\right)\right\} + \theta_3\pi_1\pi_2\left(\pi_3 + \pi_4\right) + \frac{4\pi_3\pi_4}{\left(1 - \pi_1 - \pi_2\right)^2} f_2\left(\pi_1, \pi_2\right)$$

$$= 4\sum_{i=1}^{4} \pi_i^{\,2}\left(1 - \pi_i\right) + 4\pi_3\pi_4\left\{g_2\left(\pi_1, \pi_2\right) + 3\left(\pi_1 + \pi_2\right) - 1\right\} + \theta_3\pi_1\pi_2\left(\pi_3 + \pi_4\right), \tag{4.24}$$

where $g_2\left(\pi_1, \pi_2\right) = \dfrac{f_2\left(\pi_1, \pi_2\right)}{\left(1 - \pi_1 - \pi_2\right)^2}$. Equation (4.24), on rearrangement, becomes

$$P\left(\pi_1, \pi_2, \pi_3, \pi_4\right) = 4\sum_{i=1}^{4} \pi_i^{\,2}\left(1 - \pi_i\right) + 4\pi_3\pi_4 h_2\left(\pi_1, \pi_2\right) + \theta_3\pi_1\pi_2\left(\pi_3 + \pi_4\right), \tag{4.25}$$

with $h_2\left(\pi_1, \pi_2\right) = \left\{g_2\left(\pi_1, \pi_2\right) + 3\left(\pi_1 + \pi_2\right) - 1\right\}$. By symmetry we have, $4\pi_3\pi_4 h_2\left(\pi_1, \pi_2\right) + \theta_3\pi_1\pi_2\left(\pi_3 + \pi_4\right) = 4\pi_2\pi_4 h_2\left(\pi_1, \pi_3\right) + \theta_3\pi_1\pi_3\left(\pi_2 + \pi_4\right) = 4\pi_2\pi_3 h_2\left(\pi_1, \pi_4\right) + \theta_3\pi_1\pi_4\left(\pi_2 + \pi_3\right)$.

The first two expressions of the above equation, after rearrangement and simplification give

$$\frac{4h_2\left(\pi_1, \pi_2\right) - \theta_3\left(\pi_1 + \pi_2\right)}{\pi_1\pi_2} = \frac{4h_2\left(\pi_1, \pi_3\right) - \theta_3\left(\pi_1 + \pi_3\right)}{\pi_1\pi_3}.$$

Thus, $\dfrac{4h_2\left(\pi_1, \pi_2\right) - \theta_3\left(\pi_1 + \pi_2\right)}{\pi_1\pi_2}$ is independent of π_2. Similarly, dealing with the last two expressions we can conclude that $\dfrac{4h_2\left(\pi_1, \pi_2\right) - \theta_3\left(\pi_1 + \pi_2\right)}{\pi_1\pi_2}$ is independent of π_1.

Hence, $\frac{4h_2(\pi_1,\pi_2)-\theta_3(\pi_1+\pi_2)}{\pi_1\pi_2}=\theta_4$, a constant. Substituting the value of $h_2(\pi_1,\pi_2)$ in (4.25) we get, $P(\pi_1,\pi_2,\pi_3,\pi_4)=4\sum_{i=1}^{4}\pi_i^2(1-\pi_i)+\theta_4\pi_1\pi_2\pi_3\pi_4+\theta_3\{\pi_1\pi_2(\pi_3+\pi_4)+\pi_3\pi_4(\pi_1+\pi_2)\}$

$$=4\sum_{i=1}^{4}\pi_i^2(1-\pi_i)+\sum_{i=3}^{4}\theta_i S_i(\underline{\pi}). \tag{4.26}$$

To conclude the proof of the lemma, we use induction on *k*. So, let us assume that $P(\pi_1,\pi_2,\ldots,\pi_k)=4\sum_{i=1}^{k}\pi_i^2(1-\pi_i)+\sum_{i=3}^{k}\theta_i S_i(\underline{\pi})$, for arbitrary $(\pi_1,\pi_2,\ldots,\pi_k)\in\Delta$. Consider $(\pi_1,\pi_2,\ldots,\pi_k,\pi_{k+1})\in\Delta$. As before, without loss of generality, we can assume that $(\pi_k+\pi_{k+1})\le\eta_0$. Then

$$P(\pi_1,\pi_2,\ldots,\pi_k,\pi_{k+1})=P(\pi_1,\pi_2,\ldots,\pi_k+\pi_{k+1},0)+\frac{4\pi_k\pi_{k+1}}{(\pi_k+\pi_{k+1})^2}f_k(\pi_1,\pi_2,\ldots,\pi_{k-1})$$

$$=4\left[\sum_{i=1}^{k+1}\pi_i^2(1-\pi_i)+\pi_k\pi_{k+1}\{2-3(\pi_k+\pi_{k+1})\}\right]+\sum_{i=3}^{k-1}\theta_i S_i(\underline{\pi}')+\theta_3\sum_{i_1<i_2\le k-1}\pi_{i_1}\pi_{i_2}(\pi_k+\pi_{k+1})+$$

$$\theta_4\sum_{i_1<i_2<i_3\le k-1}\pi_{i_1}\pi_{i_2}\pi_{i_3}(\pi_k+\pi_{k+1})+\ldots..+\theta_k\prod_{i=1}^{k-1}\pi_i(\pi_k+\pi_{k+1})+\pi_k\pi_{k+1}g_k(\pi_1,\pi_2,\ldots.,\pi_{k-1}),$$

where $S_l(\underline{\pi}')=\sum_{1\le i_1<i_2<\ldots.<i_l\le(k-1)}\pi_{i_1}\pi_{i_2}\ldots\pi_{i_l}$ for $3\le l\le(k-1)$. Simplifying the above equation we get,

$$P(\pi_1,\pi_2,\ldots,\pi_{k+1})=4\sum_{i=1}^{k+1}\pi_i^2(1-\pi_i)+\sum_{i=3}^{k}\theta_i S_i(\pi_1,\pi_2,\ldots,\pi_{k+1})+\pi_k\pi_{k+1}h_k(\pi_1,\pi_2,\ldots.,\pi_{k-1}),$$

for some function $h_k(\pi_1,\pi_2,\ldots.,\pi_{k-1})$. By symmetry

$$\pi_k\pi_{k+1}h_k(\pi_1,\pi_2,\ldots.,\pi_{k-1})=\pi_i\pi_j h_k(\pi_1,\pi_2,\ldots.,\pi_{i-1},\pi_{i+1},\ldots,\pi_{j-1},\pi_{j+1},\ldots,\pi_k) \tag{4.27}$$

for all $i < j$. Hence, $\dfrac{h_k(\pi_1, \pi_2, ..., \pi_{k-1})}{\pi_1 \pi_2 .... \pi_{k-1}}$ = constant $= \theta_{k+1}$, say. Consequently, $P(\pi_1, \pi_2, ..., \pi_{k+1}) = 4 \sum_{i=1}^{k+1} \pi_i^2 (1 - \pi_i) + \sum_{i=3}^{k+1} \theta_i S_i (\pi_1, \pi_2, ..., \pi_{k+1})$, thereby proving our induction hypothesis. □

Proof of Theorem 4.1: Suppose $P : \Delta \to R$ satisfies Axioms 2-5 and Property 1c. First observe that Property 1c implies Property 1. Thus, all the conditions of Lemma 4.1 are met. So, *P* must be of the form (4.6) with arbitrary constants $\theta_l \in R$ for $3 \le l \le k$. But by Property 1c, $P\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = P\left(\frac{2}{3}, \frac{1}{3}\right)$ so that $RQ\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = RQ\left(\frac{2}{3}, \frac{1}{3}\right) + \frac{1}{27}\theta_3$, which implies that $\theta_3 = 0$.

We complete the proof using induction on l . Suppose we have shown that $\theta_i = 0$ for all $i < l$. We shall show that $\theta_l = 0$. Consider the uniform distribution $\left(\frac{1}{l}, \frac{1}{l}, ..., \frac{1}{l}\right)$, which, after the merger of last $(l - 1)$ groups gives rise to the distribution $\left(\frac{1}{l}, 1 - \frac{1}{l}\right)$. Property 1c then demands that $RQ\left(\frac{1}{l}, \frac{1}{l}, ..., \frac{1}{l}\right) + \sum_{i=3}^{l} \theta_i S_i \left(\frac{1}{l}, \frac{1}{l}, ..., \frac{1}{l}\right) = RQ\left(\frac{1}{l}, 1 - \frac{1}{l}\right)$, which in view of induction hypothesis becomes $RQ\left(\frac{1}{l}, \frac{1}{l}, ..., \frac{1}{l}\right) + \theta_l \binom{k}{l} \frac{1}{l^l} = RQ\left(\frac{1}{l}, 1 - \frac{1}{l}\right)$. It is easy to see that

$$RQ\left(\frac{1}{l}, \frac{1}{l}, ..., \frac{1}{l}\right) = RQ\left(\frac{1}{l}, 1 - \frac{1}{l}\right). \tag{4.28}$$

Consequently, we must have $\theta_l = 0$, thereby proving our claim. Hence, by induction, $\theta_i = 0$ for all i, $3 \le i \le k$ whence we catch hold of RQ .

The fact that *RQ* satisfies Axioms 2-5 is demonstrated in the 'converse' part of the proof of Lemma 4.2 stated below. Satisfaction of Property 1c by *RQ* follows from Theorem 2 of Montalvo and Reynal Querol (2008) and (4.28). □

Proof of Lemma 4.2 Suppose an EPI satisfies Axioms 1–5. In addition, if *P* satisfies Axiom $4'$, then

$P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) == g_1(\pi_i)P\left(1-\frac{\varepsilon}{\pi_i}, \frac{\varepsilon}{\pi_i}\right)$ which implies that

$$P\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - P\left(\underline{\pi}^{(i)}\right) = f_2(\pi_i)\varepsilon(\pi_i - \varepsilon), \tag{4.29}$$

for some function $f_2 : [0,1] \to R$.

Proceeding exactly in the same way as in the proof of Lemma 4.1 and applying condition (4.29) we get: $\theta_l = 0$ for all $l \geq 4$. This, in turn, yields RQ_θ , as defined in (4.11).

Now, to find the required bounds on θ in (4.11), we note that for any fixed $k \in \Gamma \setminus \{2\}$, for the vector $\underline{\gamma}_k = \left(\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}\right) \in \Delta_k$, $P\left(\underline{\gamma}_k\right) = \left\{\frac{4}{k^2}\left(1-\frac{1}{k}\right)k + \binom{k}{3}\frac{\theta}{k^3}\right\}$ = $\frac{4}{k^2}(k-1) + \frac{\theta(k-1)(k-2)}{6k^2}$. By Axiom 1 we have, $0 \leq P\left(\underline{\gamma}_k\right) \leq 1$, which implies that $0 \leq \frac{4}{k^2}(k-1) + \frac{\theta(k-1)(k-2)}{6k^2} \leq 1$. From the last inequality we get, $4 + \frac{\theta(k-2)}{6} \geq 0$ and $\frac{\theta(k-2)}{6} \leq \left(\frac{k^2}{(k-1)} - 4\right)$, from which it follows that $\theta \geq \frac{-24}{(k-2)}$ and $\theta \leq \frac{6(k-2)}{(k-1)}$. This holds for all $k \geq 3$. So, we must have, $0 \leq \theta \leq 3$. Therefore, we are through with one half of the proof of the lemma.

Now, we proceed for a proof of the converse. It is trivial to verify that RQ_θ satisfies Axioms 2, 3 and 5. Take $\underline{\pi}^{(i)}$ and $\underline{\pi}_{\varepsilon}^{(i)}$ as in Axiom $4'$. Then it turns out that $RQ_\theta\left(\underline{\pi}_{\varepsilon}^{(i)}\right) - RQ_\theta\left(\underline{\pi}^{(i)}\right) = 4\varepsilon(\pi_i - \varepsilon)(3\pi_i - 2) + \theta\varepsilon(\pi_i - \varepsilon)(1-\pi_i) = \varepsilon(\pi_i - \varepsilon)\{(12-\theta)\pi_i - (8-\theta)\}$. Hence Axiom $4'$ follows.

To prove Axiom 1, observe that the non-negativity of RQ_θ is quite clear. Next, we have to show that $RQ_\theta \leq 1$. Clearly, $RQ_\theta = RQ$ if $k = 2$ and so in this case the proof follows directly from (4.1). Then, for various values of $k \geq 3$, we employ the method of Lagrange Multipliers to find the extreme values for RQ_3 at the interior of Δ_k. Note that for a given $(\pi_1, \pi_2, \pi_3, \ldots, \pi_k) \in \Delta_k$, RQ_θ is an increasing function of θ. This implies that for all $\theta \in [0,3]$, $RQ_\theta \leq RQ_3$. Therefore, to show that $RQ_\theta \leq 1$ for all $\theta \in [0,3]$, it is enough to show that $RQ_3 \leq 1$.

Case 1: k=3: Applying the method of Lagrange multipliers, the extreme points of RQ_3 in Δ_3 turn out to be $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ and all those points which are permutations of $\left(\frac{4}{9}, \frac{4}{9}, \frac{1}{9}\right)$.

Now, $RQ_3\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = 1$ and $RQ_3\left(\frac{4}{9}, \frac{4}{9}, \frac{1}{9}\right) = 4\left[2.\left(\frac{4}{9}\right)^2.\frac{5}{9} + \left(\frac{1}{9}\right)^2.\frac{8}{9}\right] + 3.\left(\frac{4}{9}\right)^2.\frac{1}{9} = \frac{80}{81} < 1$. Hence, for all (π_1, π_2, π_3) in the interior of Δ_3, we have, $RQ_3(\pi_1, \pi_2, \pi_3) \leq 1$.

Next, we need to show that the value of RQ_3 on the boundary $\partial\Delta_3$ of Δ_3 is less than or equal to 1. Observe that $\partial\Delta_3 = \bigcup_{i=1}^{3}\{(\pi_1, \pi_2, \pi_3) \in \Delta_3 : \pi_i = 0\}$. Then it is easily seen that $\max_{(\pi_1, \pi_2, \pi_3) \in \Delta_3} RQ_3(\pi_1, \pi_2, \pi_3) = 1$. This shows that $RQ_3 \leq 1$ in this case also.

Case 2: k$\geq$ 4: Using the method of Lagrange multipliers again, extreme points of RQ_3 are found as $\left(\frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}\right)$ and all those points which are permutations of $\left(\frac{4}{9(k-2)}, \frac{4}{9(k-2)}, \ldots\ldots, \frac{4}{9(k-2)}, 1 - \frac{4(k-1)}{9(k-2)}\right)$.

Now, observe that $RQ_3\left(\frac{1}{k}, \frac{1}{k}, \ldots., \frac{1}{k}\right) = \left\{4\frac{1}{k^2}\left(1 - \frac{1}{k}\right)k\right\} + \left\{3\frac{1}{k^3}\binom{k}{3}\right\} = 4\frac{(k-1)}{k^2} + \frac{(k-1)(k-2)}{2k^2} = \frac{(k-1)(k+6)}{2k^2} < 1$ since, given $k > 3$, $2k^2 - (k-1)(k+6) = (k^2 - 5k + 6) = (k-2)(k-3) > 0$.

Next, consider

$$RQ_3\left(\frac{4}{9(k-2)},\frac{4}{9(k-2)},\ldots\ldots,\frac{4}{9(k-2)},1-\frac{4(k-1)}{9(k-2)}\right) = \frac{8(k-1)}{729\,(k-2)^3}\left(69k^2-300k+324\right) = \psi(k),$$

say. To see that $\psi(k) \le 1$ for all $k \ge 4$, let us now define a map $\phi:[0,\infty)\to R$ by $\phi(x) = 729\,(x-2)^3 - 8(x-1)(69x^2-300x+324) = (177x^3-1422x^2+3756x-3240)$. Clearly, $\phi(4) = 360 > 0$. Also $\phi'(x) = (531x^2-2844x+3756)$ and $\phi''(x) = (1062x-2844)$, where ϕ' and ϕ'' denote respectively the first and second order derivatives of ϕ. Therefore, $\phi''(x) > 0$ for all $x \ge 3$, which indicates that ϕ' is increasing on $[3,\infty)$. Also note that $\phi'(4) = 876 > 0$. Consequently, $\phi'(x) \ge \phi'(4) > 0$ for all $x \ge 4$. So, ϕ is increasing on $[4,\infty)$, from which it follows that $\phi(x) > 0$ for all $x \ge 4$. This proves that the denominator of $\psi(k)$ is greater than its numerator. Thus, $\psi(k) \le 1$ for all $k \ge 4$.

Hence, $\max\limits_{(\pi_1,\pi_2,\pi_3,\ldots,\pi_k)\in\Delta_k^0} RQ_3(\pi_1,\pi_2,\pi_3,\ldots,\pi_k)$ <1, where Δ_k^0 denotes the interior of Δ_k. It follows that $\max\limits_{(\pi_1,\pi_2,\pi_3,\ldots,\pi_k)\in\Delta_k} RQ_3(\pi_1,\pi_2,\pi_3,\ldots,\pi_k)$, which is $\ge RQ_3\left(\frac{1}{2},\frac{1}{2},0,\ldots,0\right)$ =1, cannot be attained at any interior point of Δ_k. But since Δ_k is compact and RQ_3 is continuous on it, the maximum has to be attained at some point on the boundary $\partial\Delta_k$ (Rudin, 1987, p. 89). It is easily seen that $\partial\Delta_k = \bigcup_{i=1}^{k}\{(\pi_1,\pi_2,\pi_3,\ldots,\pi_k)\in\Delta_k : \pi_i = 0\}$. So finding $\max RQ_3$ in $\partial\Delta_k$ amounts to finding the maximum in Δ_{k-1}. Repeating the same argument, this maximum can again be found only on the boundary $\partial\Delta_{k-1}$.

