Economic Studies in Inequality, Social Exclusion and Well-Being

Series Editor: Jacques Silber

Indraneel Dasgupta
Manipushpak Mitra *Editors*

Deprivation, Inequality and Polarization

Essays in Honour of Satya Ranjan Chakravarty

Springer

Economic Studies in Inequality, Social Exclusion and Well-Being

Series Editor

Jacques Silber, Ramat Gan, Israel

Traditionally, economists have identified well-being with market command over goods rather than with the “state” of a person. The books in this series go precisely beyond the traditional concepts of consumption, income or wealth and offer a broad, inclusive view of inequality and well-being. Topics include, but are not limited to: Capabilities and Inequalities, Discrimination and Segregation in the Labour Market, Equality of Opportunities, Globalization and Inequality, Human Development and the Quality of Life, Income and Social Mobility, Inequality and Development, Inequality and Happiness, Inequality and Malnutrition, Income and Social Mobility, Inequality in Consumption and Time Use, Inequalities in Health and Education, Multidimensional Inequality and Poverty Measurement, Polarization, Poverty among Children and Elderly People, Social Policy and the Welfare State, and Wealth Distribution.

More information about this series at http://www.springer.com/series/7140

Indraneel Dasgupta · Manipushpak Mitra
Editors

Deprivation, Inequality and Polarization

Essays in Honour of Satya Ranjan Chakravarty

Springer

Editors
Indraneel Dasgupta
Economic Research Unit
Indian Statistical Institute
Kolkata, West Bengal, India

Manipushpak Mitra
Economic Research Unit
Indian Statistical Institute
Kolkata, West Bengal, India

ISSN 2364-107X ISSN 2364-1088 (electronic)
Economic Studies in Inequality, Social Exclusion and Well-Being
ISBN 978-981-13-7943-7 ISBN 978-981-13-7944-4 (eBook)
https://doi.org/10.1007/978-981-13-7944-4

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

Satya Ranjan Chakravarty was educated at the Indian Statistical Institute (ISI), Kolkata. From there, he received a bachelor's degree in statistics (1976), a master's degree in statistics (1977), and a Ph.D. in economics (1981). He subsequently joined the faculty of ISI, where he taught till his retirement in 2019. His academic career at ISI was complemented by visiting positions at various universities around the globe, including the University of British Columbia, the University of Karlsruhe, Bar-Ilan University, Kagawa University, the Paris School of Economics, the Chinese University of Hong Kong, Bocconi University, Yokohama National University, and the University of International Business and Economics, Beijing. He received the prestigious Mahalanobis Memorial Award of the Indian Econometric Society in 1994. He currently serves on the editorial boards of the Journal of Economic Inequality, the Review of Income and Wealth and Social Choice and Welfare.

Issues relating to the measurement of poverty (or, more generally, deprivation) and inequality constitute abiding themes in Chakravarty's research oeuvre. In recent years, his work in these areas has expanded to include the measurement of polarization. His writings on these themes remain seminal and continue to exert a major influence on current research. Industrial organization and cooperative game theory are the other broad areas in which he has made important interventions. His research papers have come out in some of the most prestigious and high-impact journals in Economics, including *Econometrica, Journal of Economic Theory, International Economic Review, Journal of Development Economics, Social Choice and Welfare, Canadian Journal of Economics, Games and Economic Behavior, Economic Theory, Journal of Economic Inequality, Mathematical Social Sciences, Theory and Decision, Health Economics*, and *Review of Income and Wealth*. He has also authored a number of books, of which 'Inequality, Polarization and Poverty' (Springer, 2009), 'Inequality, Polarization and Conflict' (Springer, 2015), and 'Analyzing Multidimensional Well-Being' (Wiley, 2017) constitute essential reading for researchers in these areas.

Chakravarty produced some of his most seminal research work in the 1980s and 1990s. This was a time when basic infrastructural and resource constraints in Indian academic institutions could, and did, often prove daunting enough to permanently demoralize even the most brilliant and motivated young scholars. Far too many of his generation did get so demoralized and burnt out, to the extent of ceasing to produce meaningful research altogether, within a few years of their doctoral degrees. Many others relocated to institutions outside India in search of a better work environment. Chakravarty was among a handful of Indian economists of his generation who not only stayed back, but, somehow, managed to keep producing research that was globally recognized, despite their debilitating and often Kafkaesque work environment. The editors of this volume, who were graduate students in the late '90s, remember those times well enough to realize that Chakravarty's sheer dogged perseverance set a template for an intellectual and institutional commitment that they can idolize, but never even attempt to emulate. One of us became a colleague of Chakravarty in 2004 and the other in 2012. Along with that privilege, we also acquired an academic role model, a mentor, a great colleague, and a true elder. "Satya da" was always available to younger colleagues for encouragement, advice, and support. But equally, he taught us how it was never beneath one's dignity to treat even much younger colleagues as equals—to look to them in turn for advice and support. And most importantly, he taught us never to compromise our academic judgment, no matter how strong the institutional pressures or how persuasive the extra-academic reasons to do so. Putting together this volume has been an act of writing up an elaborate, yet entirely inadequate, personal thank you note from us for him.

This volume is a tribute to Satya Chakravarty, the man as well as his work, from his students, colleagues, research collaborators, and intellectual admirers with cognate research interests. It brings together eleven essays, written by altogether twenty researchers based in eight different countries—Canada, France, India, Israel, Italy, Luxembourg, the UK, and the USA. The essays cover various issues in the formal analysis of deprivation, inequality, and polarization. Many of the contributions explicitly build on ideas introduced in the writings of Chakravarty, thereby attesting to their continuing impact on the research frontier. Others are influenced broadly by his general research concerns and attest to their continuing relevance. All contribute to recurrent themes in Chakravarty's research oeuvre.

The first three papers fall in the area of deprivation analysis. Pattanaik and Xu study the measurement of multidimensional well-being and deprivation when every attribute is discretely and ordinally measurable. They axiomatically characterize a class of measures of social well-being as well as a class of social deprivation measures. Bossert and D'Ambrosio adapt a class of indices on the measurement of poverty over time to that of material deprivation and apply their framework to the analysis of material deprivation within EU countries. Gajdos, Weymark, and Zoli theorize on the comparative evaluation of social risk distributions, which are probability distributions over potential sets of fatalities associated with different public policy options.

The next six papers address different aspects of inequality analysis. Kanbur starts from the observation that often greater progressivity in income taxation requires a higher volume of gross redistributive flows across income levels, which may be costly to manage (administratively or politically). Progressivity is reduced in consequence. This problem is aggravated if redistribution across income levels implies redistribution across sociopolitically salient groups. He develops a theoretical framework in which these issues are captured and formally elucidated. Banerjee examines the problem of obtaining a crisp approximation of a fuzzy inequality ranking relation over a set of income distributions. Peled and Silber propose a new definition of pro-middle class growth and, using Israeli data for the period 1995–2011, argue that Israeli growth was indeed pro-middle class in the period under examination. Jayadev and Reddy introduce concepts and measures relating to inequality between identity groups. They proceed to discuss how these concepts can be deployed to interpret segregation, clustering, and polarization in societies. Tyagarupananda and Chattopadhyay offer a general measure of inequality that, in the limit, converges to the Theil index. They discuss the properties of this measure, along with an empirical illustration using Indian consumer expenditure survey data. Bhattacharjee and Sarkar use Pearson's correlation coefficient between the weight vector and the power profile to measure the similarity between weight and power in weighted majority voting games and use standard inequality measures to quantify the inequality in the weight vector as well as in the power profile. They find that it is possible to choose a value of the winning threshold which maximizes the similarity score and also minimizes the difference in the inequality scores of the weight vector and the power profile. They use their methodology to examine the voting games arising in the decision-making processes of the International Monetary Fund and the European Union.

The last two papers in the volume relate to the measurement of polarization. Maharaj and Chattopadhyay derive the general nature of a polarization map and extend the setup to allow for pooling of two populations. They also present an empirical illustration of polarization indices using Indian National Sample Survey data. Motiram and Vakulabharanam introduce a measure of "grayness," where by grayness they mean a combination of spatial integration based upon group identity and income. They consider some desirable properties of a measure of grayness and develop a measure that satisfies these properties. They also provide an illustration using data from the Indian city of Hyderabad and from selected American cities.

Collectively, the eleven essays in this volume constitute both a significant contribution to frontier research on distributive questions and a fitting tribute to the lifework of one of the foremost global authorities in the area of deprivation and inequality measurement.

Kolkata, India

Indraneel Dasgupta
Manipushpak Mitra

Contents

Part I Deprivation

Measuring Multidimensional Well-Being and Deprivation with Discrete Ordinal Data 3
Prasanta Pattanaik and Yongsheng Xu

Intertemporal Material Deprivation: A Proposal and an Application to EU Countries 15
Walter Bossert and Conchita D'Ambrosio

Feasible Shared Destiny Risk Distributions 37
Thibault Gajdos, John A. Weymark and Claudio Zoli

Part II Inequality

On the Volume of Redistribution: Across Income Levels and Across Groups 53
Ravi Kanbur

Fuzzy Inequality Ranking Relations and Their Crisp Approximations 67
Asis Kumar Banerjee

On Pro-Middle Class Growth 83
Osnat Peled and Jacques Silber

Inequalities and Identities 109
Arjun Jayadev and Sanjay G. Reddy

A Generalization of the Theil Measure of Inequality 147
Swami Tyagarupananda and Nachiketa Chattopadhyay

Correlation and Inequality in Weighted Majority Voting Games 161
Sanjay Bhattacherjee and Palash Sarkar

Part III Polarization

Reduced Form Polarization Index: Alternative Formulations and Extensions 195
Bhargav Maharaj and Nachiketa Chattopadhyay

On the Measurement of "Grayness" of Cities 211
Sripad Motiram and Vamsi Vakulabharanam

Editors and Contributors

About the Editors

Indraneel Dasgupta is Professor of Economics at the Indian Statistical Institute, Kolkata and research fellow at the Institute for the Study of Labor (IZA), Bonn. He received his PhD in Economics from the University of California, Riverside in 1997, and taught at Deakin University, Australia (lecturer); Nottingham University, UK (lecturer, senior lecturer and reader); Durham University, UK (professor); and the Centre for Studies in Social Sciences Calcutta, India (professor) prior to joining the Indian Statistical Institute. His research interests lie in the areas of microeconomic theory, development economics and public economics. His research papers have been published in journals such as *Economic Theory, European Economic Review, Journal of Economic Inequality, Journal of Economic Theory, Journal of Public Economics, Oxford Economic Papers, Public Choice* and *Social Choice and Welfare*. He currently serves on the editorial board of the *Journal of Development Studies*.

Manipushpak Mitra is Professor of Economics at the Indian Statistical Institute, Kolkata. He received his PhD in Economics from the Indian Statistical Institute, New Delhi, India in 2000. His research interests lie in the areas of game theory, industrial organization, mechanism design under incomplete information, and social choice theory. His research papers have been published in journals such as *Economica, Economic Theory, Economics Letters, Games and Economic Behavior, Journal of Institutional and Theoretical Economics* and *Theoretical Economics*. He has also written a textbook on cooperative game theory jointly with Satya Ranjan Chakravarty and Palash Sarkar. He received the Mahalanobis Memorial Medal in 2012.

Contributors

Asis Kumar Banerjee is Honorary Visiting Professor at the Institute of Development Studies Kolkata and Former Professor and Vice Chancellor, University of Calcutta, India.

Sanjay Bhattacherjee is Lecturer of Computer Science and Informatics at the University of East London, UK.

Walter Bossert is Professor of Economics at the University of Montreal, Canada.

Nachiketa Chattopadhyay is Associate Professor, Sampling and Official Statistics Unit, Indian Statistical Institute, Kolkata, India.

Conchita D'Ambrosio is Professor of Economics, FNR PEARL Chair, Université du Luxembourg, Luxembourg.

Thibault Gajdos is Research Director, CNRS and Laboratoire de Psychologie Cognitive, Aix-Marseille University, Marseille, France.

Arjun Jayadev is Professor of Economics at Azim Premji University, India.

Ravi Kanbur is T. H. Lee Professor of World Affairs, International Professor of Applied Economics and Management, and Professor of Economics at Cornell University, USA.

Bhargav Maharaj is Vice Principal, Ramakrishna Mission Vidyamandira, Belur Math, West Bengal, India.

Sripad Motiram is Associate Professor of Economics at the University of Massachusetts Boston, USA.

Prasanta Pattanaik is Emeritus Professor in the Department of Economics, University of California, Riverside, USA.

Osnat Peled works in the Research Department of the Bank of Israel.

Sanjay G. Reddy is Associate Professor of Economics at The New School for Social Research, USA.

Palash Sarkar is Professor, Applied Statistics Unit, Indian Statistical Institute, Kolkata, India.

Jacques Silber is Professor Emeritus in the Department of Economics, Bar-Ilan University, Israel.

Swami Tyagarupananda is Secretary, Ramakrishna Mission Ashrama, Malda, West Bengal, India.

Vamsi Vakulabharanam is Associate Professor of Economics at the University of Massachusetts Amherst, USA.

John A. Weymark is Gertrude Conaway Vanderbilt Chair in Social and Natural Sciences and Professor of Economics at Vanderbilt University, USA.

Yongsheng Xu is Professor of Economics at Georgia State University, USA.

Claudio Zoli is Professor of Economics at the University of Verona, Italy.

Part I
Deprivation

Measuring Multidimensional Well-Being and Deprivation with Discrete Ordinal Data

Prasanta Pattanaik and Yongsheng Xu

Abstract This paper studies the measurement of multidimensional well-being and deprivation when every attribute is discretely and ordinally measurable. In this context, we propose an axiom, Non-dominance, and use it together with some other standard axioms from the existing literature to characterize a class of measures of social well-being and also a class of social deprivation measures.

Keywords Multidimensional well-being and deprivation · Discrete and ordinal data · Axioms · Non-dominance · Measurement

JEL Classification Numbers D6 · D7 · I3

1 Introduction

One of the difficulties that economists often face in measuring, in a multidimensional framework, the well-being or deprivation of a group of individuals arises from the nature of the data relating to the individuals' achievements along the different dimensions. From an analytical point of view, we have perhaps the most tractable basis for measuring living standards and deprivation when we have cardinal information about the individuals' achievement in terms of the different attributes or dimensions involved. Sometimes, we do have such cardinal information for some attributes. Thus, not only is the consumption of, say, calories cardinally measurable, but the cardinal information regarding such consumption is also often not too difficult to gather. In contrast, the quality of housing is much more difficult to measure cardinally, even in principle, and often we have to fall back on ordinal measurement where the achieve-

P. Pattanaik
Department of Economics, University of California, Riverside, CA 92521, USA
e-mail: prasanta.pattanaik@ucr.edu

Y. Xu (✉)
Department of Economics, Georgia State University, Atlanta, GA 30302, USA
e-mail: yxu3@gsu.edu

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*, Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_1

ments may be given as good, poor, very poor, and "no house." Similar difficulties also arise in the case of health where one may have to resort to ordinal measurement in terms of achievement levels such as excellent health, reasonable health, and poor health. Thus, one can then think of several different types of informational basis for the measurement of multidimensional well-being or deprivation. First, we have the case where we have cardinal information for individual achievement in terms of every attribute. Second, there is the case where we have only ordinal data for every attribute. Finally, we have the "mixed" case where there is cardinal information for some of the attributes and only ordinal information for the rest of the attributes. A large number of contributions have investigated the problem of measuring well-being and deprivation of societies when cardinal information is available for all attributes (see, among others, Atkinson (2003), Bourguignon and Chakravarty (2003, 2009), Tsui (2002), Alkire et al. (2015)). The second case where the measurement of every attribute is ordinal and discrete has been studied by several recent contributions which assume that the ordinal information for every attribute is "binary" in nature so that, for each attribute, the only two achievement levels are "satisfactory" (1) and "unsatisfactory" (0); Alkire and Foster (2011), Bossert et al. (2013), and Dhongde et al. (2016) are a few examples of such contributions. We are not, however, aware of any contributions which, while remaining within the framework of ordinal data, go beyond the Spartan framework of the binary 1–0 measurement to permit more than two achievement levels for attributes. Finally, we are not aware of contributions dealing with the mixed case where the information for some attributes is cardinal but the information for the rest of the attributes is ordinal.

The purpose of this paper is to develop measures of living standards and deprivation when the measurement of achievements is ordinal and discrete for all attributes but is not constrained to be binary in nature so that more than two achievement levels are permitted for each attribute. The plan of the paper is as follows. In Sect. 2, we introduce our basic notation and definitions for the measurement of the well-being of a society or a group of individuals. In Sect. 3, we introduce several properties and prove a preliminary result. Using the result proved in Sect. 3 and a further property named Non-dominance, Sect. 4 presents a measure of a society's well-being. Section 5 deals with the measurement of social deprivation. We conclude in Sect. 6.

2 Some Basic Notation and Definitions

Let there be m attributes: $f_1, \ldots, f_m$, and let $M = \{1, \ldots, m\}$. Suppose the measurement of each attribute is ordinal and discrete. For each attribute f_j, let V_j denote the set of discrete and ordinal values which f_j can take. For instance, if attribute f_j is health, then V_j may be {excellent health, good health, poor health}. We assume that, for every $j \in M$, we have an antisymmetric ordering $[\geq_j]$ ("at least as high as, in terms of f_j") over V_j, with $[>_j]$ denoting the asymmetric factor of $[\geq_j]$: a higher value for f_j indicates a higher achievement in terms of the attribute f_j. For each attribute f_j, we assume that V_j is a finite set.

As a convention, we shall assume that, for each $j \in M$, the lowest value in V_j is 0 and the highest value in V_j is v_j^*. Let $V = V_1 \times \ldots \times V_m$, and let $v^* = \left(v_1^*, \ldots, v_m^*\right)$. Then, v^* is the greatest element in V and is to be interpreted as the achievement vector in which the achievement in every dimension is at the highest level. For any two vectors $v, v' \in V$, we shall write $v \geq v'$ to mean $[\left(v_j\left[\geq_j\right]v_j' \text{ for all } j \in M\right]$.

Let the group of individuals under consideration have n individuals; for convenience we shall call the group the society and denote it by $N = \{1, \ldots, n\}$. For each $i \in N$, let $a_i = \left(a_{i1}, \ldots, a_{ij}, \ldots, a_{im}\right)$ be individual i's achievement vector. Let

$$\mathcal{A} = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} : a_i \in V, i \in N \right\}$$

be the set of all $n \times m$ matrices such that, for all $i \in N$, $a_i \in V$ is individual i's achievement vector. Each $A \in \mathcal{A}$ is interpreted as the society's achievement matrix.

Let $\mathscr{o}$ denote the zero achievement vector and let $\mathbb{O}$ denote the zero achievement matrix. For all $v \in V$, all $i \in N$, and all $A \in \mathcal{A}$, $(v; A_{-i})$ will denote the matrix $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{A}$ such that $b_i = v$ and, for all $i' \in N/\{i\}$, $b_{i'} = a_{i'}$; thus, $(v; A_{-i})$ is the achievement matrix derived by replacing a_i by v in A, other things remaining the same.

A *social well-being measure* is a function from $\mathcal{A}$ to [0, 1] with the interpretation that, for all $A, B \in \mathcal{A}$, $g(A) \geq g(B)$ indicates that the society's well-being level under A is at least as high as the society's well-being level under B, $g(A) > g(B)$ indicates that the society's well-being level under A is higher than the society's well-being level under B, and $g(A) = g(B)$ indicates that the society's well-being level under A is the same as the society's well-being level under B.

3 A Preliminary Measure of Social Well-Being

We consider the following axioms to be imposed on a social well-being measure, g.

Normalization: For all $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathcal{A}$, if $A = \mathbb{O}$, then $g(A) = 0$, and if $a_i = v^*$ for all $i \in N$, then $g(A) = 1$.

Anonymity: Let σ be a bijection from N to N. Then, for all $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{A}$, if $[a_i = b_{\sigma(i)}$ for all $i \in N]$, then $g(A) = g(B)$.

Monotonicity: For all $A, B \in \mathcal{A}$, if $[a_i \geq b_i$ for all $i \in N]$ and $A \neq B$, then $g(A) > g(B)$.

Independence: For all $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, $A' = \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix}$, $B' = \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix} \in \mathcal{A}$, and for all $i' \in N$, if [(for all $i \in N\backslash\{i'\} : a_i = b_i$ and $a'_i = b'_i$), and ($a_{i'} = a'_{i'}$ and $b_{i'} = b'_{i'}$)], then $g(A) - g(B) = g(A') - g(B')$.

Normalization, Anonymity, and Monotonicity are fairly standard axioms used in the literature on measurement of multidimensional well-being and deprivation. Independence is another straightforward axiom used in the literature and stipulates that, starting with a given achievement matrix, when the achievement vector of one individual changes while all other individuals' achievement vectors remain unchanged, the resulting change in the society's well-being is independent of the achievement vectors of those other individuals. A variant of Independence was proposed in a different context by Chakraborty et al. (2008), and was used in Dhongde et al. (2016) in the context of measuring multidimensional deprivation with binary data.

The implication of the above axioms is summarized in the following proposition, Proposition 1.

Proposition 1 *Let g be a social well-being measure. Then, g satisfies Normalization, Anonymity, Monotonicity, and Independence if and only if:*

for some increasing function $\rho : V \to [0, 1]$ *with* $\rho(\sigma) = 0, \rho(v_1^*, \ldots, v_m^*) = 1$, *we have* $g(A) = \frac{1}{n}\sum_{i=1}^{n} \rho(a_i)$ *for all* $A \in \mathcal{A}$.

Proof Let g be a well-being measure of the society satisfying Normalization, Anonymity, Monotonicity, and Independence. Let $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ be any given matrix in $\mathcal{A}$.

For every $i \in N$, let $A^i = (a_i; \mathbb{O}_{-i})$. Consider $B \equiv (\sigma; A_{-i})$. Consider the achievement matrices, A, B, A^1 and $\mathbb{O}$. Then, by Independence, we have

$$g(A) - g(B) = g\left(A^1\right) - g(\mathbb{O}). \tag{1}$$

By Normalization, $g(\mathbb{O}) = 0$. We then obtain

$$g(A) - g(B) = g\left(A^1\right) = g(a_1; \mathbb{O}_{-1}). \tag{2}$$

Consider $C = (o; B_{-2})$, and the achievement matrices B, C, A^2 and $\mathbb{O}$. By Independence and Normalization, we have

$$g(B) - g(C) = g\left(A^2\right) - g(\mathbb{O}) = g\left(A^2\right) = g(a_2; \mathbb{O}_{-2}). \tag{3}$$

From (2) and (3), we then have

$$g(A) - g(C) = g(a_1; \mathbb{O}_{-1}) + g(a_2; \mathbb{O}_{-2}). \tag{4}$$

By repeating the above procedures with C and beyond and from Independence and Normalization, we can obtain

$$g(A) - g(a_n; \mathbb{O}_{-n}) = g(a_1; \mathbb{O}_{-1}) + \cdots + g\left(a_{n-1}; \mathbb{O}_{-(n-1)}\right). \tag{5}$$

Then

$$g(A) = g(a_1; \mathbb{O}_{-1}) + \cdots + g\left(a_{n-1}; \mathbb{O}_{-(n-1)}\right) + g(a_n; \mathbb{O}_{-n}). \tag{6}$$

For each $i \in N$, let $\rho_i : V \to [0, 1]$ be such that, for all $v \in V$, $\rho_i(v) = g(v; \mathbb{O}_{-i})/n$. Clearly, for each $i \in N$, ρ_i is a real-valued function. By Monotonicity, for each $i \in N$, ρ_i is increasing. By Anonymity, $\rho_1 = \rho_2 = \cdots = \rho_n$. Denote each ρ_i by ρ. By Normalization, $\rho(v^*) = 1$. Then,

$$g(A) = \frac{1}{n}[\rho(a_1) + \cdots + \rho(a_n)]. \tag{7}$$

On the other hand, if, for some increasing function $\rho : V \to [0, 1]$ with $\rho(\mathbb{O}) = 0$, $\rho\left(v_1^*, \ldots, v_m^*\right) = 1$, $g(A) = \frac{1}{n}\sum_{i=1}^{n} \rho(a_i)$ for all $A \in \mathcal{A}$, then it can be verified that g satisfies Normalization, Anonymity, Monotonicity, and Independence. ■

For each $i \in N$ and each achievement vector a_i of i, $\rho(a_i)$ figuring in Proposition 1 can be regarded as a measure of i's well-being when i's achievement vector is a_i.

4 A Specific Class of Measures of Social Well-Being

Using our preliminary measures of social well-being developed in Sect. 3, in this section, we study a class of measures of social well-being, which have some interesting and attractive features. We note that the measurement of each attribute f_j is ordinal and discrete. Because of its discrete nature, we shall introduce a further property, Non-dominance, of the function g, which restricts the form of the individual well-being measure ρ.

Consider any $\left(a^1, \ldots, a^k\right)$, $(b^1, \ldots, b^k) \in V$, where $k > 1$.[1] We say that $\left(a^1, \ldots, a^k\right)$ and $\left(b^1, \ldots, b^k\right)$ are *achievement-equivalent* iff, for every $j \in M$, there exists a permutation ψ_j on the set $\{1, 2, \ldots, k\}$ such that $a_j^t = b_j^{\psi_j(t)}$ for all $t \in \{1, 2, \ldots, k\}$. Further, we say that $(a^1, \ldots, a^k)$ *dominates* $\left(b^1, \ldots, b^k\right)$ iff for all $i \in N$, $[g(a^t; \mathbb{O}_{-i}) \geq g\left(b^t; \mathbb{O}_{-i}\right)$ for all $t \in \{1, \ldots, k\}]$ and $[g(a^t; \mathbb{O}_{-i}) > g\left(b^t; \mathbb{O}_{-i}\right)$ for some $t \in \{1, \ldots, k\}]$.

Non-dominance. For all $k > 1$ and all $\left(a^1, \ldots, a^k\right)$, $\left(b^1, \ldots, b^k\right) \in V^k$, if $(a^1, \ldots, a^k)$ and $\left(b^1, \ldots, b^k\right)$ are achievement-equivalent, then neither $(a^1, \ldots, a^k)$ dominates $\left(b^1, \ldots, b^k\right)$ nor $\left(b^1, \ldots, b^k\right)$ dominates $(a^1, \ldots, a^k)$.

To see the intuition of Non-dominance, consider any $\left(a^1, \ldots, a^k\right)$, $\left(b^1, \ldots, b^k\right) \in V^k$ $(k > 1)$ such that $\left(a^1, \ldots, a^k\right)$ and $\left(b^1, \ldots, b^k\right)$ are achievement-equivalent. Since $\left(a^1, \ldots, a^k\right)$ and $\left(b^1, \ldots, b^k\right)$ are achievement-equivalent, it is easy to check that, for every attribute f_j and every value v_j which attribute f_j can take, the number of vectors in $(a^1, \ldots, a^k)$, in which v_j figures as the value of f_j, is the same as the number of vectors in $\left(b^1, \ldots, b^k\right)$, in which v_j figures as the value of f_j. Intuitively, the "aggregate data" regarding dimensional achievements are the same for $(a^1, \ldots, a^k)$ and $(b^1, \ldots, b^k)$. Given this, for all $i \in N$, Non-dominance rules out the possibility of having $[g(a^t; \mathbb{O}_{-i}) \geq g\left(b^t; \mathbb{O}_{-i}\right)$ for all $t \in \{1, \ldots, k\}]$ and $[g(a^t; \mathbb{O}_{-i}) > g\left(b^t; \mathbb{O}_{-i}\right)$ for some $t \in \{1, \ldots, k\}]$.

With the help of Non-dominance, we can show that the individual well-being function ρ figuring in Proposition 1 has a particular structure. This result is presented in Proposition 2 below.

Proposition 2 *Let g be a well-being measure of the society. Then, g satisfies Normalization, Anonymity, Monotonicity, Independence, and Non-dominance if and only if*

for some m positive constants, $w_1, \ldots, w_m$, with $w_1 + \ldots + w_m = 1$, there exists, for each $j \in M$, an increasing function $\varphi_j : V_j \to [0, w_j]$, with $\varphi_j(0) = 0$ and $\varphi_j\left(v_j^*\right) = w_j$, such that, for some increasing function $\sigma : [0, 1] \to [0, 1]$, with $\sigma(0) = 0$ and $\sigma\left(\sum_{j=1}^{m} w_j\right) = 1$, we have

$$g(A) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \sigma\left(\sum\nolimits_{j=1}^{m} \varphi_j(a_{ij})\right) \text{for all } A \in \mathcal{A}. \tag{8}$$

Proof Let g be a well-being measure of the society satisfying Normalization, Anonymity, Monotonicity, Independence, and Non-dominance. From Proposition 1, there exists an increasing function $\rho : V \to [0, 1]$ with $\rho(\sigma) = 0, \rho(v_1^*, \ldots, v_m^*) = 1$ such that, for all $A \in \mathcal{A}$,

[1]Note that the vectors $a^1, \ldots, a^k$ here are some achievement vectors each of which is possible for an individual and do not refer to the achievement vectors of any specific individuals in our society, and similarly for $b^1, \ldots, b^k$. In particular, k does not bear any relation to n, the number of individuals in our society.

$$g(A) = \frac{1}{n}\sum_{i=1}^{n}\rho(a_i).$$

In what follows, we shall show that the following holds:

for some m positive constants, $w_1, \ldots, w_m$, with $w_1 + \ldots + w_m = 1$, there exists, for each $j \in M$, an increasing function $\varphi_j : V_j \to [0, w_j]$, with $\varphi_j(0) = 0$ and $\varphi_j\left(v_j^*\right) = w_j$, such that for all $v, v' \in V$,

$$\rho(v) \geq \rho\left(v'\right) \Leftrightarrow \sum\nolimits_{j=1}^{m}\varphi_j(v_j) \geq \sum\nolimits_{j=1}^{m}\varphi_j(v'_j). \tag{9}$$

Note that $\rho(v_1^*, \ldots, v_m^*) = 1 > 0 = \rho(\sigma)$. Let $k > 1$ and let $\{a^1, \ldots, a^k\}$ and $\{b^1, \ldots, b^k\}$ be two achievement-equivalent collections of individual achievement vectors. We shall verify that ρ satisfies the following:

$$\text{if } \rho\left(a^p\right) \geq \rho\left(b^p\right) \text{ for all } p = 1, \ldots, k-1, \text{ then } \rho\left(b^k\right) \geq \rho\left(a^k\right). \tag{10}$$

Suppose $\rho(a^p) \geq \rho(b^p)$ for all $p = 1, \ldots, k-1$, but not $[\rho\left(b^k\right) \geq \rho\left(a^k\right)]$. Then, we must have $\rho(a^p) \geq \rho(b^p)$ for all $p = 1, \ldots, k-1$, and $\rho\left(b^k\right) < \rho\left(a^k\right)$. Let $r \in N$. Consider $\left(a^1; \mathbb{O}_{-r}\right), \ldots, g\left(a^k; \mathbb{O}_{-r}\right), g\left(b^1; \mathbb{O}_{-r}\right), \ldots, g\left(b^k; \mathbb{O}_{-r}\right)$. Then,

$$g\left(a^p; \mathbb{O}_{-r}\right) = \frac{1}{n}\rho\left(a^p\right) \geq g\left(b^p; \mathbb{O}_{-r}\right) = \frac{1}{n}\rho\left(b^p\right),\ p = 1, \ldots, k-1$$

and

$$g\left(a^k; \mathbb{O}_{-r}\right) = \frac{1}{n}\rho\left(a^k\right) > g\left(b^k; \mathbb{O}_{-r}\right) = \frac{1}{n}\rho\left(b^k\right).$$

This implies that $(a^1, \ldots, a^k)$ dominates $(b^1, \ldots, b^k)$, which contradicts Non-dominance. Therefore, (10) holds for all achievement-equivalent $(a^1, \ldots, a^k)$, $(b^1, \ldots, b^k) \in V^k$, where $k > 1$. By Theorem 4.1 of Fishburn (1970), for each $j \in M$, there exists a function $h_j : V \to [0, \infty)$ such that

$$\text{for all } v, v' \in V,\ \rho(v) \geq \rho\left(v'\right) \Leftrightarrow \sum_{j=1}^{m} h_j(v_j) \geq \sum_{j=1}^{m} h_j(v'_j). \tag{11}$$

Since ρ is increasing, each h_j $(j \in M)$ is increasing as well. Then, for each $j \in M$, $h_j\left(v_j^*\right) > h_j(0) \geq 0$. For each $j \in M$, let $w_j = \frac{h_j\left(v_j^*\right) - h_j(0)}{\sum_{k \in M}\left(h_k\left(v_k^*\right) - h_k(0)\right)}$ and define the function φ_j as follows: for all $v_j \in V_j$,

$$\varphi_j\left(v_j\right) = \frac{h_j\left(v_j\right) - h_j(0)}{\sum_{k \in M}\left(h_k\left(v_k^*\right) - h_k(0)\right)}.$$

Then, for each $j \in M$, $w_j > 0$, φ_j is increasing with $\varphi_j(0) = 0$ and $\varphi_j\left(v_j^*\right) = w_j$, and $\varphi_j : V_j \to \left[0, w_j\right]$. Note also that $w_1 + \cdots + w_m = 1$, and for all $v, v' \in V$,

$$\rho(v) \geq \rho\left(v'\right) \Leftrightarrow \sum_{j=1}^{m} \varphi_j(v_j) \geq \sum_{j=1}^{m} \varphi_j(v'_j).$$

Thus, we have established (9).