Descending thus, finally we come down to Δ_3 wherein, by Case1, the maximum is 1. This establishes the fact that RQ_3 satisfies Axiom 1. The proof of the lemma is now complete.□

Proof of Theorem 4.2: By Lemma 4.2, we know that RQ_θ given by (4.10) is the only EPI that satisfies Axioms 1-3, $4'$, 5. We now show that the only value of θ, $0 \le \theta \le 3$, for which RQ_θ satisfies Property 1 is $\theta = 0$. Note that $RQ_\theta(p,q,r) - RQ_\theta(p,\tilde{q},0) =$

$qr\{(12-\theta)(q+r)-(8-\theta)\}<0$ if and only if $(q+r)<\left(\frac{8-\theta}{12-\theta}\right)$, where $\tilde{q}=(q+r)$ and $0\le\theta\le 3$. This should hold for all $q,r\ge 0$ satisfying $q\ge r$ and $\{1-(q+r)\}>q$. Let $\rho\in(0,2/3)$ be fixed. Choosing $p=\left(\frac{1}{3}+\rho\right), q=r=\left(\frac{1}{3}-\frac{\rho}{2}\right)$, we require $\left(\frac{2}{3}-\rho\right)<\left(\frac{8-\theta}{12-\theta}\right)$. Now, letting $\rho\to 0$, we demand that $\left(\frac{8-\theta}{12-\theta}\right)\ge\frac{2}{3}$, which implies that $\theta\le 0$. So the only possibility is $\theta=0$, in which case RQ_θ coincides with *RQ*. Theorem 1 of Montalvo and Reynal-Querol (2008) shows that *RQ* satisfies Property 1. With this we complete the proof of the theorem. □

Proof of Theorem 4.3: Lemma 4.2 shows that RQ_θ given by (4.10) is the only EPI that satisfies Axioms 1-4. We now apply Property 1b to RQ_θ with *m*=2 and proceed exactly in the same way as in the proof of Theorem 1 to conclude that $\theta=0$. Theorem 2 of Montalvo and Reynal-Querol (2008) shows that *RQ* satisfies Property 1b. This completes the proof of the theorem. □

Proof of Theorem 4.4: In the proof of Lemma 4.2 it was shown that the only EPI satisfying Axioms 2-3, $4'$, 5 is of the form RQ_θ, where $\theta\in R$. (Axiom 1 was necessary and sufficient to prove that $0\le\theta\le 3$.) Let $\underline{\pi}^1=(p,p,1-2p)$, where $0<p<1/2$. Then $RQ_\theta(\underline{\pi}^1)=8p(1-3p+3p^2)+\theta p^2(1-p)$. Now, let RQ_θ satisfy Property 2. So, we require $\frac{\partial}{\partial p}RQ_\theta(\underline{\pi}^1)\ge 0$, that is, $\{(72p^2-48p+8)+\theta(2p-6p^2)\}\ge 0$. Equivalently, we require $E(p)=\{(36-3\theta)p^2-(24-\theta)p+4\}\ge 0$ for all $p\in\left(0,\frac{1}{2}\right)$. Now, first let $\theta>4$. Then choosing a sequence $\{p_n\}$ in $\left(0,\frac{1}{2}\right)$ such that $\{p_n\}\to\frac{1}{2}$ we require, $E(p_n)\ge 0$ for all $n\in N$ and hence we demand that $E\left(\frac{1}{2}\right)\ge 0$. But as a matter of fact we find: $E\left(\frac{1}{2}\right)=\left\{(36-3\theta)\frac{1}{4}-(24-\theta)\frac{1}{2}+4\right\}=\left(1-\frac{\theta}{4}\right)<0$, a contradiction. So, we necessarily have $\theta\le 4$. But then $(36-3\theta)>0$ and the discriminant of $E=(24-\theta)^2-16(36-3\theta)=\theta^2$. Hence the minimum value of E is $\frac{-\theta^2}{4(36-3\theta)}$, which is negative unless $\theta = 0$. Consequently, the only possibility is $\theta = 0$. This generates *RQ*. Theorem 3 of Montalvo and Reynal-Querol (2008) shows that the *RQ* index satisfies Property 2. This completes the proof of the theorem. □

Proof of Theorem 4.5: It has already been observed that *RQ* verifies Axioms 2, 3, $4'$ and 5. To check satisfaction of Axiom 6, simply note that $RQ(1/k, 1/k, \ldots., 1/k) = \frac{4}{k}\left(1-\frac{1}{k}\right) \to 0$ as $k \to \infty$. Conversely, by Lemma 4.2, the only EPI $P:\Delta \to R$ satisfying Axioms 1-3, $4'$, 5 is of the form (4.11). So, $RQ_\theta(1/k, 1/k, \ldots., 1/k) = \frac{4}{k}\left(1-\frac{1}{k}\right) + \theta \frac{1}{k^3}\binom{k}{3} \to \theta$ as $k \to \infty$. So, by Axiom 6, $\theta = 0$, which yields *RQ*.□

Proof of Corollary 4.1: Let $P:\Delta \to R$ satisfy Axioms 1, $2'$, $3'$ and $4'$. Obviously, then P satisfies Axioms 1-3, $4'$. Hence by Lemma 2, P must be of the form RQ_θ, with $\theta \in [0,3]$. However, the value $\theta = 3$ is not admissible since in that case, $P\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \ldots., 0\right) = 1$, contrary to Axiom 3. Thus, $\theta \in [0,3)$.

To prove the converse, we have to verify that RQ_θ verifies Axioms 1, $2'$ - $4'$. In the proof of Lemma 4.2, we have already checked that RQ_θ satisfies Axioms 1-3 and $4'$. So all that remains to be shown are 'only if' parts of Axioms $2'$ and $3'$. First observe that $RQ_\theta(\underline{\pi})$ can vanish only if $\pi_i^2(1-\pi_i) = 0$, $1 \le i \le k$, that is, only if $\pi_j = 1$ for some j, $1 \le j \le k$ and $\pi_i = 0$ for all $i \ne j$, $1 \le i \le k$, which means that $(\pi_1, \pi_2, \ldots, \pi_k)$ is a permutation of $(1, 0, \ldots, 0)$. Hence, Axiom $2'$ holds.

To prove validity of Axiom $3'$, consider first the case $k = 2$. Then $RQ_\theta = RQ$ and so, from (4.1), $RQ_\theta(\pi_1, \pi_2) = 1$ only if $(\pi_1, \pi_2) = \left(\frac{1}{2}, \frac{1}{2}\right)$. Next, let $k = 3$. As seen in the proof of Lemma 4.2, for all (π_1, π_2, π_3) in the interior of Δ_3, we have, $RQ_3(\pi_1, \pi_2, \pi_3) \leq 1$ and consequently, for all $\theta \in [0,3)$, $RQ_\theta(\pi_1, \pi_2, \pi_3) < 1 = RQ_\theta\left(\frac{1}{2}, \frac{1}{2}, 0\right)$.

Since Δ_3 is closed and bounded, it is compact. Continuity of RQ_θ on Δ_3 ensures that it achieves its global maximum at some point inside Δ_3 (Rudin, 1987, p.89). However, it is evident from our discussion in the previous paragraph that this point is not in the interior of Δ_3. Therefore, it must lie somewhere on the boundary $\partial\Delta_3$ of Δ_3, which is given by $\partial\Delta_3$ = $\bigcup_{i=1}^{3}\{(\pi_1, \pi_2, \pi_3) \in \Delta_3 : \pi_i = 0\}$. So, it clearly follows that $\max_{(\pi_1,\pi_2,\pi_3)\in\Delta_3} RQ_\theta(\pi_1, \pi_2, \pi_3) = 4.\frac{1}{4} = 1$, the maximum being attained only at those points which are permutations of $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$.

Finally, the case $k \geq 4$ can be dealt with in exactly the same way as in the proof of Lemma 4.1. All we have to do is to replace RQ_3 by RQ_θ. The conclusion is: $\max_{(\pi_1,\pi_2,\pi_3,\ldots,\pi_k)\in\Delta_k} RQ_\theta(\pi_1, \pi_2, \pi_3, \ldots, \pi_k)$ will be attained only at those points, which are permutations of $\left(\frac{1}{2}, \frac{1}{2}, \underbrace{0, \ldots, 0}_{(k-2)times}\right)$. This completes the proof of the corollary.□

Proof of Theorem 4.7: $(i) \Rightarrow (ii)$ **:** By Lemma 4.2, the only EPI $P : \Delta \to R$ satisfying Axioms 1-3, $4'$ and 5 is of the form (10). Suppose (i) holds. Then $RQ(\underline{\pi}) \leq RQ(\underline{\pi'})$ for all $\theta \in [0,3]$. Putting $\theta = 0$ and $\theta = 3$ in this inequality we get (*a*) and (*b*) respectively.

$(ii) \Rightarrow (i)$ **:** Assume (*a*) and (*b*). Given $\theta \in [0,3]$, let $\tau = \theta/3$ so that $\tau \in [0,1]$. Then

$RQ_\theta(\underline{\pi}) = RQ(\underline{\pi}) + \theta\Sigma_3(\underline{\pi}) = (1-\tau)RQ(\underline{\pi}) + \tau\{RQ(\underline{\pi}) + 3\Sigma_3(\underline{\pi})\}$

$\leq (1-\tau)RQ(\underline{\pi'}) + \tau\{RQ(\underline{\pi'}) + 3\Sigma_3(\underline{\pi'})\} = RQ(\underline{\pi'}) + \theta\Sigma_3(\underline{\pi'}) = RQ_\theta(\underline{\pi'})$. This holds for all $\theta \in [0,3]$

Therefore, we have $\underline{\pi} \leq_{POL} \underline{\pi'}$.□

Proof of Theorem 4.8: Note that the EPIs $P_l : \Delta_k^* \to R$, $2 \le l \le k$, defined by $P_1(\pi_1, \pi_2, ..., \pi_k) = 2\pi_2$ and $P_l(\pi_1, \pi_2, ..., \pi_k) = 2\pi_2 + \sum_{i=3}^{l} \pi_i$ for $3 \le l \le k$ satisfy Axioms 1-3 and Properties A and B. Hence the result follows by the defining condition on $\le_P$. □

Proof of Theorem 4.9: Let $P : \Delta_k^* \to R$ be an EPI satisfying Axioms 1-3 and Properties A and B. Also let,

$$\pi_1 \ge \pi_1' \tag{4.30}$$

$$\text{and} \quad \sum_{i=1}^{l} \pi_i \le \sum_{i=1}^{l} \pi_i' \tag{4.31}$$

for all $2 \le l \le (k-1)$. Note that (4.31) holds for all $2 \le l \le k$ since the last sum is simply unity in both the cases. Now, transform $\underline{\pi}$ into $\underline{\pi^0} = (\pi_1', \pi_2 + \pi_1 - \pi_1', \pi_3, ..., \pi_k)$ by a shift of mass $(\pi_1 - \pi_1')$ from E_1 to E_2. This transformation is rank-preserving since $\{\pi_1' - (\pi_2 + \pi_1 - \pi_1')\} = 2\pi_1' - (\pi_2 + \pi_1) \ge (\pi_1' + \pi_2') - (\pi_2 + \pi_1) \ge 0$, by (4.31) and $\{(\pi_2 + \pi_1 - \pi_1') - \pi_3\} \ge (\pi_2 + \pi_1 - \pi_1 - \pi_3) = (\pi_2 - \pi_3) \ge 0$, by (4.30). Hence, by Property B, this shift cannot decrease polarization; that is,

$$P(\underline{\pi}) \le P(\underline{\pi^0}). \tag{4.32}$$

Again, $\underline{\pi^0}$ and $\underline{\pi'} = (\pi_1', \pi_2', \pi_3', ..., \pi_k')$ have the same mean $(1/k)$ and $\sum_{i=1}^{l} \pi_i^0 = \sum_{i=1}^{l} \pi_i \le \sum_{i=1}^{l} \pi_i'$ (where π_i^0 denotes the i^{th} element of $\underline{\pi^0}$) for all $2 \le l \le k$. Since $\pi_1^0 = \pi_1'$, it follows that $\sum_{i=1}^{l} \pi_i^0 \le \sum_{i=1}^{l} \pi_i'$ for all $1 \le l \le k$. Hence by a Theorem of Hardy et al. (1934; see Chakravarty, 2009, p. 6), it follows that $\underline{\pi'}$ can be transformed into $\underline{\pi^0}$ by a sequence of rank-preserving population shifts from higher to lower groups. Clearly, these shifts keep E_1 unchanged. Consequently by Property A, $\underline{\pi^0}$ cannot be more polarized than $\underline{\pi'}$. So,

$$P\left(\underline{\pi}^{0}\right) \leq P\left(\underline{\pi}'\right). \tag{4.33}$$

(4.32) and (4.33) together yield: $P(\underline{\pi}) \leq P(\underline{\pi}')$ for all EPI $P : \Delta_k^* \to R$ satisfying Properties A and B. Hence, $\underline{\pi} \leq_P \underline{\pi}'$. □

Proof of Proposition 4.1: Let $\underline{\pi} = (0.4, 0.3, 0.3)$ and $\underline{\pi}' = (0.34, 0.34, 0.32)$. We observe that $RQ(\underline{\pi}) < RQ(\underline{\pi}')$ as well as $RQ_3(\underline{\pi}) < RQ_3(\underline{\pi}')$ so that $\underline{\pi} \leq_{POL} \pi'$. But, as we have seen already, $\underline{\pi}$ and $\underline{\pi}'$ are not comparable with respect to $\leq_P$. Thus, $\leq_{POL}$ does not imply $\leq_P$.

To demonstrate the other part of the proposition, consider $\underline{\alpha} = (.35, .30, .25, .1)$ and $\underline{\beta} = (.4, .25, .25, .1)$. Then, using condition (12), $\underline{\beta} \leq_P \underline{\alpha}$. However, a simple calculation shows that $RQ(\underline{\alpha}) < RQ(\underline{\beta})$ and $RQ_3(\underline{\alpha}) > RQ_3(\underline{\beta})$ so that $\underline{\alpha}, \underline{\beta}$ are not comparable with respect to $\leq_{POL}$. Hence $\leq_P$ does notimply $\leq_{POL}$. □

Chapter 5

GENERALIZED GINI POLARIZATION INDICES FOR AN ORDINAL DIMENSION OF HUMAN WELL-BEING[20]

5.1. Introduction

This chapter is going to be a natural continuation of the previous two chapters. In Chapter 3, we worked on income polarization while in Chapter 4 we discussed polarization in case of an ethnic data (more generally, in a categorical set up). Henceforth one may be interested in developing the theory of polarization in an ordinal context.

It is now well-known that human well-being is a multidimensional phenomenon (Sen, 1985, 1987). Examples of such dimensions are income, wealth, health, literacy and so on. While some of the dimensions correspond to ratio scale variables (e.g., income, wealth), dimensions like health and literacy are represented by ordinal variables.

To recall from Chapter 1, considering the data on self-assessed health status, Apouey (2007) characterized a family of bipolarization indices in case of ordinal data. This chapter is going to be an extension of her work.

The Gini index of income inequality is the most popular index of inequality. (For a discussion, see, among others, Sen, 1973; Blackorby and Donaldson, 1978; Lambert, 2001 and Chakravarty, 2009.) It has also been used in other income distribution-based measurement problems, including poverty (Sen, 1976; Takayama, 1979), deprivation (Yitzhaki, 1979; Kakwani, 1984; Chakravrty, 2009) and polarization (Wang and Tsui, 2000; Chakravarty and Majumder, 2001; Foster and Wolfson, 2010). Generalized versions of the Gini index have been analyzed, among others, by Mehran (1976), Donaldson and Weymark (1980, 1983), Weymark (1981), Yaari (1987, 1988), Bossert (1990) and Ben Porath and Gilboa (1994). It is a natural question whether a family of generalized Gini indices of polarization can as well be constructed for an ordinally significant dimension.

---

[20] A major portion of this chapter has been published in Chakravarty and Maharaj (2015).

The objective of this chapter is to make an attempt along this direction and suggest a family of the generalized Gini indices of polarization for an ordinal dimension using an axiomatic framework. An advantage of this family is that its members can be used to make inter-population comparisons of polarization. While making such comparisons we assume that the number of ranked categories is the same across the populations. Some implications of the axioms are also investigated. Apart from analyzing the properties of the suggested family and characterizing it axiomatically, we develop a quasi-ordering induced by the family of indices for ranking two alternative distributions of the ordinal dimension. This type of investigation is quite common in the literature. For instance, Foster and Shorrocks (1988) studied the variable-line poverty orderings for the members of the Foster-Greer-Thorbecke (1984) family. Foster and Jin (1998) characterized similar orderings for the Dalton utility-gap poverty measures. Examples of indices that are included under this class are the Chakravarty (1983) and Hagenaars (1987) indices. Partial orderings obtained from specific indices may offer new insights into what a particular index is trying to measure.

Generalized Gini indices of income inequality aggregate incomes of different individuals using some sequence of weights of real numbers. However, in the present case in order to formulate the axioms, following the literature, we start with the cumulative proportions of persons in different categories of the dimension under consideration. This forces us to use the distribution function instead of the population proportions. Although the aggregation procedures in the generalized Gini indices of income inequality and in our case are essentially the same, the restrictions imposed on the sequences turn out to be different. (See Section 5.4.)

The chapter is organized as follows. In the next section we present the axioms for an index of polarization of a dimension with ordinal characteristic. Section 5.3 looks at some implications of the axioms. The generalized Gini indices are analyzed and characterized in Section 5.4, while the induced quasi-ordering is discussed in Section 5.5. Finally, Section 5.6 concludes.

5.2 Axioms for an Index of Polarization of a Dimension with Ordinal Significance

Consider a population comprising *n* categories $E_1, E_2, ..., E_n$, ranked in ascending order of some ordinal characteristic, where $n \in N$, the set of positive integers. Let π_i denote the proportion of individuals in E_i. Therefore, $0 \le \pi_i \le 1$, $1 \le i \le n$ and $\sum_{i=1}^{n} \pi_i = 1$, $n \in N$ being arbitrary. This generates a probability distribution $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_n)$, which we will refer to as 'ordinal distribution'. We write Δ_n for the set of all (ordered) discrete probability distributions of dimension n and $\mathbf{\Delta}$ for theset of all probability distributions. Evidently, $\Delta = \bigcup_{n=2}^{\infty} \Delta_n$. The cumulative proportion of persons who are in category c and lower ones in the distribution $\underline{\pi}$ is defined as $F_c^{\underline{\pi}} = \sum_{i=1}^{c} \pi_i$. We write $m(F)$ for the median category so that for any $\underline{\pi} \in \Delta$, $F_{m(F)-1}^{\underline{\pi}} < \frac{1}{2}$ and $F_{m(F)}^{\underline{\pi}} \ge \frac{1}{2}$.

Note that corresponding to each $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_n) \in \Delta_n$, there is a vector of cumulative proportions $F^{\underline{\pi}} = \left(F_1^{\underline{\pi}}, F_2^{\underline{\pi}}, ..., F_n^{\underline{\pi}}\right) \in \Delta_n^*$, where $\Delta_n^* = \left\{(x_1, x_2, ..., x_n) \in \Delta_n : 0 \le x_1 \le x_2 \le ... \le x_n = 1\right\}$ (and vice versa). Define $\Delta^* = \bigcup_{n=2}^{\infty} \Delta_n^*$.

By a polarization index for an ordinal dimension (PIO), which we denote by P, we mean a continuous real-valued function defined on Δ^*, that is, $P : \Delta^* \to R$ .Thus, for any $\underline{\pi} \in \Delta$, $P\left(F^{\underline{\pi}}\right)$ indicates the level of polarization associated with the ordinal distribution $\underline{\pi}$ or equivalently, with the distribution function $F^{\underline{\pi}}$.

Intrinsic to the notion of polarization for an ordinally measurable dimension are Increased Spread (**IS**) and Increased Bipolarity (**IB**). The former is a monotonicity condition. It says that polarization should not decrease if there is a spread in the distribution away from the median category. In other words, greater distancing between the categories below and not below the median category should not make the distribution less polarized. Since such changes in proportions of individuals in categories below and not below the median category widen the distribution, polarization does not decrease. Apouey (2007) adapted the following form of IS from Allison and Foster (2004).

Increased Spread (IS): If the ordinal distributions $\underline{\pi}$ and $\underline{\sigma}$ have the same number of categories (n) and the same median category m; $F_c^{\underline{\sigma}} \geq F_c^{\underline{\pi}}$ for all $c < m$ and $F_c^{\underline{\sigma}} \leq F_c^{\underline{\pi}}$ for all $c \geq m$, with at least one category $c < m$ such that $F_c^{\underline{\sigma}} > F_c^{\underline{\pi}}$ or one category $c \geq m$ such that $F_c^{\underline{\sigma}} < F_c^{\underline{\pi}}$, then $P(\underline{\sigma}) > P(\underline{\pi})$.

Increased Bipolarity is a bunching or a clustering principle. In order to state this postulate formally, we need to define a transfer.

Definition 5.1: Given the ordinal distribution $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_n)$, we say that the distribution $\underline{\pi'} = (\pi'_1, \pi'_2, ..., \pi'_n)$ is obtained from $\underline{\pi}$ through a transfer if for two given categories c and $\hat{c}$ such that $\hat{c} - 1 \geq c + 1$ there are shifts of population proportion δ from c to $c + 1$ and from $\hat{c}$ to $\hat{c} - 1$, where $\pi_c \geq \delta$, $\pi_{\hat{c}} \geq \delta$ and $\pi'_i = \pi_i$ for all $i \neq \hat{c}, \hat{c} - 1, c, c + 1$; $\pi'_{\hat{c}-1} = \pi_{\hat{c}-1} + \delta$, $\pi'_{\hat{c}} = \pi_{\hat{c}} - \delta$, $\pi'_c = \pi_c - \delta$, $\pi'_{c+1} = \pi_{c+1} + \delta$. In terms of distribution functions, $F_i^{\underline{\pi'}} = F_i^{\underline{\pi}}$ for all $i \neq c,\ \hat{c} - 1$; $F_c^{\underline{\pi'}} = F_c^{\underline{\pi}} - \delta$ and $F_{\hat{c}-1}^{\underline{\pi'}} = F_{\hat{c}-1}^{\underline{\pi}} + \delta$.

The notion of transfer considered here, which is taken from Apouey (2007), says that there is a shift of a population mass δ from a category c to the next higher category $(c + 1)$ and a shift of a population mass of the same size δ from a category $\hat{c}$,which is higher than $(c + 1)$, to the preceding category $(\hat{c} - 1)$. If all the categories affected by the transfer are below (not below) the median category, we say the post-transfer distribution is obtained from the pre-transfer one by a transfer below (not below) the median category. Thus, for the transfer to be below the median category we must have $\hat{c} \leq m(F)$. Likewise, the transfer is not below the median category only if $c \geq m(F)$. In the case of health polarization if we take 'good' as the median category then a transfer from 'very poor' to 'poor' and from 'fair' to 'poor' categories can be taken as a transfer below the median category. Formally,

Definition 5.2: Given any two distributions $F^1 = \left(F_1^1, F_2^1, \ldots., F_n^1\right)$ and $F^2 = \left(F_1^2, F_2^2, \ldots., F_n^2\right)$ with the same median category *m*, we say that F^2 is derived from F^1 by transfers below the median category if and only if the highest category involved in the transfer is $\hat{c} \leq m(F)$. Similarly, we say that there are transfers not below the median category if and only if the lowest category involved in the transfer is $c \geq m(F)$.

Since transfers on the same side of the median category brings the individuals closer to each other, homogeneity among the individuals within the categories increases. This increase in within-group homogeneity does not decrease polarization. This is what is demanded by Apouey's (2007) Increased Bipolarity axiom, which is a modified version of the corresponding Wang-Tsui (2000) axiom.

Increased Bipolarity (IB): If the ordinal distributions $\underline{\pi}$ and $\underline{\sigma}$ have the same median category *m* andif $F^{\underline{\pi}}$ is transformed into $F^{\underline{\sigma}}$ by at least one transfer below the median category or at least one transfer not below the median category, then $P\left(F^{\underline{\sigma}}\right) > P\left(F^{\underline{\pi}}\right)$.

Apouey (2007) also suggested some additional postulates for a PIO. Before presenting these desiderata, we propose the following axiom which says that a PIO is bounded between zero and one.

Boundedness (BO): For all $n \in N$, $\underline{\pi} \in \Delta_n$, $0 \leq P\left(F^{\underline{\pi}}\right) \leq 1$.

The next axiom which we refer to as weak homogeneity is Apouey's (2007) 'Minimum Polarization'.

Weak Homogeneity (WH): $P\left(F^{\underline{\pi}}\right) = 0$ if $F^{\underline{\pi}}$ is of the form $(0, \ldots, 0, 1, \ldots, 1)$, where the first 1 appears in the i^{th} place for some $1 \leq i \leq n$.

This axiom says that the PIO achieves its minimum value, zero, if there is complete homogeneity in the sense that all the individuals belong to a particular categorical group identified by the ordinal dimension. The following axiom is a stronger version of WH.

Strong Homogeneity (SH): $P\left(F^{\pi}\right)=0$ if and only if F^{π} is of the form $(0,\ldots,0,1,\ldots,1)$.

However, it is worth mentioning here that the definition of the minimum polarization is not uncontroversial. Cowell and Flachaire (2014) points out some limitations on the corresponding definition of minimum level of inequality.

The next axiom, which we call ‘perfect bipolarity’, is ‘Maximum Polarization’ axiom of Apouey (2007).

Perfect Bipolarity (PB): $P\left(\frac{1}{2},\ldots\ldots,\frac{1}{2},1\right)=1$.

According to this axiom, when the number of categories (n) is kept fixed, the PIO is maximized if there is an equal splitting of the entire population into two extreme categories, the lowest and the top. We can also consider the following stronger form of PB as an axiom.

Strong Bipolarity (SB): $P\left(F^{\pi}\right)=1$ if and only if $F^{\pi}=\left(\frac{1}{2},\ldots\ldots,\frac{1}{2},1\right)$.

Perfect Bipolarity is sufficient for the PIO to take on the value 1. However, Strong Bipolarity gives us a necessary and sufficient condition showing that only in the perfect bipolar case the PIO can assume the value 1. It may be mentioned here that the inequality measures (for ordinally measurable variables) suggested by Reardon (2009) satisfy axioms resembling SH and PB. In other words, they take the maximum value when half the population belongs to the poorest income category and the remaining half to the richest income category. As regards the other extreme, the minimum is attained by these measures only when all the individuals belong to the same category.

5.3 Some Implications of the Axioms

In this section we look at certain implications of some of our axioms proposed in the earlier section. The first result shows that it is impossible to find a non-degenerate convex PIO that satisfies **WH**. Convexity of PIO demands that for any $\underline{\pi},\underline{\hat{\pi}}\in\Delta_n$,

$P\left(\lambda F^{\underline{\pi}} + (1-\lambda) F^{\underline{\hat{\pi}}}\right) \leq \lambda P\left(F^{\underline{\pi}}\right) + (1-\lambda) P\left(F^{\underline{\hat{\pi}}}\right)$, where $\lambda \in [0,1]$ is arbitrary. This means that the polarization of the distribution obtained by a smoothing of two distributions will not increase the value of the index, where smoothing refers to any convex combination of the two distributions. Thus, given a distribution, convexity of a PIO also indicates that we can reduce the level of polarization by smoothing it with another distribution with lower level of polarization than the initial distribution.

Proposition 5.1: There does not exist anynon-degenerate convex PIO satisfying **WH**.

Proof: See Appendix.

The following results, which are easy to establish, drop out as corollaries to Proposition 5.1.

Corollary 5.1: No non-degenerate convex PIO satisfies **SH**.

Corollary 5.2: No convex PIO satisfies **SH** and **SB** simultaneously.

The next proposition indicates an implication of **IS**.

Proposition 5.2: Consider a PIO $P : \Delta \rightarrow R$ satisfying **IS**. Let there be a shift of mass $\varepsilon > 0$, sufficiently small, from a higher category to a lower one. (a) (i) If both the categories are below the median category the shift increases polarization. (ii) The higher is the higher category; the greater is the amount of increase. (b) (i) If both the categories are not below the median category, then the shift decreases polarization. (ii) The lower is the lower category, the greater is the amount of reduction.

Proof: See Appendix.

The higher amount of increase (reduction) in polarization following from a shift of population mass from a higher category bears similarity with the diminishing transfers principle of Kolm (1976) and the positional transfers principle for inequality and poverty indices considered by Mehran (1976) and Kakwani (1980) respectively. The former requires reduction in inequality resulting from a progressive transfer between two persons with a given income difference by a higher amount, if the incomes are lower than when they are higher. On the other hand, the latter demands that a progressive transfer will reduce inequality by a larger quantity the lower the income of the donor is, given that the number of individuals between the donor and the recipient is fixed. In the present case a shift of a population mass from a higher to a lower category increases /reduces polarization by a larger amount as the rank difference between the categories, given the rank of the higher category, increases. This finding bears a striking similarity with the relevant observation in Fusco and Silber (2014). Considering a population partitioned into ordered categories and unordered population subgroups, these authors suggested a principle of 'swap' of individuals between ordered categories. The principle demands that "Swaps of individuals who are farther apart (as far as their income class is concerned) should have a greater impact on polarization than swaps of individuals who are closer (as far as their income class is concerned)". Proposition 5.2 is in perfect agreement with the above result.

The next proposition looks at an implication of **IS** and **IB**.

Proposition 5.3: Consider a PIO $P:\Delta \to R$ satisfying **IS** and **IB**. Let there be a transfer of mass $\varepsilon > 0$, sufficiently small, below the median category. Then the resulting increase in polarization is less than the increase in polarization due to a shift of the same amount of mass from the highest category to the corresponding lowest category involved in the transfer.

Proof: See Appendix.

The postulates **IB** and **IS** of a PIO are definitely not substitutes of each other. The reason is that they are concerned with two different notions of alterations in the original distribution and the PIO looks at the changes in the level of polarization resulting from these distributional alterations. But as shown in Proposition 5.3, we can make an unambiguous comparison between the impacts of **IS** and **IB** in a specific situation and clearly the theorem indicates that in this particular case IS has a higher impact on polarization than **IB**.

Remark 5.1: In the case of transfer not below the median category, the shift of the population mass takes place from the lowest category to the highest category involved in the transfer and the impact of IS turns out to be higher than that of **IB**.

5.4 Generalized Gini Polarization Indices for an Ordinally Significant Dimension

In this section we propose a generalized Gini index of polarization for an ordinally measurable dimension and investigate its properties, including a characterization. Apouey (2007) argued that polarization can be measured in terms of the 'distance between the observed situation' $F^{observed} = (F_1, F_2, \cdots F_{n-1}, 1)$ and 'the distribution of maximum polarization' $F^{max} = \left(\frac{1}{2}, \frac{1}{2}, \cdots, \frac{1}{2}, 1\right)$ (op. cit., p.882). The idea is to look at the distance from the symmetric bipolar distribution, which has the maximum polarization equal to unity. In view of this we can define a polarization index directly on *F*.

We measure polarization in terms of the deviations of the components of the actual distribution from those of the perfectly homogeneous distribution $F^{min} = (0, \ldots, 0, 1, \ldots, 1)$ (where the first 1 comes in the $m(F)^{th}$ slot), which has the minimum polarization (equal to 0). [Note that Apouey (2007) assigned the minimum level of polarization to all distributions in which the entire population is clubbed into a single category. However, for our characterization exercise, we begin with the distribution where the corresponding category is the median category. This does not contradict Apouey's axiom on minimum polarization as we don't insist that this is the only category with the minimum value.] In contrast, the polarization measure suggested by Apouey (2007) uses deviation of the observed situation from the situation of maximum polarization. However, it is evident that the central ideas underlying the two notions are essentially the same.

In view of the above consideration, structure of the polarization index can be of the form:

$$P(F) = \sum_{i=1}^{m(F)-1} w_i\left(m(F)\right) F_i + \sum_{i=m(F)}^{n-1} w_i\left(m(F)\right)\left(1 - F_i\right) \qquad , \qquad (5.1)$$

where $w_i(m(F))$ is the weight assigned to category i and is actually dependent on the median category $m(F)$ of the distribution.

For a non-decreasingly ordered vector of incomes $\underline{z} = (z_1, z_2, \cdots, z_n)$, the generalized Gini index of inequality can be expressed as $I_{GG}(\underline{z}) = 1 - \frac{1}{\mu}\sum_{i=1}^{n} a_i z_i$, where $\mu > 0$ is the mean income and the sequence of positive weights $\{a_i\}$ is decreasing and satisfies $\sum_{i=1}^{n} a_i = 1$ (see Weymark, 1981). We can rewrite $I_{GG}(\underline{z})$ as $I_{GG}(\underline{z}) = \frac{1}{\mu}\sum_{i=1}^{n}(\mu - z_i)a_i$, which is a linear normalized weighted sum of deviations of the actual distribution of incomes from the perfectly equal distribution where everyone enjoys the mean income. If $a_i = \frac{2(n-i)+1}{n^2}$, then I_{GG} coincides with the Gini index.

The generalized Gini structure is similar to (5.1) because the latter is the linear sum of weighted deviations between components of $F^{observed}$ and $F^{\min}$. In other words, (5.1) uses a generalized Gini type aggregation and hence we can refer to it as a generalized Gini index of PIO.

The family of generalized Gini polarization indices can be characterized as follows. Consider a shift of mass $\varepsilon > 0$ (sufficiently small) from category *i* to category $i+1$. By IS we can expect a decrease in polarization if $i < m(F)$ while an increase results in case $i \geq m(F)$. In the subsequent axiom, we assume a simple specification for this change in polarization.

Proportional Increment/ Decrement (PI): Any shift of mass $\varepsilon > 0$ (sufficiently small) from a category i to the next higher category $(i+1)$ results in a change of polarization proportional to ε (and vice versa). Formally,

$$P(F) - P(F_1, F_2, \ldots, F_i - \varepsilon, F_{i+1}, \ldots, F_{n-1}, 1) = c(i, m(F))\varepsilon\ , \tag{5.2}$$

where the constant $c\left(i, m(F)\right)$ depends on i as well as the median category $m(F)$.

Under the assumption that the median category is given, this axiom specifies that the level of change in polarization resulting from a shift of a population mass of size ε from a category i to the next higher category $(i+1)$ depends on the amount of the shift and a constant dependent on the affected categories. Clearly if $(i+1) \leq m(F)$, then under IS, $P\left(F_1, F_2, \ldots, F_i - \varepsilon, F_{i+1}, \ldots, F_{n-1}, 1\right) < P(F)$, whereas the opposite happens if $(i+1) > m(F)$. So, given IS, (5.2) can be rewritten as

$$P(F) - P\left(F_1, F_2, \ldots, F_i - \varepsilon, F_{i+1}, \ldots, F_{n-1}, 1\right) = c\left(i, m(F)\right)\varepsilon$$

where $c\left(i, m(F)\right) > 0$ if $(i+1) \leq m(F)$ and $c\left(i, m(F)\right) < 0$ for $m(F) \leq i \leq (n-1)$.

Since under the shift all categories other than *i* and $(i+1)$ remain unaffected, PI assumes that these changes are independent of unaffected categories. While there can be many more general formulations of these changes, the linear specification offered by PI is simple and easy to understand. The assumption made for the polarization change under PI is similar in nature to the inequality change, as demonstrated by several inequality indices, under a rank preserving transfer of income from a person to a richer one. For instance, for the squared coefficient of variation this change is directly proportional to the product of the size of the transfer and the income difference of the persons concerned.

Theorem 5.1: A PIO *P* satisfies **WH, IS** and **PI** if and only if *P* is given by (5.1), where $w_i\left(m(F)\right) > 0$ for all $1 \leq i \leq (n-1)$.

Proof: See Appendix.

Corollary 5.3: A generalized Gini PIO satisfying **IS, PI** and **WH** necessarily satisfies **SH**.

Proof: See Appendix.

We thus have

Theorem 5.2: For a PIO $P : \Delta_n \to R$ the following statements are equivalent:

(i) P satisfies **IS, WH** and **PI.**

(ii) P satisfies **IS, SH** and **PI**.

(iii) P is the generalized Gini-index given by (5.1). □

In Theorem 5.1 **IB** was not imposed as an axiom. The following theorem identifies the properties of $\{w_i\}$ for the generalized Gini PIO family, characterized in Theorem 5.1, to satisfy **IB**.

Theorem 5.3: The generalized Gini PIO in (5.1) satisfies **IB** if and only if $w_{i+1}(m(F)) > w_i(m(F))$ for all $i \le (m(F)-2)$ and $w_{i+1}(m(F)) < w_i(m(F))$ for all $m(F) \le i \le (n-1)$.

Proof: See Appendix.

Theorem 5.3 establishes that the generalized Gini PIO agrees with the Increased Bipolarity axiom if and only if the weight sequence $\{w_i\}$ is increasing up to the category $(m(F)-1)$ and decreasing thereafter. It may be worthwhile to compare Theorems 5.1 and 5.3 with Proposition 3 of Wang and Tsui (2000) who suggested the use of $P_{WT}(\underline{z}) = \frac{1}{m(\underline{z})}\sum_{i=1}^{n} b_i z_i$ as an income polarization index, where the income distribution $\underline{z} = (z_1, z_2, \cdots, z_n)$ is non-decreasingly ordered, $m(\underline{z})$ is the median income associated with $\underline{z}$ and $\{b_i\}$ is a sequence of real numbers. Clearly, this is a generalized Gini type index (see Weymark, 1981 and Bossert, 1990). Wang and Tsui (2000) have demonstrated that P_{WT} satisfies the Increased Spread criterion, that is, increasingness of P_{WT} under a reduction (an increment) in any $z_i < m(\underline{z})(z_i \ge m(\underline{z}))$ if and only if $b_i < 0(b_i > 0)$ for all $z_i < m(\underline{z})(z_i \ge m(\underline{z}))$. The condition stipulated in our Theorem 5.1 is clearly different from this. We need unambiguous positivity of the sequence $\{w_i\}$. Note also that while the incomes in $\underline{z}$ are non-decreasingly ordered, in the index (5.1) the terms F_i are non-decreasingly ordered up to the category $(m(F)-1)$, while with effect from the category $m(F)$, the terms $(1-F_i)$ s are non-increasingly ordered. Wang and Tsui (2000) also showed that P_{WT} satisfies the Increased Bipolarity condition, that is, P_{WT}

increases under a transfer of income from a person to someone with a lower income, where the recipient and the donor of the transfer must be on the same side of the median category, if and only if the weights are decreasingly ordered. Clearly, this finding is different from what we showed in Theorem 5.3.

The conditions on the weight sequence $\{w_i\}$ stipulated in Theorems 5.1-5.3 do not make the generalized Gini PIO bounded. Wang and Tsui (2000) did not investigate boundedness property of their index. The following theorem identifies the necessary and sufficient condition for making the PIO bounded.

Theorem 5.4: Given **IS**, the generalized Gini PIO in (5.1) satisfies **BO** if and only if $\sum_{i=1}^{n-1} w_i\left(m(F)\right) \leq 2$. Further, equality holds if the PIO satisfies **PB**.

Proof: See Appendix.

The weights assigned in the index (5.1) are dependent on the category-ranks. However, they are independent of the population size. But for P_{WT}, they depend explicitly on the population size. Since the distribution-function is population replication invariant, given the number of categories, the generalized Gini family considered in (5.1) remains invariant under replication of population. Hence, this index can be employed for making cross-population comparison of ordinal polarization when the numbers of categories across the populations are the same and ranked in the same way. However, exact replication of the population does not leave the value of P_{WT} unchanged and hence, in general, this index is unsuitable for inter-population comparison of polarization. This is another major difference between the two families.

An example of the sequence $\{w_i\left(m(F)\right)\}$ that satisfies the conditions identified in Theorems 5.1-5.3 is

$$w_i\left(m(F)\right) = \begin{cases} \dfrac{2i-1}{n^2}, & 1 \leq i \leq m(F)-1, \\ \dfrac{2(n-i)+1}{n^2}, & i \geq m(F). \end{cases} \tag{5.3}$$

This weighting scheme is in fact the Gini weighting scheme when incomes are ordered respectively non-increasingly and non-decreasingly.