From (9), there exists an increasing function $\sigma : [0, 1] \to [0, 1]$ such that, for all $v \in V$, $\rho(v) = \sigma\left(\sum_{j=1}^{m} \varphi_j(v_j)\right)$. By Normalization, it follows that $\sigma(0) = 0$ and $\sigma\left(\sum_{j=1}^{m} w_j\right) = 1$.

To complete the proof of Proposition 2, we note that if the statement (8) holds, then it can be verified that the measure g specified in the statement satisfies Normalization, Anonymity, Monotonicity, Independence, and Non-dominance. ■

The expression $\sum_{j=1}^{m} \varphi_j(a_{ij})$ figuring in (8) can be interpreted as the "nominal overall achievement" of individual i; $\sigma\left(\sum_{j=1}^{m} \varphi_j(a_{ij})\right)$ then has the obvious interpretation as i's well-being given as a function of i's overall nominal achievement. Therefore, Non-dominance restricts the form of the individual well-being measure ρ to be a positive transformation of the individual's nominal overall achievement. It may be noted that this specific form for measures of an individual's well-being in a multidimensional framework has been studied by several authors including Bossert et al. (2013) and Dhongde et al. (2016). In their studies, they consider very different and more direct axioms to be imposed on a measure of an individual's well-being, and derive their results. Viewed this way, Non-dominance is a more primitive property imposed on a well-being measure of the society.

Remark 1 It may be noted that, in Proposition 2, an individual's well-being measure, ρ, has an additive representable structure. It is well known that, in a framework such as ours, where the measurement of each attribute is ordinal and discrete, the standard independence properties (on an individual's well-being measure) are not, in general, sufficient to obtain an additive representable structure of a measure of an individual's well-being. In the presence of Normalization, Anonymity, Monotonicity, and Independence, however, our Non-dominance implies a stronger independence property (see (10) in the proof of Proposition 2) of an individual's well-being measure that proves to be necessary and sufficient for an additive representable structure of an individual's well-being measure.

5 From Achievement to Deprivation

Given an achievement matrix $A \in \mathcal{A}$, in the previous section, we showed that the well-being measure g of the society satisfying Normalization, Anonymity, Monotonicity, Independence, and Non-dominance takes the following form:

$$g(A) = \frac{1}{n}\sum_{i=1}^{n}\sigma\left(\sum_{j=1}^{m}\varphi_j(a_{ij})\right).$$

As discussed there, for any individual i with an achievement vector a_i, $\sum_{j=1}^{m}\varphi_j(a_{ij})$ can be interpreted as i's nominal overall achievement, and $\sigma\left(\sum_{j=1}^{m}\varphi_j(a_{ij})\right)$ can be interpreted as i's well-being, or "real overall achievement."

Let $\underline{w} \in (0, 1]$ be the given and fixed level of nominal overall individual achievement such that if individual i's nominal overall achievement, $\sum_{j=1}^{m}\varphi_j(a_{ij})$, falls below it, then individual i is said to be deprived, and if $\sum_{j=1}^{m}\varphi_j(a_{ij}) \geq \underline{w}$, then individual i is said to be non-deprived.[2] This method of classifying deprived and non-deprived individuals has been discussed in the literature on measurements of multidimensional well-being and deprivation, see, among others, Bourguignon and Chakravarty (2003), Pattanaik and Xu (2018), and Tsui (2002). Given our discussions on measures of social well-being in the previous two sections, our approach seems natural and becomes even more so in our current context of discrete and ordinal data.

For each individual $i \in N$ and any achievement matrix $A \in \mathcal{A}$, let

$$d_i(A) = \begin{cases} 0 & \text{if } \sum_{j=1}^{m}\varphi_j(a_{ij}) \geq \underline{w} \\ \frac{\underline{w}-\sum_{j=1}^{m}\varphi_j(a_{ij})}{\underline{w}} & \text{if } \sum_{j=1}^{m}\varphi_j(a_{ij}) < \underline{w} \end{cases}$$

$d_i(A)$ can be interpreted as the normalized overall nominal deprivation of individual i given the achievement matrix A. It may be noted that, since each attribute can take discrete values, the possible values that normalized shortfalls of an individual are discrete as well. Let $\mathcal{U} = \{d_i(A) : A \in \mathcal{A}\}$.

For every achievement matrix $A \in \mathcal{A}$, let $d(A) = (d_1(A), \ldots, d_i(A), \ldots, d_n(A)) \in \mathcal{U}^n$. We will refer $d(A)$ as a profile of individual normalized deprivations associated with the achievement matrix $A \in \mathcal{A}$. Let

$$\mathcal{D} = \{d(A) : A \in \mathcal{A}\}.$$

A deprivation measure of the society is a function h from $\mathcal{D}$ to $[0, 1]$ with the interpretation that, for all $A, B \in \mathcal{A}$, $h(d(A)) \geq h(d(B))$ indicates that the society's deprivation under A is at least as high as the society's deprivation under B, $h(d(A)) > h(d(B))$ indicates that the society's deprivation under A is higher than the society's

[2]It may be noted that we could have used individual i's well-being (or "real overall achievement"), $\sigma\left(\sum_{j=1}^{m}\varphi_j(a_{ij})\right)$, and an appropriate well-being benchmark to classify if individual i is deprived or non-deprived. Given that an individual's well-being is a positive transformation of her nominal overall achievement $\sum_{j=1}^{m}\varphi_j(a_{ij})$, the two classificatory methods are equivalent.

deprivation under B, and $h(d(A)) = h(d(B))$ indicates that the society's deprivation under A is the same as the society's deprivation under B.

We consider the following axioms to be imposed on a deprivation measure of the society. Each of these axioms is straightforward, and they are standard axioms used in the literature.

D-Normalization: For all $A \in \mathcal{A}$, if $d(A)$ is the zero vector, then $h(d(A)) = 0$, and if $d(A)$ is the 1 vector, then $h(d(A)) = 1$.

D-Anonymity: Let σ be a bijection from N to N. Then, for all $A, B \in \mathcal{A}$, if $[d_i(A) = d_{\sigma(i)}(B)$ for all $i \in N]$, then $h(d(A)) = h(d(B))$.

D-Monotonicity: For all $A, B \in \mathcal{A}$, if $[d_i(A) \geq d_i(B)$ for all $i \in N$ and $d(A) \neq d(B)]$, then $h(d(A)) > h(d(B))$.

D- independence: For all $A = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, A' = \begin{pmatrix} a'_1 \\ \vdots \\ a'_n \end{pmatrix}, B' = \begin{pmatrix} b'_1 \\ \vdots \\ b'_n \end{pmatrix} \in$ $\mathcal{A}$ and all $i' \in N$, if [for all $i \in N \backslash \{i'\}, d_i(A) = d_i(B)$ and $d_i(A') = d_i(B')$], and $[d_{i'}(A) = d_{i'}(A')$ and $d_{i'}(B) = d_{i'}(B')]$, then $h(A) - h(B) = h(A') - h(B')$.
With the help of the above axioms, we can prove the following result.

Proposition 3 *Let h be a deprivation measure of the society. Then, h satisfies D-Normalization, D-Anonymity, D-Monotonicity, and D-Independence if and only if there exists an increasing function $\tau : \mathcal{U} \to [0, 1]$, with $\tau(0) = 0$, $\tau(1) = 1$, such that, for all $A \in \mathcal{A}$, we have*

$$h(d(A)) = \frac{1}{n} \sum\nolimits_{i=1}^{n} \tau(d_i(A)). \tag{12}$$

Proof It can be easily checked that, if a deprivation measure h is given by (12), then it satisfies D-Normalization, D-Anonymity, D-Monotonicity, and D-Independence. We omit the proof of the fact that, if a deprivation measure h satisfies D-Normalization, D-Anonymity, D-Monotonicity, and D-Independence, then (12) holds since that proof is very similar to the proof that we gave earlier (see the proof of Proposition 1) to show that, if g satisfies Normalization, Anonymity, Monotonicity, and Independence, then, for some increasing function $\rho : V \to [0, 1]$ with $\rho(\sigma) = 0, \rho(v_1^*, \ldots, v_m^*) = 1$, we have $g(A) = \frac{1}{n} \sum_{i=1}^{n} \rho(a_i)$ for all $A \in \mathcal{A}$. ∎

$\tau(d_i(A))$, which figures in the social deprivation measure in Proposition 3, can be interpreted as a measure of the "real" deprivation of individual i. It has a feature which does not conform to the intuition which underlies many contributions in the literature. That intuition has two components. The first component is the idea that for every attribute f_j, there is a benchmark level, z^j such that an individual has any deprivation along the dimension f_j if and only if her achievement in terms of f_j

falls below z^j. The second component is that, for every attribute f_j, an increase in an individual's achievement in terms of f_j can never reduce her overall deprivation if, to start with, the individual is not deprived in terms of f_j, i.e., in measuring an individual's deprivation, there cannot be any tradeoff between "overachievement" (as compared to the relevant benchmark) in one dimension and "underachievement" (as compared to the appropriate benchmark) in some other dimension. The individual deprivation measure $\tau(d_i(A))$ cannot accommodate this second component: under this measure of individual deprivation, if, to start with, $d_i(A)$ is positive, then an increase in the achievement of individual i along any dimension f_j, i's other achievements remaining the same, will always reduce i's deprivation irrespective of whether i's initial achievement along f_j exceeds, falls short of, or equals the benchmark, z^j, for f_j. It is not, however, clear to us that, from an intuitive point of view, this feature constitutes a flaw of the individual deprivation measure $\tau(d_i(A))$ (see Pattanaik and Xu (2018) for further discussion of the basic issue involved here).

6 Conclusion

In this paper, we have studied measures of social well-being and social deprivation in a multidimensional framework in which the levels of an individual's achievement in terms of each attribute are discrete and ordinal. Though there are a fairly large number of studies on measures of social well-being and social deprivation in a multidimensional framework when every attribute is cardinally measurable and a few studies when every attribute is binarily measurable, there seem to be no studies for such measures when attributes are ordinally and discretely measurable. This paper has attempted to fill this gap. Our main contributions are to introduce the Non-dominance axiom in this context and axiomatically derive a class of measures of social well-being and a closely linked class of measures of social deprivation. Another important feature of our study is the way in which we use the measure of individual well-being to distinguish between deprived and non-deprived individuals. This approach has brought together the analytical framework for measuring social well-being and that for measuring social deprivation.

As we note in the Introduction, there seem to be no contributions dealing with the mixed case where some attributes are cardinally measurable while other attributes are only ordinally measurable. It would be interesting to study measures of social well-being and social deprivation for such mixed cases. We leave this for future research.

Acknowledgements We are grateful to a referee for helpful comments on an earlier version of the paper.

References

Alkire S, Foster J (2011) Counting and multidimensional poverty measurement. J Public Econ 95:476–487

Alkire S, Foster J, Seth S, Santos M, Roche J, Ballon P (2015) Multidimensional poverty measurement and analysis. Oxford University Press, Oxford

Atkinson AB (2003) Multidimensional deprivation: contrasting social welfare and counting approaches. J Econ Inequal 1:51–65

Bossert W, Chakravarty SR, D'Ambrosio C (2013) Multidimensional poverty and material deprivation with discrete data. Rev Income Wealth 59:29–43

Bourguignon F, Chakravarty SR (2003) Measurement of multidimensional poverty. J Econ Inequal 1:25–49

Bourguignon F, Chakravarty SR (2009) Multidimensional poverty orderings: theory and applications. In: Basu K, Kanbur R (eds) Arguments for a better world, vol 1. Oxford University Press, Oxford, pp 337–361

Chakraborty A, Pattanaik PK, Xu Y (2008) On the mean of squared deprivation gaps. Econ Theory 34:181–187

Dhongde S, Y. Li Y, Pattanaik PK, Xu Y (2016) Binary data, hierarchy of attributes, and multidimensional deprivation. J Econ Inequal 14: 363–378

Fishburn PC (1970) Utility theory for decision making. John Wiley & Sons Inc, New York

Pattanaik PK, Xu Y (2018) On measuring multidimensional deprivation, mimeograph. J Econ Lit 56: 657–672

Tsui KY (2002) Multidimensional poverty indices. Soc Choice Welfare 19:69–93

Intertemporal Material Deprivation: A Proposal and an Application to EU Countries

Walter Bossert and Conchita D'Ambrosio

Abstract This paper analyzes the effects of the inclusion of the past experiences in measuring current material deprivation. The method followed generalizes the proposal of Bossert et al. (Intertemporal material deprivation. Routledge, London, pp. 128–142, 2014) by adapting the class of indices on the measurement of poverty over time of Dutta et al. (Soc Choice Welf 41:741–762, 2013). An application to the analysis of material deprivation within EU countries is then provided. Following the path of material deprivation experienced by each individual over time yields a picture which differs from that in the annual results. Since the measurement of material deprivation is used by the EU member states and the European Commission to monitor national and EU progress in the fight against poverty and social exclusion, the results suggest that time cannot be neglected. Countries should not only be compared based on their year-by-year results, but additional information is gained by following individuals over time and producing an aggregate measure once dynamic considerations are taken into consideration.

Keywords Material deprivation · Intertemporal social index numbers · Persistent deprivation

Journal of Economic Literature Classification No D63

This is a substantially revised version of a previously unpublished Discussion Paper (D'Ambrosio, 2013). We thank a referee for helpful comments. Financial support from the Fonds National de la Recherche Luxembourg and the Fonds de Recherche sur la Société et la Culture of Québec is gratefully acknowledged.

W. Bossert
University of Montreal, Montreal, Canada
e-mail: walter.bossert@videotron.ca

C. D'Ambrosio (✉)
Université du Luxembourg, Luxembourg City, Luxembourg
e-mail: conchita.dambrosio@uni.lu

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*,
Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_2

1 Introduction

Material deprivation has been a key indicator of individual well-being within the EU since 2010. In June of that year, the European Council adopted as part of the Europe 2020 Strategy the aim to lift at least 20 million people in the EU from the risk of poverty and exclusion by 2020. Material deprivation is one of the EU Social Inclusion indicators which are used to monitor national and EU-wide progress in the achievement of the 2020 target. The Employment, Social Policy, Health and Consumer Affairs (EPSCO) EU Council of Ministers required improved measures of material deprivation. This paper contributes to this effort by analyzing the effects of the inclusion of past experiences in the measurement of current deprivation. The method followed generalizes the proposal of Bossert et al. (2014). We extend the class of intertemporal poverty measures of Dutta et al. (2013) to the measurement of material deprivation. The indices are applied to analyze material deprivation within EU countries. If we follow the path of material deprivation experienced by each individual over time, we obtain a picture which differs from that in the annual results (even though the three intertemporal indices do rank the countries very similarly). Although the 4-year panel employed here is relatively short, our analysis serves as a useful illustration for the application of the measures that we propose.

The theoretical work on which material deprivation measures are grounded comes from the literature on the measurement of poverty. From a theoretical point of view, material deprivation is multidimensional poverty. The difference between the two concepts is due to the aspects of well-being which are included in the empirical analysis. In particular, a multidimensional poverty measure takes into consideration all the dimensions of well-being that may be of relevance (including nonmaterial attributes, such as health status and political participation). In contrast, an index of material deprivation restricts attention to the functioning failures with respect to material living conditions. Chakravarty and Chattopadhyay (2018) compare the two approaches and provide an excellent survey of the proposed measures. According to EU policy, indices of material deprivation are to be combined with income-based poverty measures and indicators of low employment. See also Guio (2018) for a discussion of material deprivation and multidimensional poverty from an empirical perspective.

The axiomatic literature has proposed many indices of multidimensional poverty and explored their underlying properties; see, for example, Chakravarty et al. (1998), Tsui (2002), Bourguignon and Chakravarty (2003), Diez et al. (2008), Alkire and Foster (2011), and Bossert et al. (2013).

The intertemporal aspect of multidimensional poverty has received a modest amount of attention up to now—most of the work in this area has been atemporal. At the same time, many of the contributions in the field of unidimensional poverty have shown that chronic poverty and persistent periods of poverty are worse, in a number of ways, for individuals than are sporadic episodes. For surveys of this literature, see, among others, Rodgers and Rodgers (1993) and Jenkins (2000). These considerations have provided the impetus for some recent theoretical contributions on measuring income poverty over time, such as Calvo and Dercon (2009), Foster (2009), Hojman and Kast (2009), Hoy and Zheng (2011), Bossert et al. (2012), and Dutta et al. (2013). See also Hoy et al. (2012), Gradin et al. (2012) and Mendola and Busetta (2012). We refer the reader to Hoy and Zheng (2018) and Gradin et al. (2018) for an exhaustive summary of the theoretical and applied literature.

The indices proposed by Foster (2009), Bossert et al. (2012) and Dutta et al. (2013) share a similar structure. Together, they allow different aspects of past experiences to be brought into the analysis of the phenomenon under consideration. The goal of the current paper is to propose an application of these latter contributions to the measurement of poverty over time to material deprivation using the EU-SILC panel data set, which includes information on different aspects of well-being over time.

The only other papers similar in spirit that we are aware of are Nicholas and Ray (2012), Bossert et al. (2014), Nicholas et al. (2017), and Alkire et al. (2017). The first of these proposes generalizations of the contributions of Foster (2009) and Bossert et al. (2012) and applies the resulting indices to the analysis of multidimensional deprivation in Australia during the period between 2001 and 2008. The second contribution extends the analysis to aspects of the past considered in Hojman and Kast (2009) in the measurement of material deprivation among EU countries using the same dataset as this paper but with a focus on earlier periods. Hojman and Kast's (2009) index of poverty dynamics trades off poverty levels and changes (gains and losses) over time and is consistent with loss aversion. The results in the second paper based on Hojman and Kast (2009) convey a different picture of material deprivation within EU countries and tend to favor countries in which individuals experience improvements in their material deprivation scores. Nicholas et al. (2017) develop a multidimensional poverty measure that is sensitive to the distribution of deprivations within individuals, allowing to take into account not only whether the same individuals are becoming more deprived over time but also whether they are doing so in the same dimensions. Lastly, Alkire et al. (2017) combine the method proposed by Foster (2009) with the static index of Alkire and Foster (2011). This is the only contribution where some of the deprived individuals would not be included in the measure due to the identification step which will be discussed below.

In this paper, we expand the analysis of intertemporal material deprivation by including some mitigating effects of affluent periods, that is, of periods in which the individual is not deprived in any dimension; see Sect. 2 for details.

The measures proposed by Foster (2009) are generalizations of the Foster–Greer–Thorbecke (1984) class and allow time to play a role. The individual-level Foster index is the arithmetic mean over time of the per-period Foster–Greer–Thorbecke indices. In a similar spirit, the corresponding individual intertemporal index of material deprivation applied in this paper is the average material deprivation experienced by the individual over time.

Bossert et al. (2012) take persistence in the state of poverty into account. Their measure pays attention to the length of individual poverty spells by assigning a higher poverty weight to situations in which, *ceteris paribus,* poverty is experienced in consecutive rather than separated periods. The individual index is calculated as the weighted average of the individual per-period poverty values where, for each period, the weight is given by the length of the spell to which this period belongs. Similarly, the corresponding individual intertemporal index of material deprivation is calculated as the weighted average of the individual indices of material deprivation where, for each period, the weight is given by the length of the spell to which this period belongs.

Dutta et al. (2013) generalize Bossert et al. (2012) contribution by taking into account not only the debilitating impact of persistence in the state of poverty but also the mitigating effect of periods of affluence on subsequent poverty. The class of the proposed individual measures is a weighted sum over time of per-period Foster–Greer–Thorbecke indices where the weights reflect the damaging impact of consecutive periods in poverty and the mitigating effects of affluence periods. In a similar spirit, the corresponding individual intertemporal index of material deprivation is calculated as the weighted average of the individual indices of material deprivation.

In the multidimensional framework, each person is assigned a vector of several attributes that represent different dimensions of well-being. For the measurement of multidimensional poverty, it then becomes necessary to check whether a person has "minimally acceptable levels" of these attributes; see Sen (1992, p. 139). These minimally acceptable quantities of the attributes represent their threshold values or cutoffs that are necessary for an adequate standard of living. Therefore, a person is treated as deprived or poor in a dimension if the requisite observed level falls below this cutoff level. In this case, the individual is experiencing a functioning failure. Material deprivation at the individual level is an increasing function of these failures.

The identification of the poor in a multivariate framework can be carried out using different methods. One way of considering a person as poor is if the individual experiences a functioning failure in every dimension; this identifies the poor as those who are poor in all dimensions. This is known as the *intersection* method of identification of the poor. An alternative is the *union* method where the poor are identified as those experiencing at least one functioning failure. In between these two extremes lies the *intermediate* identification method, which regards a person as poor if she is deprived in at least $m \in \{1, \ldots, M\}$ dimensions, where M is the number of dimensions on which human well-being is considered to depend. The approach to identification in the current paper employs the union method. Additional results could be obtained by adopting other identification strategies, for example, by focusing exclusively on individuals who are severely materially deprived defined as those deprived for at least four items (see Eurostat 2012). Alkire et al. (2017) is the only contribution applying an intermediate identification method in an intertemporal framework. The same approach is followed by the EU Social protection committee which considers persistently materially deprived individuals to be those who are deprived in the current year and in at least 2 out of the preceding 3 years.

The different dimensions of well-being are incorporated using what Atkinson (2003) refers to as the *counting* approach. The counting measure of individual poverty consists of the number of dimensions in which a person is poor, that is, the number of the individual functioning failures. Since some of the dimensions may be more important than others, an alternative counting measure can be obtained by assigning different weights to different dimensions and then summing these weights for the dimensions in which functioning failure is observed. We follow both suggestions and produce results for two different weighting schemes: equal weights and Eurobarometer weights, where the latter reflect EU citizens' views on the importance of the dimension of well-being under consideration. For a discussion of weighting schemes in EU indicators, see Guio et al. (2009). A survey on the use of weights in multidimensional indices of well-being can be found in Decancq and Lugo (2013).

The remainder of the chapter proceeds as follows. Section 2 contains a description of the intertemporal indices of material deprivation. The application of these measures to illustrate the evolution of material deprivation in the European Union using the EU-SILC dataset appears in Sect. 3. Section 4 provides some brief concluding remarks.

2 Measuring Material Deprivation

Suppose there are $N \in \mathbb{N} \setminus \{1\}$ individuals in a society, $M \in \mathbb{N} \setminus \{1\}$ characteristics (or dimensions of material deprivation) and $T \in \mathbb{N} \setminus \{1\}$ time periods. For each individual $n \in \{1, \ldots, N\}$, for each time period $t \in \{1, \ldots, T\}$, and for each characteristic $m \in \{1, \ldots, M\}$, we observe a binary variable $x_m^{nt} \in \{0, 1\}$. A value of one indicates that individual n is poor with respect to dimension m in period t, and a value of zero identifies a characteristic with respect to which the individual is not poor in that period. For all $n \in \{1, \ldots, N\}$ and for all $t \in \{1, \ldots, T\}$, we let $x^{nt} = (x_1^{nt}, \ldots, x_M^{nt}) \in \{0, 1\}^M$. For all $n \in \{1, \ldots, N\}$, we define the deprivation profile $x^n = (x^{n1}, \ldots, x^{nT}) \in \left(\{0, 1\}^M\right)^T$. Furthermore, we let $x = (x^1, \ldots, x^N) \in \left(\left(\{0, 1\}^M\right)^T\right)^N$. Let $n \in \{1, \ldots, N\}$ and $x^n \in \left(\{0, 1\}^M\right)^T$. We say that n is deprived in period $t \in \{1, \ldots, T\}$ in x^n if and only if there exists $m \in \{1, \ldots, M\}$ such that $x_m^{nt} = 1$. That is, in order to be deprived in period t in x^n, individual n must be deprived with respect to at least one dimension in this period. This corresponds to the union method of identifying the deprived. Thus, individual n is not deprived in period t in x^n if and only if $x_m^{nt} = 0$ for all $m \in \{1, \ldots, M\}$.

For each individual $n \in \{1, \ldots, N\}$ and each time period $t \in \{1, \ldots, T\}$, individual n's material deprivation in t is given by

$$\sum_{m=1}^{M} x_m^{nt} \alpha_m,$$

where $\alpha_m \in \mathbb{R}_{++}$ is a parameter assigned to dimension $m \in \{1, \ldots, M\}$. In the applied part of the paper, we examine two different weighting schemes—one with identical weights for all dimensions, and another with weights that are derived from the Eurobarometer survey. See Sect. 3 for details.

A measure of intertemporal material deprivation for individual $n \in \{1, \ldots, N\}$ is a function $D^n\colon \left(\{0,1\}^M\right)^T \to \mathbb{R}_+$ which assigns a nonnegative individual intertemporal material deprivation value to each x^n in its domain. A measure of aggregate intertemporal material deprivation is a function $D\colon \left(\left(\{0,1\}^M\right)^T\right)^N \to \mathbb{R}_+$ that assigns a nonnegative intertemporal material deprivation value to each x in its domain.

The first approach analyzed here is inspired by Foster (2009). For each individual n, intertemporal material deprivation F^n is the average material deprivation experienced throughout the T periods. That is, for all $x^n \in \left(\{0,1\}^M\right)^T$,

$$F^n(x^n) = \frac{1}{T}\sum_{t=1}^{T}\sum_{m=1}^{M} x_m^{nt}\alpha_m.$$

Aggregate intertemporal material deprivation F is the arithmetic mean of the individual intertemporal material deprivation values. Thus, for all $x \in \left(\left(\{0,1\}^M\right)^T\right)^N$,

$$F(x) = \frac{1}{N}\sum_{n=1}^{N} F^n(x^n) = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{m=1}^{M} x_m^{nt}\alpha_m.$$

In order to discuss the adaptation of Bossert et al. (2012) approach to the intertemporal setting, some additional definitions are required.

To capture the notion of persistence in a state of material deprivation, we introduce functions $P^{nt}\colon \left(\{0,1\}^M\right)^T \to \{1, \ldots, T\}$ for each $n \in \{1, \ldots, N\}$ and for each $t \in \{1, \ldots, T\}$. If n is deprived in period t in x^n, we let $P^{nt}(x^n)$ be the maximal number of consecutive periods including t in which n is deprived. Analogously, if n is not deprived in period t in x^n, $P^{nt}(x^n)$ is the maximal number of consecutive periods including t in which n is not deprived. To illustrate this definition, suppose $T = 7$ and x^n is such that n is deprived in periods one, four, five, and seven. The length of the first spell of material deprivation is one and, thus, $P^{n1}(x^n) = 1$. This is followed by a spell out of deprivation of length two (in periods two and three), which implies $P^{n2}(x^n) = P^{n3}(x^n) = 2$. The next two periods are periods with deprivation and we obtain $P^{n4}(x^n) = P^{n5}(x^n) = 2$. Period six is a single period without deprivation and, thus, $P^{n6}(x^n) = 1$. Finally, there is a one-period spell of material deprivation and we have $P^{n7}(x^n) = 1$.

Following Bossert et al. (2012), intertemporal material deprivation BCD^n for individual $n \in \{1, \ldots, N\}$ is a weighted mean of the individual material deprivation values where, for each period t, the weight is given by the length of the spell to which this period t belongs. Thus, according to this approach, individual intertemporal material deprivation BCD^n is given by

$$BCD^n(x^n) = \frac{1}{T}\sum_{t=1}^{T} P^{nt}(x^n)\sum_{m=1}^{M} x_m^{nt}\alpha_m$$

for all $x^n \in \left(\{0,1\}^M\right)^T$. Aggregate intertemporal material deprivation BCD is the arithmetic mean of the individual intertemporal material deprivation values. Thus, for all $x \in \left(\left(\{0,1\}^M\right)^T\right)^N$,

$$BCD(x) = \frac{1}{N}\sum_{n=1}^{N} BCD^n(x^n) = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} P^{nt}(x^n)\sum_{m=1}^{M} x_m^{nt}\alpha_m.$$

To introduce the measures proposed by Dutta et al. (2013), the following definitions are of use. For a deprivation profile x^n, let s_t be the number of consecutive non-deprived periods immediately prior to a deprived period t, and let k_t be the number of preceding periods of uninterrupted positive levels of deprivation, up to and including the deprived period t. Formally,

$$s_t = \begin{cases} 0 & \text{if } t = 1 \text{ or } x^{n(t-1)} > 0 \\ t - \min\{s \mid s < t \text{ and } x^{ns} = \ldots = x^{n(t-1)} = 0\} & \text{otherwise} \end{cases}$$

and

$$k_t = \begin{cases} 1 & \text{if } t = 1 \text{ or } x^{n(t-1)} = 0 \\ t - \min\{s - 1 \mid s < t \text{ and } x^{nt'} > 0 \ \ \forall t' = s, \ldots, t\} & \text{otherwise.} \end{cases}$$

For example, for $T = 4$, the deprivation profile $x^n = (x^{n1}, 0, x^{n3}, x^{n4})$ has $s_1 = 0$, $k_1 = 1$, $s_3 = 1$ and $k_3 = 1$, and $s_4 = 0$ and $k_4 = 2$.

Dutta et al. (2013) propose to include the debilitating impact of persistence in the state of poverty and the mitigating effect of periods of affluence on subsequent poverty. Their individual measure DRZ^n is a weighted mean of the individual material deprivation values where, for each period, the weight considers the number of preceding periods of uninterrupted positive levels of deprivation, up to and including the deprived period t (see also BCD^n for an alternative weighing scheme), and the number of consecutive nonpoor periods immediately prior to a poor period, s_t. Thus, according to this approach, individual intertemporal material deprivation DRZ^n is given by

$$DRZ^n(x^n) = \frac{1}{T}\sum_{t=1}^{T}\frac{k_t}{1+s_t}\sum_{m=1}^{M} x_m^{nt}\alpha_m$$

for all $x^n \in \left(\{0, 1\}^M\right)^T$. Again, aggregate intertemporal material deprivation DRZ is the arithmetic mean of the individual intertemporal material deprivation values. Thus, for all $x \in \left(\left(\{0, 1\}^M\right)^T\right)^N$,

$$DRZ(x) = \frac{1}{N}\sum_{n=1}^{N} DRZ^n(x^n) = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{k_t}{1+s_t}\sum_{m=1}^{M} x_m^{nt}\alpha_m.$$

Following Gradin et al. (2012), generalized versions of the indices above could be computed by applying a general mean in the second aggregation stage as opposed to the arithmetic mean. This reflects the extent of aversion to inequality of intertemporal material deprivation across individuals.

3 Data and Results

In this section, we apply the indices defined above to measure material deprivation over time in the EU. The dataset we use is EU-SILC, which is employed by European Union member states and the Commission to monitor national and EU progress towards key objectives for the social inclusion process and the Europe 2020 growth strategy. We use the version of the data EUSILC LONGITUDINAL UDB 2010, version-2 of March 2013. Our analysis covers the years from 2007 to 2010 and, since we are interested in intertemporal material deprivation, we focus on the longitudinal component of the dataset. The sample is restricted to households who have been interviewed in each of the four years. The peculiarity of the sample used in this paper may give rise to differences between results obtained based on the entire sample. The variables that may be used in the measurement of material deprivation are available mainly at the household level. A conservative approach is followed in the sense that the households reporting a missing value are treated in the same way as those reporting not experiencing the functioning failure. As a result, we may be underestimating material deprivation, since we are attributing a functioning failure exclusively to households who explicitly claim to have the failure. The unit of analysis is the individual, that is, the household failure is attributed to each household member and the distribution of functioning failures among individuals is examined. The variables at the basis of the measures of material deprivation are listed in Table 1.

These variables are grouped according to three domains of quality of life: financial difficulties, housing conditions and durables, for a total of 12 indicators. These are the same variables chosen by Fusco et al. (2010). For other EU studies on material deprivation on different dimensions of well-being see, among others, Guio (2009) and Guio et al. (2009). Different dimensions of material deprivation could be applied in future work focusing exclusively on economic strain variables and excluding durable goods which are found to be more stable over time.

Table 1 Material deprivation variables

Financial difficulties		
	1.	Has been in arrears at any time in the last 12 months on: – mortgage or rent payments (hs010) – utility bills (hs020) – hire-purchase installments or other loan payments (hs030)
	2.	Cannot afford paying for 1-week annual holiday away from home (hs040)
	3.	Cannot afford a meal with meat, chicken, or fish (or vegetarian equivalent) every other day (hs050)
	4.	Lacks the capacity to face unexpected required expenses (hs060)
Durables		
	5.	Cannot afford a telephone (including mobile phone) (hs070)
	6.	Cannot afford a color tv (hs080)
	7.	Cannot afford a computer (hs090)
	8	Cannot afford a washing machine (hs100)
	9.	Cannot afford to have a car (hs110)
Housing conditions		
	10.	Lacks the ability to keep the home adequately warm (hh050)

Source EU-SILC dataset
N.B. For a selected number of countries in years 2007, 2008, and 2009, variables hs010, hs020, and hs030 have been replaced by new variables labeled hs011, hs021 and hs031, respectively. The two sets of variables measure the same dimensions. While hs010, hs020, and hs030 are binary variables (1-yes, 2-no), variables hs011, hs021, and hs031 take on three values (1-yes, once; 2-yes, two or more times; 3-no). We recode hs011, hs021, and hs031 as binary and use them in the place of hs010, hs020, and hs030

Two different weighting schemes are employed. The first consists of identical weights for all dimensions and the second is given by weights that are constructed from the views of EU citizens as surveyed in 2007 in the special Eurobarometer 279 on poverty and social exclusion (see TNS Opinion & Social 2007). This latter weighting method was first proposed by Guio et al. (2009). For each variable, we use as the weight the percentage of the EU27 citizens answering “absolutely necessary, no one should have to do without” to the requisite question as expressed by these instructions: “In the following questions, we would like to understand better what, in your view, is necessary for people to have what can be considered as an acceptable or decent standard of living in (OUR COUNTRY). For a person to have a decent standard of living in (OUR COUNTRY), please tell me how necessary do you think it is... (if one wants to).” The possible answers also included “necessary,” “desirable but not necessary,” and “not at all necessary.” See Table 2 for the relevant percentages.