We now relate our index with Apouey's index, which is defined as

$$P_{\alpha}(F) = 1 - \frac{2^{\alpha}}{n-1}\sum_{c=1}^{n-1}\left|F_c - \frac{1}{2}\right|^{\alpha}, \tag{5.4}$$

where $\alpha > 0$ is a parameter that reflects the importance given to the median category. It is however not clear how a unique value of α can be chosen. This index satisfies the axioms **IS, SB, IB, BO** and **WH.** However, Apouey (2007) did not look for any polarization quasi-ordering that becomes consistent with this index. In the next section, we look at a quasi-ordering that shows consistency with our index.

5.5 A Polarization Quasi-Ordering

In this section we would like to develop a quasi-ordering that becomes helpful in ranking two alternative distributions of an ordinally significant dimension using indices from the generalized Gini family satisfying the basic properties such as IS, IB and BO. More precisely, we wish to identify some necessary and sufficient conditions which are equivalent to the following:

Definition 5.3: For two distributions *F* and *G* having the same median category (*m*), we say that $F \leq_{PIO}^{GG} G$ if $P(F) \leq P(G)$ for all generalized Gini PIO of the form (5.1) satisfying **IS, IB** and **BO.**

The following theorem can now be stated.

Theorem 5.5: Fix $n \in N$. For two distributions *F* and *G* having the same median category *m*, the following conditions are equivalent:

(i) $F \leq_{PIO}^{GG} G$.

(ii) $F_{m-1} \leq G_{m-1}$, $F_{m-1} + F_{m-2} \leq G_{m-1} + G_{m-2}$,….., $\sum_{i=1}^{m-1} F_i \leq \sum_{i=1}^{m-1} G_i$; (5.5)

$$F_m \geq G_m,\; F_m + F_{m+1} \geq G_m + G_{m+1}, \ldots, \sum_{i=m}^{n-1} F_i \geq \sum_{i=m}^{n-1} G_i . \tag{5.6}$$

Proof: See Appendix.

Condition (ii) of Theorem 5.5 involves $(n-1)$ inequalities using sums of cumulative population proportions at different categories under the two distributions and they are very easy to check. The ordering identified in Theorem 5.5 is a quasi-ordering – it is transitive, but not complete. If one or more of the inequalities in (ii), on the either side of the median category, is violated, then the two distributions cannot be ordered. Given the ordered categories, the set of generalized Gini indices, where each member of the set corresponds to a particular weight sequence and the weights satisfy the conditions identified in Theorems 5.1-5.3, is uncountable. Our quasi-ordering is consistent with all members of the set. Thus, our quasi-ordering covers a large class of indices. Because of independence of the weights on the population size, with a given number of categories, we can rank two distributions over differing population sizes using this quasi-relation.

5.6 Concluding Remarks

We began by assuming that a polarization index for a dimension of human well-being with ordinal significance can be defined as a weighted sum of absolute deviations of the components of the observed distribution from those of the distribution that generates minimum polarization. We refer to this as the generalized Gini family of polarization indices for an ordinally measurable dimension. It is proven that the weights can take on any possible values consistent with the axioms Increased Spread, Increased Bipolarity and Boundedness. We also develop an axiomatic characterization of the family. A quasi-ordering for ranking two alternative distributions of the ordinal dimension for generalized family of indices is investigated. The partial ordering can be easily implemented by checking some elementary inequalities.

5.7 Appendix

Proof of Proposition 5.1: If possible, let $P : \Delta_n^* \to R$ be one such PIO. Then, by **WH**, $P\left(F^{\underline{\tau}^{(i)}}\right) = 0$, where $F^{\underline{\tau}^{(i)}} = (0, ..., 0, 1, 1, ..., 1)$, the first1 being in the i^{th} slot, $1 \le i \le n$. Now, for any $\underline{\pi} \in \Delta_n$ we have, $P\left(F^{\underline{\pi}}\right) = P\left(\sum_{i=1}^{n} \pi_i F^{\underline{\tau}^{(i)}}\right)$. Since P is convex, by Jensen's inequality (Marshall and Olkin, 1979 p. 454), we have, $P\left(\sum_{i=1}^{n} \pi_i F^{\underline{\tau}^{(i)}}\right) \le \sum_{i=1}^{n} \pi_i P\left(F^{\underline{\tau}^{(i)}}\right)$. Hence, $P\left(F^{\underline{\pi}}\right) =$ $P\left(\sum_{i=1}^{n} \pi_i F^{\underline{\tau}^{(i)}}\right) \le \sum_{i=1}^{n} \pi_i P\left(F^{\underline{\tau}^{(i)}}\right) = 0$. This proves that *P* is necessarily degenerate.□

Proof of Proposition 5.2: (a) (i) Consider $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_n) \in \Delta_n$. Let $\underline{\pi'} = \left(\pi_1', \pi_2', ..., \pi_n'\right)$ be obtained from $\underline{\pi}$ by a shift of mass $\varepsilon > 0$, sufficiently small, from category i to category j, where $j < i < m$, m being the median category. So, $\pi_l' = \pi_l$ for all $l \ne i, j$ and $\pi_i' = \pi_i - \varepsilon$, $\pi_j' = \pi_j + \varepsilon$. Then $F_l' = F_l + \varepsilon$ for all $j \le l < i$ and $F_l' = F_l$ for all other values of *l*. So, $\underline{\pi'}$ has a greater spread than $\underline{\pi}$. Consequently, by **IS**, $\underline{\pi'}$ is more polarized than $\underline{\pi}$.

To prove the next part, consider a shift of the same mass $\varepsilon > 0$, sufficiently small, from category i' to category j, where $j < i < i' < m$. Let $\underline{\pi''} = \left(\pi_1'', \pi_2'', ..., \pi_n''\right)$ be the new distribution. Then $\pi_l'' = \pi_l$ for all $l \ne i', j$ and $\pi_{i'}'' = \pi_{i'} - \varepsilon$, $\pi_j'' = \pi_j + \varepsilon$. Then $F_l'' = F_l + \varepsilon$ for all $j \le l < i'$ and $F_l' = F_l$ for all other values of *l*. So, $\underline{\pi''}$ has a greater spread than $\underline{\pi'}$. As a result, by **IS**, $\underline{\pi''}$ is more polarized than $\underline{\pi'}$. Hence, the result follows. Proof of part (b) of the proposition is similar to that of part (a). □

Proof of Proposition 5.3: Let $\underline{\pi} = (\pi_1, \pi_2, ..., \pi_n) \in \Delta_n$. A transfer of mass $\varepsilon > 0$ below the median category generates $\underline{\pi}_{i+1,\hat{i}-1} = \left(\pi_1, ..., \pi_{i-1}, \pi_i - \varepsilon, \pi_{i+1} + \varepsilon, \pi_{i+2}, ..., \pi_{\hat{i}-2}, \pi_{\hat{i}-1} + \varepsilon, \pi_{\hat{i}} - \varepsilon, \pi_{\hat{i}+1}, ..., \pi_n\right)$ from $\underline{\pi}$. By **IB**,

$P\left(F^{\underline{\pi}_{i+1,\hat{i}-1}}\right) > P\left(F^{\underline{\pi}}\right)$. Also, let $\underline{\pi}_{i,\hat{i}} = \left(\pi_1, \ldots, \pi_{i-1}, \pi_i + \varepsilon, \pi_{i+1}, \ldots, \pi_{\hat{i}-1}, \pi_{\hat{i}} - \varepsilon, \pi_{\hat{i}+1}, \ldots\ldots, \pi_n\right)$. Then we have to show that $P\left(F^{\underline{\pi}_{i+1,\hat{i}-1}}\right) - P\left(F^{\underline{\pi}}\right) < P\left(F^{\underline{\pi}_{i,\hat{i}}}\right) - P\left(F^{\underline{\pi}}\right)$, that is, $P\left(F^{\underline{\pi}_{i+1,\hat{i}-1}}\right) < P\left(F^{\underline{\pi}_{i,\hat{i}}}\right)$. Clearly, $\underline{\pi}_{i,\hat{i}}$ is obtained from $\underline{\pi}_{i+1,\hat{i}-1}$ by two shifts, each of mass ε: first from category $\hat{i}-1$ to category i and then from category $(i+1)$ to category i. So, by part (a) (i) of Proposition 5.2, $P\left(F^{\underline{\pi}_{i+1,\hat{i}-1}}\right) < P\left(F^{\underline{\pi}_{i,\hat{i}}}\right)$. This proves the proposition. □

Proof of Theorem 5.1: Suppose *P* satisfies **WH, IS** and **PI**.

For any distribution $F^{observed} = (F_1, F_2, \ldots, F_{n-1}, 1)$ with median category $m(F)$ we have, $0 \le F_i < \frac{1}{2}$ for all $i < m(F)$ and $\frac{1}{2} \le F_i \le 1$ for all $m(F) \le i \le (n-1)$.

Since both the distributions $\left(0, \ldots, 0, F_{m(F)-1}, 1, \ldots, 1\right)$ (where $F_{m(F)-1}$ comes at the $(m(F)-1)^{th}$ position) and $F^{\min} = (0, \ldots., 0, 1, \ldots., 1)$ have category $m(F)$ as their median category, by IS we have,

$$P\left(0, \ldots, 0, F_{m(F)-1}, 1, \cdots, 1\right) \ge P\left(0, \ldots, 0, 1, 1, \cdots, 1\right) \tag{5.7}$$

with equality only if $F_{m(F)-1} = 0$.

Then **PI** implies that

$$P\left(0, \ldots, 0, F_{m(F)-1}, 1, \cdots, 1\right) - P\left(0, \ldots, 0, 1, 1, \cdots, 1\right) = c\left(m(F)-1, m(F)\right) F_{m(F)-1}, \tag{5.8}$$

where $c\left(m(F)-1, m(F)\right) > 0$ is a constant (see the discussion in the paragraph preceding to Theorem 5.1).

Hence, by **WH** it follows that

$$P\left(0, \ldots, 0, F_{m(F)-1}, 1, \cdots, 1\right) = c\left(m(F)-1, m(F)\right) F_{m(F)-1} \tag{5.9}$$

for some $c\left(m(F)-1, m(F)\right) > 0$.

Next we observe that both the distributions $\left(0,\ldots,0,F_{m(F)-2},F_{m(F)-1},1,\ldots,1\right)$ and $\left(0,\ldots,0,F_{m(F)-1},1,\ldots,1\right)$ have category $m(F)$ as their median category so that by another application of **IS** we have,

$$P\left(0,\ldots,0,F_{m(F)-2},F_{m(F)-1},1,\ldots,1\right)\geq P\left(0,\ldots,0,F_{m(F)-1},1,\ldots,1\right), \tag{5.10}$$

with equality only if $F_{m(F)-2}=0$.

Therefore, by **PI**,

$$P\left(0,\ldots,0,F_{m(F)-2},F_{m(F)-1},1,\ldots,1\right)-P\left(0,\ldots,0,F_{m(F)-1},1,\ldots,1\right)=c\left(m(F)-2,m(F)\right)F_{m(F)-2}, \tag{5.11}$$

for some constant $c\left(m(F)-2,m(F)\right)>0$.

Combining (5.11) with (5.9) we get

$$P\left(0,\ldots,0,F_{m(F)-2},F_{m(F)-1},1,\cdots,1\right)=\sum_{j=1}^{2}c\left(m(F)-j,m(F)\right)F_{m(F)-j}. \tag{5.12}$$

By induction it follows that

$$P\left(F_1,\ldots.,F_{m(F)-1},1,\ldots,1,1\right)=\sum_{j=1}^{m(F)-1}c\left(m(F)-j,m(F)\right)F_{m(F)-j}=\sum_{i=1}^{m(F)-1}c\left(i,m(F)\right)F_i \tag{5.13}$$

for some $c\left(i,m(F)\right)>0, 1\leq i\leq m(F)-1$.

Now under a shift of mass $\left(1-F_{m(F)}\right)$ from the $m(F)^{th}$ category of the distribution $\left(F_1,\ldots,F_{m(F)-1},1,\ldots,1,1\right)$ to the next category we arrive at the distribution for which the cumulative proportion at the $m(F)^{th}$ category is $F_{m(F)}$. Then the resulting distribution is $\left(F_1,\ldots.,F_{m(F)-1},F_{m(F)},1,\ldots,1\right)$, which has the same median category as the previous one. By **IS**, we have an increment of polarization and consequently, PI implies that

$$P\left(F_1,\ldots.,F_{m(F)-1},F_{m(F)},1,..,1\right)-P\left(F_1,\ldots.,F_{m(F)-1},1,\ldots,1\right)=c\left(m(F),m(F)\right)\left(1-F_{m(F)}\right) \quad (5.14)$$

for some $c\left(m(F),m(F)\right)>0$. Hence, by (5.13) and (5.14) we have,

$$P\left(F_1,\ldots.,F_{m(F)-1},F_{m(F)},1,\ldots,1\right)=\sum_{i=1}^{m(F)-1}c\left(i,m(F)\right)F_i+c\left(m(F),m(F)\right)\left(1-F_{m(F)}\right) \quad (5.15)$$

(see again the discussion in the paragraph preceding to Theorem 5.1).

Repeating the process, that is, by effecting a shift of mass $\left(1-F_{m(F)+1}\right)$ from the $\left(m(F)+1\right)^{th}$ category to the $\left(m(F)+2\right)^{th}$ category we get

$$P\left(F_1,\ldots.,F_{m(F)},F_{m(F)+1},1,\ldots,1\right)=\sum_{i=1}^{m(F)-1}c\left(i,m(F)\right)F_i+\sum_{j=m(F)}^{m(F)+1}c\left(j,m(F)\right)\left(1-F_j\right) \quad (5.16)$$

for some $c\left(m(F),m(F)\right)$, $c\left(m(F)+1,m(F)\right)>0$.

Continuing the process analogously $\left(n-m(F)-1\right)$ times we deduce,

$$P\left(F_1,\ldots.,F_{m(F)-1},F_{m(F)},\ldots.,F_{n-1},1\right)=\sum_{i=1}^{m(F)-1}c\left(i,m(F)\right)F_i+\sum_{i=m(F)}^{n-1}c\left(i,m(F)\right)\left(1-F_i\right) \quad (5.17)$$

for some $c\left(i,m(F)\right)>0, 1\le i\le n-1$. Substituting $c\left(i,m(F)\right)$ by $w_i\left(m(F)\right)$ in the above equation we finally get,

$$P\left(F_1,F_2,\ldots,F_{m(F)-1},F_{m(F)},\ldots,F_{n-1},1\right)=\sum_{i=1}^{m(F)-1}w_i\left(m(F)\right)F_i+\sum_{i=m(F)}^{n-1}w_i\left(m(F)\right)\left(1-F_i\right)$$

where $w_i\left(m(F)\right)>0$ for $1\le i\le n-1$. This proves necessity part of the theorem. Converse is easy to check.□

Proof of Corollary 5.3: The proof is immediate, given Theorem 5.1. For, $P(F_1, F_2, \ldots\ldots, F_n) = 0$ if and only if $F_i = 0$ for all $i < m(F)$ and $F_i = 1$ for all $m(F) \le i \le (n-1)$. This is the case if and only if the underlying distribution is $F^{\min} = (0, \ldots, 0, 1, \ldots, 1)$. □

Proof of Theorem 5.3: Suppose P satisfies **IB**. Consider shifts of a sufficiently small mass $\varepsilon > 0$ from the i^{th} category to the $(i+1)^{th}$ and from the $(i+2)^{th}$ to the $(i+1)^{th}$, where $(i+2) < m(F)$ and look at $F' = (F_1', F_2', \ldots, F_n')$ with $F_j' = F_j$ for all $j \ne i, i+1$ and $F_i' = F_i - \varepsilon$, $F_{i+1}' = F_{i+1} + \varepsilon$. Then it is easy to see that $P(F')$ $= P(F) + \{w_{i+1}(m(F)) - w_i(m(F))\}\varepsilon$.

By **IB** it is demanded that $P(F') > P(F)$, which implies that $w_{i+1}(m(F)) > w_i(m(F))$. This is true for all $i < (m(F) - 2)$.

To check validity of the inequality when $i = (m(F) - 2)$, we have to consider a transfer of a sufficiently small mass $\varepsilon > 0$ from the $(m(F)-2)^{th}$ category to the $(m(F)-1)^{th}$ and from the $m(F)^{th}$ (below the median individual, an individual in the median category) to the $(m(F)-1)^{th}$. The new distribution F' has $F_j' = F_j$ for all $j \ne (m(F)-2), (m(F)-1)$ and $F'_{m(F)-2} = F_{m(F)-2} - \varepsilon$, $F'_{m(F)-1} = F_{m(F)-1} + \varepsilon$. The rest of the proof is quite similar to the non-extremal case. Hence, $w_{i+1}(m(F)) > w_i(m(F))$ for all $i \le (m(F) - 2)$. Similarly, considering transfer on the other side of the median category we conclude that $w_{i+1}(m(F)) < w_i(m(F))$.

To check the converse, consider shifts of a sufficiently small mass $\varepsilon > 0$ from the i^{th} category to the $(i+1)^{th}$ and from the $(j+1)^{th}$ to the j^{th}, where $j \ge (i+1)$ and $(j+1) < m(F)$. For $F' = (F_1', F_2', \ldots, F_n')$ we have, $F_l' = F_l$ for all $l \ne i, j$ and $F_i' = F_i - \varepsilon$, $F_j' = F_j + \varepsilon$. It then clearly follows that

$$P(F') = w_i(m(F))(F_i - \varepsilon) + w_j(m(F))(F_j + \varepsilon) + \sum_{\substack{1 \le l \le m(F)-1 \\ l \ne i,j}} w_l(m(F))F_l + \sum_{l=m(F)}^{n-1} w_l(m(F))(1 - F_l)$$

$= P(F) + \{-w_i(m(F)) + w_j(m(F))\}\varepsilon > P(F)$, since $w_i(m(F)) < w_j(m(F))$. Thus, **IB** holds for a transfer below the median category. A similar calculation shows the same for a transfer above the median category. This completes the proof.□

Proof of Theorem 5.4: By Theorem 5.1, satisfaction of **IS** forces: $w_i(m(F)) > 0$ for all $i < n$. So, by definition, $P(F) = \left\{\sum_{i=1}^{m(F)-1} w_i(m(F))F_i + \sum_{i=m(F)}^{n-1} w_i(m(F))|1 - F_i|\right\} \ge 0$ for all *F*. Suppose $P(F) = \left\{\sum_{i=1}^{m(F)-1} w_i(m(F))F_i + \sum_{i=m(F)}^{n-1} w_i(m(F))|1 - F_i|\right\} \le 1$ for all *F*. Then considering the distribution *F* for which $F_i = \left(\frac{1}{2} - \varepsilon\right)$ for all $i < m(F)$ (where $\varepsilon > 0$ is sufficiently small) and $F_i = \frac{1}{2}$ for all $m(F) \le i \le (n-1)$ we have, $\left\{\left(\frac{1}{2} - \varepsilon\right)\sum_{i=1}^{m(F)-1} w_i(m(F)) + \frac{1}{2}\sum_{i=m(F)}^{n-1} w_i(m(F))\right\} \le 1$.