The results of the intertemporal indices are reported in Tables 3 and 4 for the two weighting schemes. Each table includes the value of the index and the rankings of

Table 2 Answers in percentages to: "In the following questions, we would like to understand better what, in your view, is necessary for people to have what can be considered as an acceptable or decent standard of living in (OUR COUNTRY). For a person to have a decent standard of living in (OUR COUNTRY), please tell me how necessary do you think it is...(if one wants to)"

EU27	Absolutely necessary, no one should have to do without (%)	Necessary (%)	Desirable but not necessary (%)	Not at all necessary (%)
A place to live without a leaking roof, damp walls, floors, and foundation	68	28	3	1
To be able to keep one's home adequately warm	62	35	3	0
A place to live with its own bath or shower	63	31	6	0
An indoor flushing toilet for sole use of the household	69	27	4	0
To be able to pay rent or mortgage payments on time	62	34	3	0
To be able to pay utility bills (electricity, water, gas, etc.) on time	68	30	2	0
To be able to repay loans (such as loans to buy electrical appliances, furniture, a car or student loans, etc.) on time	48	40	9	2
Paying for 1 week annual holiday away from home	15	29	43	13
A meal with meat, chicken or fish at least once every 2 days	43	37	17	3
To be able to cope with an unexpected financial expense of X (NATIONAL CURRENCY)	32	43	21	2
A fixed telephone, landline	18	37	32	13
A mobile phone	12	26	37	25
A color TV	19	36	35	10
A computer	9	21	41	28
A washing machine	48	41	10	1
A car	17	34	36	13
A place to live without too much noise from neighbors or noise from the street (traffic, businesses, factories, etc.)	28	43	27	2
A place to live without too much pollution or other environmental problems (such as air pollution, grime, or rubbish)	42	44	13	1
A place to live without crime, violence, or vandalism in the area	49	38	12	1

Table 3 Intertemporal Material Deprivation and Ranking of EU Member States in the years 2007–2010 with Unitary Weights

Country	Foster	rank_Foster	DRZ	rank_DRZ	BCD	rank_BCD
AT	0.877	7	1.816	7	1.866	7
BE	0.792	6	1.689	6	1.733	6
BG	3.394	20	8.056	20	8.092	20
CY	1.665	14	3.519	13	3.610	13
CZ	1.092	9	2.381	9	2.439	9
EE	1.318	12	2.843	12	2.925	12
ES	0.921	8	1.897	8	1.971	8
FI	0.655	4	1.401	5	1.438	5
HU	2.339	19	5.581	19	5.619	19
IT	1.111	11	2.404	11	2.459	10
LT	1.885	16	4.257	16	4.335	16
LU	0.368	2	0.730	2	0.763	2
LV	2.203	18	5.049	18	5.132	18
NL	0.495	3	1.029	3	1.062	3
PL	1.943	17	4.506	17	4.537	17
PT	1.666	15	3.803	15	3.845	15
SE	0.341	1	0.623	1	0.658	1
SI	1.103	10	2.404	10	2.476	11
SK	1.661	13	3.673	14	3.730	14
UK	0.656	5	1.305	4	1.363	4

the countries (where 1 indicates the country with minimum deprivation). Figures 1 and 2 plot, for each weighting scheme, the rankings of the intertemporal material deprivation indices. The countries are ordered according to the values of the Foster index. As a benchmark, we also compute the indices of material deprivation for each year. These are contained in Table 5 (results with equal weights) and Table 6 (results with Eurobarometer weights), and the ranks are plotted in Figs. 3 and 4 respectively. Figures 5 and 6 compare the rankings of the countries resulting from the three intertemporal indices with those of yearly material deprivation in 2009 used as a benchmark for the two weighting schemes. Material deprivation over time is also compared with standard income poverty results based on the headcount index. The adopted income poverty line is set to 60% of the national median of the distribution of yearly equivalized household income using the OECD modified equivalence scale in order to account for different household size and composition. A note of caution is necessary with the analysis of income poverty: the results may differ from those obtained with the same dataset due to the restriction of the sample adopted in this paper.

Table 4 Intertemporal Material Deprivation and the Ranking of EU Member States in the years 2007–2010 with Eurobarometer Weights

Country	Foster_EU	rank_Foster_EU	DRZ_EU	rank_DRZ_EU	BCD_EU	rank_BCD_EU
AT	0.490	7	1.070	7	1.101	7
BE	0.506	6	1.055	6	1.082	6
BG	2.966	20	6.772	20	6.804	20
CY	1.253	17	3.030	16	3.108	16
CZ	0.668	8	1.277	9	1.307	8
EE	0.906	12	2.090	12	2.151	12
ES	0.612	9	1.277	8	1.324	9
FI	0.441	4	0.883	5	0.906	5
HU	1.658	19	4.012	19	4.040	19
IT	0.676	10	1.591	10	1.626	10
LT	1.378	15	3.025	15	3.079	15
LU	0.231	2	0.449	2	0.470	2
LV	1.761	18	3.731	18	3.791	18
NL	0.229	3	0.528	3	0.545	3
PL	1.569	16	3.162	17	3.184	17
PT	1.189	14	2.837	14	2.867	14
SE	0.231	1	0.372	1	0.394	1
SI	0.763	11	1.669	11	1.719	11
SK	1.384	13	2.692	13	2.735	13
UK	0.431	5	0.819	4	0.856	4

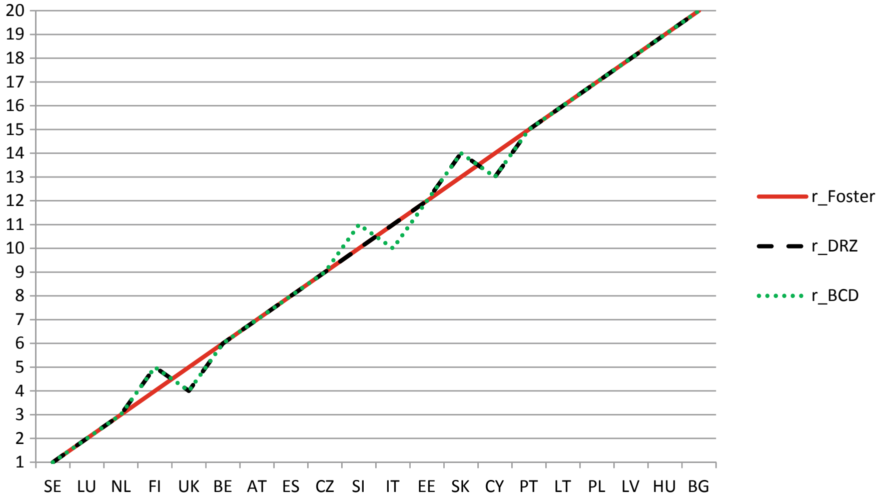


Fig. 1 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Unitary Weights

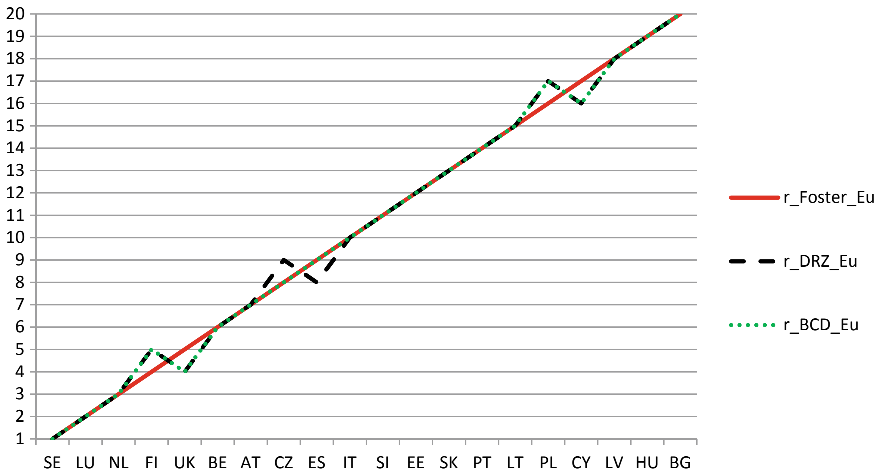


Fig. 2 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Eurobarometer Weights

Table 5 Yearly Material Deprivation and Ranking of EU Member States in the years 2007–2010 with Unitary Weights

Country	I2007	rank_2007	I2008	rank_2008	I2009	rank_2009	I2010	rank_2010
AT	0.962	7	1.062	8	0.854	6	0.778	6
BE	0.906	6	0.895	6	0.859	7	0.782	7
BG	4.431	20	3.648	20	4.167	20	4.301	20
CY	2.287	16	1.633	13	1.754	14	1.830	15
CZ	1.187	9	1.184	10	1.240	9	1.200	8
EE	1.331	11	1.166	9	1.458	12	1.705	12
ES	1.134	8	1.026	7	1.222	8	1.248	10
FI	0.648	4	0.650	4	0.636	4	0.577	4
HU	2.557	18	2.480	19	2.675	18	2.657	18
IT	1.348	12	1.292	11	1.244	10	1.236	9
LT	2.063	15	1.899	16	1.981	16	2.362	17
LU	0.592	3	0.593	3	0.558	3	0.480	3
LV	2.636	19	2.478	18	2.815	19	3.104	19
NL	0.458	1	0.414	1	0.385	1	0.393	2
PL	2.484	17	2.179	17	2.109	17	2.069	16
PT	1.860	13	1.821	15	1.727	13	1.726	14
SE	0.479	2	0.461	2	0.423	2	0.337	1
SI	1.280	10	1.409	12	1.387	11	1.422	11
SK	2.009	14	1.789	14	1.844	15	1.710	13
UK	0.733	5	0.759	5	0.839	5	0.768	5

Table 6 Yearly Material Deprivation and Ranking of EU Member States in the years 2007–2010 with Eurobarometer Weights

Country	I_Eu2007	rank_Eu2007	I_Eu2008	rank_Eu2008	I_Eu2009	rank_Eu2009	I_Eu2010	rank_Eu2010
AT	0.572	6	0.664	8	0.547	7	0.503	7
BE	0.601	7	0.578	6	0.545	5	0.500	6
BG	3.670	20	3.080	20	3.590	20	3.734	20
CY	1.996	19	1.407	16	1.541	17	1.619	16
CZ	0.669	8	0.648	7	0.695	8	0.668	8
EE	0.954	12	0.859	10	1.106	12	1.320	14
ES	0.780	9	0.705	9	0.869	9	0.895	10
FI	0.422	4	0.436	4	0.433	4	0.397	4
HU	1.848	17	1.815	18	1.985	18	2.005	18
IT	0.944	11	0.904	11	0.871	10	0.868	9
LT	1.468	14	1.336	14	1.482	15	1.703	17
LU	0.381	3	0.367	3	0.354	3	0.300	3
LV	1.925	18	1.852	19	2.140	19	2.409	19
NL	0.251	1	0.220	1	0.206	1	0.210	1
PL	1.803	16	1.544	17	1.518	16	1.498	15
PT	1.402	13	1.364	15	1.303	13	1.302	13
SE	0.293	2	0.301	2	0.273	2	0.219	2
SI	0.909	10	1.014	12	1.012	11	1.043	11
SK	1.487	15	1.311	13	1.385	14	1.278	12
UK	0.471	5	0.485	5	0.545	6	0.494	5

The three intertemporal indices rank the countries very similarly: see Figs. 1 and 2 and Tables 3 and 4. The short length of the panel is not sufficient to distinguish the different aspects of past experiences of material deprivation. For both weighting schemes, the least-deprived country is Sweden followed by Luxembourg and the Netherlands. Finland and the UK swap positions when material deprivation over time takes into consideration persistence in the state as opposed to an average value. At the opposite end of the rankings are Bulgaria, Hungary, and Latvia. The order among relatively highly deprived countries such as Cyprus, Lithuania, Poland, and Portugal depends on the weighting scheme and index used.

As clearly depicted in Figs. 1 and 2, the rankings of the countries change only little when the different intertemporal considerations are included. Part of the explanation for the robustness of these rankings may be that our data covers a relatively small number of periods, hence the effect of persistence may not be fully captured.

When time is not taken into consideration, in all the years analyzed (but in 2010) and for both weighting schemes, the Netherlands is the least-deprived country, followed by Sweden, Luxembourg, and Finland. See Figs. 3 and 4 and Tables 5 and 6. At the other end of the rankings, the worst position with respect to yearly material deprivation is occupied by Bulgaria, for both weighting schemes. Many relevant changes occur during the years starting from the fifth position onwards especially when Eurobarometer weights are applied.

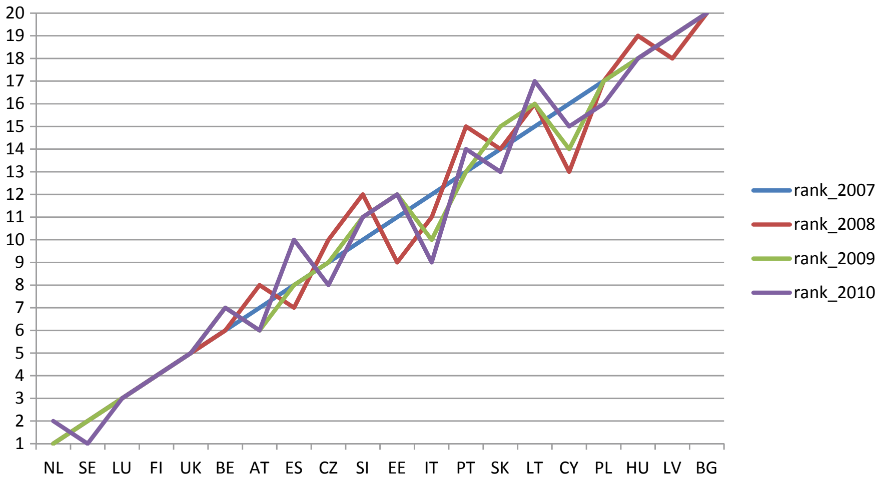


Fig. 3 Changes in the ranking of Yearly Material Deprivation among EU Member States with Unitary Weights

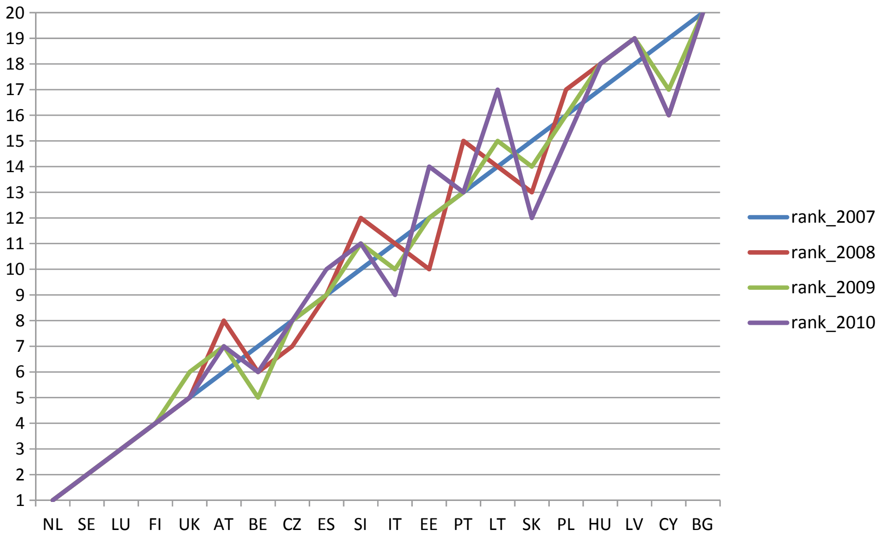


Fig. 4 Changes in the ranking of Yearly Material Deprivation among EU Member States with Eurobarometer Weights

When time is taken into account, the picture that emerges is different. Figures 5 and 6 compare the rankings of the countries with the yearly values for 2010. According to intertemporal material deprivation, the ranking of the least deprived countries is (for both weighting schemes) given by Sweden, Luxembourg and the Netherlands. Focusing only on 2010, the country with the best performance is Sweden for

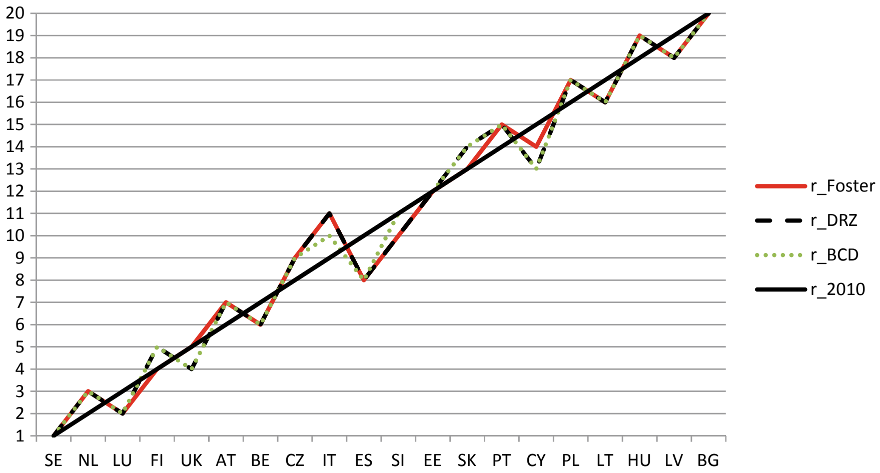


Fig. 5 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Unitary Weights versus the ranks of Yearly Material Deprivation in 2010

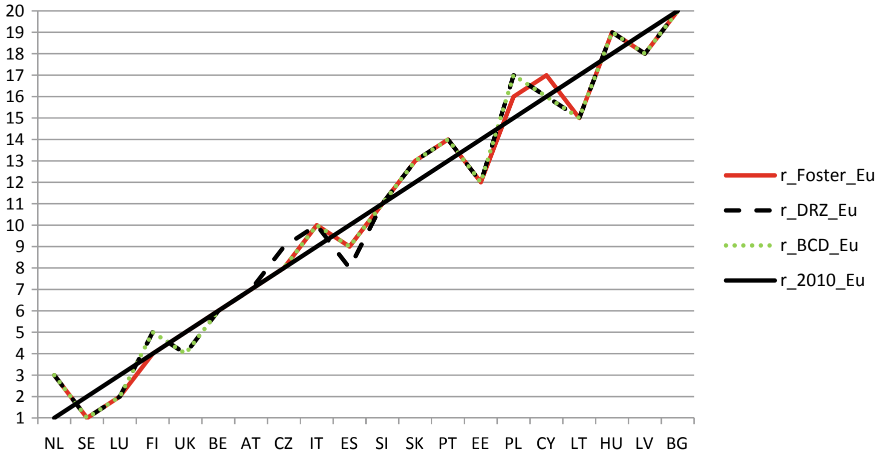


Fig. 6 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Eurobarometer Weights versus the ranking of Yearly Material Deprivation in 2010

the situation of equal weights and the Netherlands for Eurobarometer weights. This indicates that the materially deprived Swedes enjoy more affluent periods and are less persistently deprived than are the corresponding Dutch. Spain improves by two positions when intertemporal material deprivation is measured and by one position for Eurobarometer weights according to the Foster and BCD indices. For some other countries, we observe a movement of one position in both directions. The Netherlands, Italy, and Poland are the countries whose position is the most deteriorated

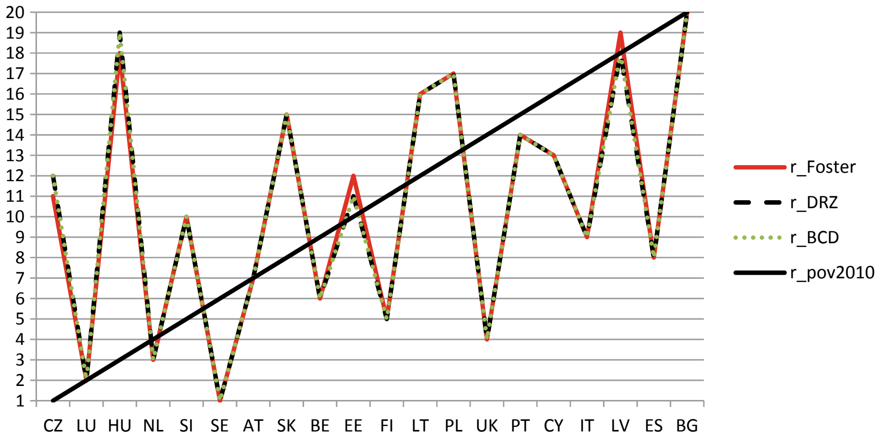


Fig. 7 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Unitary Weights versus the ranks of Yearly Income Poverty in 2010

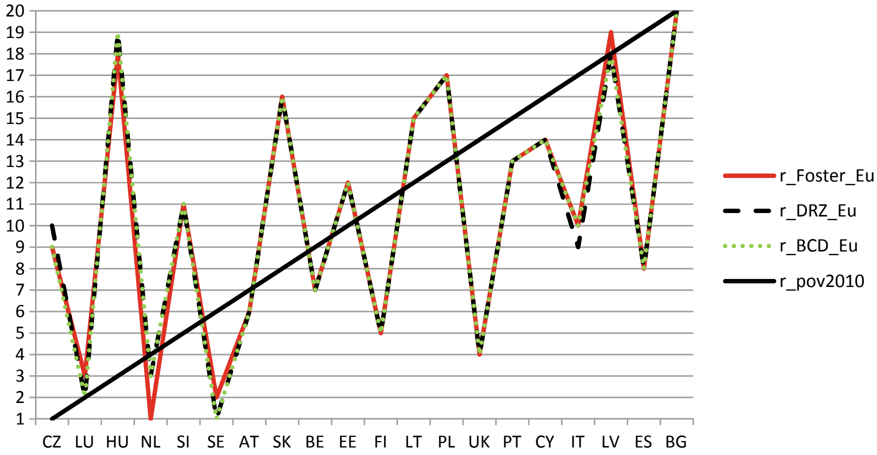


Fig. 8 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Eurobarometer Weights versus the ranking of Yearly Income Poverty in 2010

when intertemporal considerations are included but these results differ depending on the weighting schemes and the index.

Figures 7, 8, 9, and 10 and Table 7 contain results for intertemporal material deprivation and income poverty. In the first two figures, we compare the ranking of the countries with poverty during the last year of analysis, while the comparisons with poverty in 2007 are plotted in Figs. 9 and 10. It is apparent that intertemporal material deprivation is not related to income poverty in any of the 2 years. As recommended by

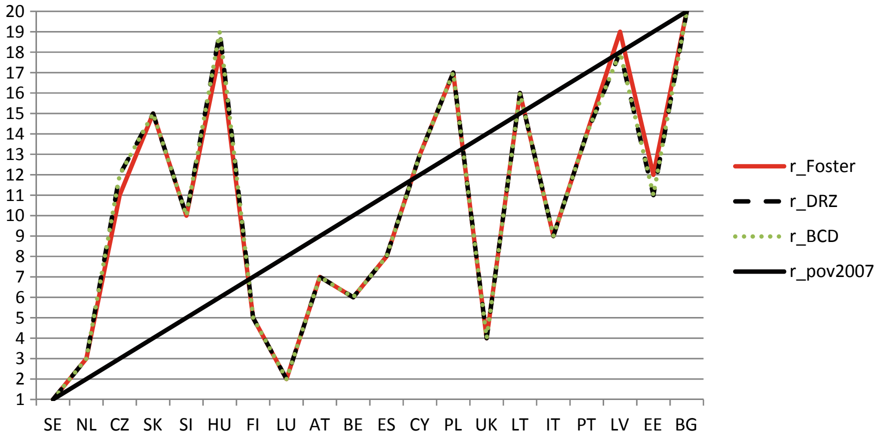


Fig. 9 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Unitary Weights versus the ranks of Yearly Income Poverty in 2007

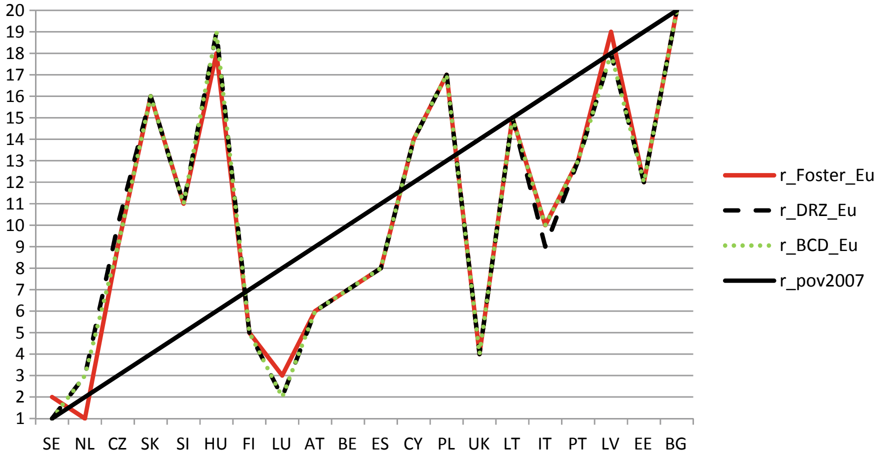


Fig. 10 Changes in the ranking of Intertemporal Material Deprivation among EU Member States with Eurobarometer Weights versus the ranking of Yearly Income Poverty in 2007

EU policy makers, indices of material deprivation are to be combined with income-based poverty measures since the two capture different aspects of individual well-being.

Table 7 Yearly Income Poverty and Ranking of EU Member States in the years 2007–2010

Country	P_2007	Rank_2007	P_2008	rank_2008	P_2009	rank_2009	P_2010	rank_2010
AT	0.134	9	0.117	6	0.117	8	0.112	7
BE	0.153	10	0.141	10	0.150	10	0.120	9
BG	0.225	20	0.212	18	0.233	19	0.223	20
CY	0.175	12	0.172	12	0.155	12	0.180	16
CZ	0.084	3	0.091	3	0.071	1	0.080	1
EE	0.220	19	0.219	19	0.197	18	0.132	10
ES	0.174	11	0.175	13	0.171	14	0.194	19
FI	0.117	7	0.140	9	0.137	9	0.134	11
HU	0.114	6	0.096	5	0.086	3	0.100	3
IT	0.200	16	0.197	16	0.185	16	0.182	17
LT	0.198	15	0.198	17	0.191	17	0.155	12
LU	0.120	8	0.118	7	0.107	6	0.094	2
LV	0.213	18	0.270	20	0.240	20	0.184	18
NL	0.083	2	0.092	4	0.114	7	0.102	4
PL	0.178	13	0.178	14	0.163	13	0.160	13
PT	0.212	17	0.183	15	0.179	15	0.175	15
SE	0.070	1	0.089	2	0.102	4	0.111	6
SI	0.113	5	0.120	8	0.102	5	0.106	5
SK	0.088	4	0.088	1	0.082	2	0.113	8
UK	0.181	14	0.170	11	0.152	11	0.163	14

4 Concluding Remarks

This paper analyzes the role of intertemporal considerations in material deprivation and compares EU countries according to this additional information. If we follow the path of material deprivation experienced by each individual over time, we obtain a different picture from that given by the yearly results. The contribution of this paper is mainly of a methodological nature. In addition, it provides some guidance for thorough empirical analyses of material deprivation. A possible subject for further research is to extend the results to patterns of deprivation by population subgroups to better understand the risk factors generating deprivation. Since the measurement of material deprivation is used by the EU member states and the European Commission to monitor national and EU progress in the fight against poverty and social exclusion, the basic results reported here suggest that time cannot be neglected. Countries should not only be compared according to their yearly results but also on data that follows individuals over several periods in order to produce a time-sensitive aggregate measure of material deprivation. Intertemporal material deprivation indices can be thought of as indicators of extreme forms of poverty along the lines of the recommendations of the Indicators Sub Group of the Social Protection Committee (see Bradshaw and Mayhew 2011).

References

Alkire S, Foster J (2011) Counting and multidimensional poverty measurement. J Public Econ 95:476–487

Alkire S, Apablaza M, Chakravarty S, Yalonetzky G (2017) Measuring chronic multidimensional poverty: a counting approach. J Policy Model 37:983–1006

Atkinson AB (2003) Multidimensional deprivation: contrasting social welfare and counting approaches. J Econ Inequal 1:51–65

Bossert W, Ceriani L, Chakravarty SR, D'Ambrosio C (2014) Intertemporal material deprivation. In: Betti G, Lemmi A (eds) Poverty and social exclusion, 128–142. Routledge, London

Bossert W, Chakravarty SR, D'Ambrosio C (2012) Poverty and time. J Econ Inequal 10:145–162

Bossert W, Chakravarty SR, D'Ambrosio C (2013) Multidimensional poverty and material deprivation with discrete data. Rev Income Wealth 10:29–43

Bourguignon F, Chakravarty SR (2003) The measurement of multidimensional poverty. J Econ Inequal 1:25–49

Bradshaw J, Mayhew E (2011) The measurement of extreme poverty in the European Union. Directorate-General for Employment, Social Affairs and Inclusion, European Commission

Calvo C, Dercon S (2009) Chronic poverty and all that: the measurement of poverty over time. In: Addison T, Hulme D, Kanbur R (eds) Poverty dynamics: interdisciplinary perspectives, 29–58. Oxford University Press, Oxford

Chakravarty SR, Chattopadhyay N (2018) Multidimensional poverty and material deprivation: theoretical approaches. In D'Ambrosio C (ed) Handbook of research on economic and social well-being, forthcoming

Chakravarty SR, Mukherjee D, Ranade R (1998) On the family of subgroup and factor decomposable measures of multidimensional poverty. Res Econ Inequal 8:175–194

D'Ambrosio C (2013) The indicators of intertemporal material deprivation: a proposal and an application to EU countries. ImPRovE Discussion Paper No. 13/08

Decancq K, Lugo M (2013) Weights in multidimensional indices of well-being: an overview. Econom Rev 32:7–34

Diez H, Lasso de la Vega C, Urrutia A (2008) Multidimensional unit- and subgroup consistent inequality and poverty measures: some characterization results. Res Econ Inequal 16:189–211

Dutta I, Roope L, Zank H (2013) On intertemporal poverty measures: the role of affluence and want. Soc Choice Welf 41:741–762

Eurostat, Measuring material deprivation in the EU. Methodologies and Working papers (2012)

Foster J (2009) A class of chronic poverty measures. In: Addison T, Hulme D, Kanbur R (eds) Poverty dynamics: interdisciplinary perspectives, 59–76. Oxford University Press, Oxford

Foster J, Greer J, Thorbecke E (1984) A class of decomposable poverty indices. Econometrica 52:761–766

Fusco A, Guio A-C, Marlier E (2010) Income poverty and material deprivation in European countries. Population and social conditions, methodologies and working papers, Eurostat

Gradin C, del Rio C, Cantó O (2012) Measuring poverty accounting for time. Rev Income Wealth 58:330–354

Gradin C, del Rio C, Canto O (2018) Poverty over time: empirical approaches. In D'Ambrosio C (ed) Handbook of research on economic and social well-being, forthcoming

Guio A-C (2009) What can be learned from deprivation indicators in Europa. Population and social conditions, methodologies and working papers, Eurostat

Guio A-C (2018) Multidimensional poverty and material deprivation: empirical approaches. In D'Ambrosio C (ed) Handbook of research on economic and social well-being, forthcoming

Guio A-C, Fusco A, Marlier E(2009) An EU approach to material deprivation using EU-SILC and Eurobarometer data. Working Paper 2009-19, IRISS

Hojman D, Kast F (2009) On the measurement of poverty dynamics. Faculty Research Working Paper Series RWP09-35, Harvard Kennedy School

Hoy M, Thompson BS, Zheng B (2012) Empirical issues in lifetime poverty measurement. J Econ Inequal 10:163–189

Hoy M, Zheng B (2011) Measuring lifetime poverty. J Econ Theory 146:2544–2562

Hoy M, Zheng B (2018) Poverty over time: theoretical approaches. In D'Ambrosio C (ed) Handbook of research on economic and social well-being, forthcoming

Jenkins SP (2000) Modelling household income dynamics. J Popul Econ 13:529–567

Mendola D, Busetta A (2012) The importance of consecutive spells of poverty: a path-dependent index of longitudinal poverty. Rev Income Wealth 58:355–374

Nicholas A, Ray R (2012) Duration and persistence in multidimensional deprivation: methodology and an Australian application. Econ Rec 88:106–126

Nicholas A, Ray R, Sinha K (2017) Differentiating between dimensionality and duration in multidimensional measures of poverty: methodology with an application to China. Review of Income and Wealth, forthcoming. https://doi.org/10.1111/roiw.12313

Rodgers JR, Rodgers JL (1993) Chronic poverty in the United States. J Human Resour 28:25–54

Sen A (1992) Inequal Re-examined. Harvard University Press, Cambridge, MA

TNS Opinion & Social (2007) Poverty and Exclusion, Report on the Special Eurobarometer No. 279/Wave 67.1, http://ec.europa.eu/public_opinion/archives/ebs/ebs_279.pdf

Tsui K-Y (2002) Multidimensional poverty indices. Soc Choice Welfare 19:69–93

Feasible Shared Destiny Risk Distributions

Thibault Gajdos, John A. Weymark and Claudio Zoli

Abstract Social risk equity is concerned with the comparative evaluation of social risk distributions, which are probability distributions over the potential sets of fatalities. In the approach to the evaluation of social risk equity introduced by Gajdos, Weymark, and Zoli (Shared destinies and the measurement of social risk equity, Annals of Operations Research 176:409–424, 2010), the only information about such a distribution that is used in the evaluation is that contained in a shared destiny risk matrix whose entry in the kth row and ith column is the probability that person i dies in a group containing k individuals. Such a matrix is admissible if it satisfies a set of restrictions implied by its definition. It is feasible if it can be generated by a social risk distribution. It is shown that admissibility is equivalent to feasibility. Admissibility is much easier to directly verify than feasibility, so this result provides a simple way to identify which matrices to consider when the objective is to socially rank the feasible shared destiny risk matrices.