Now, letting $\varepsilon \to 0$ it follows that

$$\sum_{i=1}^{n-1} w_i(m(F)) \le 2. \tag{5.18}$$

Conversely, if (5.18) holds, then we claim that $P(F) \le 1$ for all *F*. To prove this, note that $F_i < \frac{1}{2}$, for $i \le (m(F) - 1)$ and $F_i \ge \frac{1}{2}$ for $i \ge m(F)$ so that $(1 - F_i) \le \frac{1}{2}$ for all $i \ge m(F)$. Consequently, $P(F) = \left\{\sum_{i=1}^{m(F)-1} w_i(m(F))F_i + \sum_{i=m(F)}^{n-1} w_i(m(F))|1 - F_i|\right\} \le \frac{1}{2}\sum_{i=1}^{n-1} w_i(m(F)) \le 1$, by (5.18). The remaining part of the theorem is easy to check. □

Proof of Theorem 5.5: $(i) \Rightarrow (ii)$: Suppose (i) holds. Then for all sequences $\{w_i\}$ such that $w_i > 0$ for all *i*, $\sum_{i=1}^{n-1} w_i \leq 2$ and $w_i > w_{i-1}$ for all $i < m$ and $w_{i+1} < w_i$ for all $i \geq m$ we require: $\sum_{i=1}^{m-1} w_i F_i + \sum_{i=m}^{n-1} w_i (1 - F_i) \leq \sum_{i=1}^{n-1} w_i G_i + \sum_{i=m}^{n-1} w_i (1 - G_i)$, that is, $\sum_{i=1}^{m-1} w_i F_i - \sum_{i=m}^{n-1} w_i F_i \leq \sum_{i=1}^{n-1} w_i G_i - \sum_{i=m}^{n-1} w_i G_i$ for all $\{w_i\}$ satisfying the above set of conditions. Take $w_{m-1} = 2 - \varepsilon$, $w_i > 0$ for $i \neq m-1$ small enough so that $\sum_{i \neq m-1} w_i \leq \varepsilon$. This gives

$(2 - \varepsilon)(F_{m-1} - G_{m-1}) \leq \sum_{i \neq m-1} w_i (F_i - G_i) \leq \sum_{i \neq m-1} w_i \left|F_i - G_i\right|$. Since $F_i \leq 1$ and $G_i \leq 1$, it follows that $\left|F_i - G_i\right| \leq (1 + 1) = 2$ so that $\sum_{i \neq m-1} w_i \left|F_i - G_i\right| \leq 2 \sum_{i \neq m-1} w_i \leq 2\varepsilon$.

Now, letting $\varepsilon \to 0$ we get: $F_{m-1} \leq G_{m-1}$.

To establish the next condition, simply take $w_{m-1} = 1 - \varepsilon/2$, $w_{m-2} = 1 - \varepsilon$, $\sum_{i \neq m-1, m-2} w_i \leq \frac{3}{2}\varepsilon$. Apply the same logic and get the result. Thus, gradually we get all the conditions for movement below the median category. Proof of the other set of conditions is analogous.

$(ii) \Rightarrow (i)$: Assume (ii). Then for any generalized Gini PIO of the form (5.1) we have,

$$\sum_{i=1}^{m-1} w_i F_i - \sum_{i=m}^{n-1} w_i F_i = \sum_{i=1}^{m-1} w_1 F_i + \sum_{i=2}^{m-1} (w_2 - w_1) F_i + \sum_{i=3}^{m-1} (w_3 - w_2) F_i + \ldots.. + (w_{m-1} - w_{m-2}) F_{m-1} -$$

$$\sum_{i=m}^{n-1} w_{n-1} F_i - \sum_{i=m}^{n-2} (w_{n-2} - w_{n-1}) F_i - \sum_{i=m}^{n-3} (w_{n-3} - w_{n-2}) F_i - \ldots.. - (w_m - w_{m+1}) F_m .$$

Now using the given set of conditions on w_i 's and (5.5)-(5.6), we get, $\sum_{i=1}^{m-1} w_i F_i - \sum_{i=m}^{n-1} w_i F_i \leq \sum_{i=1}^{m-1} w_i G_i - \sum_{i=m}^{n-1} w_i G_i$ which in turn implies $P(F) \leq P(G)$. Since the generalized Gini PIO *P* is arbitrary, $F \leq_{PIO}^{GG} G$. □

Chapter 6

CONTEST SUCCESS FUNCTIONS: SOME NEW PROPOSALS[21]

6.1 Introduction

Closely related to the notion of polarization is the theory of contests. The connection can be traced back to Esteban and Ray (1999), wherein putting forth a behavioral model of conflict, conflict has been presented as a contest game. In this concluding chapter, we try to make a rigorous study on the structural properties of Contest Success Function (CSF), which specifies a contestant's probability of winning the contest and obtaining a prize.

As we mentioned in Chapter 1, Skaperdas (1996) characterized this probability for any contestant as the ratio between the level of effective investment made by the contestant and the sum of effective investments across all the contestants. Using this basic structure, Skaperdas (1996) also developed axiomatic characterizations of the Tullock (1980)-Hirschleifer (1989) functional forms of CSFs.

The basic structure of Skaperdas (1996) points out how to derive general CSFs that satisfy five basic axioms, namely, Efficiency, Monotonicity, Anonymity, Consistency and Independence of Outsiders' Efforts (see Section 6.2). However, without invoking any further condition characterizations of the general consistent class of contest success functions will not yield any specific form of contest success function.Skaperdas (1996) invokes two alternative axioms of invariance. The first axiom, the scale invariance postulate, demands that an equi-proportionate change in the efforts of all the agents will keep the winning probabilities unchanged. In contrast, the second axiom, which is known as the translation invariance postulate, requires invariance of winning probabilities under equal absolute changes in the efforts of all the agents. The underlying effective investment functions turn out to be of power function and logit function type respectively.

[21] A major portion of this chapter is available at Chakravarty and Maharaj (2014).

A natural generalization of scale and translation invariance axioms is an intermediate condition,which stipulates that a convex mixture of an equi-proportionate change and an equal absolute change in the efforts should keep winning probabilities unchanged. (See Section 6.2 for more discussion.) One objective of this chapter is to characterize the entire class of CSFs that satisfies this generalized invariance concept. It is explicitly shown that the Tullock and Hirschleifer functional forms characterized by Skaperdas (1996) become particular cases of the CSF that fulfils intermediate equivalence.

We then analyze the likelihood of occurrence of Nash equilibrium for the CSF derived using this generalized invariance concept. It is known that the Tullock CSF has Nash equilibrium in pure strategies and the Hirschleifer CSF has no Nash equilibrium in pure strategies. We demonstrate that the CSF satisfying the generalized invariance axiom has a unique Nash equilibrium in pure strategies and this equilibrium can as well be a corner solution in a pure intermediate situation, which coincides neither with the relative nor with the absolute invariance case. It may be noted that the existence of a Nash equilibrium as a corner solution is not possible for the Tullock CSF.

Given two contests CI and CII, investors may be interested in ranking them in terms of their probabilities of winning. This is a general ordinal postulate. However, in order to pin down some specific functional forms of CSFs, one needs to impose some value judgement postulate. In fact, in the last few years attempts have been made to provide foundations of commonly used CSFs.[22] One such postulate that ensures ordinal property of CSFs is the scale consistency axiom, which says that if all the agents are participating in two contests and for some agents the probabilities of winning one contest are less than or equal to that of winning the other, then an equi-proportionate change in the efforts of the agents in both contests will not alter the agents' ordering of chances of winning the contests. To understand this, suppose the investments are measured in money units, say euro. Then suppose some individuals' chances of winning CI are more than that of CII. Now, if investments are converted into dollars from euro, the inequality between chances of winning CI and CII should not alter. Scale consistency demands this

[22] See the survey papers referred to in footnote 1.Some authors have also attempted to develop econometric estimation of several CSFs. (See Jia and Skaperdas, 2012 and Jia, Skaperdas and Vaidya, 2013 for detailed discussions.)

condition. Note that since the sum of probabilities of winning a contest across the agents is one if for some agents the probabilities of winning one contest over another are lower, then there will be at least one agent for whom the reverse inequality for probabilities of winning the contests will hold. CSFs satisfying scale invariance are definitely scale consistent.

Likewise, we can have a translation consistency axiom, which specifies that inequality between winning probabilities for two contestants should remain invariant under equal absolute changes in all the efforts. Translation consistency implies translation invariance. However, as we will demonstrate, if the number of contestants is only 2, there can be CSFs that satisfy scale (translation) consistency but not scale (translation) invariance.

A second objective of the chapter is to axiomatize the classes of CSFs that are scale and translation consistent respectively. It is fairly interesting to observe that if the number of contestants is greater than 2, the Tullock and the Hirschleifer CSFs turn out to be the only CSFs that verify scale and translation consistency axioms respectively. Thus, both the Tullock and the Hirschleifer CSFs can be supported by ordinal axioms. This is another attractive feature of our chapterer.

Finally, we define an intermediate consistency condition, which may be viewed as the ordinal counterpart of intermediate invariance. Alternatively, it can be seen as a convex mixture of translation consistency and scale consistency. We demonstrate that if the number of contestants is greater than 2, the only class of intermediate consistent CSFs is necessarily intermediate invariant.

6.2 The Formal Framework

Let $N = \{1, 2, \ldots, n\}$ be a set of agents participating in a contest and let y_i stand for effort or investment of agent $i \in N$ in the contest. We denote the vector of investments $(y_1, y_2, .., y_n)$ $\in [0, \infty)^n$ by y , where $[0, \infty)^n$ is the $n-$ fold Cartesian product of $[0, \infty)$ .The success of any contestant is probabilistic. For any $y \in [0, \infty)^n$, each contestant i 's probability of winning the contest is denoted by $p^i(y)$. Evidently, $p^i : [0, \infty)^n \rightarrow [0,1]$. The non-negative function p is called the Contest Success Function (CSF).

The following axioms for a CSF have been suggested by Skaperdas (1996).

(A1) (Efficiency) $\sum_{i=1}^{n} p^i(y) = 1$ and for all $y \in [0,\infty)^n$, if $y_i > 0$ then $p^i(y) > 0$.

(A2 (Monotonicity) $p^i(y)$ is strictly increasing in y_i and strictly decreasing in y_j for all $j \neq i$.

(A3) (Anonymity) For any permutation $\pi : N \to N$, $p^{\pi(i)}(y) = (y_{\pi_1}, y_{\pi_2}, ..., y_{\pi_N})$.

(A4) (Consistency) For all $M \subseteq N$ with at least two elements, the probability of success of agent $i \in M$ in a contest among the members of M is $p_m^i(y) = \dfrac{p^i(y)}{\sum_{j \in M} p^j(y)}$; provided that there is at least one $j \in M$ such that $p^j(y) > 0$.

(A5) (Independence of Outsiders' Efforts) $p_m^i(y)$ is independent of the efforts of the players not included in the subset $M \subseteq N$ or $p_m^i(y)$ can be written as $p_m^i(y_m)$, where $y_m = (y_j : j \in M)$.

(A $5'$) $p^i(y) = \dfrac{f(y_i)}{\sum_{j \in N} f(y_j)}$ for all $i \in N$ and $p_m^i(y) = \dfrac{f(y_i)}{\sum_{j \in M} f(y_j)}$ for all $i \in M$ $(\subseteq N)$, provided that there is $j \in M$ with $y_j > 0$, where $f : [0,\infty) \to [0,\infty)$ is strictly increasing in its argument.

(A1) states that the sum of winning probabilities across the participants in a contest is 1 and if some participant's outlay is positive he has a positive chance of winning the contest. (A2) says that a participant's probability of success is increasing in his own effort but decreasing in the efforts of the other participants. According to (A3), the probability of success remains invariant under any reordering of the participants. This anonymity condition demands that any characteristic other than individual outlays is irrelevant to the determination of success probabilities. The consistency condition (A4) says that for any subgroup of participants, the probabilities of success of the members of the subgroup are the conditional probabilities obtained by restricting the original probability distribution to the subgroup. For (A4) to be well-defined, it is implicitly assumed, under (A1), that $y \neq 01^n$, where 1^n is the $n-$ coordinated vector of ones**.** Otherwise the denominator on the right hand side of $p_m^i(y)$ may vanish. (A5) means that for any subgroup of participants, the success probabilities are independent of the outlays of the participants who are not members of the subgroup. Finally, (A $5'$) provides a particular specification of the winning probabilities using a positive valued strictly increasing function of efforts. We can refer to $f(y_i)$ as the effective investment made by contestant i. Strict increasingness of f reflects the view that an increase in the actual investment strictly increases effective investment. Skaperdas (1996) demonstrated that (A1)-(A5) hold simultaneously if and only if the CSF is of the form specified in (A$5'$). Since our characterizations employ the basic axioms (A1)-(A5), we will deal with the general form given by (A$5'$).

Note that the expression of $p_m^i(y)$ given by (A$5'$) is undefined at all those points where

$$\sum_{j\in M} f\left(y_j\right) = 0 . \tag{6.1}$$

By strict increasingnessof f it follows that $f\left(y_j\right) > f(0) \geq 0$ whenever $y_j \in (0,\infty)$. Thus, (6.1) is an impossibility if there is $j \in M$ such that $y_j > 0$. Moreover, given the structure of the function f, $p^i(y)$ will be defined and continuous everywhere on $[0,\infty)^n \setminus \{(0,\dots,0)\}$. So, if $f(0) = 0$, then the domain of the CSF defined in Skaperdas (1996) excludes the origin.

To get rid of the problem of definition at the origin, Corchon (2007) suggests the use of the following functional form of the CSF:

$$p^i(y) = \begin{cases} \dfrac{f(y_i)}{\sum_{j\in N} f\left(y_j\right)}, if \sum_{j\in N} f\left(y_j\right) > 0 \\ \dfrac{1}{n},\ otherwise. \end{cases} \tag{6.2}$$

The function $f:[0,\infty) \to [0,\infty)$ is assumed to obey the following properties:

i) f is twice continuously differentiable in $(0,\infty)$.

ii) f is concave.

iii) f is strictly increasing.

iv) $f(0) = 0$ and $\lim_{z\to\infty} f(z) = \infty$.

v) $\dfrac{zf'(z)}{f(z)}$ is bounded for all $z \in (0,\infty)$.

The functional form (6.2) along with properties (i) – (v) will be required in the sequel in our quest for the existence of a Nash equilibrium in a particular situation.

As stated in the introduction, some additional axiom(s) have to be invoked in order to identify specific functional forms of CSFs. Skaperdas (1996) imposed the following axioms:

(A6) (Scale invariance) $p^i(y) = p^i(\lambda y)$ for all $\lambda > 0$ and for all $i \in N$.

(A7) (Translation invariance) $p^i(y) = p^i(y + c1^n)$, where 1^n is the n – coordinated vector of ones and c is a scalar such that $y_i + c \geq 0$ for all $i \in N$.

The scale invariance axiom (A6) is a homogeneity condition, which says that proportional changes in the efforts of all the contestants do not change the winning probabilities. In contrast, (A7) is a translation invariance axiom, which demands that winning probabilities remain unchanged when all the efforts are augmented or diminished by the same absolute quantity.

It has been shown in Skaperdas (1996) that a CSF defined (and continuous) on $[0,\infty)^n \setminus \{(0,\ldots,0)\}$ satisfies (A1) – (A6) if and only if it is of the power function type, that is, of the form $p^i(y) = \dfrac{y_i^{\delta}}{\sum_{j\in N} y_j^{\delta}}$, where $\delta > 0$ is a constant. This is the Tullock(1980) form of CSF. It has a Nash equilibrium in pure strategies for $\delta \in (0,1]$. The particular case $\delta = 1$ was considered by Esteban and Ray (2011) in a behavioural model of conflict that provides a link between conflict, inequality and polarization (see also Chakravarty, 2015). On the other hand, as Skaperdas (1996) established, the logit function, that is, $p^i(y) = \dfrac{e^{\theta y_i}}{\sum_{j\in N} e^{\theta y_j}}$ is the only continuous CSF that satisfies (A1) – (A5) and (A7), where $\theta > 0$ is a constant. This Hirschleifer (1989) CSF has no Nash equilibrium in pure strategies. (A systematic comparison of the properties of these two functional forms is available in Hirschleifer (1989).) It is easy to verify that the only CSF that satisfies (A6) and (A7) is the constant function $p^i(y) = \dfrac{1}{n}$. But constancy of a CSF is ruled out by the assumption that $p^i(y)$ is strictly increasing in y_i and is strictly decreasing in y_j for all $j \neq i$.

However, adoption of either (A6) or (A7) reflects a particular notion of value judgment. Investors may not be unanimous in their choice between these two invariance notions. If we replace p^i by an inequality index and y by the income distribution in an n – person society, then these two invariance concepts are referred to as rightist and leftist notions of inequality invariance (Kolm, 1976). In fact, experimental questionnaire studies provide ample evidence for a middle position between these two views (Amiel and Cowell, 1992).

In the current context, the following represents a diversity of views concerning invariance of CSFs:

(A8) $\quad p^i\left(y + c(\mu y + (1-\mu)1^n)\right) = p^i(y)$,

where μ , $0 \le \mu \le 1$, is a parameter which reflects a contestant's view on winning probability equivalence, c is a scalar such that $y + c\left(\mu y + (1-\mu)1^n\right) \in [0,\infty)^n$ and 1^n , the *n*-coordinated vector of ones, is expressed in the unit of measurement of efforts, so that $y' = y + c(\mu y + (1-\mu)1^n$ becomes well defined. The scale and translation invariance criteria given by (A6) and (A7) emerge as polar cases of the intermediate notion (A8) when μ takes on the values 1 and 0 respectively. As the value of μ increases (decreases) to one (zero) the contestant becomes more concerned about scale (translation) invariance[23].

The following theorem isolates the CSF that satisfies (A8). We first identify the CSF for the parametric range $0 < \mu < 1$. The two extreme cases will be discussed later. We make the following assumption at the outset.

Assumption (A): In (A$5'$), we assume that $f(0) > 0$ and f is continuously differentiable on $[0,\infty)$ with $f'(0) > 0$.

Theorem 6.1: Assume that the CSF meets A. Then it satisfies axioms (A$5'$) and (A8) if and only if it is of the following form

[23] In the context of income inequality measurement this axiom is the Bossert-Pfingsten (1990) intermediate inequality equivalence axiom. See also Chakravarty (2015) for a recent discussion.

$$p^i(y) = \frac{[1+\mu(y_i - 1)]^{\frac{\eta}{\mu}}}{\sum_{j\in N}[1+\mu(y_j - 1)]^{\frac{\eta}{\mu}}} \quad , \tag{6.3}$$

where $\eta > 0$ is a constant and $0 < \mu < 1$.

Proof: See Appendix.

As $\mu \to 0$, $p^i(y)$ in (6.3) approaches $\frac{e^{\eta y_i}}{\sum_{j\in N} e^{\eta y_j}}$, the Hirshleifer CSF associated with (A7) (given that $\theta = \eta$). (Here for evaluating the limit we use the fact that $\lim_{z\to 0+}(1+z)^{\frac{1}{z}} = e$.) On other hand, for $\mu = 1$, $p^i(y)$ given by (6.3) coincides with the Tullock (1980) CSF corresponding to (A6) (given that $\eta = \delta$). Thus, p^i in (6.3) may be regarded as a generalization of scale and translation invariant CSFs.

It will now be worthwhile to investigate whether this CSF supports a Nash equilibrium in efforts. Let $V_i = V_i(y_1, y_2, ..., y_n)$ be the value of the prize obtained by the i^{th} contestant and $C_i(y_i)$ be the cost attributed by i to his action y_i. Following Corchon (2007), we make the following assumptions:

a) All agents have the same cost function *C*, that is, $C_i = C$ for all *i*.

b) The common functional form of V_i is the following:

   $V = V_0 + a\sum_{i=1}^{n} f(y_i)$, where $V_0 > 0,\ a \geq 0$.

c) $af'(0) - C'(0) > 0$ and there exists $(\bar{z}, \tau)$ such that for all $z > \bar{z}$ we have, $af'(z) - C'(z) < \tau < 0$.

It may be noted that there are no well-founded criteria to guide the choice of a cost function here[24]. The quantity V_0 in (b) may be regarded as fixed cost. We, therefore, develop the analysis using the above common cost function.

[24]This discussion does not apply to characterizations of CSFs where groups are contestants, since there is no relation between individual effort and group performance (see Münster, 2009).

Following Proposition 3.1 of Corchon (2007), we maintain that there is a Nash equilibrium if and only if the equation

$$f'(z)\left(a+V_0\frac{n-1}{f(z)n^2}\right)-C'(z)=0 \tag{6.4}$$

has a solution.

In our case, by (6.19) (in appendix) we have,

$$f(0)=\xi(1-\mu)^{\frac{\eta}{\mu}}>0\,. \tag{6.5}$$

Differentiating (6.19) twice with respect to z we get,

$$f'(z)=\xi\eta\left\{\mu(z-1)+1\right\}^{\frac{\eta}{\mu}-1} \tag{6.6}$$

and

$$f''(z)=\xi\eta(\eta-\mu)\left\{\mu(z-1)+1\right\}^{\frac{\eta}{\mu}-2}\,. \tag{6.7}$$

From (6.6) it is immediate that $f'(z)>0$ for all $z\in(0,\infty)$. Also, for $\eta\le\mu$ we observe that $f''(z)<0$ for all $z\in(0,\infty)$ so that f is concave. Finally, it is clear that

$$\lim_{z\to\infty}f(z)=+\infty\,. \tag{6.8}$$

Denoting LHS of (6.4) by $\psi(z)$ we see that

$$\psi(0)=f'(0)\left(a+V_0\frac{n-1}{f(0)n^2}\right)-C'(0)>0\,. \tag{6.9}$$

In view of (6.7) and the assumption that $af'(z)-C'(z)<\tau<0$ for sufficiently large z, $\psi(z)$ becomes –ve as $z\to+\infty$ .So, there exists a solution to (6.4), which in turn implies that there is a Nash equilibrium, if $\eta\le\mu$. Differentiation of the left hand side of (6.4) shows that ψ is strictly decreasing so that the solution to (6.4), that is, the Nash equilibrium is unique. We summarize these findings as follows:

Proposition 6.1: Under the assumptions (a) – (c) stated above, the contest game with CSF given by equation (6.3) has a unique Nash equilibrium.

Note that in (6.3) for $y_i=0$ we have, $p^i(y)>0$. Proposition 6.1 shows that in such a situation, there is a possibility of existence of a corner solution. Clearly, this is not the case with the power function $p^i(y) = \frac{y_i^{\delta}}{\sum_{j\in N} y_j^{\delta}}$, since in this case we have, $p^i(y) = 0$ for $y_i = 0$. It is worth mentioning here that all the properties of the function f specified in Corchon (2007) are satisfied whenever $\delta \in (0,1)$. By Proposition 3.1 of Corchon (2007) , this clearly establishes that a unique Nash equilibrium exists. However, for the logit function $p^i(y) = \frac{e^{\theta y_i}}{\sum_{j\in N} e^{\theta y_j}}$ we have, $f''(z) = \theta^2 e^{\theta z} > 0$ for all $z \in (0,\infty)$. Thus, f is never concave and hence from the proof of Proposition 3.1 of Corchon (2007) it follows that there is no Nash equilibrium.

In view of the above discussion we can now state the following:

Remark 6.1: In a pure intermediate situation, that is, when $0 < \eta \le \mu < 1$, Nash equilibrium may emerge as a corner solution.

Note that the scale invariance condition (A6) can very well be relaxed to the following more general ordinal property.

(A9) Scale Consistency: For $x, y \in [0,\infty)^n$, if for some $i \in N$, $p^i(y) \ge p^i(x)$ holds, then $p^i(\lambda y) \ge p^i(\lambda x)$ for all $\lambda > 0$.

Evidently, scale invariance implies scale consistency but the converse is not true. For example, consider the CSF $p^i(y) = \frac{2^{y_i}}{2^{y_1} + 2^{y_2}}$ for $y = (y_1, y_2) \in [0,\infty)^n$. Then $p^i(\lambda y) \ne p^i(y)$ for any $\lambda > 0, \lambda \ne 1$. However, $p^1(y) \ge p^1(x)$ implies $2^{y_1 - y_2} \ge 2^{x_1 - x_2}$, which gives, $2^{\lambda(y_1 - y)_2} \ge 2^{\lambda(x_1 - x_2)}$, that is, $p^1(\lambda y) \ge p^1(\lambda x)$ for any $\lambda > 0$. Thus, if we restrict ourselves to the dimension $N = 2$, then $p^i(y)$ is scale consistent, but not scale invariant.

Note that satisfaction of $p^i(\lambda y) \ge p^i(\lambda x)$ for all $\lambda > 0$ implies fulfilment of $p^i(y) \ge p^i(x)$. Note also that if $p^i(y) > p^i(x)$ holds, then there is at least one contestant $j \ne i$ such that $p^j(y) < p^j(x)$ holds. The reason for this is that $\sum_{i=1}^{n} p^i(y) = \sum_{i=1}^{n} p^i(y) = 1$. (A9) is an ordinal property in the sense that the inequality remains invariant under any ordinal transformation Ω of p^i s. Furthermore, $\Omega\left(p^i\right)$ s given by $\Omega\left(p^i(y)\right)=\dfrac{\Omega\left(p^i(y)\right)}{\sum_{j\in N}\Omega\left(p^j(y)\right)}$, $i \in N$, are probabilities[25].

The next theorem demonstrates that the CSF of the power function type is the only one that fulfils (A9). For this characterization, we omit the origin from the domain of the CSF.

Theorem 6.2: Assume that the number of contestants is greater than 2 and the function f is continuously differentiable in $(0,\infty)$. Then the CSF satisfies axioms (A1) – (A5) and (A9) if and only if it is of the Tullock (1980) form given by

$$p^i(y)=\frac{y_i^{\delta}}{\sum_{j\in N} y_j^{\delta}}, \tag{6.10}$$

where $\delta > 0$ is a constant, $y \neq 01^n$.

Proof: See Appendix.

Remark 6.2: However, it is easy to check that in dimension $N = 2$, both the forms of $p^i(y)$ (specified in (6.10)) satisfy (A9). Thus, in this case we get a CSF distinct from the Tullock form.

Combining Theorem 2 of Skaperdas (1996) and Theorem 6.2 of this chapter we arrive at the following result:

Theorem 6.3: Assume that the number of contestants is greater than 2. Then the following statements are equivalent:

(i) The CSF satisfies axioms (A1) – (A6).

(ii) The CSF satisfies axioms (A1) – (A5) and (A9).

(iii) The CSF is of the Tullock form given by (6.10).

We next consider the following ordinal counterpart to (A7):

---

[25](A9) becomes Zheng's (2007) unit consistency axiom if we replace p^i by an inequality index, *y* and *x* by income distributions in two *n*-person societies and the weak inequality $\geq$ by the strict inequality > in (see also Chakravarty 2015).

(A10) Translation Consistency: For $x, y \in [0,\infty)^n$, if for some $i \in N$, $p^i(y) \geq p^i(x)$ holds, then $p^i\left(y + c1^n\right) \geq p^i\left(x + c1^n\right)$, where 1^n is the *n*-coordinated vector of ones and c is a scalar such that $y_i + c \geq 0$ for all $i \in N$.

Evidently, (A7) is sufficient but not necessary for (A10). Like (A9), (A10) is also an ordinal property.

Remark 6.3: Fix $x \in (0,\infty)^n$ and define

$Y_x^* = \left\{ y \in R^n : there\ is\ i \in N\ such\ that\ y_i \geq x_i\ and\ y_j \leq x_j\ for\ all\ j \neq i \right\}$. Then for all $y \in Y_x^*$ we have, $p^i(y) \geq p^i(x)$. Also, $y \in Y_x^*$ implies: $\lambda y \in Y_x^*$ for all $\lambda > 0$ and $y + c1^n \in Y_x^*$ for all $c > 0$. From this it follows that $p^i(\lambda y) \geq p^i(\lambda x)$ and $p^i\left(y + c1^n\right) \geq p^i\left(x + c1^n\right)$. This observation, however, implies neither axiom (A9) nor (A10). For, $p^i(y) \geq p^i(x)$ never implies that $y \in Y_x^*$.

In the following theorem we characterize the entire class of CSFs that are translation consistent.

Theorem 6.4: Assume that the number of contestants is greater than 2 and the function f meets assumption (A).Then the CSF satisfies axioms (A$5'$) and (A10) if and only if it is of the Hirschleifer (1989) form given by:

$$p^i(y) = \frac{e^{\theta y_i}}{\sum_{j \in N} e^{\theta y_j}}, \tag{6.11}$$

where θ is a positive constant.

Proof: See Appendix.

Remark 6.4: However, it is easy to check that for $N = 2$, both the forms of $p^i(y)$ (mentioned in (6.48)) satisfy (A10). Thus, in this case we get a CSF distinct from the Hirschleifer form.

Theorem 3 of Skaperdas (1996) and Theorem 6.4 of this chapter can now be combined to yield the following result:

Theorem 6.5: Assume that the number of contestants is greater than 2. Then the following statements are equivalent:

(i) The CSF satisfies axioms (A1) – (A5) and (A7).

(ii) The CSF satisfies axioms (A1) – (A5) and (A10).

(iii) The CSF is of the Hirshleifer form given by (6.11).

Instead of considering scale consistency (A9) or translation consistency (A10), we can also consider the following intermediate form of consistency, which is clearly an ordinal counterpart of (A8).

(A11) **Intermediate Consistency:** For $x, y \in [0,\infty)^n$, if for some $i \in N$, $p^i(y) \geq p^i(x)$ holds, then

$$p^i\left\{x + c\left(\mu x + (1-\mu)1^n\right)\right\} \geq p^i\left\{y + c\left(\mu y + (1-\mu)1^n\right)\right\}$$

where $\mu \in [0,1]$ is a parameter and $c \in R$ is a scalar such that

$$x + c\left(\mu x + (1-\mu)1^n\right),\ y + c\left(\mu y + (1-\mu)1^n\right) \in [0,\infty)^n .$$

We now characterize all CSFs satisfying intermediate consistency.

Theorem 6.6: Assume that the number of contestants is greater than 2 and let the function f meet assumption (A). Then the CSF satisfies axioms (A$5'$) and (A11) if and only if it is of the intermediate form given by (6.3).

Proof: See Appendix.

We are now in a position to state the following:

Theorem 6.7: Assume that the number of contestants is greater than 2 and the function *f* meets assumption (A), is twice continuously differentiable with positive second order derivatives on $[0,\infty)$. Then the following statements are equivalent:

(i) The CSF satisfies axioms (A1) – (A5) and (A8).

(ii) The CSF satisfies axioms (A1) – (A5) and (A11).

(iii) The CSF is of the functional form given by (6.3).

Remark 6.5: However, it is easy to check that in dimension $N = 2$, all the forms of $p^i(y)$ resulting from (6.66), (6.70) and (6.72) satisfy (A9). Thus, in this case there are CSFs other than the one given by (6.3).

6.3 Conclusion

Axiomatic characterizations of contest success functions enable us to understand them in an intuitively reasonable way in the sense that necessary and sufficient conditions are identified to isolate them uniquely. Skaperdas (1996) characterized the Tullock and Hirschleifer forms of contest success functions. In this chapter we have substantially extended the characterizations of Skarpedas (1996) by considering a general axiom (on intermediate invariance) and three more axioms viz. scale, translation and intermediate consistencies, which are ordinal in nature, a characteristic that has not been explored earlier in the literature. It has been shown that if the number of contestants in the game is at least 3, the Tullock and Hirschleifer functional forms are the only functional forms satisfying respectively scale and translation consistencies. The consistency axioms, which are simple and elegant, may be considered as the most fundamental contributions of the chapter. We also look at the possibility of existence of Nash equilibria, including the ones that may turn out as corner solutions, in different situations.

6.4 Appendix

Proof of Theorem 6.1: Consider $(y_1, y_2) \in (0, \infty)^2$ and note that $p^i(y) = \dfrac{f(y_i)}{f(y_1) + f(y_2)}$, $i = 1, 2$.

Then by (A8) we get,

$$\frac{f[c(1+\mu)y_1 + c(1-\mu)]}{f(y_1)} = \frac{f[c(1+\mu)y_2 + c(1-\mu)]}{f(y_2)} , \tag{6.12}$$

where for simplicity it is assumed that $c > 0$. From (6.12) it follows that $\dfrac{f[c(1+\mu)z + c(1-\mu)]}{f(z)}$ is independent of the effort level z. Differentiating $\dfrac{f[c(1+\mu)z + c(1-\mu)]}{f(z)}$ with respect to z we get,

$$\frac{d}{dz}\left(\frac{f[c(1+\mu)z + c(1-\mu)]}{f(z)}\right) = 0 , \tag{6.13}$$

which implies that

$$(c\mu + 1)\, f(z)\, f'\{(c\mu + 1)z + c(1-\mu)\} = f'(z)\, f\{(c\mu + 1)z + c(1-\mu)\} , \tag{6.14}$$

where f' stands for the derivative of f.

Equation (6.14) holds for all finite $z > 0$. Letting $z \to 0$ on each side of (6.14) and applying continuity of f' we get

$$(c\mu + 1) f(0) f'\{c(1-\mu)\} = f'(0) f\{c(1-\mu)\} \qquad (6.15)$$

from which it follows that

$$\frac{f'\{c(1-\mu)\}}{f\{c(1-\mu)\}} = \frac{\eta}{(c\mu + 1)}, \qquad (6.16)$$

where $\eta = \frac{f'(0)}{f(0)} > 0$ (since $f(0) > 0$ and $f'(0) > 0$, by assumption (A)). Integrating both sides of (6.16) we get,

$$lnf\{c(1-\mu)\} = \frac{\eta}{\mu} ln(c\mu + 1) + k, \qquad (6.17)$$

which yields:

$$f\{c(1-\mu)\} = (c\mu + 1)^{\frac{\eta}{\mu}} e^{k}. \qquad (6.18)$$

This holds for all $c > 0$ and for all $\mu \epsilon (0,1)$. Thus,

$$f(z) = \xi\{\mu(z-1) + 1\}^{\frac{\eta}{\mu}}, \qquad (6.19)$$

where $\xi = \frac{e^{k}}{(1-\mu)^{\eta/\mu}}, \eta > 0$ are constants. By continuity of f, the solution extends to the case where $z = 0$. Substituting this form of f into $p^{i}(y) = \frac{f(y_i)}{\sum_{j \in N} f(y_j)}$ we get the desired form of the CSF. This establishes the necessity part of the theorem. The sufficiency is easy to verify. □

Proof of Theorem 6.2: By Theorem 1 of Skaperdas (1996), axioms (A1) – (A5) are satisfied if and only if the CSF is given by (A$5'$). Observe that for any $y = (y_1, y_2) \in (0, \infty)^2$ we have,

$p^1(y) = \dfrac{f(y_1)}{f(y_1)+f(y_2)}$. Consider $(y_1', y_2'), (\bar{y}_1, \bar{y}_2) \in (0,\infty)^2$. Then $p^1(y') \geq p^1(\bar{y})$ is same as $\dfrac{f(y_1')}{f(y_1')+f(y_2')} \geq \dfrac{f(\bar{y}_1)}{f(\bar{y}_1)+f(\bar{y}_2)}$, that is, if and only if $\dfrac{f(y_2')}{f(y_1')} \leq \dfrac{f(\bar{y}_2)}{f(\bar{y}_1)}$. Thus, by (A9) we have,

$$\frac{f(y_2')}{f(y_1')} \leq \frac{f(\bar{y}_2)}{f(\bar{y}_1)} \text{ if and only if } \frac{f(\lambda y_2')}{f(\lambda y_1')} \leq \frac{f(\lambda \bar{y}_2)}{f(\lambda \bar{y}_1)} \text{ for all } \lambda > 0. \tag{6.20}$$

Now, we claim that $\dfrac{f(\lambda y_2)}{f(\lambda y_1)} = F_\lambda\left(\dfrac{f(y_2)}{f(y_1)}\right)$ for some non-decreasing function F_λ. To demonstrate this, consider, as before, two distinct effort vectors $(y_1', y_2'), (\bar{y}_1, \bar{y}_2) \in (0,\infty)^2$. Then we have,

$$\frac{f(\lambda \bar{y}_2)}{f(\lambda \bar{y}_1)} = \frac{f(\lambda y_2')}{f(\lambda y_1')} \text{ if and only if } \frac{f(\bar{y}_2)}{f(\bar{y}_1)} = \frac{f(y_2')}{f(y_1')}.$$

This implies that $\dfrac{f(\lambda y_2)}{f(\lambda y_1)}$ is a function of $\dfrac{f(y_2)}{f(y_1)}$. Non-decreasingness of this function is a consequence of (6.20).

Define

$$u_\lambda(y_1, y_2) = \frac{f(\lambda y_2)}{f(\lambda y_1)} \tag{6.21}$$

and

$$q(y_1, y_2) = \frac{f(y_2)}{f(y_1)}. \tag{6.22}$$

Since u_λ and q are functionally related, the Jacobian of u_λ and q with respect to y_1 and y_2 must vanish. More precisely,

$$\begin{vmatrix} \dfrac{\partial u_\lambda}{\partial y_1} & \dfrac{\partial u_\lambda}{\partial y_2} \\ \dfrac{\partial q}{\partial y_1} & \dfrac{\partial q}{\partial y_2} \end{vmatrix} = 0. \tag{6.23}$$

This implies that

$$\frac{f(\lambda y_2) f'(\lambda y_1) f'(y_2)}{f(\lambda y_1)} = \frac{f'(y_1) f(y_2) f'(\lambda y_2)}{f(y_1)}. \tag{6.24}$$

Equation (6.24) can be rearranged as

$$\frac{f'(\lambda y_1)}{f(\lambda y_1)}\cdot\frac{f'(y_2)}{f(y_2)}=\frac{f'(\lambda y_2)}{f(\lambda y_2)}\cdot\frac{f'(y_1)}{f(y_1)}. \tag{6.25}$$

Now, (6.25) holds for all $(y_1, y_2) \in (0,\infty)^2$. Putting $y_1 = z > 0, y_2 = 1$ in (6.25) and letting $h(z)=\frac{f'(z)}{f(z)}$ we get,

$$h(\lambda z)\,h(1) = h(z)\,h(\lambda). \tag{6.26}$$

Given that f is positive valued on $(0,\infty)$ and increasing, h is positive. It is continuous as well. Since (6.26) holds for all positive z and λ, it is a fundamental Cauchy equation, of which the only continuous solution is given by

$$h(z) = K_1 z^{\alpha} \tag{6.27}$$

for some $K_1 > 0$ and α is a real number (Aczel, 1966, p. 41, Theorem 3).

Case I: $\alpha \neq -1$

Then (6.27) yields:

$$\frac{f'(z)}{f(z)} = K_1 z^{\alpha}. \tag{6.28}$$

Integrating both sides of (6.28) we get,

$$\ln(f(z)) = Kz^{\alpha+1} + K', \tag{6.29}$$

where K and K' are real numbers. Equation (6.29) is equivalent to

$$f(z) = AB^{z^{\beta}}, \tag{6.30}$$

where $A = e^{K'} > 0, B = e^{K} > 0$ and $\beta = 1+\alpha$ is a non-zero real number.

Case II: $\alpha = -1$.