Keywords Social risk evaluation · Social risk equity · Public risk · Shared destinies

JEL classification numbers D63 · D81 · H43

T. Gajdos
CNRS and Laboratoire de Psychologie Cognitive, Aix-Marseille University, Bâtiment 9 Case D, 3 place Victor Hugo, 13331 Marseille Cedex 3, France
e-mail: thibault.gajdos@univ-amu.fr

J. A. Weymark (✉)
Department of Economics, Vanderbilt University, VU Station B #35189, 2301 Vanderbilt Place, Nashville, TN 37235-1819, USA
e-mail: john.weymark@vanderbilt.edu

C. Zoli
Department of Economics, University of Verona, Via Cantarane 24, 37129 Verona, Italy
e-mail: claudio.zoli@univr.it

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*, Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_3

1 Introduction

Governments routinely implement policies that affect the risks that a society faces. For example, barriers are installed to lessen the risk of a terrorist driving a vehicle into pedestrians, dikes are built to reduce the risk of flooding, and carbon taxes are imposed to slow down the rise in the temperature of the Earth's atmosphere so as to reduce the likelihood of the serious harms that result from climate change. Policies differ in the degree to which they change the expected aggregate amount of a harm and how it is distributed across the population. A consequentialist approach to evaluating the relative desirability of different policies that affect these kinds of social risks does so by ranking the possible distributions of the resulting harms. If this ranking, or an index representing it, takes account of the equity of the resulting distribution of risks, it is a measure of *social risk equity*.

The measurement of social risk equity has its origins in the work of Keeney (1980a, b, c). The analysis of social risk equity has been further developed by Broome (1982), Fishburn (1984), Fishburn and Sarin (1991), Fishburn and Straffin (1989), Gajdos et al. (2010), Harvey (1985), Keeney and Winkler (1985), and Sarin (1985), among others. While these analyses apply to any kind of social harm, for the most part, the harm that they consider is death. Our analysis also applies to any socially risky situation in which a harm may affect some, but not necessarily all, of the society in question but, for concreteness, we, too, suppose that this harm is death.

The set of individuals who die as a result of their exposure to the risk is a fatality set and a social risk distribution is a probability distribution over all of the possible fatality sets. Social risk distributions are ranked using a social risk equity preference ordering. Not all of the information about a social risk distribution may be regarded as being relevant when determining the social preference relation. For example, Fishburn and Straffin (1989), Keeney and Winkler (1985), and Sarin (1985) only take account of the risk profiles for individuals and for fatalities. The former lists the likelihoods of each person dying, whereas the latter is the probability distribution over the number of fatalities. These statistics can be computed from a social risk distribution, but in doing so, some information is lost.

Gajdos et al. (2010) propose also taking into account a concern for shared destinies; specifically, with the number of other individuals with whom someone perishes. Chew and Sagi (2012) describe this concern as being one of ex-post fairness. For example, for a given probability of there being k fatalities, it might be socially desirable to have this risk spread more evenly over the individuals. As Example 3 in Gajdos et al. (2010) demonstrates, it is possible for the distribution of how many people someone dies with to differ in two social risk distributions even though the risk profiles for individuals and for fatalities are the same in both distributions. As a consequence, a concern for shared destinies cannot be fully captured if one restricts attention to the information provided by the likelihoods of each person dying and the probability distribution over the number of fatalities.[1]

[1]There are other dimensions of social risk equity that may be of concern, such as dispersive equity and catastrophe avoidance. There is a concern for dispersive equity if account is taken of individual

The distribution of shared destiny risks can be expressed using a shared destiny risk matrix whose entry in the kth row and ith column is the probability that person i dies in a group containing k individuals. In the approach developed by Gajdos et al. (2010), if two social risk distributions result in the same shared destiny risk matrix, they are regarded as being socially indifferent. Because the risk profiles for individuals and for fatalities can be computed from the information contained in a shared destiny risk matrix, their approach to evaluating social risks can take account of these two risk profiles, not just a concern for shared destinies. In effect, in their approach to social risk evaluation, ranking social risk distributions is equivalent to ranking shared destiny risk matrices.

The entries of a social destiny risk matrix are probabilities, and so all lie in the interval [0, 1]. There are three other independent properties that such a matrix must necessarily satisfy as a matter of definition: (i) nobody's probability of dying can exceed 1, (ii) the probability that there are a positive number of fatalities cannot exceed 1, and (iii) nobody can have a probability of dying in a group of size k that exceeds the probability that there are k fatalities. A social destiny risk matrix that satisfies these properties is said to be admissible. Starting with a social risk distribution, we can compute the entries in the corresponding shared destiny risk matrix. A shared destiny risk matrix that can be generated in this way from a social risk distribution is said to be feasible. A feasible shared destiny risk matrix is necessarily admissible. The question we address is whether there are admissible shared destiny risk matrices that are not feasible. We show that there are not. Thus, a shared destiny risk matrix is admissible if and only if it is feasible.

In order to establish this result, we develop an algorithm that shows how to construct a social risk distribution from a shared destiny risk matrix in such a way that the resulting distribution can be used to generate the matrix. It is easy to determine if a shared destiny risk matrix is admissible but, as our algorithm makes clear, confirming that it is also feasible by finding a social risk distribution that generates it may be a formidable undertaking. However, if our objective is only to socially rank the feasible shared destiny risk matrices, our result tells us that this is equivalent to socially ranking the admissible shared destiny risk matrices. We do not need to know how to generate these matrices from social risk distributions in order to know that they are feasible; we only need to know that they are admissible.

In Sect. 2, we introduce the formal framework used in our analysis. The algorithm employed to determine a social risk distribution that generates a given admissible shared destiny risk matrix is presented in Sect. 3. We illustrate the operation of this algorithm in Sect. 4. We prove that a shared destiny risk matrix is admissible if and only if it is feasible in Sect. 5.

---

characteristics such as gender, race, or geographic location in addition to the individuals' exposures to social risks. See Fishburn and Sarin (1991), for an analysis of the evaluation of social risks that allows for dispersive equity. Bommier and Zuber (2008), Fishburn (1984), Fishburn and Straffin (1989), Harvey (1985), and Keeney (1980a) consider social preferences for catastrophe avoidance. We do not examine dispersive equity or catastrophe avoidance here.

2 Shared Destiny Risk Matrices

There is a society of $n \geq 2$ individuals who face a social risk. Let $N = \{1, \ldots, n\}$ be the set of these individuals. A *fatality set* is a subset $S \subseteq N$ consisting of the set of individuals who expost die as a consequence of the risk that this society faces. There are 2^n possible fatality sets, including $\varnothing$ (nobody dies) and N (everybody dies). A *social risk distribution* is a probability distribution p on 2^n, with $p(S)$ denoting the ex ante probability that the fatality set is S. We suppose that only this probability distribution is relevant for the purpose of social risk evaluation. The set of all such probability distributions is $\mathscr{P}$.

For each $k \in N$, let $M(k) \in [0, 1]^n$ denote the vector whose ith component $M(k, i)$ is the ex ante probability that person i will die when there are exactly k fatalities. A *shared destiny risk matrix* is an $n \times n$ matrix M whose kth row is $M(k)$. Let $\bar{n}(k)$ be the number of positive entries in $M(k)$. The *risk profile for individuals* is the vector $\alpha \in [0, 1]^n$, where

$$\alpha(i) = \sum_{k=1}^{n} M(k, i), \quad \forall i \in N, \tag{1}$$

which is the ex ante probability that person i will die. The *risk profile for fatalities* is the vector $\beta \in [0, 1]^n$, where

$$\beta(k) = \frac{1}{k} \sum_{i=1}^{n} M(k, i), \quad \forall k \in N, \tag{2}$$

which is the ex ante probability that there will be exactly k fatalities. Note that a risk profile for fatalities does not explicitly specify the probability that nobody dies. The probability that there are no fatalities is $1 - \sum_{k=1}^{n} \beta(k)$.

By definition, each of the entries of M is a probability and so must lie in the interval $[0, 1]$. Hence, each of the components of α and β must be nonnegative as they are sums of entries in M. There are three other restrictions on M. They are

$$\alpha(i) \leq 1, \quad \forall i \in N, \tag{3}$$

$$\sum_{k=1}^{n} \beta(k) \leq 1, \quad \forall k \in N, \tag{4}$$

and

$$M(k, i) \leq \beta(k), \quad \forall (k, i) \in N^2. \tag{5}$$

The first of these requirements is that no person can die with a probability greater than 1. The second is that the probability that there are a positive number of fatalities cannot exceed 1. The third is that nobody's probability of dying in a group of size k

can exceed the probability of there being k fatalities. Of course, it must also be the case that

$$\beta(k) \le 1, \quad \forall k \in N. \tag{6}$$

That is, the probability that there are a particular number of fatalities cannot exceed 1. However, (6) follows from (4) because all probabilities are nonnegative. A social risk equity matrix M is *admissible* if it satisfies (3), (4), and (5).

It is obvious that M must satisfy (3) and (4), but the necessity of (5) is less so because there is more than one way that someone can die with $k-1$ other individuals when $k > 1$. To see why (5) is required, suppose, on the contrary, that $M(k,i) > \beta(k)$ for some $i \in N$. Then, because $M(k,i)$ is the probability that person i perishes with $k-1$ other individuals, it must be the case that $\sum_{j \neq i} M(k,j) > (k-1)\beta(k))$. Hence, $\sum_{i=1}^{n} M(k,i) > k\beta(k)$. It then follows that $\beta(k) = \frac{1}{k}\sum_{i=1}^{n} M(k,i) > \beta(k)$, a contradiction.

For each $k \in N$, $\mathscr{T}(k) = \{S \in 2^n \| S| = k\}$ is the set of subgroups of the society in which exactly k individuals die. For each $(k,i) \in N^2$, $\mathscr{S}(k,i) = \{S \in \mathscr{T}(k) | i \in S\}$ is the set of subgroups in which exactly k people die and i is one of them. A shared destiny risk matrix M is *feasible* if there exists a social risk distribution $p \in \mathscr{P}$ and an $n \times n$ matrix M_p such that $M(k,i) = M_p(k,i)$, where $M_p(k,i) = \sum_{S \in \mathscr{S}(k,i)} p(S)$. That is, $M_p(k,i)$ is the probability that there are k deaths and i is one of them when the social risk distribution is p.

3 The Decomposition Algorithm

By construction, for any $p \in \mathscr{P}$, M_p is an admissible shared destiny risk matrix. In other words, any feasible shared destiny risk matrix is admissible. The question then arises as to whether feasibility imposes any restrictions on M other than that it be admissible. We show that it does not.

For any admissible shared destiny risk matrix M, we need to show that there exists a social risk distribution $p \in \mathscr{P}$ such that $M_p = M$. This is done by considering each value of k separately. For each $k \in N$, we know that the probability of having this number of fatalities is $\beta(k)$. We need to distribute this probability among the subgroups in $\mathscr{T}(k)$ (the subgroups for which there are k fatalities) in such a way that the probability that person i dies in a group of size k is $M(k,i)$. The resulting probabilities for the subgroups in $\mathscr{T}(k)$ are called a *probability decomposition*. Put another way, for each $i \in N$, we need to distribute the probability $M(k,i)$ among the subgroups in $\mathscr{S}(k,i)$ (the subgroups containing person i for which there are k fatalities) in such a way that the amount π^S allocated to any $S \in \mathscr{S}(k,i)$ is the same for everybody in this group. The value π^S is then the probability that the set of individuals who perish is S.

If $M(k,i) = 0$ for all $i \in N$ (so $\bar{n}(k) = 0$), then $\beta(k) = 0$, so we assign probability 0 to each $S \in \mathscr{T}(k)$. If $k = n$, only N is in $\mathscr{T}(n)$, so no decomposition is needed; N is simply assigned the probability $\beta(n)$. When $\beta(k) \neq 0$ and $k < n$, we construct an algorithm that produces the requisite probability decomposition. The algorithm proceeds through a number of steps, which we denote by $t = 0, 1, 2, \ldots$. We show that the algorithm terminates in no more than $\bar{n}(k)$ steps. The relevant variables in each step are distinguished using a superscript whose value is the step number.

The vector $\hat{M}(k)$ is a *nonincreasing rearrangement* of $M(k)$ if $\hat{M}(k, i) \geq \hat{M}(k, i + 1)$ for all $k = 1, \ldots, n - 1$. Whenever a vector of probabilities for the n individuals is rearranged in this way, ties are broken in such a way that the original order of the individuals is preserved. For example, if $n = 3$, in the rearrangement $(2, 1, 1)$ of $(1, 2, 1)$, the first 1 is associated with person 1 and the second 1 with person 3. Without loss of generality, we suppose that $M(k)$ is initially ranked in nonincreasing order. We now describe our algorithm.

Probability Decomposition Algorithm. The initial values of the relevant variables are

$$M^0(k) = \hat{M}^0(k) = M(k) = \hat{M}(k)$$

and

$$\beta^0(k) = \beta(k).$$

Step 1. In Step 1, we assign a probability π^1 to the first k individuals, which is the set of individuals with the k highest probabilities in $\hat{M}^0(k)$. After π^1 is subtracted from each of the first k components of $\hat{M}^0(k)$, we are left with the fatality probability

$$\beta^1(k) = \beta^0(k) - \pi^1$$

to distribute among the groups of size k using the probabilities in

$$M^1(k) = \hat{M}^0(k) - (\pi^1, \ldots, \pi^1, 0, \ldots, 0).$$

Letting $\rho^0(k, i)$ denote the rank of individual i in $\hat{M}^0(k)$, we define the vector

$$\pi^1(k) = \pi^1 \cdot (I_1^0, I_2^0, \ldots, I_n^0),$$

where $I_i^0 = 1$ if $\rho^0(k, i) \leq k$ and $I_i^0 = 0$ otherwise. Using $\pi^1(k)$, $M^1(k)$ can be equivalently written as

$$M^1(k) = M^0(k) - \pi^1(k).$$

We need to ensure that each of the probabilities in $M^1(k)$ is nonnegative. Because $\hat{M}^0(k)$ is a nonincreasing rearrangement of $M^0(k)$ and $\hat{M}^1(k, i) = \hat{M}^0(k, i)$ for $i = k + 1, \ldots, n$, it must therefore be the case that $\pi^1 \leq \hat{M}^0(k, k)$. We also need to ensure that none of these probabilities exceeds the fatality probability $\beta^1(k)$ left to distribute. This condition is satisfied by construction for the first k individuals. Hence, because $\hat{M}^0(k)$ is a nonincreasing rearrangement of $M^0(k)$, in order to satisfy this condition, it is only necessary that $\hat{M}^0(k, k + 1) \leq \beta^1(k) = \beta^0(k) - \pi^1$. Both of these requirements are satisfied by setting

$$\pi^1 = \min\{\hat{M}^0(k,k), \beta^0(k) - \hat{M}^0(k,k+1)\}.$$

By (5), $\hat{M}^0(k,k) \leq \beta^0(k)$. Therefore, $\pi^1 \leq \beta^0(k)$ and, hence, $\beta^1(k) \leq \beta^0(k)$. Because $\beta^1(k) = \frac{1}{k}\sum_{i=1}^n M^1(k,i)$ and $M^1(k,i) \geq 0$ for all $i \in N$, $\beta^1(k) \geq 0$.

Let S^1 denote the first k individuals in $\hat{M}^0(k)$. We choose $p(S^1)$ to be π^1.

If $M^1(k) = (0,0,0,\ldots,0)$, the algorithm terminates. Otherwise, it proceeds to the next step.

Step t ($t \geq 2$). The operation of the algorithm in this step follows the same basic logic as in Step 1. The value of π^t is chosen by setting

$$\pi^t = \min\{\hat{M}^{t-1}(k,k), \beta^{t-1}(k) - \hat{M}^{t-1}(k,k+1)\}. \tag{7}$$

Letting $\rho^{t-1}(k,i)$ denote the rank of individual i in $\hat{M}^{t-1}(k)$, we define the vector

$$\pi^t(k) = \pi^t \cdot (I_1^{t-1}, I_2^{t-1}, \ldots, I_n^{t-1}), \tag{8}$$

where $I_i^{t-1} = 1$ if $\rho^{t-1}(k,i) \leq k$ and $I_i^{t-1} = 0$ otherwise.

We define $M^t(k)$ and $\beta^t(k)$ by setting

$$M^t(k) = M^{t-1}(k) - \pi^t(k) \tag{9}$$

and

$$\beta^t(k) = \beta^{t-1}(k) - \pi^t. \tag{10}$$

Analogous reasoning to that used in Step 1 shows that $0 \leq \beta^t(k) \leq \beta^{t-1}(k)$.

Let S^t denote the first k individuals in $\hat{M}^{t-1}(k)$. We choose $p(S^t)$ to be π^t.

If $M^t(k) = (0,0,0,\ldots,0)$, the algorithm terminates. Otherwise, it proceeds to the next step.

If the algorithm terminates and a group S with k members has not been assigned a probability by the algorithm, we set $p(S) = 0$.

4 Examples of the Probability Decomposition Algorithm

The operation of the probability decomposition algorithm is illustrated with three examples. In each of these examples, it is assumed that M is admissible. In the first example, the algorithm is applied to the case in which nobody dies with anybody else.

Example 1 Let $k = 1$ with $M(1,i) > 0$ for some $i \in N$. If M is feasible, for each $i \in N$, we must have $p(\{i\}) = M(1,i)$. We show that the algorithm produces this result. We first consider the case in which $M(1,i) > 0$ for all $i \in N$.

In Step 1, person 1 is the highest ranked individual in $\hat{M}^0(1)$. Therefore, we have $\beta^0(1) - \hat{M}^0(1,2) = \frac{1}{1}\sum_{i=1}^n \hat{M}^0(1,i) - \hat{M}^0(1,2) = \sum_{i\neq 2}^n \hat{M}^0(1,i) \geq \hat{M}^0(1,1)$. It then follows that $\pi^1 = \hat{M}^0(1,1)$ and, hence, $p(\{1\}) = \hat{M}^0(1,1) = M(1,1)$. We have $\pi^1(1) = (M(1,1), 0, \ldots, 0)$ and so $M^0(1)$ is now replaced with $M^1(1) = (0, M^0(2), \ldots, M^0(n))$. Because $M^1(1,1) = 0$, person 1 is never considered again by the algorithm. There is fatality probability $\beta^1(1) = \beta^0(1) - \pi^1 = \sum_{i=1}^n M(1,i) - M(1,1) = \sum_{i=2}^n M(1,i)$ left to allocate.

Step 2 uses the vector $\hat{M}^1(1)$. Person 2 is the second highest ranked individual in $M^0(1)$ and so is first ranked in $\hat{M}^1(1)$. As in Step 1, $\pi^2 = \hat{M}^1(1,1)$ and, hence, $p(\{2\}) = \hat{M}^1(1,1) = M(1,2)$. We have $\pi^1(1) = (0, M(1,2), 0, \ldots, 0)$ and so $M^1(1)$ is replaced with $M^2(1) = (0, 0, M^0(3), \ldots, M^0(n))$ and person 2 is never considered again. There is fatality probability $\beta^2(1) = \beta^1(1) - \pi^2 = \sum_{i=2}^n M(1,i) - M(1,2) = \sum_{i=3}^n M(1,i)$ left to allocate.

More generally, person $i \in N$ is singled out in Step i and assigned the probability $p(\{i\}) = M(1,i)$. In Step n, $M^n(1) = (0, 0, 0, \ldots, 0)$, and so the algorithm terminates.

If $M(1,i) = 0$ for some $i \in N$, then the algorithm proceeds as above but terminates in Step $\bar{n}(1) < n$, where it is recalled that $\bar{n}(1)$ is the number of individuals for whom $M(1,i)$ is positive. For any individual i for whom $M(1,i) = 0$, when the algorithm terminates, $p(\{i\})$ is set equal to 0.

In the next two examples, individuals do not die alone. In these examples, all probabilities are expressed in terms of percentages, so, for example, 5 is the probability 0.05.

Example 2 Let $n = 7$ and $k = 4$. We suppose that $M^0(4) = M(4) = \hat{M}(4) = \hat{M}^0(4) = (5, 4, 4, 4, 4, 2, 1)$. Consequently, $\beta^0(4) = \beta(4) = \frac{1}{4}\sum_{i=1}^n M(4,i) = 6$.

Step 1. We have $\pi^1 = \min\{\hat{M}^0(4,4), \beta^0(4) - \hat{M}^0(4,5)\} = 2$. Therefore, $\pi^1(4) = (2, 2, 2, 2, 0, 0, 0)$, $M^1(4) = M^0(4) - \pi^1(4) = (3, 2, 2, 2, 4, 2, 1)$, and $\beta^1(4) = \beta^0(4) - \pi^1 = 6 - 2 = 4$. Hence, $p(\{1, 2, 3, 4\}) = \pi^1 = 2$.

Step 2. There are four individuals with the third highest probability in $M^1(4)$. Using our tie-breaking rule, individuals 1, 2, 3, and 5 have the first four probabilities in $\hat{M}^1(4)$. We thus have $\hat{M}^1(4) = (4, 3, 2, 2, 2, 2, 1)$, so $\pi^2 = \min\{\hat{M}^1(4,4), \beta^1(4) - \hat{M}^1(4,5)\} = 2$. Therefore, $\pi^2(4) = (2, 2, 2, 0, 2, 0, 0)$, $M^2(4) = M^1(4) - \pi^2(4) = (1, 0, 0, 2, 2, 2, 1)$, and $\beta^2(4) = \beta^1(4) - \pi^2 = 4 - 2 = 2$. Hence, $p(\{1, 2, 3, 5\}) = \pi^2 = 2$.

Step 3. There are two individuals with the fourth highest probability in $M^2(4)$. The tie is broken in favor of person 1, so the first four individuals in $\hat{M}^2(4)$ are 1, 4, 5, and 6. We have $\hat{M}^2(4) = (2, 2, 2, 1, 1, 0, 0)$, so $\pi^3 = \min\{\hat{M}^2(4,4), \beta^2(4) - \hat{M}^2(4,5)\} = 1$. Therefore, $\pi^3(4) = (1, 0, 0, 1, 1, 1, 0)$, $M^3(4) = M^2(4) - \pi^3(4) = (0, 0, 0, 1, 1, 1, 1)$, and $\beta^3(4) = \beta^2(4) - \pi^3 = 2 - 1 = 1$. Hence, $p(\{1, 4, 5, 6\}) = \pi^3 = 1$.

Step 4. The four individuals with the highest probabilities in $M^3(4)$ are 4, 5, 6, and 7. We have $\hat{M}^3(4) = (1, 1, 1, 1, 0, 0, 0)$, so $\pi^4 = \min\{\hat{M}^3(4,4), \beta^3(4) - \hat{M}^3(4,5)\} = 1$. Therefore, $\pi^4(4) = (0, 0, 0, 1, 1, 1, 1)$, $M^4(4) = M^3(4) - \pi^4(4) = (0, 0, 0, 0, 0, 0, 0)$, and $\beta^4(4) = \beta^3(4) - \pi^4 = 1 - 1 = 0$. Hence, $p(\{4, 5, 6, 7\}) = \pi^4 = 1$.

Because $M^4(4) = (0, 0, 0, 0, 0, 0, 0)$, the algorithm terminates in Step 4. The four groups identified in Steps 1–4 are assigned positive probability. For any other set of individuals S with four members, $p(S) = 0$. There are $\frac{n!}{k!(n-k)!} = \frac{7!}{4!3!} = 35$ possible groups of this size, so 31 of them are assigned a zero probability.

In Examples 1 and 2, in Step t, π^t is set equal to $\hat{M}^t(k, k)$. In Example 3, it is instead sometimes set equal to $\beta^{t-1}(k) - \hat{M}^{t-1}(k, k+1)$.

Example 3 Let $n = 3$ and $k = 2$. We suppose that $M^0(2) = M(2) = \hat{M}(2) = \hat{M}^0(2) = (7, 5, 4)$. Consequently, $\beta^0(2) = \beta(2) = \frac{1}{2}\sum_{i=1}^{n} M(2, i) = 8$.

Step 1. We have $\pi^1 = \min\{\hat{M}^0(2, 2), \beta^0(2) - \hat{M}^0(2, 3)\} = 4$. Therefore, $\pi^1(2) = (4, 4, 0)$, $M^1(2) = M^0(2) - \pi^1(2) = (3, 1, 4)$, and $\beta^1(2) = \beta^0(2) - \pi^1 = 8 - 4 = 4$. Hence, $p(\{1, 2\}) = \pi^1 = 4$.

Step 2. The two individuals with the highest probabilities in $M^1(2)$ are 1 and 3. We have $\hat{M}^1(2) = (4, 3, 1)$, so $\pi^2 = \min\{\hat{M}^1(2, 2), \beta^1(2) - \hat{M}^1(2, 3)\} = 3$. Therefore, $\pi^2(2) = (3, 0, 3)$, $M^2(2) = M^1(2) - \pi^2(2) = (0, 1, 1)$, and $\beta^2(2) = \beta^1(2) - \pi^2 = 4 - 3 = 1$. Hence, $p(\{1, 3\}) = \pi^2 = 3$.

Step 3. There only two individuals (2 and 3) left with positive probabilities, and these probabilities are the same. Hence, this group of individuals must be assigned the unallocated fatality probability, so $p(\{2, 3\}) = 1$. We confirm that the algorithm produces this result. We have $\hat{M}^2(2) = (1, 1, 0)$, so $\pi^3 = \min\{\hat{M}^2(1, 2), \beta^2(2) - \hat{M}^2(2, 2)\} = 1$. Therefore, $\pi^3(1) = (0, 1, 1)$, $M^3(2) = M^2(2) - \pi^3(2) = (0, 0, 0)$, and $\beta^3(2) = \beta^2(2) - \pi^3 = 1 - 1 = 0$. Hence, $p(\{2, 3\}) = 1$, as was to be shown.

The algorithm terminates in Step 3. All subgroups with two members are assigned positive probability.

As these examples illustrate, each step of the algorithm identifies a subgroup with k members and determines the probability that it is this group that perishes. For each individual i in this group, this probability must be subtracted from whatever part of the probability $M(k, i)$ that remains unallocated at the end of the previous step. In all three of the examples, at the end of the penultimate step of the algorithm, there is a group of size k whose members all have the same probability left to distribute. In the next section, we show that this is a general feature of the algorithm. When this amount has been allocated as the probability of this group perishing together, we have $M^t(k) = (0, \ldots, 0)$, and so the algorithm terminates because, for each $i \in N$, the probability $M(k, i)$ that person i dies in a group of size k has been distributed among each of the groups of size k that include i.

The distribution of the probability in $\beta(k)$ across the groups with k members need not be unique. This is the case in Example 2 because there is more than one way to rearrange the vector of fatality probabilities being considered in a nonincreasing way in some of the steps. For example, if in Step 2 in this example, with the tie-breaking rule used in our algorithm, individuals 1, 2, 3, and 5 are regarded as having the four highest probabilities in $M^1(4)$. However, we could have used a tie-breaking rule that selects individuals 1, 2, 5, and 6 instead, in which case $p(\{(1, 2, 5, 6)\}) > 0$, which is not the case with the tie-breaking rule used in the algorithm. Feasibility of a shared destiny risk matrix M only requires that there exists a social risk distribution p such that $M_p = M$, not that this distribution be unique.

5 The Equivalence of Admissibility and Feasibility

In order to show that an admissible shared destiny risk matrix is feasible, we first establish a number of lemmas that identify some important properties of the probability decomposition algorithm. In each of our lemmas, we suppose that $k \neq n$ and that the probability decomposition algorithm is being applied to the kth row $M(k)$ of an admissible shared destiny risk matrix M for which the probability $\beta(k)$ that there are k fatalities is positive.

Lemma 1 shows that in each step of this algorithm, analogues of (2) and the admissibility restriction in (5) hold.

Lemma 1 *In any Step t of the algorithm,*

$$\beta^t(k) = \frac{1}{k}\sum_{i=1}^{n} M^t(k,i) \tag{11}$$

and

$$0 \leq M^t(k,i) \leq \beta^t(k), \quad \forall i \in N. \tag{12}$$

Proof For any $k \neq n$, at the end of Step $t-1$ of the algorithm, from the probability $\beta(k)$ that there will be exactly k fatalities, there is still $\beta^{t-1}(k)$ left to allocate. In Step t, π^t is subtracted from the first k components of $\hat{M}^{t-1}(k)$ and 0 from the other $n-k$ components. Hence, by (2), (9), and (10), at the end of Step t, the amount from $\beta(k)$ left to allocate is (11).

Because $\pi^t \leq \hat{M}^{t-1}(k,k)$, $M^t(k,i) \geq 0$ for all $i \in N$. The argument used to show that $M^t(k,i) \leq \beta^t(k)$ for all $i \in N$ is the same as the argument used in Sect. 2 to show that (5) holds but with $M^t(k,i)$ substituting for $M(k,i)$ and $\beta^t(k)$ substituting for $\beta(k)$. □

In order for the probability decomposition algorithm to distribute *all* of the probability $\beta(k)$ that there are k fatalities among the subgroups of size k, the algorithm must terminate in a finite number of steps. Lemma 2 shows that this is the case if the algorithm reaches a step in which there are k positive entries left to distribute.

Lemma 2 *The algorithm terminates in Step $t+1$ if there are k positive entries in $M^t(k)$.*

Proof By Lemma 1, (11) and (12) hold. If $M^t(k)$ contains k positive entries, (11) and (12) imply that they are all equal to $\beta^t(k)$. Thus, the algorithm terminates in the next step because $\pi^t = \beta^t(k)$ is subtracted from $\hat{M}^t(k, i)$ for each $i = 1, \ldots, k$, and so $M^{t+1}(k) = (0, \ldots, 0)$. □

There are $\bar{n}(k)$ individuals who have a positive probability of dying in a group of size k. Lemma 3 shows that the probability decomposition algorithm terminates in a finite number of steps that does not exceed this value.

Lemma 3 *The algorithm terminates in at most $\bar{n}(k)$ steps.*

Proof If $k = \bar{n}(k)$, then Lemma 2 applies with $t = 0$, so the algorithm terminates in Step 1.

Now, suppose that $k < \bar{n}(k)$. From (7), we know that in Step t of the algorithm, π^t is either $\hat{M}^{t-1}(k, k)$ or $\beta^{t-1}(k) - \hat{M}^{t-1}(k, k+1)$, whichever is the smallest. We consider two cases distinguished by whether the first of these possibilities holds for all t or not.

Case 1. For each Step t of the algorithm, $\pi^t = \hat{M}^{t-1}(k, k)$. Then, by (7)–(9), $M^t(k)$ has at least one more 0 entry than $M^{t-1}(k)$. Thus, $M^t(k)$ has at least $n - \bar{n}(k) + t$ entries equal to 0 and, hence, has at most k positive entries in Step $\bar{n}(k) - k$. It follows from (11) and (12) (which hold by Lemma 1) that there is no Step t such that the number of positive entries in $M^t(k)$ is positive but less than k. Therefore, because the algorithm subtracts a common positive amount of probability from k individuals in each step, for some $t \le \bar{n}(k) - k$, $M^t(k)$ has exactly k positive entries, which, by Lemma 2, implies that the algorithm terminates in at most $\bar{n}(k) - k + 1$ steps. Because $k < \bar{n}(k)$, this upper bound is at most $\bar{n}$.

Case 2. In some Step t of the algorithm, $\pi^t \ne \hat{M}^{t-1}(k, k)$. Let t^* be the first step for which this is the case. By (7), we then have that $\pi^{t^*} = \beta^{t^*-1}(k) - \hat{M}^{t^*-1}(k, k+1)$. Let i^* be the individual for whom $\rho^{t^*-1}(k, i^*) = k + 1$. That is, i^* is the individual for whom $M^{t^*-1}(k, i^*) = \hat{M}^{t^*-1}(k, k+1)$. Because $\pi^{t^*} = \beta^{t^*-1}(k) - \hat{M}^{t^*-1}(k, k+1)$, by (10), $\beta^{t^*}(k) = \hat{M}^{t^*-1}(k, k+1)$. Because $M^{t^*}(k, i^*) = M^{t^*-1}(k, i^*)$, it follows that $M^{t^*}(k, i^*) = \beta^{t^*}(k)$.

By (11) and (12), there cannot be more that k entries in $M^t(k)$ which are at least as large as $\beta^t(k)$. Hence, i^* must occupy one of the first k ranks in $M^{t^*}(k)$ and so i^*'s probability is reduced by π^{t^*} in Step t^*. By (10), for all t, $\pi^t = \beta^{t-1}(k) - \beta^t(k)$. Therefore, $M^{t^*+1}(k, i^*) = \beta^{t^*+1}(k)$. Iteratively applying the same reasoning in each of the subsequent nonterminal steps of the algorithm, we conclude that $M^\tau(k, i^*) = \beta^\tau(k)$ for any Step τ for which $\tau \ge t^*$ which is not a terminal step.

Because there cannot be more that k entries in $M^\tau(k)$ which are at least as large as $\beta^\tau(k)$, we now know that for each $\tau \ge t^*$, i^* has a rank not exceeding k in $M^\tau(k)$. Hence, in any Step t^{**} for which $t^{**} > t^*$, the individual who occupies rank $k+1$ in $\hat{M}^{t^{**}-1}(k)$ is someone, say i^{**}, who is different from i^*. Reasoning as above, if $\pi^{t^{**}} \ne \hat{M}^{t^{**}-1}(k, k)$, then $M^\tau(k, i^{**}) = \beta^\tau(k)$ for any Step τ for which $\tau \ge t^{**}$ which is not a terminal step. Furthermore, both i^* and i^{**} have ranks not exceeding k in $M^\tau(k)$ for any such τ.

By an iterative application of the preceding argument, we conclude that there can be at most k steps in which $\pi^t \ne \hat{M}^{t-1}(k, k)$. Because $M^t(k)$ has at least one more 0 entry than $M^{t-1}(k)$ in each Step t for which $\pi^t = \hat{M}^{t-1}(k, k)$, there are at most $\bar{n}(k) - k - 1$ values of t for which (i) $\pi^t = \hat{M}^{t-1}(k, k)$ and (ii) there are at least $k + 1$ positive entries in $M^t(k)$. Thus, the algorithm terminates in at most $\bar{n}(k)$ steps. □

In each step of the algorithm, a group of size k is identified and assigned a probability. Lemma 4 shows that no group is considered in more than one step of the algorithm and, therefore, no group is assigned more than one probability.

Lemma 4 *No group of individuals with k members is assigned a probability in more than one step of the algorithm.*

Proof We need to show that for all Steps t and t' of the algorithm for which $t \neq t'$, $(I_1^t, I_2^t, \ldots, I_n^t) \neq (I_1^{t'}, I_2^{t'}, \ldots, I_n^{t'})$. On the contrary, suppose that there exist $t < t'$ for which $(I_1^t, I_2^t, \ldots, I_n^t) = (I_1^{t'}, I_2^{t'}, \ldots, I_n^{t'})$. Let S be the set of individuals for whom the value of these indicator functions is 1. Because both π^t and $\pi^{t'}$ are positive, by (7), we must have $\pi^t \geq \pi^t + \pi^{t'}$, which is impossible. That is, both π^t and $\pi^{t'}$ must be subtracted in the same step from the probabilities of the members of S that have yet to be allocated when this group is the one being considered. □

With these lemmas in hand, we can now prove our equivalence theorem.

Theorem *A shared destiny risk matrix M is admissible if and only it is feasible.*

Proof Because a feasible shared destiny risk matrix is necessarily admissible, we only need to show the reverse implication. Suppose that M is an admissible shared destiny risk matrix. For each $k \neq n$ for which $\beta(k) > 0$, Lemmas 3 and 4 imply that the probability decomposition algorithm assigns a probability $p(S) \in [0, 1]$ to each group $S \in \mathscr{T}(k)$ (the set of groups with k members) in such a way that $\sum_{S \in \mathscr{T}(k)} p(S) = \beta(k)$. If $k \neq n$ and $\beta(k) = 0$, we let $p(S) = 0$ for all $S \in \mathscr{T}(k)$. Because $\mathscr{T}(n) = \{N\}$, we set $p(N) = \beta(n)$. Finally, we set $p(\varnothing) = 1 - \sum_{i=1}^{n} \beta(k)$. The function $p\colon 2^n \to [0, 1]$ is therefore a social risk distribution. By construction, the corresponding shared destiny risk matrix M_p is the same as M because $M_p(k, i) = \sum_{S \in \mathscr{S}(k,i)} p(S) = M(k, i)$ for all (k, i). Hence, M is feasible. □

References

Bommier A, Zuber S (2008) Can preferences for catastrophe avoidance reconcile social discounting with intergenerational equity? Soc Choice Welf 31:415–434

Broome J (1982) Equity in risk bearing. Oper Res 30:412–414

Chew SH, Sagi J (2012) An inequality measure for stochastic allocations. J Econ Theory 147:1517–1544

Fishburn PC (1984) Equity axioms for public risks. Oper Res 32:901–908

Fishburn PC, Sarin RK (1991) Dispersive equity and social risk. Manag Sci 37:751–769

Fishburn PC, Straffin PD (1989) Equity considerations in public risks evaluation. Oper Res 37:229–239

Gajdos T, Weymark JA, Zoli C (2010) Shared destinies and the measurement of social risk equity. Ann Oper Res 176:409–424

Harvey CM (1985) Preference functions for catastrophe and risk inequity. Large Scale Syst 8:131–146

Keeney RL (1980a) Equity and public risk. Oper Res 28:527–534

Keeney RL (1980b) Evaluating alternatives involving potential fatalities. Oper Res 28:188–205

Keeney RL (1980c) Utility functions for equity and public risk. Manag Sci 26:345–353

Keeney RL, Winkler RL (1985) Evaluating decision strategies for equity of public risks. Oper Res 33:955–970

Sarin RK (1985) Measuring equity in public risk. Oper Res 33:210–217

Part II
Inequality

On the Volume of Redistribution: Across Income Levels and Across Groups

Ravi Kanbur

Abstract The optimal income taxation literature focuses on the tradeoff between the equity gains of higher progressivity versus its greater incentive costs at the individual level. This paper highlights a neglected aspect of redistribution—greater progressivity requires a higher volume of gross redistributive flows, across income levels. If these flows are costly to manage, administratively, or politically, then progressivity will be lower. Moreover, if redistribution across income levels implies redistribution across sociopolitically salient groups because of the way in which these groups line up relative to the income distribution, this can be an added cost in the objective function and progressivity is further disadvantaged. The paper develops a simple framework in which these questions can be addressed. Among the many interesting results is that when the capacity for the volume of redistributive flows, across income levels or across sociopolitical groups, is reached, an increase in market inequality can lead to a fall in progressivity in the tax-transfer regime without any change in the government's preferences for equity. A focus on the volume of redistribution thus opens up an important set of theoretical and empirical questions for analysis and for policy.

Keywords Volume of redistribution · Administrative costs of redistribution · Progressivity · Sociopolitically salient groups · Political costs of redistribution

JEL Codes D31 · D63 · H21 · H24

Paper written as contribution to Festschrift for Satya Chakravarty upon his retirement from the Indian Statistical Institute. This chapter has been published by the author under (i) the ECINEQ Society for the Study of Economic Inequality Working Paper Series as ECINEQ WP 2018–462, (ii) the CEPR Discussion Paper Series as DP 12816, and (iii) Dyson Cornell Working Papers Series as 18–03.

R. Kanbur (✉)
Applied Economics and Management, Cornell University, Ithaca, NY, USA
e-mail: sk145@cornell.edu
URL: https://www.kanbur.dyson.cornell.edu

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*, Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_4

1 Introduction

Why has post-tax and transfer[1] inequality increased in many countries around the world? In simple accounting terms, to get the post-tax distribution we start with the market distribution of income and superimpose on that the redistribution implemented by the government, to arrive at the distribution of post-tax or "take home" income. Thus, if post-tax inequality rises, it must be because of the net effect of change in the inequality of market income and change in progressivity of redistribution. For example, an often discussed narrative is that while in the US and the UK both changes went in the direction of raising inequality, in Latin America redistribution overcame increasing market inequality to reduce post-tax inequality.

In the standard Mirrleesian model of optimal nonlinear income taxation (Mirrlees 1971), the optimal degree of tax progressivity depends on three key parameters—the degree of market or "inherent" inequality[2]; the preference for equality (the government's inequality aversion); and the strength of incentive effects (captured in the model by the elasticity of labor supply). It can then be shown in such Mirrleesian models that, holding fixed government's inequality aversion and individuals' incentive effects, an increase in "inherent" inequality will increase the optimal progressivity of the tax system, although the net effect on post-tax inequality will be to increase it (Kanbur and Tuomala 1994). It is now generally agreed that the skill premium in labor markets is on the rise, the result of skill-biased technical progress,[3] and this is leading to rising market inequality at any given level of progressivity of redistribution. However, in the US and UK at least, it seems as though the tax system has become less progressive not more, compounding the effects of rising market inequality. This would seem to suggest either that incentive effects have become stronger, or that the preference for equity has declined, or both.

It can indeed be argued that in the era of globalization incentive effects have become stronger, certainly for capital and for skilled labor as relocation prospects have improved for them. However, incentive effects will not be the focus of this paper. It could also be argued that the political system has been captured by the wealthy, with the result that the tax system reflects this capture, with lowered inequality aversion in the government's objective function. Indeed, there could be a vicious spiral, whereby rising inequality leads to greater political capture and thence greater inequality still.[4] However, this type of mechanism will also not be the primary focus of this paper.

Rather, I wish to highlight the effects of what I call "the volume of redistribution." The idea is simple. Redistribution involves taking resources away from some and giving these resources to others. Both the taking and the giving will have individual-level incentive effects and these are well modeled in the economics literature. But the taking and the giving requires administrative and other mechanisms for transfer. One can visualize these as the "pipes" which take the flow of redistribution from

[1] Henceforth, we will use "tax" to mean "tax and transfer".

[2] It will be recalled that this is the inequality of the Mirrlees "ability" parameter n.

[3] See, for example, Autor (2014).

[4] See Stiglitz (2017).

one set of incomes to others. These pipes, these mechanisms, do not just exist—they have to be built. And if the pipes have been laid for an earlier period, they may not be able to take a greatly increased flow of redistribution, and begin to impose costs which militate against redistribution. We can think of these mechanisms and pipes in physical, administrative, terms; but another interpretation is the flexibility of current political economy, having arrived at a given political equilibrium level of redistribution, to now adjust to a far greater flow required by new circumstances. Although I have this at the back of my mind, I will not model the political economy, preferring at this stage to stay with the physical analogy of pipes and their ability to withstand the force of greater flow.

Suppose now that individuals differ not only in levels of income but also in characteristics which define sociopolitically salient groups, such as ethnic groups, immigrant versus natives, young versus old, regional groupings, and so on. Then, except in particular special cases, redistribution flows across income levels through an income tax-transfer system will also imply flows across these groups. If there are political costs to flows across groups, these have to be further accounted for in the social objective function and their implications for progressivity need to be worked out. How should we think about the dependence of costs on the volume of flows (across income levels or across groups)? The pipes analogy helps. For a given width of piping, more flow can be accommodated up to capacity with marginal cost of additional flow. But once this capacity is reached, there is a fixed cost in building new capacity to take the next level of flow. Such fixed costs, as might be expected, also affect the levels and patterns of progressivity in response to increasing market inequality.

The plan of the paper is as follows. Section 2 sets out the basic idea of the volume of redistribution and analyzes its dependence on distributional parameters. In particular, it traces a possible line of linkage between greater inherent inequality leading to the desire for greater progressivity in the tax and transfer system, but this being blocked by the inability of the system to handle the higher volume of flows of redistribution implied by greater progressivity. Section 3 plays out the flows perspective through the lens of the implications for transfers between politically salient groups such as ethnic groups, or natives and immigrants. Section 4 concludes with an extended discussion of the metaphor of the "volume of redistribution" which the model of the paper tries to set out in a simple and precise way. It argues that the concept opens up an interesting line of theoretical and empirical research.

2 The Volume of Redistribution

Let the market distribution of income y be represented by its density f(y). Let the tax (and transfer) function be denoted t(y), and the post-tax income by x

$$x = y - t(y). \tag{1}$$

Note that t(y) will be positive if it is a net tax and negative if it is a net transfer. The volume of redistribution is simply the aggregate of absolute values of the difference between x and y, whether positive or negative—the total flow through the redistribution pipes.

$$V = \int |x - y| f(y)dy = \int |t(y)| f(y)dy. \tag{2}$$

A particular simplification which will prove useful for us is the linear tax and transfer regime:

$$\mathrm{t(y) = -a + by}. \tag{3}$$

We have a demogrant of a for every individual and a constant marginal tax rate of b. Normalizing population size to unity, total tax revenue T is

$$\mathrm{T = -a + b\mu}, \tag{4}$$

where μ is the mean income. Sticking to a pure redistributive role for taxation and setting $\mathrm{T} = 0$, we get

$$\mathrm{a = b\mu}, \tag{5}$$

which leaves us with one free parameter in the linear tax function. We choose this to be b, the marginal tax rate, which captures the degree of progressivity of the tax system.

There is a single switch point of market income, where t changes from negative to positive. Denoting this by s, it is clear that

$$\mathrm{s} = \mu. \tag{6}$$

Thus, all those with market incomes below the mean receive a net transfer; all those with incomes above are taxed positively on net to finance those transfers. This is an obvious feature of a linear tax system with no net revenue requirement. In this setting, the volume of redistribution, as defined by (2), is given by

$$\mathrm{V} = \int_{\ddot{y}}^{\bar{y}} |y - \mu| f(y)dy = \int_{\ddot{y}}^{\mu} b(\mu - y) f(y)dy + \int_{\mu}^{\bar{y}} b(y - \mu) f(y)dy, \tag{7}$$

where $\ddot{y}$ and $\bar{y}$ are the minimum and maximum levels of income, respectively. However, since we have assumed revenue neutrality the two components of the right-hand side of (5) must be identical. Thus, in the linear case, we have the following expression for the volume of distribution:

$$V = 2b \int_{\mu}^{\bar{y}} (y - \mu) f(y) dy. \tag{8}$$

Further analytical tractability is provided by the case where f(y) is the uniform density lying between $\mu + d/2$ as maximum and $\mu - d/2$ as the minimum, so that $f(y) = (1/d)$ and d is a measure of the inequality of market income. In this case, simple integration of (8) shows that

$$V = bd/4. \tag{9}$$

Expression (9) captures in tractable form the relationship between the volume of redistribution needed when the attempted progressivity is b, and market inequality is d. We focus on the case where mean is constant, in other words, pure redistribution. In this linear, uniform, fixed mean case, market inequality measured by the variance of income is given by

$$I_y = (1/12)d^2, \tag{10}$$

and final inequality is given by

$$I_x = (1/12)(1 - b)^2 d^2. \tag{11}$$

Expressions (9), 10), and (11) provide the links we need between market inequality, tax progressivity, and volume of redistribution.

Taking d as the proxy for market inequality and b as the proxy for attempted redistribution to achieve the desired post-tax inequality, we see from (11) that an increase in market inequality requires an increase in progressivity to hold final inequality I_x constant. But from (9) we see that an increase in progressivity for any given d will increase the required volume of redistribution. In fact, there is an interaction between market inequalities in determining the volume of required redistribution:

$$\frac{\partial V}{\partial b} = d/4. \tag{12}$$

Thus, the higher the degree of market inequality, the greater is the redistribution volume increase required for a given increase in progressivity.

If there were no other costs, then an inequality averse government would simply choose to equalize all incomes with a 100% marginal tax rate and a demogrant equal to mean income. But the key assumption of this paper is that redistribution volume is not simply available to policy-makers but needs costly construction of administrative and political infrastructure—the "pipes." Let the per capita cost of volume V of redistribution be γV. Then one specification of social welfare combines mean income, variance of final income, and cost of the volume of redistribution:

$$W = \mu - \beta I_x - \gamma V. \tag{13}$$

Substituting from (9) and (11) and maximizing with respect to b gives us an expression for the optimal degree of progressivity.

$$b^* = 1 - (3\gamma)/(2\beta d). \tag{14}$$

From (7), the volume of distributive effort for this level of progressivity is given by

$$V^* = (d/4) - (3/8)(\gamma/\beta), \tag{15}$$

and with this response, final inequality is given by

$$I_x^* = (3/16)\gamma^2/\beta^2. \tag{16}$$

Final inequality is higher, the higher is the cost of redistribution. Optimal progressivity increases with market inequality but decreases with the costs of redistribution. In this case, an increase in market inequality leads to just enough increase in progressivity to leave final inequality unchanged.

Suppose now that the cost of the volume for redistribution is nonlinear, γV^2. Then, the expressions corresponding to (13), (14), (15), and (16) are as follows:

$$W = \mu - \beta I_x - \gamma V^2 \tag{17}$$

$$b^* = 1/[1 + 3\gamma/4\beta], \tag{18}$$

$$V^* = d/[4(1 + 3\gamma/4\beta)], \tag{19}$$

$$I_x^* = (1/12)(d^2)[(3\gamma/4\beta)/(1 + 3\gamma/4\beta)]^2. \tag{20}$$

In this case, optimal progressivity is independent of market inequality but decreases in the cost of redistribution. Final inequality is increasing in market inequality and in the cost of redistribution.

The argument above shows that use of progressivity as a measure of "redistributive effort" may be misleading. When market inequality increases, even with progressivity unchanged, the volume of distributive effort increases. Indeed, it must do so to keep progressivity constant. Unchanged redistribution volume will imply a decrease in progressivity. Thus, in many ways, an appropriate measure of redistributive effort (within the progressive taxation regime) is in fact the volume of redistribution. From (18) and (19), it is seen that with rising market inequality the flows through the pipes have to be greater to maintain progressivity at the optimal level—one has to run harder to keep still.

This feature, in which the degree of progressivity and volume of redistribution may not move together when market inequality increases, appears even more sharply when the cost function for volume of redistribution takes a different form. Up to now, we have supposed that the costs of increasing distributive flow are all marginal costs—a little bit more redistributive effort can be achieved at a little bit more cost. But what if some of these costs are in the nature of fixed costs? If the increased flow required is substantial, or crosses a critical threshold, then new investment may be needed, new pipes need to be installed, for the increased flow.

Let the cost function be γV *for* $V \leq \hat{V}$ *and* $F + \gamma V$ *for* $V > \hat{V}$. Then from (15), as d increases up to

$$\hat{d} = 4\hat{V} - 3\gamma/2\beta, \tag{21}$$

the cost function remains at γV, and the optimal volume and progressivity are given by (15) and (14). However, as d crosses the threshold to values higher than $\hat{d}$, the fixed cost component F kicks in. If this additional cost was not present, optimal policy would simply follow along (14) and (15). However, with the additional cost, the impact on welfare is quite different. To see this more clearly, rewrite the problem as one of choosing V rather than b, and rewrite the objective function (13) in terms of V by writing I_x in terms of V using (9):

$$\begin{aligned} \mathrm{W} &= \mu - \beta(1/12)[1 - (4\mathrm{V/d})]^2\mathrm{d}^2 - \gamma\mathrm{V} \quad \text{for} \quad V \leq \hat{V} \\ \mathrm{W} &= \mu - \beta(1/12)[1 - (4\mathrm{V/d})]^2\mathrm{d}^2 - \gamma\,\mathrm{V} - \mathrm{F} \quad \text{for} \quad V > \hat{V}. \end{aligned} \tag{22}$$

Differentiating each portion of this with respect to V and setting equal to zero gives the solution (15). Now, beginning with $d = \hat{d}$ and $V = \hat{V}$, the corresponding optimal volume of redistribution, it is clear that the optimal policy when d increases marginally is to stay at $V \leq \hat{V}$, since the marginal benefits from increasing V are zero but in doing so the fixed cost F is incurred.

But consider now the implications of the result above that V stays fixed at $\hat{V}$ as d increases. From (9), this must mean that b decreases. In other words, the degree of progressivity of the tax and transfer system as we usually measure it, the marginal tax rate, falls, and so from (11) inequality of post-tax income rises for two reasons—because market inequality rises and progressivity falls. As argued in the Introduction, this is the narrative that has played out in the US and the UK over the past three decades. Notice, however, that distributive effort as measured by the volume of redistribution remains constant.

Eventually, as market inequality becomes so high that the gain to the social welfare function from increasing the value of redistribution dominates the fixed (and marginal) cost of the higher volume beyond the current level. It can be shown that this level of market inequality is the (higher) solution in d to the quadratic formed by equating the optimized value of (20) for $V > \hat{V}$ with its value when $V = \hat{V}$

$$\mu - \beta(1/12)\left[1 - \left(4\hat{V}/\mathrm{d}\right)\right]^2\mathrm{d}^2 - \gamma\hat{V} = \mu + (3/16)\left(\gamma^2/\beta\right) - (\gamma\mathrm{d}/4) - \mathrm{F}. \tag{23}$$

If such a solution exists, denoting it $\tilde{d}$, we get a further interesting phenomenon. Progressivity and the volume of redistribution both jump up to the values given by (14) and (15) for the now higher value of $\tilde{d}$. We can thus see a cycle emerging, which is intuitively clear once the fixed costs of adjustment are factored in. Starting with a given system of pipes for redistribution, as market inequality increases, progressivity and volume both increase to mitigate the market inequality (in the linear marginal costs case, post-tax inequality is held constant as in (16)). However, once these pipes become strained and new pipes have to built, the fixed cost of this keeps the volume constant, progressivity declining, and compounding the rise in market inequality. However, once market inequality gets sufficiently high, the fixed cost is worth paying and there is a big jump in volume and progressivity, and the cycle can start again from this point onward. Such processes could perhaps explain long cycles and sudden jumps in tax progressivity and redistributive effort.

3 Group Divisions

The metaphor I have used in motivating the costs of the volume of redistribution is primarily an "administrative/infrastructure" one. The visualization of flows through pipes has been useful here. I have also indicated, but not developed, a political economy metaphor. Starting from the thinking that a "political settlement" is needed for redistribution of those with market income to those without, the costs of greater redistribution can be thought of, in a very reduced form way, as the costs of achieving the new political settlement. Although I have not and will not model this political economy here, I believe it has considerable political appeal. But the political settlement metaphor also raises another important issue, that of transfers across politically salient groups.

The rise of far-right Xenophobic parties in Europe and elsewhere is often predicated on the appeal to the notion that some groups, usually ethnic minorities, are takers from society. In particular, the platform of these political entrepreneurs is that some long-established ethnic groups, or newly arrived immigrant groups, get transfers from the existing political settlement. This is not just the argument of demagogues. Albeit in more measured tones, academics like Miller (2016) and Collier (2013) also argue that greater heterogeneity has the potential to undermine the social redistributive contract which has characterized the post-war political settlement in much of Western Europe and to some extent the US. The issue here is not redistribution from rich to poor, which was or is the current social contract, but redistributing in favor of certain identifiable groups who are poor.

It should be noted that this perspective on transfers across groups as an impediment to redistribution is very different from the literature which views group-specific information and group—contingent tax and transfer policies as being an advantage. At least since Akerlof (1978), the idea of "tagging" an individual with easily observable characteristics and implementing separate tax and transfer schedules for each tag is seen as overcoming informational disadvantages and providing the policymaker with more instruments. The idea has been applied to targeting of anti-poverty transfers (Kanbur 1987, 2017), and to nonlinear income taxation more generally (Immonen et al. 1998; Kanbur and Tuomala 2016). But these very same transfers across groups, whether intentional or not, are seen in the new dispensation as politically problematic and undermining agreement on redistribution in general.

The simple model developed in the previous section can be used to highlight and sharpen some of these concerns in a precise way. Let there be two groups in society A and B. These two "tags" can be found at different points in the income distribution, but the tags have salience in and of themselves, irrespective of the income of the individual to whom they attach. Specifically, let us suppose that flows across these two groups are of political salience and, in effect, impose a cost on attempts to redistribute income generally. The costs of these cross-group flows G have to be added to the costs of the flows across income levels, V.

Clearly, for any income tax-transfer regime, the implication for cross-group flows will depend on how the groups are distributed across the income distribution. To fix ideas, take the basic model of Sect. 2, suppose that the two groups are of equal size, and suppose that all those above the mean μ are of group A and all those below the mean are of group B. Then the cross-group flow G is simply the volume of redistributive flow V. At the other end, suppose that the groups are to be found equally at every income level. Then the cross-group flow G is zero. In effect, each group is representative of the whole society so redistribution can be seen as taking place within each group and none across groups. This is true even when the groups are not of equal size, so long as their representation at each income level is the same as their representation in the whole population. In between, as the representation of group B below the mean increases relative to its population share, cross-group flow increases from zero to V.

For simplicity, return to the case of equal group size overall, but let representation of group B be θ below the mean and (1-θ) above the mean. In other words, at each income level below the mean, a fraction θ of the population is of group B and at each income level above the mean, the fraction is (1-θ). We focus on the case where $\theta \geq$ (1/2). In this case, group A is taxed by an amount θ(1/2)V and receives an amount (1-θ)(1/2)V, so the net flow from this group out is (2θ-1)(1/2)V, whereas the net receipt for group B is the mirror of this, (1-2θ)(1/2)V. Thus, the sum of the net outflow and the net inflow in absolute terms is

$$G = (2\theta - 1)V. \tag{24}$$

This is the offending cross-group flow with political salience and a political cost for policy-makers.

If we represent the cost of cross-group flow in the usual linear manner with marginal cost δ, we have a social welfare function analogous to (13)

$$W = \mu - \beta I_x - \gamma V - \delta G = \mu - \beta I_x - [\gamma + \delta(2\theta - 1)]V = \mu - \beta I_x - \gamma^* V. \tag{25}$$

This is simply the social welfare function in (11) with an augmented marginal cost of the volume of cross-income flow, denoted by γ^*, which is the marginal cost of volume γ plus the additional term (2θ-1) which comes from the cost of cross-group flow:

$$\gamma^* = \gamma + \delta(2\theta - 1) \tag{26}$$

All of the previous analysis now goes through with γ replaced by (the higher) γ^*. Thus, from (14), (15), and (16), progressivity and volume are lower, and final inequality is higher with the cross-group factor added in. Notice that the more heavily represented is group B in the lower half of the population (the higher is θ), the higher will be the cost of a unit of flow across income classes, because this will now involve more flows between politically salient groups, and the lower will be the progressivity and higher will be the final inequality. Clearly, the gist of the analysis also goes through when the groups are of unequal size; what will matter then is the representation of group B individuals in the below mean income population, relative to their representation in the population as a whole. The key point is that if an ethnic minority or an immigrant group, say, is concentrated at lower income levels, then this will reduce progressivity. This matches the analysis of Tabellini (2017, pp. 38–39) which shows for US jurisdictions, “the inflow of immigrants” led “cities to cut tax rates and limit redistribution.”

Analogously to the discussion on fixed costs for volume of transfer across income levels, we can now consider what happens when a political settlement reaches its limit in terms of the amount of cross-group transfers it will permit. Let this be denoted $\hat{G}$. Then, from (24), there is a corresponding critical value for volume of redistribution across income levels:

$$\hat{V} = \hat{G}/[2\theta - 1]. \tag{27}$$

Going beyond $\hat{G}$ requires a new settlement, which can be represented by a fixed cost. A similar analysis can then be carried out as in (22) and (23) with $\hat{V}$ given by (27). Then beyond a critical value of market inequality $\hat{d}$, as in (21) but with γ replaced by γ^*, the volume of distribution will stay fixed at (27) and as market inequality rises, progressivity will decline and final inequality will rise. But these effects are now coming not from the fixed costs of managing increasing redistribution across income levels, but from the fixed costs of a new political settlement to manage the redistribution across politically salient groups to which redistribution across income levels gives rise as a corollary.

Is (25) the right way to represent political tensions in flows across groups? It is based on the idea that it is the total flows that matter. But what if per capita flows matter; in other words, it is the amount given by the typical person and the amount received by the typical person which matter. In this case, the right correction factor is not (2θ-1) but (2θ-1)/θ = [2 – (1/θ)]. The same arguments still go through. A higher θ, i.e., a greater representation of group B among below mean income individuals, still increases the correction factor on γ. However, the correction factor is greater for the relevant range of $1 > \theta > (1/2)$. Per capita perceptions will lead to greater perceived cost of redistribution and thus lower progressivity and higher final inequality for any given degree of market inequality.

4 Conclusion

The optimal income taxation literature focuses on the tradeoff between the equity gains of higher progressivity versus its greater incentive costs at the individual level. This paper highlights a neglected aspect of redistribution—greater progressivity requires a higher volume of gross redistributive flows, across income levels. If these flows are costly to manage, administratively or politically, then progressivity will be lower. Moreover, if redistribution across income levels implies redistribution across sociopolitically salient groups because of the way in which these groups line up relative to the income distribution, this can be an added cost in the objective function and progressivity is further disadvantaged. When the capacity for the volume of redistributive flows, across income levels or across sociopolitical groups, is reached, increase in market inequality can lead to a fall in progressivity in the tax-transfer regime without any change in the government's preferences for equity.

The term "capacity of redistribution" has been used in the literature, but in a different sense to the one used in this paper. Thus, Ravallion (2010, p. 1) defines it as "the marginal tax rate (MTR) on the 'rich'—defined as those living in a developing country who would not be considered poor by rich country standards—that is needed to provide the revenue for a specific redistribution." Hoy and Sumner (2016) also apply the same measure to updated and more extensive data. Kanbur and Mukherjee (2007, pp 52–53) have a similar perspective when they characterize poverty reduction failure as "is the extent of poverty relative to the resources available in the society to eradicate it?" Thus, they all highlight the resources available for redistribution. But none of these papers focus on the gross flows needed to achieve a given redistribution and the cost associated with these flows.

How might we think of the costs of the volume of gross flows needed for redistribution across income levels? The easiest interpretation is in terms of administrative costs. Not surprisingly perhaps, these costs are often highlighted by economists more oriented to the free market:

> Some fraction of each dollar taxed will always be absorbed in wages and salaries of the administrative bureaucracy, costs of purchasing, powering, maintaining and replacing equipment, buildings, etc., and other overhead costs. Only the remainder will actually be received by the target population in the form of cash or in kind payments…… Using government data, Woodson (1989, p. 63) calculated that, on average, 70 cents of each dollar budgeted for government assistance goes not to the poor, but to the members of the welfare bureaucracy and others serving the poor. Tanner (1996, p. 136 n. 18) cites regional studies supporting this 70/30 split. (Edwards 2007, pp. 3–4).

One issue with simply calculating the manpower costs of the "welfare bureaucracy" is to separate out the simple cash shifting function of administration from that part of the function which provides direct services—the first is more like our costs of redistribution. Nevertheless, even if the administrative costs were significantly lower than the 70/30 split, they are not negligible.

In the developing country context, the literature on targeting of transfers for poverty reduction has often remarked on the administrative costs of "fine targeting." Caldes et al. (2006), for example, calculate the cost of making a one-unit transfer to a beneficiary, the "cost–transfer ratio" for a range of Latin American transfer programs. They find a wide range, with a low of 4% but a high of 25%, the range depending on how finely the program attempted to target the poor.

One perspective on the costs of an attempted volume of redistribution is provided by corruption. The former Indian Prime Minister Rajiv Gandhi is famously said to have remarked that only 15% of the outlay on the public food distribution system reached the poor. More formal estimates are provided by Olken (2006) for a particular program in Indonesia:

> I find that corruption is substantial—the central estimate is that at least 18% of the subsidized rice in the Indonesian program I study went missing.... The estimates suggest that corruption in developing countries such as Indonesia may substantially inhibit a government's ability to carry out redistributive programs, particularly in rural areas. (Olken 2006, p. 867).

Correspondingly on the taxation side, there is a literature on the "compliance gap" in tax revenue raising (Keen and Slemrod 2017).

The issue of fixed versus marginal costs of redistributive flows is not addressed very much in the administrative costs literature, partly because of the lack of sufficiently disaggregated data to allow allocation of costs. Caldes et al. (2006) do mention that some programs have low average costs of transfer because of economies of scale. But the general idea that managing redistribution can hit capacity constraints, and creation of new capacity will incur fixed costs before additional redistribution can be handled, needs deeper empirical investigation.

Another area for deeper investigation, this time theoretical, is the political economy interpretation of the capacity for redistribution. We need models which can make precise and test the intuition advanced in this paper that (i) greater redistribution incurs greater political cost even within a given political settlement and (ii) once that capacity is reached, a new political settlement, with its higher costs, is needed to increase the capacity for redistribution. The political economy interpretation is clearly the appropriate one for the costs of cross-group flows introduced in this paper. There is of course a significant literature on the impact of population heterogeneity on economic and distributional outcomes. Alesina et al. (1999) find lower public spending in more ethnically diverse jurisdictions in the US; Dahlberg et al. (2012) find negative effects of increase immigration on support for redistribution in Sweden; and Tabellini (2017) finds that immigrant inflow led US cities to "cut tax rates and limit redistribution." From a different perspective, Dasgupta and Kanbur (2007) advance a theory of why cross-group flows might induce group tensions.

A central point made in this paper is that redistribution of income across income levels is also redistribution of income across sociopolitically salient groups when these groups are spread unrepresentatively across the income distribution. A focus purely on the costs of flows across income levels—micro-level individual incentive effects as in the optimum income taxation literature, or macro-level costs of managing and administering gross flows as emphasized in this paper—may prove to be incomplete and thus misleading. If greater progressivity in taxes and transfers across income levels also leads to, say, redistribution across natives and immigrants, the political costs of this will have to be borne in mind by economists in analyzing and in designing tax and transfer regimes. The framework in this paper provides a start.

A focus on the *volume of redistribution, across income levels and across groups*, thus opens up an important set of theoretical and empirical questions for analysis and for policy.