Then (6.28) becomes:

$$\frac{f'(z)}{f(z)} = K_1 z^{-1}\ , \tag{6.31}$$

which, on integration, gives

$$\ln\left(f(z)\right) = K_1 \ln(z) + K'. \tag{6.32}$$

That is,

$$f(z) = Az^B, \tag{6.33}$$

where $A = e^{K'} > 0$ and B is a real number. Since f is strictly increasing, we further require the restriction $B > 0$.

(6.27) and (6.33) solves $f(z)$ for $z > 0$. By continuity of f, the solution extends to $z = 0$.

Plugging the forms of f given by (6.30) and (6.33) into $p^i(y) = \frac{f(y_i)}{\sum_{j\in N} f(y_j)}$, we get the following forms of $p^i(y)$:

$$p^i(y) = \begin{cases} \dfrac{B^{y_i^\beta}}{\sum_{j\in N} B^{y_j^\beta}}, \\ \dfrac{y_i^B}{\sum_{j\in N} y_j^B}. \end{cases} \tag{6.34}$$

Out of these two, only the latter satisfies (A9) if $N > 2$. To see a counterexample for $N = 3$, consider the CSF given by the first functional form in (6.34). Take $B = 2$ and $\beta = 1$. Let $y = \left(1, \frac{\ln 14}{\ln 2}, \frac{\ln 6}{\ln 2}\right)$ and $x = \left(1, 4, \frac{\ln 3}{\ln 2}\right)$. Note that $p^1(y) = \frac{2}{2+14+6} = \frac{1}{11}$ and $p^1(x) = \frac{2}{2+16+3} = \frac{2}{21}$ so that $p^1(y) < p^1(x)$. But $p^1(2y) = \frac{2^2}{2^2+14^2+6^2} = \frac{4}{236}$ and $p^1(2x) = \frac{2^2}{2^2+16^2+3^2} = \frac{4}{269}$ implying that $p^1(2y) > p^1(2x)$. Thus the CSF fails to satisfy (A9).

Putting $B = \delta$ in the second functional form in (6.34), we get the Tullock form of CSF given by (6.10). This completes the necessity part of the proof of the theorem. The sufficiency can be easily verified by checking that the CSF given by (6.10) fulfils (A1)-(A5) and (A9). □

Proof of Theorem 6.4: Take, as in the proof Theorem 6.2, $(y_1', y_2'), (\bar{y}_1, \bar{y}_2) \in (0, \infty)^2$. Then $p^1(y') \geq p^1(\bar{y})$ is same as $\frac{f(y_2')}{f(y_1')} \leq \frac{f(\bar{y}_2)}{f(\bar{y}_1)}$.

By (A10),

$$\frac{f(y_2')}{f(y_1')} \leq \frac{f(\bar{y}_2)}{f(\bar{y}_1)} \text{ if and only if } \frac{f(y_2' + c)}{f(y_1' + c)} \leq \frac{f(\bar{y}_2 + c)}{f(\bar{y}_1 + c)} \text{ for all } c > 0 . \tag{6.35}$$

As in the proof of Theorem 6.2, one can easily see that there exists a continuous and non-decreasing function G_c such that

$$\frac{f(y_2 + c)}{f(y_1 + c)} = G_c\left(\frac{f(y_2)}{f(y_1)}\right). \tag{6.36}$$

for all $y_1, y_2 > 0$.

Define

$$w_c(y_1, y_2) = \frac{f(y_2 + c)}{f(y_1 + c)}. \tag{6.37}$$

Since w_c and q in (6.22) are functionally related, the Jacobian of w_c and q with respect to y_1 and y_2 must vanish. This implies that

$$\frac{f'(y_1 + c)}{f(y_1 + c)} . \frac{f'(y_2)}{f(y_2)} = \frac{f'(y_2 + c)}{f(y_2 + c)} . \frac{f'(y_1)}{f(y_1)} . \tag{6.38}$$

Equation (6.38) holds for all $(y_1, y_2) \in (0, \infty)^2$. Putting $y_1 = z > 0, y_2 = \varepsilon > 0$ and substituting $\frac{f'(z)}{f(z)}$ by $\psi_1(z)$, which is positive on $(0, \infty)$, we get

$$\psi_1(z + c)\psi_1(\varepsilon) = \psi_1(z)\psi_1(c + \varepsilon) \tag{6.39}$$

Letting $\varepsilon \to 0$ in (6.39) and using continuous differentiability of f we get,

$$\psi_1(z+c)\psi_1(0)=\psi_1(z)\psi_1(c). \tag{6.40}$$

From (6.40) it follows that $\psi_1(0)>0$. This equation holds for all positive z and c. The only continuous solution to (6.40) is given by

$$\psi_1(z)=\nu e^{\rho z} \tag{6.41}$$

for some positive $\nu=\psi_1(0)$ and real ρ (see Aczel, 1966, p.84). By continuity of ψ_1, the solution extends to the case when $z=0$.

From (6.41) it is evident that

$$\frac{f'(z)}{f(z)}=\upsilon e^{\rho z}. \tag{6.42}$$

Case I: $\rho \neq 0$.

Integrating both sides of (6.42) we get,

$$ln(f(z))=K_3 \nu e^{\rho z}+K_4 \quad , \tag{6.43}$$

where K_3 and K_4 are real numbers. That is,

$$f(z)=EH^{e^{\rho z}} \tag{6.44}$$

where $E=e^{K_4}$ and $H=e^{K_3\upsilon}$ are positive constants.

Case II: $\rho=0$.

Then (6.42) becomes:

$$\frac{f'(z)}{f(z)}=\upsilon. \tag{6.45}$$

Integrating both sides of (6.45) we get,

$$\ln f(z)=\nu z+c \tag{6.46}$$

for some real number c.

Equation (6.46) is equivalent to:

$$f(z)=Qe^{\nu z}, \tag{6.47}$$

where $Q = e^{c} > 0$.

(6.44) and (6.47) solves $f(z)$ for $z > 0$. By continuity of f the solution extends to $z = 0$.

For strict increasingness of f we need the restriction $\upsilon > 0$. Substituting the forms of f given by (6.44) and (6.47) in $p^{i}(y) = \dfrac{f(y_i)}{\sum_{j\in N} f(y_j)}$, the resulting forms of $p^{i}(y)$ become:

$$p^{i}(y) = \begin{cases} \dfrac{H^{e^{\rho y_i}}}{\sum_{j\in N} H^{e^{\rho y_j}}}, \\ \dfrac{e^{\upsilon y_i}}{\sum_{j\in N} e^{\upsilon y_j}}. \end{cases} \tag{6.48}$$

However, it can be easily checked that the former violates (A10) if $N > 2$. To see a counterexample for dimension $N = 3$, consider the CSF given by the first functional form in (6.48). Take $H = 2$, $\rho = 1$ and $c = \ln 2$. Let $y = \left(0, \ln\left(\dfrac{\ln 14}{\ln 2}\right), \ln\left(\dfrac{\ln 6}{\ln 2}\right)\right)$ and $x = \left(0, \ln 4, \ln\left(\dfrac{\ln 3}{\ln 2}\right)\right)$. Note that $p^{1}(y) = \dfrac{2}{2+14+6} = \dfrac{1}{11}$ and $p^{1}(x) = \dfrac{2}{2+16+3} = \dfrac{2}{21}$ so that $p^{1}(y) < p^{1}(x)$. But $p^{1}\left(y + c1^{3}\right) = \dfrac{2^{2}}{2^{2}+14^{2}+6^{2}} = \dfrac{4}{236}$ and $p^{1}\left(x + c1^{3}\right) = \dfrac{2^{2}}{2^{2}+16^{2}+3^{2}} = \dfrac{4}{269}$ implying that $p^{1}\left(y + c1^{3}\right) > p^{1}\left(x + c1^{3}\right)$. Thus the CSF fails to satisfy (A10).

Putting $\upsilon = \theta$ in the second functional form specified in (6.48) we arrive at the CSF given by (6.34). Hence the necessity part of the theorem is demonstrated. The sufficiency follows easily. □

Proof of Theorem 6.6: Take, as in the proofs of Theorems 6.2 and 6.4, $(y_1', y_2'), (\bar{y}_1, \bar{y}_2) \in [0, \infty)^2$.

By (A11) it follows that

$$\frac{f(y_2')}{f(y_1')} \le \frac{f(\bar{y}_2)}{f(\bar{y}_1)} \;\; iff \;\; \frac{f\{y_2' + c(\mu y_2' + 1 - \mu)\}}{f\{y_1' + c(\mu y_1' + 1 - \mu)\}} \le \frac{f\{\bar{y}_2 + c(\mu \bar{y}_2 + 1 - \mu)\}}{f\{\bar{y}_1 + c(\mu \bar{y}_1 + 1 - \mu)\}} \tag{6.49}$$

for all $c > 0$.

Therefore, for all $y_1, y_2 \in [0, \infty)$ we have,

$$\frac{f\{y_2 + c(\mu y_2 + 1 - \mu)\}}{f\{y_1 + c(\mu y_1 + 1 - \mu)\}} = H_\mu\left(\frac{f(y_2)}{f(y_1)}\right) \tag{6.50}$$

for some continuous and non-decreasing function H_μ .

Define

$$v_c(y_1, y_2) = \frac{f\{y_2 + c(\mu y_2 + 1 - \mu)\}}{f\{y_1 + c(\mu y_1 + 1 - \mu)\}} \tag{6.51}$$

and

$$r(y_1, y_2) = \frac{f(y_2)}{f(y_1)}. \tag{6.52}$$

Since v_c and r are functionally related, the Jacobian of v_c and r with respect to y_1 and y_2 must vanish. That is,

$$\begin{vmatrix} \frac{\partial v_c}{\partial y_1} & \frac{\partial v_c}{\partial y_2} \\ \frac{\partial r}{\partial y_1} & \frac{\partial r}{\partial y_2} \end{vmatrix} = 0 \tag{6.53}$$

Simplifying and rearranging we get,

$$\frac{f'\{y_2 + c(\mu y_2 + 1 - \mu)\}}{f\{y_2 + c(\mu y_2 + 1 - \mu)\}} \frac{f'(y_1)}{f(y_1)} = \frac{f'\{y_1 + c(\mu y_1 + 1 - \mu)\}}{f\{y_1 + c(\mu y_1 + 1 - \mu)\}} \frac{f'(y_2)}{f(y_2)}. \tag{6.54}$$

For $z \in [0, \infty)$, put

$$h(z) = \frac{f'(z)}{f(z)}. \tag{6.55}$$

Then h is positive-valued (since by assumption, f is positive and strictly increasing) and is continuousy differentiable (by the assumed twice continuous diffeentiability of f).

From (6.54) it follows that

$$\frac{h\left\{y_2+c\left(\mu y_2+1-\mu\right)\right\}}{h\left\{y_1+c\left(\mu y_1+1-\mu\right)\right\}}=\frac{h\left(y_2\right)}{h\left(y_1\right)}. \tag{6.56}$$

This holds for all $y_1, y_2 \in [0, \infty)$. Putting $y_2 = z$ and $y_1 = 0$ we get,

$$\frac{h\left((1+c\mu)z+c(1-\mu)\right)}{h\left(c(1-\mu)\right)}=\frac{h(z)}{h(0)} . \tag{6.57}$$

Put

$$\ln\left(\frac{h(z)}{h(0)}\right)=\phi(z), \tag{6.58}$$

which gives

$$\phi(0)=0 . \tag{6.59}$$

Moreover, ϕ is differentiable since h is.

Then (6.57) yields:

$$\phi\left\{(1+c\mu)z+c\left(1-\mu\right)\right\}=\phi(z)+\phi\left(c(1-\mu)\right) \tag{6.60}$$

which implies that

$$\frac{\phi\{(1+c\mu)z+c(1-\mu)\}}{(1+c\mu)z}=\frac{\phi(z)-\phi(0)}{(1+c\mu)z}. \tag{6.61}$$

Proceeding to limits of both sides as $z \to 0$ we have,

$$\phi'\left\{c\left(1-\mu\right)\right\}=\frac{1}{\left(1+c\mu\right)}\phi'(0) . \tag{6.62}$$

Differentiating both sides of (6.58) we get,

$$\phi'(z)=\frac{h'(z)}{h(z)} . \tag{6.63}$$

Substituting ϕ' from (6.63) into (6.62) we get,

$$\frac{h'\left\{c\left(1-\mu\right)\right\}}{h\left\{c\left(1-\mu\right)\right\}}=\frac{\eta_0}{1+c\mu}, \tag{6.64}$$

where $\eta_0=\frac{h'(0)}{h(0)}$.

Case I: $h'(0) = 0$.

Then $\eta_0 = 0$ and from (6.64) it follows that $h'(t) = 0$ for all $t \in [0, \infty)$. Consequently, $h(t) = c_1$ for some positive constant c_1. This, in turn implies that

$$\frac{f'(t)}{f(t)} = c_1 \tag{6.65}$$

for all $t \in [0, \infty)$. Integrating both sides of (6.65) we get,

$$f(t) = c_2 e^{c_1 t}, \tag{6.66}$$

where $c_2 > 0$ is a constant.

Case II: $h'(0) \neq 0$.

Then proceeding as in the proof of Theorem 6.1 we can show that

$$h(z) = \kappa \left\{ \mu (z - 1) + 1 \right\}^{\frac{\eta_0}{\mu}} \tag{6.67}$$

for some constant $\kappa > 0$.

Using (6.55) we have,

$$\frac{f'(z)}{f(z)} = \kappa \left\{ \mu (z - 1) + 1 \right\}^{\frac{\eta_0}{\mu}} . \tag{6.68}$$

If $\frac{\eta_0}{\mu} \neq -1$, then integrating out both sides of (6.68) we get,

$$\ln f(z) = \kappa \frac{1}{\left(\frac{\eta_0}{\mu} + 1 \right)} \left\{ \mu (z - 1) + 1 \right\}^{\frac{\eta_0}{\mu}} + \chi_1 , \tag{6.69}$$

for some constant χ_1 .

Thus,

$$f(z) = \chi \exp \left[\kappa \frac{1}{\left(\frac{\eta_0}{\mu} + 1 \right)} \left\{ \mu (z - 1) + 1 \right\}^{\frac{\eta_0}{\mu}} \right], \tag{6.70}$$

where $\chi > 0$.

On the other hand, if $\frac{\eta_0}{\mu} = -1$, then (6.68), on integration w.r.t. *z* yields:

$$\ln f(z) = \kappa \frac{1}{\mu} \ln\{\mu(z-1)+1\} + \chi_1 , \tag{6.71}$$

which implies that

$$f(z) = \chi\{\mu(z-1)+1\} e^{\frac{\kappa}{\mu}} \tag{6.72}$$

for some constant $\chi > 0$.

Now if $N > 2$, then it is easy to see that out of the forms of CSF given by (6.66), (6.70) and (6.72), only the one resulting from (6.72) is in conformity with (A11). Substituting κ by η we catch hold of the CSF given by (6.3). This completes the proof of the necessity part of the Theorem. The sufficiency can be checked easily. □

BIBLIOGRAPHY

1. Aczel, J. (1966) *Lectures on Functional Equations and Their Applications,* Academic Press, New York.

2. Alesina, A., R. Baqir and W. Easterly (1999) "Public goods and ethnic divisions" *Quarterly Journal of Economics* 114, 1243-1284.

3. Alesina, A. and E. La Ferrara (2000) "Participation in heterogeneous communities" *Quarterly Journal of Economics* 115, 847-904.

4. Alesina, A and E. La Ferrara (2005) "Ethnic diversity and economic performance" *Journal ofEconomic literature* 43, 762-800.

5. Alesina, A. and E. Spolaore (1997) "On the number and size of nations", *Quarterly Journal of Economics* 112, 1027-1056.

6. Allison, R.A. and J.E. Foster (2004) "Measuring health inequality using qualitative data", *Journal of Health Economics* 23, 505-524.

7. Amegashie, J.A. (2006) "A contest success function with a tractable noise parameter", *Public Choice*126, 135 – 144.

8. Amiel, Y. and F.A Cowell (1992) "Measurement of Income Inequality: Experimental test by questionnaire welfare and monotonicity", *Journal of Public Economics* 47, 3-26.

9. Amiel, Y. and F.A. Cowell, (2003) "Inequality, welfare and monotonicity", *Research on Economic Inequality* 9, 35-46.

10. Amiel, Y, F.A. Cowell, and X. Ramos (2010): "Poles apart? An analysis to the meaning of polarization", *Review of Income and Wealth* 56, 23-46.

11. Apouey, B. (2007) "Measuring health polarization with self-assessed health data", *Journal of Health Economics* 23, 505-524.

12. Apouey, B. (2010) "On measuring and explaining socioeconomic polarization in health with an application to French data", *Review of Income and Wealth* 56, 141-170.

13. Arbatskaya, M. and H. M. Mialon (2010) "Multi-activity contests", *Economic Theory* 43, 23-43.

14. Arcagni, A. and M. Fettore (2013) "Measuring multidimensional polarization with ordinal data", mimeo, University of Milano-Bicocca, Italy.

15. Baik, K. H. (1998) "Difference form contest success functions and effort level in contests", *European Journal of Political Economy* 14, 685-701.

16. Baye, M., Kovenock, D., de Vries, C.G. (2003) "The strategic equivalence of rent-seeking, innovation and patent-race games", *Games Economic Behaviour* 44, 217-226.

17. Ben Porath, E. and I. Gilboa (1994): "Linear measures, the Gini index, and the income-equality tradeoff", *Journal of Economic Theory* 64, 443-467.

18. Blackburn, L. Mckinley and D.E. Bloom (1985) "What is happening to the middle class?", *American Demographics* 7, 18-25.

19. Blakorby, C., M. Auersperg and D. Donaldson (1981) "A new procedure for the measurement of inequality, within group and among population subgroups: ethically consistent subindices", *Canadian Journal of Economics* 14, 665-685.

20. Blackorby, C. and D. Donaldson (1978), "Measures of relative equality and their meaning in terms of social welfare", *Journal of Economic Theory* 18, 59-80.

21. Blakorby, C., W. Bossert and D. Donaldson (1999) "Income inequality measurement:the normative approach" in *The Handbook of Income Inequality Measurement* by Silber, J. (Eds.), Springer; 133-161.

22. Bosmans, K. and F.A. Cowell (2010) "The class of absolute decomposable inequality measures", *Economic Letters* 109, 154-156.

23. Bossert, W. (1990): "An axiomatization of the single-series Ginis", *Journal of Economic Theory* 50, 82-92.

24. Bossert W, C. D'Ambrosio and E. La Ferrara (2011) "A generalized index of fractionalization", *Economica* 78, 723-750.

25. Bossert W. and A. Pfingsten (1990) "Intermediate inequality: concepts, indices and welfare implications", *Mathematical Social Sciences* 19, 117-134.

26. Bossert, W and W. Schworm (2008) "A class of two-group polarization measures", *Journal of Public Economic Theory* 10, 1169-1187.

27. Bourguignon, F. (1979) "Decomposable income inequality measures", *Econometrica* 47, 901-920.

28. Chakravarty, S.R. (1983) "A new index of poverty", *Mathematical Social Sciences* 6, 307 – 313.

29. Chakravarty, S.R. (2001) "The Variance as a subgroup-decomposable measure of inequality", *Social Indicators Research* 53, 79-95.

30. Chakravarty, S.R. (2009) *Inequality, Polarization and Poverty*: *Advances in Distributional Analysis*, Springer, New York.