References

Akerlof G (1978) The economics of "Tagging" as applied to the optimal income tax, welfare programs, and manpower planning. Am Econ Rev 68(1):8–19

Alesina A, Baqir R, Easterly W (1999) Public goods and ethnic divisions. Q J Econ 114(4):1243–1284

Autor D (2014) Skills, education, and the rise of earnings inequality among the "other 99 percent". Science 344:843–851

Caldes N, Coady D, Maluccio J (2006) The cost of poverty alleviation transfer programs: a comparative analysis of three programs in Latin America. World Dev 34(5):818–837

Collier P (2013) Exodus: How migration is changing our world. Oxford University Press

Dahlberg M, Edmark K, Lundqvist H (2012) Ethnic diversity and preferences for redistribution. J Polit Econ 120(1):41–76

Dasgupta I, Kanbur R (2007) Community and class antagonism. J Public Econ 91(9):1816–1842

Edwards JR (2007) J Libert Stud 21(2):3–20

Hoy C, Sumner A (2016) Global poverty and inequality: Is there new capacity for redistribution in developing countries? J Glob Dev 7(1):117–157

Immonen R, Kanbur R, Keen M, Tuomala M (1998) Tagging and taxing: the optimal use of categorical and income information in designing tax/transfer schemes. Economica 65(258):179–192

Kanbur R (1987) Measurement and alleviation of poverty: with an application to the impact of macroeconomic adjustment, IMF Staff Pap 34(1):60–85

Kanbur R (2017) The digital revolution and targeting public expenditure for poverty reduction. In: S. Gupta et. al. (Eds.) Digital revolutions in public finance, International Monetary Fund

Kanbur R, Muherjee D (2007) Poverty, relative to the ability to eradicate it: an index of poverty reduction failure. Econ Lett Sci Direct 97(1):52–57

Kanbur R, Tuomala M (1994) Inherent inequality and the optimal graduation of marginal tax rates (with M. Tuomala). Scand J Econ 96(2):275–282

Kanbur R, Tuomala M (2016) Groupings and the gains from targeting. Res Econ 70:53–63

Keen M, Slemrod J (2017) Optimal tax administration. IMF Working Paper WP/17/8

Miller D (2016) Strangers in Our Midst: the political philosophy of immigration. Harvard University Press, Cambridge, MA

Mirrlees JA (1971) An exploration in the theory of optimum income taxation. Rev Econ Stud 38(175–208):1971

Olken BA (2006) Corruption and the costs of redistribution: micro evidence from Indonesia. J Public Econ 90:853–870

Ravallion M (2010) Do poorer countries have less capacity for redistribution? J Glob Dev 1(2), Article 1

Stiglitz J (2017) Inequality, stagnation and market power: the need for a new progressive era. Roosevelt Institute Working Paper. Roosevelt Institute

Tabellini M (2017) Gifts of the immigrants, woes of the natives: lessons from the age of mass migration. http://economics.mit.edu/files/13646

Tanner M (1996) The end of welfare. Cato Institute, Washington, D.C

Woodson RL (1989) Breaking the poverty cycle: private sector alternatives to the welfare state. Commonwealth Foundation, Harrisburg, Penn

Fuzzy Inequality Ranking Relations and Their Crisp Approximations

Asis Kumar Banerjee

Abstract This paper is concerned with inequality ranking of income distributions. Ranking a pair of distributions in terms of inequality is a difficult exercise excepting in the special case where one of these Lorenz dominates the other. To get around the difficulty, the notion of a *fuzzy inequality ranking relation* (FIRR) has been fruitfully used in the literature. However, it is natural to ask how, from a given FIRR, we can derive a *crisp approximation* that would give us an overall crisp (i.e., non-fuzzy) judgment regarding the inequality ranking of any pair of distributions. It turns out that this can be done in many different ways (leading to different crisp judgments) and, therefore, there arises the question which of these alternative ways is to be chosen. It is seen that this question is closely linked to the more general one as to how we can decide which income vectors in any given finite set A of such vectors can be considered to be least unequal in an overall crisp assessment. In this paper, we first axiomatically characterize a specific procedure for making this decision. A particular procedure for obtaining a crisp approximation of an FIRR then follows as a corollary. We also construct an illustrative example of an FIRR that does not seem to have appeared in the literature before and compare it with some other known examples.

Keywords Fuzzy inequality ranking relations · Crisp approximations · Inequality dominance function · Lorenz dominance

JEL Classification D60 · D63

This paper is dedicated to Satya Chakravarty. I am grateful to Prasanta Pattanaik for his detailed and incisive comments on an earlier version of the paper. Useful discussions were also held with Dipankar Dasgupta, Indraneel Dasgupta, Pradip Maity, Mihir Rakshit, Susmita Rakshit, and Soumyen Sikdar. The usual caveat applies.

A. K. Banerjee (✉)
University of Calcutta, 56A, B.T. Road, 700050 Kolkata, India
e-mail: asisbanerjee.cu@gmail.com

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*,
Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_5

1 Introduction

Measurement of inequality of the distribution of income in an economy is one of the oldest areas of economic research. Over the years, many important things have been learnt about the issues that arise in course of this exercise. Cowell (2000), Jenkins and VanKerm (2009) and Sen (1997) are some of the well-known expositions of these issues.

In this paper, we shall be concerned with *inequality ranking* of income distributions. Given any two such distributions **x** and **y**, we shall ask whether one of these is a less unequal distribution than the other. Similarly, given a set of distributions, we shall ask which member of the set, if any, can be considered to be least unequal. Questions of this type arise in many contexts and their practical importance constitutes the motivation behind this paper.

Inequality ranking, however, may turn out to be a difficult exercise in many cases. In fact, the only ranking principle on which there seems to be wide agreement among economists is that if two income distributions **x** and **y** are such that **x** Lorenz dominates **y** (i.e., if the Lorenz curve for **x** is not anywhere below that for **y**), then **x** is to be considered to be definitely no more unequal than **y**. However, this does not give us a *complete* ranking of all income distributions in terms of inequality. There can be distributions **x** and **y** whose Lorenz curves intersect. In such cases which of the two distributions are to be considered less unequal is not clear.[1]

The possibility of intersecting Lorenz curves points toward the fact that in many situations it is not possible to obtain *unambiguous* inequality rankings of income distributions. Indeed, as has long been long recognized in the literature, a degree of *vagueness* is inherent in the very notion of inequality ranking (see, especially, Basu 1987, Ok 1996 and Sen 1992, 1997).

The ambiguity inherent in inequality comparisons makes a case for applying the tools of *fuzzy relations* for the purpose. Basu (1987) and Ok (1996) are among the relatively few applications of such relations in this area. Basu (1987) derived a particular fuzzy inequality ranking relation (FIRR) from a set of necessary and sufficient conditions. Ok (1996) contained a useful study of the properties of a broader class of FIRRs.

Following this approach, we shall, in this paper, assume that inequality ranking relations are fuzzy. However, while the fuzzy relations' approach is motivated by the need to recognize the ambiguity that lies at the core of the inequality ranking

[1]As stated in the text, our primary concern in this paper is with inequality *ranking*. Therefore, we do not use inequality *indices*. However, it may be noted in passing that using an inequality index for the purpose of obtaining an inequality ranking does not constitute a satisfactory solution of the problem referred to in the text. While any given inequality index would induce an inequality ranking, the problem is that there is no uniquely defined inequality index. In fact, there is an infinity of such indices. A given pair of income distributions may not be ranked in the same way by different inequality indices. Which specific inequality index is chosen for the purpose of the comparison, therefore, becomes a crucial question and cannot be settled in principle. Again, however, the problem would vanish in the special case where there is Lorenz dominance. If **x** Lorenz dominates **y**, then I(**x**) $\leq$ I(**y**) for *all* inequality indices I (see Foster 1985). The converse is also true.

exercise, it is, nevertheless, natural to ask whether we can make an *overall* crisp (i.e., non-fuzzy) judgment regarding inequality ranking of the income distributions, i.e., whether there is a *crisp* ranking that can in some sense be considered to be an *approximation* to a given FIRR.

There does not, however, seem to be an obvious and unique procedure for obtaining a crisp approximation of an FIRR. Many alternative procedures can (and have been) suggested. In this paper, we seek to obtain a specific procedure from more basic considerations.

For this purpose, however, we need a formal definition of what is meant by a crisp approximation of an FIRR. Formulating such a definition turns out to be a nontrivial task. In particular, the matter is seen to be closely related to a more general question. Given a *set* of income distributions (possibly with more than two members) and given the fact that inequality ranking is fuzzy, how do we arrive at a *crisp judgment* as to which members of the set are least unequal?

We shall first seek an answer to this broader question by imposing conditions on the procedure of making such judgments. For convenience, however, we shall limit the scope of our enquiry in some respects. We shall ask how, given an FIRR, R, on the (infinite) set X of all possible distributions of a *fixed* total income among a *finite* number of individuals and given any *finite* subset A of X, we can decide which members of the set A are least unequal in an overall (crisp) assessment.

Any given *procedure* of making this decision will be called an *inequality dominance function* (IDF). For any FIRR, R, satisfying some intuitively reasonable consistency conditions, we formulate a definition of the "crisp approximation of R generated by an IDF." The main part of the paper proposes a specific IDF and characterizes it. The crisp approximation procedure that follows as a corollary is the one that assigns, to any R, the crisp approximation generated by this particular IDF. The suggested crisp approximation procedure is seen to coincide with a procedure that has already been mentioned in the literature. For any given R, its crisp approximation obtained in this way is seen to be an ordering (i.e., a reflexive, complete, and transitive relation) on X.

Section 2 below introduces the notations and some definitions. Section 3 contains a characterization of the proposed IDF. The suggested crisp approximation procedure follows as an immediate corollary. Section 4 provides some examples of fuzzy inequality ranking relations. Section 5 concludes the discussion.

2 Notations and Definitions

Consider an economy with n individuals and an amount T of total income that is distributed among the individuals. Throughout this paper, n and T are kept fixed. An income vector **x** is a nonnegative n-vector, whose ith entry x_i denotes the amount of income allocated to the ith individual, $i = 1, 2, \ldots, n$. Hence, $\sum_{i=1}^{n} x_i = T$. The set of all income vectors is denoted by X.

For convenience we assume that, for any income vector **x**, the entries in **x** are arranged in nondecreasing order (ties being broken arbitrarily): $x_1 \leq x_2 \leq \cdots \leq x_n$.

We are interested in obtaining a ranking of the income vectors in terms of their degrees of inequality. For instance, given any two income vectors **x** and **y** in X, we wish to answer the question which of the two vectors represents a less unequal distribution of the total income. (We do not rule out the possibility that they have the same degree of inequality, i.e., they are "equally unequal.") Similarly, given a set A of income vectors (A $\subseteq$ X), we wish to know which members of A are least unequal. However, we shall confine ourselves to the case where A is a *finite* subset of X.

By the weak Lorenz dominance relation, denoted by L, we mean the binary relation on X such that, for all **x** and **y** in X, **x** L **y** if and only if $\sum_{i=1}^{k} x_i \geq \sum_{i=1}^{k} y_i$ for $k = 1, 2, \ldots, n$. The strict Lorenz dominance relation is the asymmetric component of L and will be denoted by L_P.

An alternative statement of the relation uses the notion of the *Lorenz curve.* For any **x** in X let

$$L(\mathbf{x}, k/n) = \sum_{i=1}^{k} x_i/T \text{ for all } k = 1, 2, \ldots, n.$$

As per the standard Gastwirth (1971) definition of a Lorenz curve, the Lorenz curve of **x** is the curve obtained by letting L(**x**, 0) = 0 and joining adjacent points of the form (*k*/*n*, L(**x**, *k*/*n*)) by line segments. Equivalently, for any given vector **x,** let $L_{\mathbf{x}}(p)$ denote the proportion of the total available income going to the bottom (i.e., the poorest) *p* proportion of the population. The Lorenz curve of **x** is obtained by plotting $L_{\mathbf{x}}(p)$ against *p*. Needless to say, *p* ranges from 0 to 1 and, for any **x**, $L_{\mathbf{x}}(0) = 0$ and $L_{\mathbf{x}}(1) = 1$.

It is then easily seen that, for any two income vectors **x** and **y** in X, **x** L **y** if and only if the Lorenz curve of **x** is not below that of **y** for any $k = 1, 2, \ldots, n$ and **x** L_P **y** if and only if the Lorenz curve of **x** is not below that of **y** for any *k* and is above it for at least one *k*.

In our framework (where T and *n* are the same for all income distributions), **x** L **y** if and only if [either **x** is a permutation of **y** or **x** is obtained from **y** by a finite sequence of progressive income transfers]. (see Hardy, Littlewood, and Polya 1952 and Marshall and Olkin 1979). Since the order in which the incomes of the individuals are listed in an income vector should be of no consequence for distributional judgments (i.e., if **x** is a permutation of **y**, then **x** and **y** should be judged to have the same degree of inequality), this result is the basis for the usual interpretation of Lorenz dominance in terms of inequality ranking: if **x** L **y**, then **x** can be judged to be unambiguously no more unequal than **y**.

The solution to the problem of obtaining unambiguous inequality rankings that the Lorenz dominance relation provides, however, is obviously *partial*. This dominance relation is *incomplete*. There may exist income vectors **x** and **y** that are *Lorenz* *incomparable* (i.e., that are such that $\neg(\mathbf{x}\,\mathrm{L}\,\mathbf{y})$ and $\neg(\mathbf{y}\,\mathrm{L}\,\mathbf{x})$, i.e., the Lorenz curves of **x** and **y** intersect). Unambiguous inequality rankings are not possible in such cases.

To deal with such cases, attempts have been made to extend the idea of inequality ranking using the theory of fuzzy relations. Under this approach, we seek to suggest a *fuzzy* inequality ranking relation that would be complete. Recall that, from the mathematical point of view, a relation can be looked upon as a set. The notion of fuzzy sets was introduced in mathematics independently by Zadeh (1965) and Klaua (1965).[2] A fuzzy set S is a subset of some universal set U for which there is a mapping f_S (say) from U to the closed unit interval [0, 1]. Informally, for any u in U, if $f_S(u) = 1$, u definitely belongs to S. If $f_S(u) = 0$, it definitely does not belong to S. But there can be intermediate values of $f_S(u)$ indicating "degrees of membership." This approach has been fruitfully applied to other areas of economics, especially, to the theory of choice based on fuzzy preferences.

To introduce this approach in the present context, let R be a *fuzzy* binary relation on X (i.e., a mapping from $X \times X$ into the closed interval [0, 1] on the real line). For any **x** and **y** in X, we shall interpret $R(\mathbf{x}, \mathbf{y})$ as the extent to which it is true that "**x** is no more unequal than **y**." $R(\mathbf{x}, \mathbf{y})$ is 1 if and only if it is definitely true; it is 0 if and only if the statement is definitely false. However, we permit $R(\mathbf{x}, \mathbf{y})$ to take intermediate values between 0 and 1.

In view of the relation (discussed above) between Lorenz dominance and inequality rankings of income vectors, it seems natural to require that R should respect the Lorenz dominance criterion.

Definition 2.1 R respects Lorenz dominance if and only if, for all **x** and **y** in X, $R(\mathbf{x}, \mathbf{y}) = 1$ if and only if $\mathbf{x}\,\mathrm{L}\,\mathbf{y}$ and $R(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{y}\,\mathrm{L_P}\,\mathbf{x}$.

We also desire R to satisfy the fuzzy versions of the usual consistency conditions (i.e., reflexivity, completeness, and transitivity) on a weak crisp relation. R is called *reflexive* if, for all **x** in X, $R(\mathbf{x}, \mathbf{x}) = 1$. It is called *complete* if, for all **x** and **y** in X, $R(\mathbf{x}, \mathbf{y}) + R(\mathbf{y}, \mathbf{x}) \geq 1$.

However, there does not seem to be an agreed definition of *transitivity* of a fuzzy relation. Many different definitions (all of which are consistent with the notion of transitivity in crisp theory) have been proposed. A widely used notion in the mathematical theory of fuzzy sets is that of *max-min transitivity*. R is max-min transitive if, for all **x**, **y**, and **z** in X, $R(\mathbf{x}, \mathbf{z}) \geq \min\,[R(\mathbf{x}, \mathbf{y}, R(\mathbf{y}, \mathbf{z})]$. However, it has been shown that if a fuzzy inequality ranking relation respects Lorenz dominance, it cannot be max-min transitive (see Ok 1996, Theorem 5.4, p. 523).

In this paper, we shall use the following notion of f-transitivity which seems to be intuitively fairly transparent. R is *f-transitive* if, for all **x**, **y**, and **z** in X, $[R(\mathbf{x}, \mathbf{y}) \geq R(\mathbf{y}, \mathbf{x}), R(\mathbf{y}, \mathbf{z}) \geq R(\mathbf{z}, \mathbf{y})]$ implies $R(\mathbf{x}, \mathbf{z}) \geq R(\mathbf{z}, \mathbf{x})$.

Max-min transitivity and f-transitivity are independent conditions. A variant of the condition of f-transitivity is that of *strong transitivity* due to Kolodziejczyk (1986).

---

[2]Klaua's contribution was published in German. A recent analysis of the contribution in English is by Gottwald (2010).

R is strongly transitive if, for all $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ in X, [R($\mathbf{x}$, $\mathbf{y}$) > R($\mathbf{y}$, $\mathbf{x}$), and R($\mathbf{y}$, $\mathbf{z}$) > R($\mathbf{z}$, $\mathbf{y}$)] implies R($\mathbf{x}$, $\mathbf{z}$) > R($\mathbf{z}$, $\mathbf{x}$). It has been shown that strong transitivity is weaker than max-min transitivity. Several other definitions of fuzzy transitivity have been proposed in the literature. For a discussion of a number of these notions and the implicational relations between them, see Ok (1996).

Definition 2.2 A *fuzzy inequality* ranking relation (FIRR) is a fuzzy binary relation on X that respects Lorenz dominance and is reflexive, complete, and f-transitive.

Given an FIRR, R, the fuzzy relations P and J will denote its asymmetric and symmetric components, respectively. Asymmetry of P is taken to mean that, for all $\mathbf{x}$ and $\mathbf{y}$, P($\mathbf{x}$, $\mathbf{y}$) +P($\mathbf{y}$, $\mathbf{x}$) $\leq$ 1. (This is motivated by crisp theory where P($\mathbf{x}$, $\mathbf{y}$) and P($\mathbf{y}$, $\mathbf{x}$) cannot both be 1 so that the weak inequality is valid there.) Symmetry of J means that J($\mathbf{x}$, $\mathbf{y}$) = J($\mathbf{y}$, $\mathbf{x}$) for all $\mathbf{x}$ and y in X. In the present context, for all $\mathbf{x}$ and $\mathbf{y}$, P($\mathbf{x}$, $\mathbf{y}$) is interpreted to mean the extent to which it is true that $\mathbf{x}$ is less unequal than $\mathbf{y}$; J($\mathbf{x}$, $\mathbf{y}$) is the extent to which $\mathbf{x}$ and $\mathbf{y}$ are equally unequal.

However, how to decompose a fuzzy relation R into its asymmetric and symmetric components is, again, a question on which fuzzy set theorists do not seem to agree. We shall follow Barrett and Pattanaik (1989) (BP) and assume that, for all reflexive and complete R and for all $\mathbf{x}$ and $\mathbf{y}$ in X,

$$\mathrm{P}(\mathbf{x}, \mathbf{y}) = 1 - \mathrm{R}(\mathbf{y}, \mathbf{x}) \text{and}\, \mathrm{J}(\mathbf{x}, \mathbf{y}) = \mathrm{R}(\mathbf{x}, \mathbf{y}) + \mathrm{R}(\mathbf{y}, \mathbf{x}) - 1.$$

Note that this procedure implies that if $\mathbf{x}$ and $\mathbf{y}$ are distinct members of X and are such that R($\mathbf{x}$, $\mathbf{y}$) + R($\mathbf{y}$, $\mathbf{x}$) = 1, then P($\mathbf{x}$, $\mathbf{y}$) = R($\mathbf{x}$, $\mathbf{y}$) and J($\mathbf{x}$, $\mathbf{y}$) = 0.

For a discussion of other decomposition rules that have been suggested in the literature and the reasons for which the BP procedure may be considered to be intuitively more reasonable than some of the others, see Llamazares (2005).

Let $\mathbf{A}$ denote the set of all non-empty *finite* subsets of X and let $\mathbf{R}$ be the set of all FIRRs on X.

Definition 2.3 An inequality dominance function (IDF) is a function D: $\mathbf{A} \times \mathbf{R} \rightarrow \mathbf{A}$ such that, for all A in $\mathbf{A}$ and all R in $\mathbf{R}$, D(A, R) $\subseteq$ A.

For all admissible A and R, D(A, R) is interpreted to be the (crisp) set of those income vectors in A that are no more unequal than any member of A as judged by R.

An IDF can be defined in many different ways. We shall be interested in a specific IDF which we denote by D^0.

Definition 2.4 D^0 is a mapping from $\mathbf{A} \times . \mathbf{R}$ into the class of all subsets of X such that, for all admissible A and R, D^0(A, R) = {$\mathbf{x}$ ε A: R($\mathbf{x}$, $\mathbf{y}$) $\geq$ R($\mathbf{y}$, $\mathbf{x}$) for all $\mathbf{y}$ ε A**}.**

Lemma D^0 is an IDF.

Proof For any admissible A and R, D^0(A, R) is, by definition, a subset of X. Therefore, it suffices to show that, for any such A and R, D^0(A, R) is non-empty. For that purpose, let A = {$\mathbf{x}^1$, $\mathbf{x}^2$, ….., $\mathbf{x}^q$} for some positive integer q where $\mathbf{x}^i$ is in X for all i = 1,2,…,q. Let Q denote the set of integers {1, 2,…, q} If $\mathbf{x}^1$ is such that, for all i in Q, R($\mathbf{x}^1$, $\mathbf{x}^i$) ≥ R($\mathbf{x}^i$, x^1), the proof ends. If this is not the case, then there exists i in Q such that R($\mathbf{x}^i$, $\mathbf{x}^1$) > R($\mathbf{x}^1$, $\mathbf{x}^i$). Since R is reflexive, $\mathbf{x}^i \neq \mathbf{x}^1$. Without loss of generality, let $\mathbf{x}^i = \mathbf{x}^2$. Thus, we have

$$R\left(\mathbf{x}^2, \mathbf{x}^1\right) > R(\mathbf{x}^1, \mathbf{x}^2). \qquad \text{(i)}$$

If, now, R($\mathbf{x}^2$, $\mathbf{x}^i$) ≥ R($\mathbf{x}^i$, $\mathbf{x}^2$) for all i in Q, the proof ends. If this is not true, then, for some i in Q, R($\mathbf{x}^i$, $\mathbf{x}^2$) > R($\mathbf{x}^2$, $\mathbf{x}^i$). Again, reflexivity of R implies $\mathbf{x}^i \neq \mathbf{x}^2$. Moreover, if $\mathbf{x}^i = \mathbf{x}^1$, then the inequality (i) above is contradicted. Thus, $\mathbf{x}^i$ is distinct from both $\mathbf{x}^1$ and $\mathbf{x}^2$ and we can assume it to be $\mathbf{x}^3$ without loss of generality. We then have

$$R\left(\mathbf{x}^3, \mathbf{x}^2\right) > R\left(\mathbf{x}^2,\ \mathbf{x}^3\right) \qquad \text{(ii)}$$

Since R is f-transitive, (i) and (ii) imply

$$R\left(\mathbf{x}^3, \mathbf{x}^1\right) \geq R\left(\mathbf{x}^1,\ \mathbf{x}^3\right). \qquad \text{(iii)}$$

If, now, it is not the case that R($\mathbf{x}^3$, $\mathbf{x}^i$) ≥ R($\mathbf{x}^i$, $\mathbf{x}^3$) for all i in Q, then, for some such i, R($\mathbf{x}^i$, $\mathbf{x}^3$) > R($\mathbf{x}^3$, $\mathbf{x}^i$). Using (ii), (iii), and the reflexivity of R, it is now seen that $\mathbf{x}^i$ is distinct from $\mathbf{x}^1$, $\mathbf{x}^2$, and $\mathbf{x}^3$. Let i = 4.

Proceeding in this way, suppose that we reach a member of A which is distinct from each of the first (q–2) members and, therefore, can be called $\mathbf{x}^{q-1}$ without loss of generality and which is such that it is not the case that R($\mathbf{x}^{q-1}$, $\mathbf{x}^i$) ≥ R($\mathbf{x}^i$, $\mathbf{x}^{q-1}$) for all i in Q. Then, it must be the case that R($\mathbf{x}^q$, $\mathbf{x}^{q-1}$) > R($\mathbf{x}^{q-1}$, $\mathbf{x}^q$). By repeated application of the fact that R is f-transitive, it can now be seen that R($\mathbf{x}^q$, $\mathbf{x}^i$) ≥ R($\mathbf{x}^i$, $\mathbf{x}^q$) for all i in Q. □

We shall be interested in characterizing the function D^0. Given any FIRR, R, we shall then be able to answer the question which crisp relation on X can be considered to be a crisp approximation of R.

To state the next definition in this connection, we use the fact that a crisp binary relation S on X is a fuzzy binary relation such that, for all **x** and **y** in X, S(**x**, **y**) is either 0 or 1. We shall write **x** S **y** if and only if S(**x**, **y**) = 1. Let S be a crisp relation on X. **x** S **y** will be interpreted to mean that **x** is no more unequal than **y**. We shall assume that S is an ordering (although weaker assumptions such as the one that it is reflexive, complete, and acyclic would also suffice for our purposes). Let **S** be the set of all such crisp relations. Let the mapping D*: **A** × **S** → **A** be such that, for all admissible A and S, D*(A, S) = {**x** ε A: **x** S **y** for all **y** ε A}. Thus, for any such A and S, D*(A, S) is the set of those members of A which are no more unequal than any member of A as per the crisp relation S.

Definition 2.5 Given any R in **R** and given any IDF, D, a relation R^C in **S** is called "the crisp approximation of R generated by D" if and only if, for all A in **A**,

$$D * \left(A, R^C\right) = D(A, R).$$

□

In other words, for R^C to be considered to be the crisp approximation of R generated by D, it is necessary and sufficient that, for all admissible A, the (crisp) set of members of A that are considered by D to be least unequal as per the fuzzy relation R is precisely the set of least unequal members of A as per the crisp elation R^C.

Since **A** includes all finite subsets of X and, therefore, all pairs of members of X, it is easily seen that, for any admissible R and D, the crisp approximation of R generated by D is unique.

The crisp approximation of R generated by D^0 will be denoted by R^0. It is seen that this crisp approximation procedure coincides with a procedure that has already been suggested in the literature. As per this suggested procedure, for any given R, its crisp approximation is the crisp relation C_R (say) on X defined as follows.

Definition 2.6 For any R in **R** and for any **x** and **y** in X, **x** C_R **y** if and only if R(**x**, **y**) ≥ R(**y**, **x**).

In the literature, C_R has been called the "fuzzy dominance" relation generated by R and we shall use this terminology. Thus, **x** fuzzy dominates **y** in terms of R if and only if the extent to which it is true that **x** is no more unequal than **y** (as per the given R) is not less than the extent to which it is true that **y** is no more unequal than **x** (as per R). As a notion of a crisp approximation of R, the fuzzy dominance relation C_R has intuitive appeal. As Ok (1996, p. 519) observes, "Fuzzy dominance might be viewed as the originator (or, the best predictor) of actual decisions one has to reveal in some situations (like when comparing income distributions)." In view of Definition 2.5 it follows that, for any R in **R**, $R^0 = C_R$.

3 Characterization of D^0

Let D be an IDF. Consider the following conditions on D.

Condition 3.1: For all **x** and **y** in X and for all R in **R**, [R(**x**, **y**) ≠ R(**y**, **x**)] implies [D({**x**, **y**}, R) ≠ {**x**, **y**}].
Condition 3.2: For all A in **A** and for R in **R**, if **[**{**x**, **y**} ⊆ A, **x** ε D(A, R) and R(**y**, **x**) ≥ R(**x**, **y**)**]**, then **y** ε D(A, R).
Condition 3.3: For all R in **R** [**x** ε A ⊂ B and **x** ε D(B, R)] implies **x** ε D(A, R).

Condition 3.1 states that, for any pair of admissible income vectors **x** and **y**, if R(**x**, **y**) ≠ R(**y**, **x**), then the IDF D would not declare both **x** and **y** to be the least unequal members of the finite set {**x**, **y**}. Thus, for all **x** and **y** in X, and for all admissible R, [D({**x**, **y**}, R) = {**x**, **y**}] implies [R(**x**, **y**) = R(**y**, **x**)]. The converse, however, is not necessarily true. Condition 3.2 states that if both **x** and **y** are in a finite set A, if **x** is declared by D to be a least unequal member of A and if and R(**y**, **x**) $\geq$ R(**x**, **y**), then **y** also must be considered to be a least unequal member of A. This is similar to some "congruence" conditions used in revealed preference theory. However, in the present context, R is a prespecified FIRR (rather than a revealed preference relation). Condition 3.3 is a familiar "independence" condition. It is a condition of internal consistency on the IDF, D: it states that any income vector that is declared to be a least unequal member of a finite set B of vectors as per a relation R in **R** must also be considered by R to be a least unequal member in any subset A of B, provided that it belongs to A.

In the following proposition, we prove that the three conditions stated above are independent.

Proposition 3.1 Conditions 3.1–3.3 are independent.

Proof For each of the three conditions, we cite the example of an IDF that violates that particular condition while satisfying the other three.

(i) *Violation of Condition 3.1*: Consider the IDF, D, for which, for any finite subset A of X and for any R in **R**,

$$\begin{aligned} \mathrm{D(A, R)} &= \{\mathbf{x}\,\varepsilon\,\mathrm{A} : \mathrm{R}(\mathbf{x},\ \mathbf{y}) \geq 0.5 \text{ for all } \mathbf{y}\,\varepsilon\,\mathrm{A}\} \text{ if } \#\mathrm{A} \leq 2 \text{ and} \\ &= \{\mathbf{x}\,\varepsilon\,\mathrm{A} : \mathrm{R}(\mathbf{x}, \mathbf{y}) \geq \mathrm{R}(\mathbf{y}, \mathbf{x}) \text{for all } \mathbf{y}\,\varepsilon\,\mathrm{A}\} \text{otherwise.} \end{aligned}$$

D satisfies Conditions 3.2 and 3.3. To check Condition 3.2, suppose that **x** and **y** are in A, **x** ε D(A, R) and R(**y**, **x**) $\geq$ R(**x**, **y**). If #A = 2, i.e., A = {**x**, **y**}, then, by definition of D, R(**x**, **y**) $\geq$ 0.5. Hence, R(**y**, **x**) $\geq$ 0.5 so that **y** ε D(A, R) as is required by Condition 3.2. On the other hand, if #A $\geq$ 3, then [**x** ε D(A, R)] implies R(**x**, **z**) $\geq$ R(**z**, **x**) for all **z** in A. Since R is f-transitive, we have R(**y**, **z**) $\geq$ R(**z**, **y**) for all **z** in A. The definition of D for this case, therefore, again implies that **y** ε D(A, R).

To check Condition 3.3, suppose that **x** ε A $\subseteq$ B for some **x** in X and some A and B in **A** and that **x** ε D(B, R). To show that **x** ε D(A, R), note that if A has at least 3 members, then so does B so that we have R(**x**, **y**) $\geq$ R(**y**, **x**) for all **y** in B. Hence, the same inequality is valid for all **y** in A, implying the desired conclusion. If #A = 2 and #B $\geq$ 3, then the same argument again establishes that R(**x**, **y**) $\geq$ R(**y, x**) for all **y** in A. Moreover, since R is complete, R(**y**, **x**) $\geq$ 1 – R(**x**, **y**). Thus, R(**x**, **y**) $\geq$ 0.5, implying that **x** ε D(A, R).

However, Condition 3.1 is violated. For instance, if a reflexive, complete, and f-transitive R is such that, for some **x** and **y** in X, R(**x**, **y**) = 0.7 and R(**y, x**) = 0.5, then D({**x**, **y**}, R) = {**x**, **y**} but R(**x**, **y**) $\neq$ R(**y, x**).

(ii) *Violation of Condition 3.2*: Consider the following IDF. For all R in **R** and for all A in **A**,

$$\text{If } \#\mathrm{A} \leq 2, \text{ then } \mathrm{D(A, R)} = \{\mathbf{x}\,\varepsilon\,\mathrm{A} : \mathrm{R}(\mathbf{x}, \mathbf{y}) \geq \mathrm{R}(\mathbf{y}, \mathbf{x}) \text{for all } \mathbf{y}\,\varepsilon\,\mathrm{A}\} \qquad \text{(a)}$$

$$\text{If}\#A \geq 3, \text{then } D(A, R) = \{\mathbf{x}\,\varepsilon\, A : R(\mathbf{y}, \mathbf{x}) \leq 0.5 \text{ for all } \mathbf{y}\,\varepsilon\, A\} \tag{b}$$

provided that the set on the r.h.s.of (b) is non-empty; otherwise, D(A, R) is as in (a).

It is easily seen that D satisfies Condition 3.1. To prove that it satisfies Condition 3.3, let **x** ε A ⊂ B for some admissible A and B and let **x** ε D(B, R). We have to show that **x** ε D(A, R). If #B = 2, the conclusion is trivial since A is then a singleton and since R is reflexive. Hence, let #B ≥ 3.

Consider first the case where we also have #A ≥ 3. Since **x** ε D(B, R), either R(**y**, **x**) ≤ 0.5 for all **y** in B or R(**x**, **y**) ≥ R(**y**, **x**) for all such **y**. In the former case, since A ⊂ B, it follows that R(**y**, **x)** ≤ 0.5 for **y** in A so that **x** ε D(A, R). A similar remark applies to the latter case.

Suppose now that #A = 2. Note that the completeness of R implies that {**x** ε A: R(**y**, **x**) ≤ 0.5 for all **y** ε A}is a subset of {**x** ε A: R(**x**, **y**) ≥ R(**y**, **x**) for all **y** ε A} irrespective of the cardinality of A. Hence, if **x** ε D(B, R), then, in both of the cases referred to in the preceding paragraph, we shall have R(**x**, **y**) ≥ R(**y**, **x**) for all **y** ε A. Hence, **x** ε D(A, R) by Eq. (a) above.

However, consider the set A = {**x**, **y**, **z**} where **x**, **y,** and **z** are in X. Consider the R in **R** whose restriction to A is as follows: R(**x**, x) = R(**y**, **y**) = R(**z**, **z**) = 1; R(**x**, **y**) = 0.5 = R(**y**, **x**); R(**x**, **z**) = 0.6, R(**z**, **x**) = 0.5; R(**y**, **z**) = 0.7, R(**z**, **y**) = 0.6. It is seen that **x** ε D(A, R), R(**y**, **x**)≥ R(**x**, **y**) but **y** is not in DA, R) since ¬ [R(**z**, **y**) ≤ 0.5] although the set {**x** ε A: R(**z**, **x**) ≤ 0.5 for all **z** ε A} is non-empty. Thus, Condition 3.2 is violated.

(iii) *Violation of Condition 3.3*: Let **z*** in X be such that for all **x** in A other than **z*,** R(**x**, **z***) = 1 and R(**z***, **x**) = 0. Consider the IDF, D, for which, for all A in **A**,

$$\begin{aligned} D(A, R) &= \{\mathbf{x}\,\varepsilon\, A : R(\mathbf{x}, \mathbf{y}) \geq R(\mathbf{y}, \mathbf{x}) \text{for all } \mathbf{y}\,\varepsilon\, A\} \text{if } \mathbf{z}* \text{ is not in A,} \\ &= A - \{\mathbf{z}*\} \text{if } \mathbf{z}* \text{ is in A and A is not a singleton and} \\ &= A \text{ if } A = \{\mathbf{z}*\}. \end{aligned}$$

Note that, under this IDF, **z*** is not considered to be a least unequal member of any set A in **A,** whenever A includes any other income vector. Thus, if D({**x**, **y**}, R) = {**x**, **y**}, then both **x** and **y** are distinct from **z***. Hence, the definition of D for this case implies that R(**x**, **y**) = R(**y**, **x**). Hence, Condition 3.1 is satisfied.

To show that Condition 3.2 is satisfied, suppose that **x** and **y** are in an admissible A, **x** ε D(A, R) and R(**y**, **x**) ≥ R(**x**, **y**). We are to show that **y** ε D(A, R). If **z*** is not in A, then **x** ε D(A, R) implies that R(**x**, **z**) ≥ R(**z**, **y**) for all **z** ε A. Hence, R(**y**, **x**) ≥ R(**x**, **y**) implies that R(**y**, **z**) ≥ R(**z**, **y**) for all **z** ε A since R is f-transitive. Thus, **y** ε D(A, R), completing the proof. If **z*** is in A, then, since A is not a singleton, all members of A other than **z*** are in D(A, R). The desired conclusion follows from the fact that one such member of A is **y**. To see this, note that since **x** ε D(A, R), **x** ≠ **z***. Therefore, y = **z*** would imply that R(**y**, **x**) = 0 and R(**x**, **y**) = 1, contradicting the hypothesis that R(**y**, **x**) ≥ R(**x**, **y**). Hence, **y** ≠ **z***.

However, consider the set A = {**x**, **y**, **z***} where **x**, **y,** and **z*** are distinct. Consider the R in **R** whose restriction to A is as follows: R(**x**, x) = R(**y**, **y**) = R(**z**, **z**) = 1; R(**x**, **y**) = 0.6, R(**y**, **x**) = 0.4; R(**x**, **z**) = 0.6, R(**z**, **x**) = 0.4; R(**y**, **z**) = 0.6, R(**z**, **y**) = 0.4. Since D(A, R) = {**x**, **y**} but D({**x**, **y**}, R) = {**x**}, Condition 3.3 is violated. □

The following proposition gives a characterization of the IDF given by D^0 in terms of Conditions 3.1–3.3.

Proposition 3.2 D is an IDF satisfying Conditions 3.1–3.3 if and only if $D = D^0$.

Proof Only if:

Suppose that D is an IDF and that it satisfies Conditions 3.1–3.3. If $D \neq D^0$, then, for some admissible A and R, either (i) D(A, R) includes a member **y** of A such that R(**z**, **y**) > R(**y**, **z**) for some **z** in A or (ii) it excludes a member **y** of A for which R(**y**, **z**) ≥ R(**z**, **y**) for all **z** in A.

In Case (i), on the one hand, since {**y**, **z**} ⊂ A and **y** ε D(A, R), by Condition 3.3 we have **y** ε D({**y**, **z**}, R). On the other hand, since R(**z**, **y**) > R(**y,** z), we must have D({**y**, **z**}, R) = {**z**} because, otherwise, D({**y**, **z**}, R) is equal to either {**y**, **z**} or {**y**}. However, if D({**y**, **z**}, R) = {**y**, **z**}, then by Condition 3.1, R(**y**, **z**) = R(**z**, **y**) which contradicts the hypothesis while, if D{**y**, **z**}, R) = {**y**}, then Condition 3.2 is violated since R(**z**, **y**) > R(**y**, **z**). Thus, Case (i) is ruled out.

Consider Case (ii) now. By definition of an IDF, **z** ε D(A, R) for some **z** in A. Since {**y**, **z**} ⊂ A, and since, in the case under consideration, R(**y**, **z**) ≥ R(**z**, **y**), by Condition 3.2, we now have **y** ε D(A, R) which contradicts the hypothesis.

If:

Let $D = D^0$. For any **x** and **y** in X, and for any admissible R, if D({**x**, **y**}, R) = {**x**, **y**}, then, by definition of the mapping D^0, we have R(**x**, **y**) ≥ R(**y**, **x**) and R(**y**, **x**) ≥ R(**x**, **y**) so that R(**x**, **y**) = R(**y**, **x**), verifying that Condition 3.1 is satisfied. To check Condition 3.2, let **x** and **y** in X, A in **A** and R in **R** be such that [{**x**, **y**} ⊆ A, **x** ε D(A, R) and R({**y**, **x**} ≥ R(**x**, **y**)]. To show that **y** ε D(A, R), note that, since $D = D^0$, [**x** ε D(A, R)] implies R(**x**, **z**) ≥ R(**z**, **x**) for all **z** in A. Together with the hypothesis R(**y**, **x**) ≥ R(**x**, **y**) and the fact that R is f-transitive, this implies that R(**y**, **z**) ≥ R(**z**, **y**) for all **z** in A so that we have the desired conclusion. That D satisfies Condition 3.3 is easily seen. □

Having derived D^0 number of conditions on the IDF that seem to be intuitively acceptable, we shall henceforth focus on this particular IDF. In particular, in obtaining a crisp approximation of a given R in **R**, we shall concentrate on the crisp approximation that is obtained if the IDF, D, from which it is generated satisfies Conditions 3.1, 3.2, and 3.3.

We are now in a position to state the following result.

Corollary to Proposition 3.2: Let D satisfy Conditions 3.1, 3.2, and 3.3. For any R in **R**, its crisp approximation generated by D is the crisp relation R^0 on X for which, for all **x** and **y** in X, **x** R^0 **y** if and only if R(**x**, **y**) ≥ R(**y**, **x**).

In view of Definitions 2.4 and 2.5 of Sect. 2, the corollary is an immediate consequence of Proposition 3.2. Since R is reflexive, f-transitive, and real-valued, it also follows that R^0 is reflexive, complete, and transitive, i.e., that it is an ordering on X.[3]

If $D = D^0$ (as it is in the above Corollary), $D^*(A, R^C) = C(A, R)$ for all A in **A** if and only if this equality is valid for all sets A such that $\#A \leq 2$. However, this is not true in general. Consider, for instance, the IDF given by the mapping D specified by Eqs. (a) and (b) in part (ii) of the proof of Proposition 3.1. This is a well-defined IDF. However, if this mapping D is used for the purpose of generating the crisp approximation R^C of a given R, it is obvious that $R^C \neq R^0$ although $D^*(A, R^0) = D(A, R)$ for all A such that #A is at most 2. The reason is that this equality is not necessarily valid if A has three or more members.

It may also be noted that, in determining the crisp approximation of a given FIRR, R, the fuzzy transitivity property of R also plays a crucial role. For instance, if the condition of f-transitivity of R is replaced by max-min transitivity and if the IDF, D, in Definition 2.5 is taken to be such that, for all admissible A and R, $D(A, R) = \{\mathbf{x}\ \varepsilon\ A: R(\mathbf{x}, \mathbf{y}) \geq \alpha$ for all $\mathbf{y}\ \varepsilon\ A\}$ where α is any real number in (0. ½], then the crisp approximation of any given R would be the crisp relation R^C on X such that $\mathbf{x}\ R^C\ \mathbf{y}$ if and only if $R(\mathbf{x}, \mathbf{y}) \geq \alpha$. Max-min transitivity of R would ensure that R^C is an ordering. As has been remarked before, however, in the context of fuzzy inequality relations that respect Lorenz dominance, max-min transitivity is not an intuitively reasonable condition on R while f-transitivity is.

4 Examples of FIRRs

So far, we have discussed FIRRs and their crisp approximations in general terms. We end this paper by suggesting a specific FIRR. We also mention a number of other FIRRs that are to be found in the literature.

In our framework, one possible measure of the extent to which **x** Lorenz dominates **y** would be a measure of the set $\{p\ \varepsilon\ [0, 1]: L_{\mathbf{x}}(p) > L_{\mathbf{y}}(p)\}$. In the present case, a natural measure of this type would seem to be *length* of that portion of the horizontal axis of the usual Lorenz box diagram over which the Lorenz curve of **x** lies above

[3]It may be noted that our identification of "fuzzy dominance" as the crisp approximation rule has been a consequence of our identification of a specific IDF. Ok (1994) took a different (and, from the mathematical point of view, a more direct) route in the context of a very similar issue in the theory of *preference relations* (I am indebted to Efe Ok for bringing this reference to my attention.). A decision-maker's preference relation R on a set of alternatives may be fuzzy but we may wish to know whether it has a crisp approximation. Interpreting a crisp approximation of a given fuzzy relation R as a crisp relation which is "nearest" to R, Ok considered *metrics* (as measures of "nearness") on the space of fuzzy relations. Different metrics would identify different crisp relations as being nearest to a given fuzzy relation. The paper introduced a particular metric and proved that it leads to the "fuzzy dominance" relation generated by R as the crisp approximation of R. However, while this is an important proposition, it has not been reported whether its converse is true, i.e., whether it constitutes a characterization of the fuzzy dominance relation as a crisp approximation procedure.

that of $\mathbf{y}$. This will be denoted by $N_{\mathbf{xy}}$. For example, if $L_{\mathbf{x}}(p)$ and $L_{\mathbf{y}}(p)$ intersect at a point where $p = 0.7$ and if $L_{\mathbf{x}}(p) > L_{\mathbf{y}}(p)$ for p in (0, 0.7) while $L_{\mathbf{x}}(p) < L_{\mathbf{y}}(p)$ for p in (0.7, 1), then $N_{\mathbf{xy}} = 0.7$ and $N_{\mathbf{yx}} = 0.3$ so that $N_{\mathbf{xy}} > N_{\mathbf{yx}}$.

Consider now the fuzzy binary relation R on X such that, for all $\mathbf{x}$ and $\mathbf{y}$ in X,

$$\begin{aligned} R(\mathbf{x}, \mathbf{y}) &= 1 \text{ if } \mathbf{x} = \mathbf{y} \text{ and} \\ &= N_{\mathbf{xy}} / \left(N_{\mathbf{xy}} + N_{\mathbf{yx}}\right) \text{ otherwise.} \end{aligned} \tag{4.1}$$

Since population size and total income are fixed, the Lorenz curves of $\mathbf{x}$ and $\mathbf{y}$ cannot coincide if $\mathbf{x}$ and $\mathbf{y}$ are distinct. It is seen that, for all such $\mathbf{x}$ and $\mathbf{y}$, exactly one of the following three statements is true: (i) $\mathbf{x}$ strictly Lorenz dominates $\mathbf{y}$, (ii) $\mathbf{y}$ strictly Lorenz dominates $\mathbf{x}$, and (iii) the Lorenz curves of $\mathbf{x}$ and $\mathbf{y}$ intersect. In all of these cases at least one of the two nonnegative numbers $N_{\mathbf{xy}}$ and $N_{\mathbf{yx}}$ is positive and, therefore, so is their sum. Thus, for all $\mathbf{x}$ and $\mathbf{y}$, $R(\mathbf{x}, \mathbf{y})$ is well defined and is in [0, 1]. R is obviously reflexive. It is complete since, for all $\mathbf{x}$ and $\mathbf{y}$ in X, $R(\mathbf{x}, \mathbf{y}) + R(\mathbf{y}, \mathbf{x})$ is either 1 or 2. It is also seen that, for all $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ in X, if $N_{\mathbf{xy}} \geq N_{yx}$ and $N_{\mathbf{yz}} \geq N_{\mathbf{zy}}$, then $N_{\mathbf{xz}} \geq N_{\mathbf{zx}}$. Thus, R is f-transitive.

Moreover, if $R(\mathbf{x}, \mathbf{y}) = 1$, then either $\mathbf{x} = \mathbf{y}$ or [$\mathbf{x} \neq \mathbf{y}$ but $N_{\mathbf{yx}} = 0$]. In the former case, the Lorenz curves of $\mathbf{x}$ and $\mathbf{y}$ coincide and in the latter it is seen that $\mathbf{x}$ strictly Lorenz dominates $\mathbf{y}$. In both circumstances, therefore, $\mathbf{x}$ weakly Lorenz dominates $\mathbf{y}$. Conversely, if $\mathbf{x}$ weakly Lorenz dominates $\mathbf{y}$ then either $\mathbf{x} = \mathbf{y}$ or [$\mathbf{x} \neq \mathbf{y}$ but $N_{\mathbf{yx}} = 0$]. Hence, $R(\mathbf{x}, \mathbf{y}) = 1$. Similarly, it can be checked that $R(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{y}$ strictly Lorenz dominates $\mathbf{x}$. Thus, R is a fuzzy binary relation on X that respects Lorenz dominance and is reflexive, complete, and f-transitive. Hence, it is an FIRR as per Definition 2.2 of Sect. 2.

The crisp approximation R^0 of R is such that, for all $\mathbf{x}$ and $\mathbf{y}$ in X, $\mathbf{x}\ R^0\ \mathbf{y}$ if and only if $N_{\mathbf{xy}} \geq N_{\mathbf{yx}}$. It is an ordering on X.

A characterization of the specific FIRR, R, specified by (4.1) can be obtained with the help of a system of axioms similar in form to that used in Basu (1987). In that paper, for all $\mathbf{x}$ and $\mathbf{y}$ in X, $E_{\mathbf{xy}}$ was defined to be $\sum_{j=1}^{n} \max(\sum_{i=1}^{j} \left(x_i - y_i\right), 0)$. In the familiar Lorenz diagram, $E_{\mathbf{xy}}$ is the area of the dominance of $\mathbf{x}$ over $\mathbf{y}$. Basu ignored the case where $\mathbf{x}$ and $\mathbf{y}$ are nondistinct and worked with the *strict* fuzzy ranking relation P. Consider a fuzzy relation P on X representing "less unequal than." Let $\mathbf{R}_+^2$ be the nonnegative orthant of the two-dimensional Euclidean space The crucial assumption in Basu's approach was that there exists a mapping φ from $\mathbf{R}_+^2$ into [0, 1] such that, for all $\mathbf{x}$ and $\mathbf{y}$ in X that are distinct from the perfectly egalitarian distribution and also distinct from one another,

$$P(\mathbf{x}, \mathbf{y}) = \varphi\left(E_{\mathbf{xy}}, E_{\mathbf{yx}}\right). \tag{4.2}$$

By combining this condition with three others, the following fuzzy relation P was characterized[4]: For all **x** and **y** in X with the characteristics mentioned above,

$$\mathrm{P}(\mathbf{x}, \mathbf{y}) = \mathrm{E}_{\mathbf{xy}}/(\mathrm{E}_{\mathbf{xy}} + \mathrm{E}_{\mathbf{yx}}).$$

Basu (1987) also formulated a definition of a "nearest crisp relation" of a given fuzzy relation and showed that the crisp relation (P*, say) nearest to the fuzzy strict relation characterized by him is essentially the Gini ordering in the sense that, for all distinct **x** and **y**, **x** P* **y** if and only if G(**x**) < G(**y**) where G is the Gini index.[5]

It can be checked that if we replace (4.2) by the condition that for all **x** and **y** in X, if $\mathbf{x} = \mathbf{y}$, $\mathrm{R}(\mathbf{x}, \mathbf{y}) = 1$ and, otherwise,

$$\mathrm{R}(\mathbf{x}, \mathbf{y}) = \mathrm{f}\big(\mathrm{N}_{\mathbf{xy}}, \mathrm{N}_{\mathbf{yx}}\big),$$

then, together with the three other conditions used in Basu (1987) (with the relation P interpreted as the asymmetric component of R), we would obtain a characterization of the relation R in (4.1). Note that since, for all **x** and **y** in X such that $\mathbf{x} \neq \mathbf{y}$, $\mathrm{R}(\mathbf{x}, \mathbf{y}) + \mathrm{R}(\mathbf{y}, \mathbf{x}) = 1$ so that, as per our rule for defining the asymmetric component of R (discussed in Sect. 2), $\mathrm{P}(\mathbf{x}, \mathbf{y}) = \mathrm{R}(\mathbf{x}, \mathbf{y})$ for all such **x** and **y**. Since the proof of this characterization result is exactly analogous to that of Basu's result, it is omitted. Ok (1996, p. 523) studied (albeit without characterization) a broader class of fuzzy relations which includes the relation suggested by Basu as a special case.[6]

5 Summary and Conclusion

In this paper, we have been concerned with the question how to obtain a crisp approximation of any given fuzzy inequality ranking relation. In seeking to formalize the notion of a suitable approximation procedure, however, we were naturally led to the broader question as to how we can decide which income vectors in any finite set A of

[4]For the proof, Basu did not need to assume that the strict fuzzy ranking relation (or the weak relation R of which it is the asymmetric component) is f-transitive. He used a very weak transitivity condition which requires that, for any **x**, **y,** and **z** in X, $[\mathrm{R}(\mathbf{x}, \mathbf{y}) = 1$ and $\mathrm{R}(\mathbf{y}, \mathbf{z}) = 1]$ implies $\mathrm{R}(\mathbf{x}, \mathbf{z}) = 1$. However, it is known that the relation that is characterized is actually f-transitive. See Ok (1996, pp. 525–526) where it is shown that it also satisfies a different fuzzy transitivity condition called O-transitivity, a condition proposed in Ovchinnikov (1986).

[5]This particular feature of the fuzzy relation suggested by Basu (1987) is hardly surprising since the stated objective of that paper was the fuzzification of what was called the "Gini-Lorenz" framework of inequality measurement. The crisp approximation of the fuzzy relation suggested in Eq. (4.1) in the text is not related in this way to the Gini ranking. We do not enter into the question whether it is analogously related to the (crisp) inequality ranking derived from any *other* inequality index. Indeed, since in this paper we are concerned exclusively with inequality *rankings*, the notion of an inequality *index* is not of much relevance here.

[6]All members of the class are both f-transitive and O-transitive (see Note 4).

such vectors can be considered to be least unequal in an overall (crisp) assessment, given that the underlying inequality ranking relation on the set of all income vectors is fuzzy. We have axiomatically characterized a specific procedure for making this decision under the assumption that the fuzzy inequality ranking satisfies some intuitively appealing consistency conditions. It is seen that our characterization result leads to a specific crisp approximation procedure which has already been mentioned in the literature. We have also constructed an illustrative example of a fuzzy inequality ranking relation that does not seem to have appeared in the literature before and have compared it with some other known ranking relations.

References

Barrett CR, Pattanaik PK (1989) Fuzzy sets, choice and preferences: some conceptual issues. Bull Econ Res 41:229–253

Basu K (1987) Axioms for a fuzzy measure of inequality. Math Soc Sci 14:275–288

Cowell FA (2000) Measurement of inequality. In: Atkinson AB, Bourguignon F (eds) Handbook of income distribution. North-Holland, pp 67–166

Foster JE (1985) Inequality measurement. In: Young HP (ed) Fair allocation. American Mathematical Society, Providence, RI, pp 31–68

Gastwirth JL (1971) A general definition of the Lorenz curve. Econometrica 39:1037–1039

Gottwald S (2010) An early approach toward graded identity and graded membership in set theory. Fuzzy Sets Syst 161:2369–2379

Hardy GH, Littlewood JE, Poliya G (1952) Inequalities, 2nd edn. Cambridge University Press, London

Jenkins SP, VanKerm P (2009) The measurement of economic inequality. In: Nolan B, Salverda W, Smeeding T (eds) Oxford handbook of economic inequality. Oxford University Press, pp 40–68

Klaua D (1965) Über einen Ansatz zur mehrwertigen Mengenlehre. Montasb. Deutsch. Akad. Wiss. Berlin 7:859–876

Kolodziejczyk W (1986) Orlovsky's concept of decision making with fuzzy preference relations. Fuzzy Sets Syst 19:11–20

Llamazares B (2005) Factorization of fuzzy preferences. Soc Choice Welf 24:475–496

Marshall AW, Olkin I (1979) Inequalities: theory of majorization and its applications. Academic Press, New York

Ok E (1994) A note on the approximation of fuzzy preferences by exact relations. Fuzzy Sets Syst 67:173–179

Ok E (1996) Fuzzy measurement of income inequality: some possibility results on the fuzzification of the Lorenz ordering. Econ Theory 7:513–530

Ovchinnikov S (1986) On the transitivity property. Fuzzy Sets Syst 20:241–243

Sen A (1992) Inequality re-examined. Harvard University Press

Sen A (1997) On economic inequality. Oxford University Press

Zadeh LA (1965) Fuzzy sets. Inf Control 8:338–35

On Pro-Middle Class Growth

Osnat Peled and Jacques Silber

1 Introduction

Is economic growth a sufficient condition for reducing poverty or should governments enact policies that fight poverty? As stressed by Kakwani and Son (2018), in the 1950s and 1960s development economists believed in the idea of "trickle-down" which assumes that even if growth is beneficial mainly to the rich, the poor will ultimately also profit from growth, once the wealthy spend their gains. Moreover, it was also thought that the rich are those who can generate economic activities, which will increase the probability for a poor to find an employment. Such a Weltanschauung started, however, to be criticized in the 1970s (e.g., Ahluwalia 1976a,b) when development economists realized that poverty levels did not really decrease and that the rate of growth of the incomes of the poor was significantly smaller than that of the population as a whole. These observations were at the origin of the development of a vast literature dealing with the concept of pro-poor growth (see, for example, Deutsch and Silber 2011; Dollar and Kraay 2002; Foster and Szekely 2008; Kakwani and Pernia 2000; Kakwani and Son 2008; Kraay 2006; Ravallion and Chen 2003; Son 2004; Son and Kakwani 2008).

There is, however, another strand of the development literature, which stresses the fact that a sizable middle class is supposed to be an important factor in economic development (see, for example, Landes 1998). Easterly (2001) thus argued that an

---

O. Peled · J. Silber
Research Department, Bank of Israel, Jerusalem, Israel
e-mail: osnat.peled@boi.org.il

J. Silber (✉)
LISER, Esch-Sur-Alzette, Luxembourg
e-mail: jsilber_2000@yahoo.com

Centro Camilo Dagum, Tuscan Inter University Centre, Advanced Statistics for Equitable and Sustainable Development, Tuscan, Italy

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*,
Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_6

increase in the income share of the middle class is associated with a rise in the growth rate. A greater income share of the middle class is also assumed to bring higher levels of publicly provided health services, and hence better health outcomes. It is also associated with higher levels of political rights and civil liberties.

Similarly, Birdsall (2007a, b) contends that reforms, which are critical to market-based economies in the developing world, will not be sustained over long periods if the middle class does not grow. For Pressman (2007), a large middle class contributes not only to economic growth but also to social and political stability because, among other reasons, it helps reducing class warfare.

Loayza et al. (2012) concluded, using a cross-country panel dataset containing information on 128 countries, that when the size of the middle class grows, there is generally a more active social policy, governance improves, and corruption is less frequent. Similar arguments are brought by Bussolo et al. (2014), who believe that a rise in the size of the middle class in developing countries is likely to improve transparency, reinforce the fight against corruption efforts, and more generally lead to a more open society.

Boushey and Hersh (2012) emphasize other impacts of a strong middle class. They claim that it encourages the development of human capital, and a well-educated population is the basis for a stable source of demand for goods and services and favors entrepreneurship, all these effects clearly supporting economic growth.

Actually, as stressed by Bhalla and Kharas (2013, page 6), "the notion of the 'middle class' has roots that go back millennia, originating as a concept in the writings of Aristotle, who defined it as owners of property and thus the people best positioned to rule the state. According to him, they were a moderating force with both the capability and incentive for sober governance, but through its long history, the middle class has been linked to a wide range of concepts from thriftiness to democratic spirit to unchecked consumerism." Bhalla and Kharas therefore consider the middle class as an important driver of the global economy, because they have a significant amount of time and money at their disposal and, at the difference of rich people, they are numerous enough to have an impact on world trends.

This prediction is also shared by Birdsall (2013, page 11). She argues that, "perhaps in a virtuous cycle, recent growth in India, Africa, China and much of Latin America—whether driven by 'luck' (high commodity prices, natural resource windfalls), 'globalization' (trade, capital and labor movements), good policy (sound macroeconomic fundamentals, more democratic and accountable governments) or the intangible benefits of the information revolution or of changing global norms (consider the Millennium Development Goals) that have put more girls in school—will be more likely to be sustained and institutionalized, because an independent middle class has become big enough and politically powerful enough, to be a force for good government and equal opportunity growth."

Lopez-Calva and Ortiz-Juarez (2012) recommend in fact taking a broader view of the middle class and they emphasize the concept of "middle-class functioning," which, according to them, means "being protected from falling into poverty." The definition of the middle class is, hence, related to that of vulnerability since, as argued by Lopez-Calva (2013, page 15), it is economic security "that defines a person as middle class. Individuals who are above the poverty line and who have a low risk of falling into poverty may have characteristics in terms of risk-taking capabilities, investment decisions, consumption patterns and the like that differ from the characteristics of those individuals who are just above the poverty line."

Ferreira et al. (2013) take a similar view and argue that, between 1995 and 2010, at least forty percent of the households in Latin America moved upward in "socioeconomic class." They stress, however, that generally the poor do not move directly to the middle class. They rather join a group located between the poor and the middle class. This group corresponds in fact to the vulnerable class, and it is the largest class in the region. Ferreira et al. (2013) stress the fact that, although this new middle class differs from one country to the other, one may observe common characteristics. Those entering the middle class are thus more educated, more likely to live in urban areas, and work in the formal sector. It appears also that middle-class women tend to have fewer children and a higher labor force participation rate than women belonging to the poor or vulnerable groups.

These features of the middle class lead to a broader definition of this group, an approach taken also by Atkinson and Brandolini (2013) in their study of the middle class in selected OECD countries. They started by adopting definitions based on income but they also examined the role of property and wealth, as well as that of the occupational structure. They argue that classical economists linked class differences to the extent of control over resources and to the position of individuals in the division of labor. Economists today centre their analysis on income, an approach criticized by sociologists who believe that economists ignore the importance of the social stratification implied by labor market relations.

Looking at OECD and some emerging countries, OECD and the World Bank (2016, page 2) concluded that, although almost two-thirds of people live in middle-income households in OECD countries and between one-third and one-half in Emerging Economies, "middle-class self-identification has fallen significantly in recent years, much more than income trends would suggest." There is clearly a positive correlation between the share of the population with "middle incomes," and that, which identifies itself (subjectively) as middle class, but there are differences between the countries. Thus in Northern and Continental European countries, as well as in Italy and Turkey, the share of the population that identifies itself as "middle class" is higher than what the income data seem to show, while the opposite is true for Canada, Portugal, and the United Kingdom. The data seem in fact to indicate (see, OECD and the World Bank 2016, page 7) that "the economic influence of middle-income groups compared to upper-income groups has been declining and that the cost of typical middle class goods, services and assets are rising faster than median incomes and driving some middle income households into debt."

A somehow similar view is taken by Thewissen et al. (2015, pages 24–25) for whom in OECD countries "there are countries and sub-periods where the median stagnated and inequality rose rapidly, but also ones where increasing inequality accompanied rapid growth in the median and others where the median rose only modestly while inequality was stable." These findings led the authors to conclude that, to improve the living standards of middle-income households, it is not enough to promote economic growth, and even not to look at the evolution of both growth and inequality. There is a need to take a specific look at what happens in the middle of the income distribution. Such a task implies evidently that there should be a definition in income terms of the middle class. Various ways of defining the lower and upper bounds of the middle class (75–125%, or 60–225%, or 50–150% of the median household income; or $10 a day and the 90*th* percentile of the income distribution) have been proposed (see, for example, Nissanov et al. 2010, and Nissanov 2017, for a review of these definitions).

Rather than defining such bounds, Massari et al. (2009) in their analysis of the Italian income distribution, and Nissanov and Pittau (2016) in their study of the evolution of the middle class in Russia from 1992 to 2008, applied a nonparametric tool, the so-called "relative distribution."

Another interesting approach is the mixture model method (see, McLachlan and Peel 2000) which is a semi-parametric method, which enables to model unknown distributional shapes and to represent subpopulations and parameters of their densities without having to define in advance the number and characteristics of these groups. Pittau et al. (2010) applied this approach to the world distribution (close to 100 countries). They estimated a three-component mixture model, the components being, respectively, labeled "poor," "middle," and "rich." They found that, while the gap between the mean relative per capita incomes of the rich and poor group widened somewhat over time, the "hollowing out" of the middle of the distribution was largely attributable to the increased concentration of the rich and poor countries around their respective component means. Nissanov (2017) applied the same technique to Russian data and concluded that the share of the middle class decreased between 1992 and 1996 and then increased. While the differences between the poor and the two other groups (middle class and rich) are clear, when looking at most of the characteristics, that were examined, the middle class and the rich group differ mostly for three characteristics: age, education, and work status.

Foster and Wolfson (2010) introduced a simpler approach to the study of the middle class, in their famous paper, which defined the concept of polarization curves as well as an index of bipolarization.[1] While Foster and Wolfson took a relative approach to the measurement of bipolarization, Chakravarty et al. (2007) presented an absolute approach to the measurement of bipolarization that will be the basis of the analysis of pro-middle-class growth presented in this chapter.[2]

This chapter is organized as follows. In Sect. 2, starting from the notion of absolute polarization curves proposed by Chakravarty et al. (2007), we derive the concept of pro-middle-class growth, which amounts to defining the type of growth, which reduces absolute bipolarization. Then, in Sect. 3, we present an empirical illustration

[1]This paper was originally written in 1992 and was widely cited over the years until it was finally published in 2010.

[2]The contribution of Satya Chakravarty to the study of polarization goes much beyond this paper. He has published two books on the topic (Chakravarty 2009, 2015) and to the best of our knowledge seven articles (Chakravarty and Majumder 2001; Chakravarty et al. 2007, 2010; Chakravarty and D'Ambrosio 2010; Chakravarty and Maharaj 2011, 2012, 2015).

based on Israeli data and covering the period 1995–2011. Concluding comments are given in Sect. 4.

2 Deriving a Measure of Pro-middle-Class Growth

Following earlier work by Foster and Wolfson (2010), Chakravarty et al. (2007) have stressed the fact that bipolarization indices can be relative or absolute. A relative index is supposed to remain invariant under equi-proportionate changes in all incomes, while an absolute index is supposed not to vary under equal absolute changes in all incomes. The choice between these two approaches to the measurement of bipolarization is, without any doubt, a problem of value judgement, in the same way as choosing between an absolute and a relative approach to inequality measurement depends on how one views inequality.

The basic idea of Chakravarty et al. (2007) is to scale up the Foster–Wolfson (second) bipolarization curve by the median m in order to obtain what they call an "Absolute Polarization Curve (APC)." Such a curve shows, for any population proportion, how far the total income enjoyed by that proportion is from the corresponding income that it would receive under a hypothetical distribution where everyone receives the median income.

Assuming that the incomes x_i are ranked in nonincreasing order ($x_1 \geq \cdots \geq x_i \geq \cdots \geq x_n$), that n is the size of the population and n_m the rank of the median, the APC ordinate[3] corresponding to the population proportion (k/n) is then defined as

$$AP\left[x, \left(\frac{k}{n}\right)\right] = (1/n)\sum(m - x_i) \quad \text{when}\, x_i < m \tag{1}$$

and as

$$AP\left[x, \left(\frac{k}{n}\right)\right] = (1/n)\sum(x_i - m) \quad \text{when}\, x_i > m. \tag{2}$$

It is easy to observe that for a typical income distribution the APC will decrease monotonically, until the median income is reached, and then increase monotonically. One may observe that by dividing $AP(x; p)$ by the median m, one obtains the (second) bipolarization curve proposed by Foster and Wolfson (2010).

Chakravarty et al. (2007) show also that the area under the Absolute Polarization Curve is an absolute index of polarization. The link between this index and the Foster and Wolfson (2010) index P_{FW} is quite easy to derive, as will now be shown.

Let us first compute the area R under the APC on the R.H.S. of the median (incomes higher than the median income). We may note that at the rank n_m, the height of the

[3]Part of the demonstrations that follow is borrowed from Nissanov et al. (2010).

APC is zero ($x_{n_m} = m$). For simplicity in what follows, we will assume that n is even.