31. Chakravarty, S. R. (2015) *Inequality, polarization and conflict: An analytical study*, Springer, New York.

32. Chakravarty S.R., N. Chattopadhyay and B. Maharaj (2010) 'Inequality and Polarization: An Axiomatic Approach' in *The Measurement of Individual Well-Being and Group Inequalities: Essays in Memory of Z.M.Berrebi*, by J. Deutsch and J. Silber, Eds., Routledge: London.

33. Chakravarty, S.R. and C. D'Ambrosio (2006) "The measurement of social exclusion", *Review of Income and Wealth* 52, 377-398.

34. Chakravarty S.R. and C. D'Ambrosio (2010) "Polarization ordering of income distributions", *Review of Income and Wealth* 56, 47-64.

35. Chakravarty, S.R., B. Dutta and J.A. Weymark (2007) "Ethical indices of income mobility", *Social Choice and Welfare* 2, 1-21.

36. Chakravarty, S.R. and B. Maharaj (2011a): "Measuring ethnic polarization", *Social Choice and Welfare* 37, 431-452.

37. Chakravarty, S.R. and B. Maharaj (2011b) "Subgroup-decomposable inequality indices and reduced-form indices of polarization" *Keio Economic Studies* 47, 57-83.

38. Chakravarty, S.R. and B. Maharaj (2012) "Ethnic polarization orderings and indices" *Journal of Economic Interaction and Co-ordination* 7, 99-123.

39. Chakravarty, S.R. and B. Maharaj (2014) "New Contest success functions", ECINEQ Working paper 340.

40. Chakravarty, S.R. and B. Maharaj (2015) "Generalized Gini Indices of human well-being" *International Journal of Economic Theory*, 11, 231-246.

41. Chakravarty, S.R. and A. Majumder (2001) "Inequality, polarization and welfare: theory and applications", *Australian Economic Papers* 40, 1-13.

42. Chakravarty, S.R., A. Majumder and S. Roy (2007) "A treatment of absolute indices of polarization", *Japanese Economic Review* 58, 273-293.

43. Chakravarty, S.R. and S. Tyagarupananda (1998) "The subgroup decomposable absolute indices of inequality" in Chakravarty, S.R., D. Coondoo and R. Mukherjee (Eds.) *Quantitative Economics: Theory and Practice, Essays in Honor of Professor N. Bhattacharya*, Allied Publishers, New Delhi, 247-257.

44. Chakravarty, S.R. and S. Tyagarupananda (2009) "The subgroup decomposable intermediate indices of inequality", *Spanish Economic Review* 11, 83-97.

45. Chakravarty, S.R. and C. Zoli (2012): "Stochastic dominance relations for integer variables", *Journal of Economic Theory* 147, 1331-1341.

46. Chantreuil, F. and A. Trannoy (1999) "Inequality decomposition values: the trade-off between marginality and consistency", THEMA working papers, 99-124, Universite' de Cergy-Pontoise.

47. Che, Y. K. and I. Gale (2000) "Difference form contests and the robustness of all-pay auctions" , *Games and Economic Behaviour* 30, 22-43.

48. Clark, D.J. and C. Riis (1998) "Contest success functions: an extension", *Economic Theory* 11, 201-204.

49. Collier, P. and A. Hoeffler (2004) "Greed and grievances in civil war", *Oxford Economic Papers* 56, 563-595.

50. Corchon, L.C. (2000) "On the allocative effects of rent seeking", *Journal of Public Economic Theory* 2, 483-491.

51. Corchon, L.C. (2007) "The theory of contests: a survey", *Review of Economic Design* 11, 69-100.

52. Corchon, L.C. and M. Dahm (2011) "Welfare maximizing contest success functions when the planner cannot commit", *Journal of Mathematical Economics* 47, 303-317.

53. Cornes, R. and R. Hartley (2005) "Asymmetric contests with general technologies", *Economic Theory* 26, 923-946.

54. Cowell, F.A. and E. Flachaire (2014) "Inequality with ordinal data", Public Economics Programme Discussion Paper No. 16, London School of Economics and Political Sciences, London (http://darp.lse.ac.uk/pdf/IneqOrdinal.pdf).

55. D'Ambrosio, C. (2001) "Household characteristics and the distribution of income in Italy: an application of social distance measures", *Review of Income and Wealth* 47, 43–64.

56. Dalton, H. (1920) "The measurement of the inequality of incomes", *Economic Journal* 30, 348-361.

57. Desmet, K., I. Ortunoâ-Ortin and S. Weber (2008) “Stability of nations and genetic diversity" Working Paper 004-08, *International School of Economics* at TSU, Tbilisi, Republic of Georgia.

58. Desmet K., I. Ortuño-Ortín, and R. Wacziarg (2009) “The Political Economy of Ethnolinguistic Cleavages”, *CEPR Discussion Papers* 7478.

59. Deustch J. and J. Silber (2010) “Analyzing the impact of income sources on changes in bipolarization” in Deutsch, J. and J. Silber (eds.) *The Measurement of Individual Well-Being and Group Inequalities: Essays in Memory of Z.M. Berrebi*, Routledge Economics, Taylor and Francis Group, 127-152.

60. Dixit, A. (1987) “Strategic behavior in contests”, *American Economic Review* 77, 891-898.

61. Donaldson, D. and J.A. Weymark, (1980) “A single parameter generalization of the Gini indices of inequality”, *Journal of Economic Theory* 22, 67-86.

62. Donaldson, D. and J.A. Weymark, (1983): “Ethically flexible Gini indices for income distributions in the continuum”, *Journal of Economic Theory* 29, 353-358.

63. Duclos, J.-Y., J. Esteban and D. Ray (2004) “Polarization: concepts, measurement, estimation”, *Econometrica* 72, 1737-1772.

64. Duclos, J.Y. and D. Echevin (2005) “Bipolarization dominance”, *Economics Letters* 87, 249-258.

65. Easterly, W. and R. Levine (1997) “Africa’s growth tragedy: policies and ethnic divisions”, *The Quaterly Joural of Economics* 112, 1203-1250.

66. Ebert, U. (1987) “Size and distribution of income as determinants of social welfare”, *Journal of Economic Theory* 41, 25-33.

67. Ebert, U. (2010) “The decomposition of inequality reconsidered: weakly decomposable measures”, *Mathematical Social Sciences* 60, 94-103.

68. Esteban, J.M. and L. Mayoral (2011) “Ethnic and religious polarization and social conflict”, mimeo, CSIC, Barcelona.

69. Esteban, J.M., L. Mayoral and D. Ray (2012) “Ethnicity and conflict: an empirical study”, *American Economic Review* 102, 1310-1342.

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

71. Esteban, J.M. and D. Ray (1999) "Conflict and distribution", *Journal of Economic Theory* 87, 379-415.

72. Esteban, J.M., C. Gradin, and D. Ray (2007) "An extension of a measure of polarization, with an application to the income distribution of five OECD countries", *Journal of Economic Inequality* 5, 1-19.

73. Esteban, J.M. and D. Ray (2008a): "On the salience of ethnic conflict", *American Economic Review* 98, 2185-2202.

74. Esteban J.M. and D. Ray (2008b) "Polarization, fractionalization and conflict", *Journal of Peace Research* 45, 163-182.

75. Esteban, J.M. and D. Ray (2011) "Linking conflict to inequality and polarization", *American Economic Review* 101, 1345-74.

76. Esteban, J.M. and D. Ray (2012) "Comparing polarization measures" in *Oxford Handbook of Peace and Conflict* by Garfinkel, M. and S. Skaperdas (Eds.), Oxford University Press, Oxford, 127-151.

77. Fearon, J.D. (2003) "Ethnic and cultural diversity by country", *Journal of Economic Growth* 8, 195-222.

78. Fearon, J.D. and D.D. Laitin (2000) "Violence and the social construction of ethnic identity", *International Organization* 54, 845-877.

79. Foster, J.E. and Y. Jin (1998) "Poverty orderings for the Dalton utility-gap measures" in S.P. Jenkins, A. Kapteyn and B. M.S. van Praag (Eds.), *The Distribution of Welfare and Household Production*. Cambridge University Press, Cambridge.

80. Foster, J.E. and A. Shneyrov (1999) "A general class of additively decomposable inequality measures", *Economic Theory* 14, 89-111.

81. Foster, J.E. and A.F. Shorrocks (1988) "Poverty orderings", *Econometrica* 52, 761-766.

82. Foster, J.E., J. Greer and E. Thorbecke (1984) "A class of decomposable poverty measures", *Econometrica* 42, 761 – 766.

83. Foster, J.E. and M.C. Wolfson (1992) "Polarization and the decline of the middle class: Canada and the U.S.", Mimeo., Vanderbilt University and Statistics Canada.

84. Foster, J.E. and M.C. Wolfson (2010) "Polarization and the decline of the middle class: Canada and the U.S.", *Journal of Economic Inequality* 8, 247-273.

85. Fusco, A. and J. Silber (2014) "On social polarization and ordinal variables: the case of self-assessed health", *European Journal of Health Economics* 15, 841-851.

86. Gehrig, W. (1988) "On the Shannon-Theil concentration measures and its characterizations" in *Measurement in Economics* by W. Eichhorn (ed.), Physica-Verlag Heidelberg, 413-427.

87. Gersbach, H. and H. Haller (2009) "The affectionate society; does competition for partners promote friendliness?" , *Economic Theory* 40, 389-403.

88. Gradin, C. (2000) "Polarizations by subpopulations in Spain, 1973-1991", *Review of Income and Wealth* 48, 457-474.

89. Greenberg, J. (1956) " The measurement of Linguistic diversity", *Language* 32, 109-115.

90. Hagenaars, A. (1987) "A class of poverty measures", *International Economic Review* 28, 583-607.

91. Hardy G. H., Littlewood J. and Polya G. (1934) *Inequalities*, Cambridge University Press, Cambridge.

92. Harff, B. (2003) "No lessons learned from holocaust? Assessing risks of genocide and political mass murder since 1955", A*merican Political Science Review* 97, 7-73.

93. Hirshleifer, J. (1989) "Conflict and rent-seeking success functions: Ratio vs. difference models of relative success", *Public Choice* 63, 101-12.

94. Hirshleifer, J. (1991) "The paradox of power", *Economic Polity* 3, 177-200.

95. Horowitz, D. (1985) *Ethnic Groups in Conflict*, Berkeley: University of California Press.

96. Jia, H. and S. Skaperdas (2012) "Technologies of conflict" in *Oxford Handbook of the Economics of Peace and Conflict* by Garfinkel, M. R.and S. Skaperdas (Eds.), Oxford University Press, Oxford, 2012.

97. Jia, H., S. Skaperdas and S. Vaidya (2013) "Contest functions: theoretical foundations and issues in estimation" *International Journal of Industrial Organization* 31, 211-222.

98. Kakwani, N.C. (1980): "On a class of poverty measures", *Econometrica* 44, 137-148.

99. Kakwani, N.C. (1984): "The relative deprivation curve and its applications", *Journal of Business and Economic Statistics* 2, 384-405.

100. Kolm, S.C. (1976) "Unequal inequalities II", *Journal of Economic Theory* 13, 82-111.

101. Kolmar, M. and A. Wagener (2012) "Contests and the private provision of public good", *Southern Economic Journal* 79, 161-79.

102. Konrad, K.A. (2009) *Strategy and Dynamics in Contests,* Oxford University Press, New York.

103. Krtscha, M. (1994) "A new compromise measure of inequality" in W. Eichhorn (Ed.), *Models and Measurement of Welfare and Inequality*, Springer, Heidelberg.

104. Kuttner, B. (1983) "The declining middle", *Atlantic Monthly* 252, 60-71.

105. Lambert, P.J. (2001) *The Distribution and redistribution of income*, Manchester University Press, Manchester

106. Lasso de la Vega, M.C. and A.M. Urrutia, (2006) "An alternative formulation of the Esteban-Gradin-Ray extended measure of bipolarization", *Journal of Income Distribution* 15, 42-54.

107. Lasso de la Vega, M.C., A.M. Urrutia, and H. Diez (2010) "Unit consistency and bipolarization of income distributions", *Review of Income and Wealth* 56, 65-8.

108. Levy, F. (1987a) "Changes in the distribution of American family incomes", 1947 to 1984, *Science*, 236, 923-27.

109. Levy, F. (1987b) "The middle class: is it really vanishing?" *Brookings Review* 3, 17-21.

110. Lichbach, M. I. (1989) "An evaluation of 'Does economic inequality breed political conflict?' studies", *World Politics* 41, 431-470.

111. Limaye, B. V. (1996) *Functional Analysis*, New Age International, New Delhi.

112. Marshall, A. W. and I. Olkin (1979) *Inequalities: Theory of Majorization and Its Applications,* Academic Press, New York.

113. Mauro, P. (1995) "Corruption and growth" *Quarterly Journal of Economics* 110, 681-712.

114. Mehran, F. (1976): "Linear measures of income inequality", *Econometrica*44, 805-809.

115. Merz, J. and Scherg, B. (2013) "Polarization of time and income: a multi-dimensional approach with well-being gap and minimum 2DGAP German evidenc", IZA Discussion Paper 7418.

116. Miguel, E., S. Satyanath, and E. Sergenti, (2004) "Economic shocks and conflict: an instrumental variable approach", *Journal of Political Economy* 112, 725-753.

117. Mogues, T. (2008) "A two-dimensional measure of polarization", *IFPRI Discussion Paper*, 00837.

118. Montalvo, J.G. and M. Reynal-Querol (2005) "Ethnic polarization, potential conflict and civil wars" *American Economic Review* 95, 796-816.

119. Montalvo, J.G. and M. Reynal-Querol (2008) "Discrete polarization with an application to the determinants of genocides" *Economic Journal* 118, 1835-1865.

120. Münster , J. (2009) "Group contest success functions", *Economic Theory* 41, 345-57.

121. Nissanov, Z., A. Poggi and J. Silber (2011) "Measuring bipolarization and polarization" in *The Measurement of Individual Well-being and Group-Inequalities* by Deustch, J. and J. Silber (Eds.), Routledge, London and New York, 2011.

122. Nitzan, S. (1991) "Collective rent dissipation" *Economic Journal*101, 1522-34.

123. Nitzan, S. (1994) "Modelling rent seeking contests", *European Journal of Political Economy* 10, 41-60.

124. Pigou, A.C. (1912) *Wealth and Welfare*, Macmillan, London.

125. Pressman, S. (2001) "The decline of the middle class: an international perspective", Luxembourg Income Study Working Paper Series No. 280.

126. Permaneyer, I. (2012) "The conceptualization and measurement of social polarization", *Journal of Economic Inequality* 10, 45-74.

127. Permaneyer, I. (2015) "Theoretical approaches to the measurement of income and social polarization", (forthcoming) in *Handbook of Research on Economic and Social Well-being*, by D'Ambrosio, C. (Eds.), Edward Elgar Publishing Limited.

128. Rai, B.K. and R. Sarin, R. (2009) "Generalized contest success functions" *Economic Theory* 40, 139-49.

129. Reardon, S. F. (2009) "Measures of ordinal segregation," in *Occupational and Residential Segregation*, Vol. 17 of *Research on Economic Inequality*, Y. Flückiger, S. F. Reardon and J. Silber (Eds.), Emerald, Bingley, UK, 129-55.

130. Reynal-Querol, M. (2002) "Ethnicity, political systems and civil wars", *Journal of Conflict Resolution* 46, 29-54.

131. Rodriguez, J. and R. Salas (2003) "Extended bipolarization and inequality measures", *Research on Economic Inequality* 9, 69-83.

132. Rosen, S. (1986) “Prizes and incentives in elimination turnaments”, *American Economic Review* 76, 701-714.

133. Rudin, W. (1987) *Principles of Mathematical Analysis*, McGraw Hill, London.

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

135. Sen, A.K. (1976) “Poverty: an ordinal approach to measurement”, *Econometrica* 44, 219 – 231.

136. Sen, A.K. (1985) *Commodities and Capabilities*, North-Holland, Amsterdam.

137. Sen, A.K. (1987) *Standard of Living*, Cambridge University Press, Cambridge.

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

139. Shorrocks, A.F. (1984) “Inequality decomposition by population subgroups” *Econometrica* 52, 1369-1385

140. Shorrocks, A.F. (1988) “Aggregation Issues in Inequality Measurement” in *Measurement in Economics* by W. Eichhorn (Ed.), Heidelberg: Physica –Verlag.

141. Silber, J., J. Deutsch and M. Hanoka (2007) “On the link between the concepts of kurtosis and bipolarization”, *Economics Bulletin* 4, 1-6.

142. Skaperdas, S. (1992) “Cooperation, conflict, and power in the absence of property rights”, *American Economic Review* 82, 720-39.

143. Skaperdas, S. (1996) “Contest success functions”, *Economic Theory* 7, 283-290.

144. Skaperdas, S. and B. Grofman (1995) “Modeling negative campaigning”, *American Political Science Review* 89, 79-161.

145. Skaperdas, S. and M. R. Garfinkel (2012) “Economic perspectives on piece and conflict” in *The economics of peace and conflict* by Skaperdas, S. and M.R. Garfinkel (Eds.), New York, Oxford University Press, 3-19.

146. Spolaore E. and R. Wacziarg (2009) "War and Relatedness," NBER Working Papers, 15095, National Bureau of Economic Research, Inc.

147. Subramanian, S. (2011) “Inequality measurement with subgroup-decomposability and level-sensitivity”, *Economics (e-journal) discussion paper*, 2011-7.

148. Snyder, J.M. (1989) “Election goals and the allocation of campaign resources”, *Econometrica* 57, 637-660.

149. Szymanzki, S. (2003) “The economic design of sporting contests”, *Journal of Economic Literature* 61, 1137-1187.

150. Takayama, N. (1979), “Poverty income inequality and their measures: Professor Sen's axiomatic approach reconsidered”, *Econometrica* 47, 747-759.

151. Theil, H. (1967) *Economics and Information Theory*, North-Holland: Amsterdam.

152. Theil, H. (1972) *Statistical Decomposition Analysis*, North-Holland: Amsterdam.

153. Thurow, L. (1984) “The disappearance of the middle class”, *New York Times*, 5 February 1984, section 3, p. 2.

154. Tullock, G. (1980) “Efficient rent seeking” in *Toward A Theory of the Rent-seeking Society* by Buchanan, J.M., R.D. Tollison and G. Tullock (Eds.), College Station: Texas A & M University Press 1980.

155. Vigdor, J.L. (2002) “Interpreting ethnic fragmentation effects”, *Economics Letters* 75, 271-276.

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

157. Weymark, J.A. (1981) “Generalized Gini inequality indices”, *Mathematical Social Sciences* 1, 409-430.

158. Wolfson, M.C. (1994) “When inequalities diverge”, *American Economic Review, Papers and Proceedings* 84, 353-358.

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

160. Yaari, M. (1987): “The dual theory of choice under risk”, *Econometrica* 55**,** 99-115.

161. Yaari, M. (1988): “A controversial proposal concerning inequality measurement”, *Journal of Economic Theory* 44, 381-397.

162. Yitzhaki, S. (1979): “Relative deprivation and the Gini coefficient”, *Quarterly Journal of Economics* 93, 321-32.

163. Zhang, X. and R. Kanbur (2001) “What differences do polarization measures make? An application to China", *Journalof Development Studies* 37, 85-98.

164. Zheng B (2005) "*Unit-consistent decomposable inequality measures: some extensions*" Working Paper 2005-2, Department of Economics, University of Colorado at Denver and HSC.

165. Zheng, B. (2007) "Unit-consistent decomposable inequality measures", *Economica* 74, 97-111.