Using (2) and the usual formulations concerning the areas of a triangle and of a trapezium, we derive that

$$R = \left(\frac{1}{2}\right)\left(\frac{1}{n}\right)\left\{\left[\left(\frac{1}{n}\right)\left(x_{n_m-1} - m\right)\right] + \left[2\left(\frac{1}{n}\right)\left(x_{n_m-1} - m\right) + \left(\frac{1}{n}\right)\left(x_{n_m-2} - m\right)\right]\right\} + \cdots$$
$$+ \left(\frac{1}{2}\right)\left(\frac{1}{n}\right)\left\{2\left(\frac{1}{n}\right)\left[\left(x_{n_m-1} - m\right) + \left(x_{n_m-2} - m\right) + \cdots + (x_2 - m)\right] + \left(\frac{1}{n}\right)(x_1 - m)\right\} \tag{3}$$
$$\leftrightarrow R = (1/n)^2\left\{\left(\frac{1}{2}\right)x_1 + \left(\frac{3}{2}\right)x_2 + \cdots + \left(\frac{2i-1}{2}\right)x_i + \cdots + \left(\frac{2(n_m-1)-1}{2}\right)x_{n_m-1}\right\}$$
$$- (1/n)^2\left\{\left(\frac{1}{2}\right)m + \left(\frac{3}{2}\right)m + \cdots + \left(\frac{2i-1}{2}\right)m + \cdots + \left(\frac{2(n_m-1)-1}{2}\right)m\right\} \tag{4}$$
$$\leftrightarrow R = (1/2)\left[\sum_{i=1}^{\frac{n}{2}}\left(\frac{2i-1}{n^2}\right)(x_i - m)\right] = \left(\frac{1}{8}\right)\left[\sum_{i=1}^{\frac{n}{2}}\frac{(2i-1)}{\left(\frac{n^2}{4}\right)}(x_i - m)\right] = \frac{\mu^{ER} - m}{8}, \tag{5}$$

where μ^{ER} refers to the "equally distributed equivalent level of income among the "rich," defined, in the case of the Gini index, (see, Berrebi and Silber 1989)" as

$$\mu^{ER} = \sum_{i=1}^{n/2}\left[(2i-1)/\left(n^2/4\right)\right]x_i, \tag{6}$$

while it is easy to check that $\sum_{i=1}^{n/2}(2i-1)/\left(n^2/4\right) = 1$.

Let us similarly compute the area L under the APC on the L.H.S. of the median (incomes lower than the median income). Using (1), we derive that

$$L = \left(\frac{1}{2}\right)\left(\frac{1}{n}\right)\left\{\left[\left(\frac{1}{n}\right)\left(m - x_{n_m+1}\right)\right] + \left[2\left(\frac{1}{n}\right)\left(m - x_{n_m+1}\right) + \left(\frac{1}{n}\right)\left(m - x_{n_m+2}\right)\right]\right\} + \cdots$$
$$+ \left(\frac{1}{2}\right)\left(\frac{1}{n}\right)\left\{2\left(\frac{1}{n}\right)\left[\left(m - x_{n_m+1}\right) + \left(m - x_{n_m+2}\right) + \cdots + (m - x_{n-1})\right] + \left(\frac{1}{n}\right)(m - x_n)\right\} \tag{7}$$
$$\leftrightarrow L = \left(\frac{1}{n}\right)^2\left\{\left[\left(\frac{1}{2}\right)m + \left(\frac{3}{2}\right)m + \cdots + \left(\frac{(2(n-n_m)-1)}{2}\right)m\right]\right\}$$
$$- \left(\frac{1}{n}\right)^2\left\{\left[\left(\frac{1}{2}\right)x_n + \left(\frac{3}{2}\right)x_{n-1} + \cdots + \left(\frac{(2(n-n_m)-1)}{2}\right)x_{n_m+1}\right]\right\}$$
$$\leftrightarrow L = \left(\frac{1}{2}\right)\sum_{i=\left(\frac{n}{2}\right)+1}^{n}\left\{\left[\frac{[2(n-i+1)-1]}{n^2}\right](m - x_i)\right\} = \frac{\left(m - \mu^{FP}\right)}{8}, \tag{8}$$

where μ^{FP} is a weighted average of the incomes of the "poor" and is defined as

$$\mu^{FP} = \sum_{i=\left(\frac{n}{2}\right)+1}^{n}\left[(2(n-i+1)-1)]/\left(n^2/4\right)\right]x_i, \tag{9}$$

while it is easy to verify that

$$\sum_{i=(n/2)+1}^{n} [2(n-i+1)-1] = \left(\frac{n^2}{4}\right).$$

Note that the weights defining μ^{FP} are such that the further away a "poor" individual is from the median income (that is, the "poor" he is), the smaller the weight given to this "poor." In other words, μ_{FP} gives, for the subpopulation of "poor," a higher weight the less "poor" the "poor" is.

Combining (5) and (9), we conclude that the total area A under the *APC* (on both the R.H.S. and the L.H.S. of the *APC*) may be expressed as

$$A = R + L = \left(\frac{1}{8}\right)\left(\mu^{ER} - \mu^{FP}\right). \tag{10}$$

Let us now remember that the Foster and Wolfson (2010) is expressed as

$$P_{FW} = (G_B - G_W)\left(\frac{\mu}{m}\right), \tag{11}$$

where μ is the arithmetic mean of the distribution.

Remembering the link between the Gini index and the mean difference, expression (11) will be written as

$$P_{FW} = \left[\left(\frac{1}{2}\right)(\Delta_B - \Delta_W)/\mu\right]\left(\frac{\mu}{m}\right) = \left(\frac{1}{2}\right)(\Delta_B - \Delta_W)/m, \tag{12}$$

where Δ_B and Δ_W are the between and within groups mean differences.

From Berrebi and Silber (1989), we know that

$$\Delta_B = \left(\frac{1}{2}\right)\left(\mu^R - \mu^P\right), \tag{13}$$

where μ^R and μ^P are, respectively, the mean incomes of the "rich" and the "poor," and that

$$\Delta_W = \left(\frac{1}{4}\right)\left(\Delta^R + \Delta^P\right), \tag{14}$$

where Δ^R and Δ^P are, respectively, the mean differences in the populations of the "rich" and of the "poor."

Combining (12), (13), and (14), we derive that

$$P_{FW} = \left\{ \left[\left(\frac{1}{4}\right)\left(\mu^R - \mu^P\right) \right] - \left[\left(\frac{1}{8}\right)\left(\Delta^R + \Delta^P\right) \right] \right\} / m. \tag{15}$$

We know, however, that

$$\Delta^R = 2\mu^R G^R = 2\mu^R \left[\left(\mu^R - \mu^{ER}\right)/\mu^R \right] = 2\left(\mu^R - \mu^{ER}\right). \tag{16}$$

We can similarly write (see also Berrebi and Silber 1989) that

$$\Delta^P = 2\mu^P G^P = 2\mu^P \left[\left(\mu^{FP} - \mu^P\right)/\mu^P \right] = 2\left(\mu^{FP} - \mu^P\right). \tag{17}$$

Combining (15), (16), and (17), we easily derive that

$$P_{FW} = \left[\left(\frac{1}{4}\right)\left(\mu^{ER} - \mu^{FP}\right) \right] / m. \tag{18}$$

Finally, combining (10) and (18), we derive the link between the Foster and Wolfson index P_{FW} and the area A that lies under the Absolute Polarization Curve (APC), with

$$P_{FW} = 2\left(\frac{A}{m}\right). \tag{19}$$

Let us now assume that we observe a set of incomes $\{x_1, \ldots, x_i, \ldots, x_n\}$ at time 0 and a set of incomes $\{y_1, \ldots, y_i, \ldots, y_n\}$ at time 1. Using (18), the change ΔP_{FW} in the value of the index P_{FW} between times 0 and 1 will then be written as

$$\Delta P_{FW} = \left(P^1_{FW} - P^0_{FW}\right) = \left(\frac{1}{4}\right)\left[\left(\frac{(y^{ER} - y^{FP}}{m_y} \right) - \left(\frac{x^{ER} - x^{FP}}{m_x} \right) \right], \tag{20}$$

where P^1_{FW}, P^0_{FW}, y^{ER}, x^{ER}, y^{FP}, x^{FP}, m_y, and m_x refer, respectively, to the Foster and Wolfson indices, the "equally distributed equivalent level of income among the 'rich' μ^{ER}, the weighted average μ^{FP} of the incomes of the 'poor' and the median incomes at times 1 and 0."

Combining (6), (9), and (20), we derive that

$$\begin{aligned}
\Delta P_{FW} = {} & \left(\frac{1}{4}\right)\left\{ \left[\sum_{i=1}^{n/2} \left(\frac{2i-1}{n^2/4}\right)\left(\frac{y_i}{m_y}\right) \right] - \left[\sum_{i=\left(\frac{n}{2}\right)+1}^{n} \left(\frac{[2(n-i+1)-1]}{n^2/4}\right)\left(\frac{y_i}{m_y}\right) \right] \right\} \\
& - \left(\frac{1}{4}\right)\left\{ \left[\sum_{i=1}^{n/2} \left(\frac{2i-1}{n^2/4}\right)\left(\frac{x_i}{m_x}\right) \right] - \left[\sum_{i=\left(\frac{n}{2}\right)+1}^{n} \left(\frac{[2(n-i+1)-1]}{n^2/4}\right)\left(\frac{x_i}{m_x}\right) \right] \right\} \\
\leftrightarrow \Delta P_{FW} = {} & \left\{ \sum_{i=1}^{\frac{n}{2}} \left(\frac{(2i-1)}{n}\right)\left[\left(\frac{y_i}{nm_y}\right) - \left(\frac{x_i}{nm_x}\right) \right] \right\}
\end{aligned}$$

$$-\left\{\sum_{i=\left(\frac{n}{2}\right)+1}^{n}\left(\frac{2(n-i+1)-1}{n}\right)\left[\left(\frac{y_i}{nm_y}\right)-\left(\frac{x_i}{nm_x}\right)\right]\right\}. \quad (21)$$

Let us now define the shares w_i and s_i as $w_i = \frac{y_i}{nm_y}$ and $s_i = \frac{x_i}{nm_x}$.
We may then rewrite (21) as

$$\Delta P_{FW} = \left\{\left[\sum_{i=1}^{\frac{n}{2}}\left(\frac{(2i-1)}{n}\right)\{[w_i]-[s_i]\}\right]\right\}$$
$$-\left\{\left[\sum_{i=\left(\frac{n}{2}\right)+1}^{n}\left(\frac{2(n-i+1)-1}{n}\right)\{[w_i]-[s_i]\}\right]\right\}. \quad (22)$$

In what follows we will use the following definitions:

$$\Delta m = m_y - m_x \quad (23)$$

$$\eta_i = \frac{y_i - x_i}{x_i} = \frac{\Delta x_i}{x_i} \quad (24)$$

$$\bar{\eta} = \frac{\Delta m}{m_x} \quad (25)$$

$$w_i = \left(\frac{x_i + \Delta x_i}{n(m_x + \Delta m)}\right). \quad (26)$$

We then derive that

$$w_i - s_i = \left(\frac{x_i nm_x + \Delta x_i nm_x - (x_i nm_x + x_i n\Delta m)}{nm_x(n(m_x + \Delta m))}\right) = \left(\frac{\Delta x_i nm_x - x_i n\Delta m}{nm_x(n(m_x + \Delta m))}\right)$$
$$\leftrightarrow w_i - s_i = \left[\left(\frac{x_i}{nm_x}\right)\left(\frac{\Delta x_i}{x_i}\right)\left(\frac{1}{(m_x+\Delta m)/m_x}\right) - \left[\left(\frac{x_i}{nm_x}\right)\left(\frac{\Delta m/m_x}{(m_x+\Delta m)/m_x}\right)\right]\right]$$
$$\leftrightarrow w_i - s_i = s_i\eta_i\frac{1}{1+\bar{\eta}} - s_i\frac{\bar{\eta}}{1+\bar{\eta}} = s_i\frac{\eta_i - \bar{\eta}}{1+\bar{\eta}}, \quad (27)$$

so that

$$\Delta P_{FW} = \left[\sum_{i=1}^{\frac{n}{2}}\left(\frac{(2i-1)}{n}\right)\left\{s_i\frac{\eta_i-\bar{\eta}}{1+\bar{\eta}}\right\}\right] - \left[\sum_{i=\left(\frac{n}{2}\right)+1}^{n}\left(\frac{2(n-i+1)-1}{n}\right)\left\{s_i\frac{\eta_i-\bar{\eta}}{1+\bar{\eta}}\right\}\right]. \quad (28)$$

Note, however, that

$$\sum_{i=1}^{n/2}\left(\frac{2i-1}{n}\right)s_i = \sum_{i=1}^{n/2}\left(\frac{2i-1}{n}\right)\left(\frac{x_i}{nm_x}\right) = \sum_{i=1}^{n/2}\left(\frac{2i-1}{n^2}\right)\left(\frac{x_i}{m_x}\right)$$

$$\leftrightarrow \sum_{i=1}^{n/2}\left(\frac{2i-1}{n}\right)s_i = \left(\frac{1}{4}\right)\sum_{i=1}^{n/2}\left(\frac{2i-1}{n^2/4}\right)x_i\left(\frac{1}{m_x}\right) = \left(\frac{1}{4}\right)\left(\frac{x_{ER}}{m_x}\right) \tag{29}$$

and that

$$\sum_{i=1}^{n/2}\left(\frac{[2(n-i+1)-1]}{n}\right)s_i = \sum_{i=1}^{n/2}\left(\frac{[2(n-i+1)-1]}{n}\right)\left(\frac{x_i}{nm_x}\right) = \sum_{i=1}^{n/2}\left(\frac{[2(n-i+1)-1]}{n^2}\right)\left(\frac{x_i}{m_x}\right)$$

$$\leftrightarrow \sum_{i=1}^{n/2}\left(\frac{[2(n-i+1)-1]}{n}\right)s_i = \left(\frac{1}{4}\right)\sum_{i=1}^{n/2}\left(\frac{[2(n-i+1)-1]}{n^2/4}\right)x_i\left(\frac{1}{m_x}\right) = \left(\frac{1}{4}\right)\left(\frac{x_{FP}}{m_x}\right). \tag{30}$$

Combining (28), (29), and (30), we derive that

$$\Delta P_{FW} = \left\{\left[\left(\frac{1}{4}\right)\left(\frac{x_{ER}}{m_x}\right)\sum_{i=1}^{n/2}(\lambda_i)\left(\frac{\eta_i-\bar{\eta}}{1+\bar{\eta}}\right)\right] - \left[\left(\frac{1}{4}\right)\left(\frac{x_{FP}}{m_x}\right)\sum_{i=\left(\frac{n}{2}\right)+1}^{n}(v_i)\left(\frac{\eta_i-\bar{\eta}}{1+\bar{\eta}}\right)\right]\right\}, \tag{31}$$

where

$$\lambda_i = \frac{\left[\left(\frac{2i-1}{n}\right)s_i\right]}{\left[\sum_{i=1}^{\frac{n}{2}}\left(\frac{2i-1}{n}\right)s_i\right]} = \left[\left(\frac{2i-1}{n}\right)s_i\right] \Big/ \left[\left(\frac{1}{4}\right)\left(\frac{x_{ER}}{m_x}\right)\right] \tag{32}$$

with $\sum_{i=1}^{n/2}\lambda_i = 1$
and

$$v_i = \frac{\frac{[2(n-i+1)-1]}{n}s_i}{\sum_{i=\left(\frac{n}{2}\right)+1}^{n}\frac{[2(n-i+1)-1]}{n}s_i} = \frac{\frac{[2(n-i+1)-1]}{n}s_i}{\left(\frac{1}{4}\right)\left(\frac{x_{FP}}{m_x}\right)} \tag{33}$$

with $\sum\limits_{i=\left(\frac{n}{2}\right)+1}^{n} v_i = 1.$

Let us now define η_{ER} and η_{FP} as

$$\eta_{ER} = \sum_{i=1}^{n/2}\lambda_i\eta_i \tag{34}$$

$$\eta_{FP} = \sum_{i=\left(\frac{n}{2}\right)+1}^{n} v_i\eta_i. \tag{35}$$

Combining expressions (31)–(35), we then end up with

$$\Delta P_{FW} = \left(\frac{1}{4}\right)\left\{\left[\left(\frac{x_{ER}}{m_x}\right)\frac{(\eta_{ER} - \bar{\eta})}{1 + \bar{\eta}}\right] + \left[\left(\frac{x_{FP}}{m_x}\right)\frac{(\bar{\eta} - \eta_{FP})}{1 + \bar{\eta}}\right]\right\}. \quad (36)$$

We therefore can state that a sufficient condition for growth to be pro-middle class (that is, for polarization to decrease) is that the growth rate of the median income of the whole population is higher than the weighted average of the growth rates of the rich and smaller than the weighted average of the growth rates of the poor. The intuition of this proposition is indeed very simple. It simply says that when the rich get closer to the median income (because their growth rate is smaller than that of the median income) and when the poor get closer to the median income (because their growth rate is higher than that of the median income), polarization will decrease and the middle class will become more important.

3 An Empirical Illustration

3.1 The Database

Our database is a set of income surveys conducted by Israel's Central Bureau of Statistics for the years 1995–2011.[4] During this period, significant changes in welfare and tax policies took place, as well as major changes in participation and employment patterns.

We have data on several income types, for households and individuals: for households, we consider the household's income from salaried work, economic income (including work income, pensions and capital income, and excluding allowances and transfers) household total income and net disposable income. We also consider the household total and net equivalized income (adjusted for family size according to the Israeli/OECD equivalence scale.[5]) Other income sources, such as allowances or capital income were either too volatile or were only available for a small part of the sample and hence were not taken into account. For individuals, we consider income from salaried work (wage) and wage per hour worked. Individuals who did not work or did not receive a wage were excluded from the database. All the incomes were expressed at 2011 prices.

[4]Until 2011, labor force surveys were conducted every quarter. Every individual/household was asked to answer the questions in this survey during two consecutive quarters. There was then an interruption of two quarters and then again the individual/household was asked to answer the labor force survey questionnaire during two additional quarters. During the last quarter, the individual/household was also asked to fill a questionnaire on his/her income (the income survey). But this practice has been interrupted in 2011 because of a major change in the labor force survey.

[5]The Israeli equivalence scale assigns the value 1.25 to the first household member, 0.75 to the second, 0.65 to the third, 0.55 to the fourth and fifth members, 0.50 to the sixth and seventh, 0.45 to the eighth, and 0.40 to each additional member. The most recent OECD equivalence scale amounts to dividing the household income by the square root of household size.

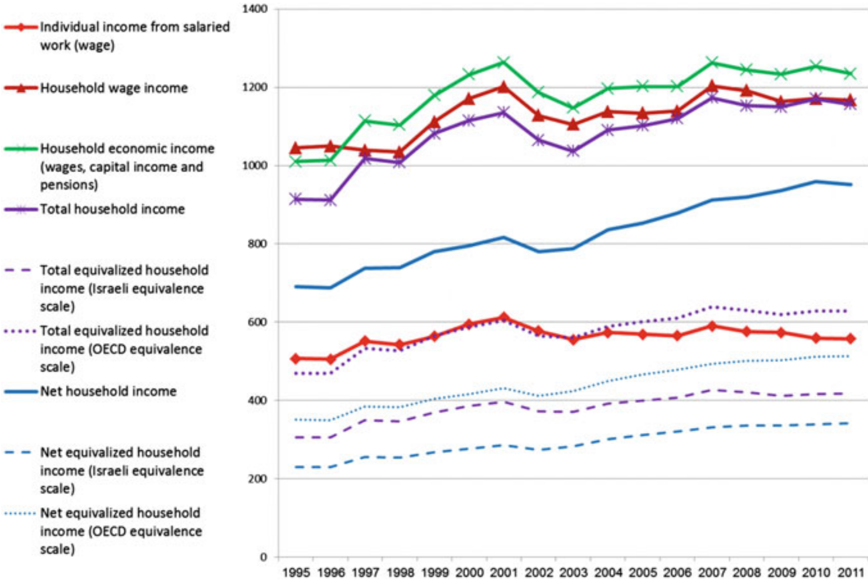


Fig. 1 Absolute bipolarization index for selected income types during the period 1995–2011. *Source* Income surveys, Central Bureau of Statistics

Each income survey includes data on approximately 32,000–35,000 individuals and 13,000–15,000 households[6]. Since our method requires a fixed number of observations, we divided the sample into 1,000 groups with similar sample weights, according to the relevant income-type variable.

3.2 The Results

At the beginning of the period (1995–2000), absolute polarization increased for all income types. Divergence started in the following periods, as shown in Fig. 1. While the absolute polarization of household wage, economic and total household income had the same patterns of change, the absolute polarization of individual income from salaried work decreased and the polarization of net household income continued its upward trend almost unimpededly. These conflicting trends reflect some major developments in labor market participation and changes in policy, as will be explained below.

Table 1a–d presents some descriptive statistics for various income types, at four points in time. As expected, household wage and economic income are the most polarized sources of income, as assortative mating exacerbates household income inequal-

[6]The coverage of the income surveys increased in 1997 so that the number of observations almost doubled. The survey for 1995 includes data for about 17,000 individuals and 7,000 households.

ity and bipolarization. Allowances and progressive taxes work to lower inequality and thus to narrow the spread around the median and reduce bipolarization. Hence, net income is the least polarized. The equivalence scale of the OECD assumes higher economies of scale in comparison with the Israeli equivalence scale, and thus the mean and median of equivalized income are higher when computed according to the OECD scale, as is the absolute polarization index (see, Table 1). There is, however, no significant difference as far as the Foster–Wolfson bipolarization index is concerned. When we compute the Gini index, it is somehow higher when using the Israeli equivalence scales.

During the period 1995–2011, significant changes in participation patterns and social policy took place. Following an increase in social welfare spending, that took place during the last two decades of the twentieth century, a major budgetary cut took place in 2003. Social welfare payments were reduced, and eligibility criteria became stricter. In parallel, there was a decrease in income tax rates for the middle- and top-income brackets. As a result, the contribution of the government's direct intervention to reduce income inequality and polarization decreased. In addition, the change in policy, together with renewed economic growth, led to a significant increase in the participation and employment rates of population subgroups that previously were under-represented in the labor market, such as the ultraorthodox Jews and Arab women. The higher labor force participation and employment rates led to an increase in wage and market incomes for the lower social strata of the population, and therefore to a reduction in the degree of inequality and relative bipolarization of these incomes. This appears clearly in Table 1c, d, when comparing the Gini indices in 2005 and 2011 of income from salaried work, of individual wage per hour worked and of household wage income. Since the tax reduction was concentrated on the middle- and top-income brackets, it offset the impact of higher participation, leading to the following combined effect of these changes: despite the significant decrease in labor income bipolarization, the bipolarization of net incomes increased.

The development of absolute bipolarization was somewhat different from that of relative bipolarization, since it also reflects changes in the median income. For example, the economic downturn between 2000 and 2003 was reflected in a decrease in median incomes (in real terms) and thus, in a decline of absolute polarization. The economic recovery in 2004–2007 led to an increase in median incomes, which intensified the increase in polarization. The increase in absolute polarization during this period was thus more pronounced than the increase in relative polarization.

Figures 2 and 3 show the evolution over time of the different types of income.

It appears that while the increase in wage income affected mainly the bottom of the income distribution, the change in net household income, was more uniform. Most of the change in wage income occurred at the beginning of the period, while the change in total and net income was more evenly spread over the whole period. Table 2 presents the distributional changes of various income types during selected time periods.

As mentioned before, a sufficient condition for growth to be pro-middle class is that the growth rate of the median income is higher than a weighted average of the growth rates of the rich and smaller than a weighted average of the growth

Table 1 **a** Descriptive statistics by income type, 1995. **b** Descriptive statistics by income type, 2001. **c** Descriptive statistics by income type, 2005. **d** Descriptive statistics by income type, 2011

	Median	Mean	Mean to median ratio	μ^{ER}	μ^{FP}	A-total area under the APC	P_{fw}-Foster Wolfson index of polarization	Gini index		
								Total	Within	Between
(*a*)										
Individual income from salaried work (wage)	5,448	7,247	1.3	8,062	4,004	507	0.19	0.422	0.141	0.281
Individual wage per hour worked	134	173	1.3	193	102	11	0.17	0.379	0.124	0.255
Household wage income	6,002	8,613	1.4	10,860	2,502	1045	0.35	0.586	0.171	0.414
Household economic income (wages, capital income and pensions)	6,841	9,709	1.4	11,789	3,709	1010	0.30	0.542	0.167	0.375
Total household income	8,559	11,455	1.3	13,178	5,867	914	0.21	0.434	0.137	0.297
Total equivalized household income (Israeli equivalence scale)	3,203	4,287	1.3	4,779	2,331	306	0.19	0.405	0.131	0.274
Total equivalized household income (OECD equivalence scale)	4,858	6,460	1.3	7,265	3,515	469	0.19	0.405	0.130	0.275
Net household income	7,656	9,180	1.2	10,968	5,442	691	0.18	0.375	0.112	0.263
Net equivalized household income (Israeli equivalence scale)	2,859	3,468	1.2	3,996	2,150	231	0.16	0.346	0.107	0.240
Net equivalized household income (OECD equivalence scale)	4,366	5,211	1.2	6,063	3,255	351	0.16	0.344	0.105	0.239

(continued)

Table 1 (continued)

	Median	Mean	Mean to median ratio	μ^{ER}	μ^{FP}	A-total area under the APC	P_{fw}-Foster Wolfson index of polarization	Gini index		
								Total	Within	Between
(b)										
Individual income from salaried work (wage)	6,342	8,840	1.4	9,642	4,745	612	0.19	0.428	0.145	0.283
Individual wage per hour worked	158	214	1.4	233	122	14	0.18	0.390	0.130	0.260
Household wage income	5,641	9,536	1.7	11,674	2,063	1201	0.43	0.624	0.186	0.438
Household economic income (wages, capital income and pensions)	8,272	12,305	1.5	14,583	4,477	1263	0.31	0.556	0.176	0.381
Total household income	10,423	14,488	1.4	16,291	7,210	1135	0.22	0.446	0.145	0.301
Total equivalized household income (Israeli equivalence scale)	4,020	5,520	1.4	6,066	2,890	397	0.20	0.424	0.140	0.284
Total equivalized household income (OECD equivalence scale)	6,074	8,277	1.4	9,173	4,342	604	0.20	0.422	0.138	0.284
Net household income	9,129	11,134	1.2	13,049	6,514	817	0.18	0.382	0.118	0.264
Net equivalized household income (Israeli equivalence scale)	3,520	4,281	1.2	4,887	2,601	286	0.16	0.361	0.114	0.248
Net equivalized household income (OECD equivalence scale)	5,326	6,404	1.2	7,373	3,925	431	0.16	0.357	0.111	0.246

(continued)

Table 1 (continued)

	Median	Mean	Mean to median ratio	μ^{ER}	μ^{FP}	A-total area under the APC	P_{fw}-Foster Wolfson index of polarization	Gini index		
								Total	Within	Between
(*c*)										
Individual income from salaried work (wage)	6,054	8,325	1.4	9,141	4,588	569	0.19	0.421	0.142	0.279
Individual wage per hour worked	154	207	1.3	225	121	13	0.17	0.379	0.127	0.252
Household wage income	5,664	9,089	1.6	11,237	2,167	1134	0.40	0.612	0.181	0.431
Household economic income (wages, capital income, and pensions)	8,341	11,941	1.4	14,261	4,643	1202	0.29	0.541	0.170	0.371
Total household income	10,015	13,843	1.4	15,735	6,917	1102	0.22	0.443	0.142	0.301
Total equivalized household income (Israeli equivalence scale)	3,949	5,341	1.4	5,971	2,774	400	0.20	0.429	0.140	0.289
Total equivalized household income (OECD equivalence scale)	5,962	7,996	1.3	8,983	4,175	601	0.20	0.426	0.138	0.288
Net household income	9,074	11,186	1.2	13,170	6,341	854	0.19	0.395	0.121	0.274
Net equivalized household income (Israeli equivalence scale)	3,535	4,339	1.2	5,024	2,524	313	0.18	0.382	0.119	0.263
Net equivalized household income (OECD equivalence scale)	5,354	6,488	1.2	7,548	3,815	467	0.17	0.377	0.117	0.261

(continued)

Table 1 (continued)

	Median	Mean	Mean to median ratio	μ^{ER}	μ^{FP}	A-total area under the APC	P_{fw}-Foster Wolfson index of polarization	Gini index		
								Total	Within	Between
(d)										
Individual income from salaried work (wage)	6,131	8,308	1.4	9,115	4,653	558	0.18	0.411	0.138	0.273
Individual wage per hour worked	154	205	1.3	222	122	12	0.16	0.369	0.124	0.246
Household wage income	6,352	9,793	1.5	12,164	2,826	1167	0.37	0.588	0.175	0.413
Household economic income (wages, capital income, and pensions)	9,293	12,837	1.4	15,421	5,545	1235	0.27	0.511	0.159	0.352
Total household income	11,162	14,728	1.3	16,970	7,716	1157	0.21	0.424	0.133	0.290
Total equivalized household income (Israeli equivalence scale)	4,391	5,711	1.3	6,463	3,114	419	0.19	0.411	0.132	0.279
Total equivalized household income (OECD equivalence scale)	6,622	8,529	1.3	9,711	4,681	629	0.19	0.406	0.129	0.277
Net household income	10,197	12,435	1.2	14,781	7,172	951	0.19	0.384	0.116	0.269
Net equivalized household income (Israeli equivalence scale)	4,020	4,835	1.2	5,632	2,890	343	0.17	0.372	0.115	0.257
Net equivalized household income (OECD equivalence scale)	6,075	7,217	1.2	8,462	4,354	514	0.17	0.366	0.112	0.254

Table 2 Distributional changes of various income types during selected time periods

	$\bar{\eta}$	η_{ER}	$\frac{(\eta_{ER}-\bar{\eta})}{1+\bar{\eta}}$	η_{ER}	$\frac{(\eta_{ER}-\bar{\eta})}{1+\bar{\eta}}$	ΔP_{FW}
	Percent change					Percentage points difference
Individual income from salaried work (*wage*)						
1995–1997	5.1	7.2	2.0	5.7	−0.6	0.6
1997–2001	10.8	11.5	0.7	12.1	−1.2	0.0
2001–2005	−4.5	−5.2	−0.7	−3.3	−1.3	−0.5
2005–2007	5.3	4.3	−0.9	4.9	0.3	−0.3
2007–2011	−3.8	−4.4	−0.6	−3.4	−0.5	−0.3
1995–2011	**12.5**	**13.1**	**0.5**	**16.2**	**−3.2**	**−0.4**
Total household income						
1995–1997	8.7	10.0	1.2	8.4	0.3	0.5
1997–2001	12.0	12.4	0.3	13.4	−1.2	−0.1
2001–2005	−3.9	−3.4	0.5	−4.1	0.2	0.2
2005–2007	9.0	7.5	−1.3	8.9	0.1	−0.5
2007–2011	2.2	0.3	−1.9	2.4	−0.2	−0.8
1995–2011	**30.4**	**28.8**	**−1.2**	**31.5**	**−0.9**	**−0.6**
Net household income						
1995–1997	8.7	10.0	1.2	8.4	0.3	0.5
1997–2001	12.0	12.4	0.3	13.4	−1.2	−0.1
2001–2005	−3.9	−3.4	0.5	−4.1	0.2	0.2
2005–2007	9.0	7.5	−1.3	8.9	0.1	−0.5
2007–2011	2.2	0.3	−1.9	2.4	−0.2	−0.8
1995–2011	**30.4**	**28.8**	**−1.2**	**31.5**	**−0.9**	**−0.6**
Net equivalized household income (*OECD equivalence scale*)						
1995–1997	8.7	10.0	1.2	8.4	0.3	0.5
1997–2001	12.0	12.4	0.3	13.4	−1.2	−0.1
2001–2005	−3.9	−3.4	0.5	−4.1	0.2	0.2
2005–2007	9.0	7.5	−1.3	8.9	0.1	−0.5
2007–2011	2.2	0.3	−1.9	2.4	−0.2	−0.8
1995–2011	**30.4**	**28.8**	**−1.2**	**31.5**	**−0.9**	**−0.6**

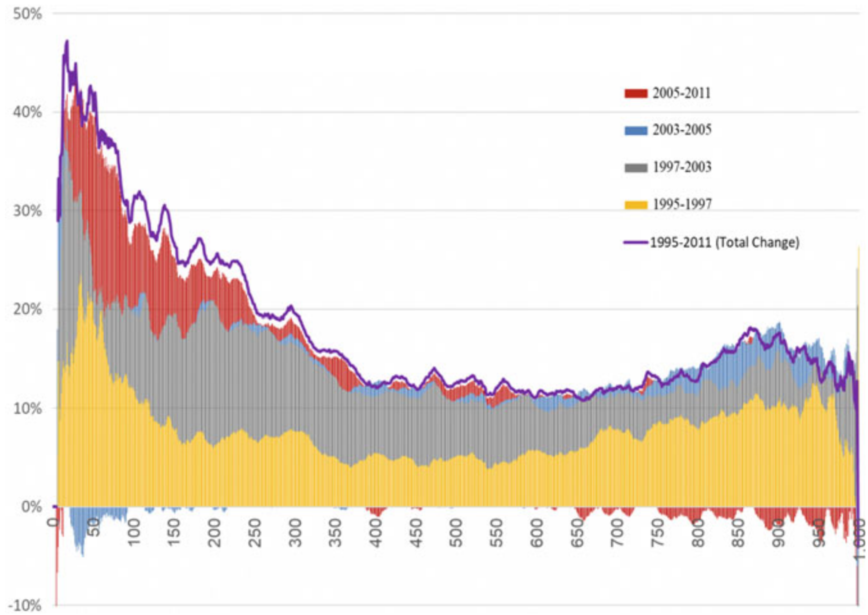


Fig. 2 Change in Individual income from salaried work (wage)

rates of the poor. Looking at what happened to wage income at the beginning of the period (1995–1997 and 1997–2001), it appears that the lowest growth rate was around the middle of the income distribution and the further away an individual was from the middle, the higher his/her growth rate was. In other words, the "poorer" and the richer enjoyed a larger increase in their incomes. We can then conclude that $\eta_{ER} > \bar{\eta}$, since the higher the income, the higher the weight assigned to its growth rate. For the "poor," however, the picture is less clear. On one hand, the growth rate was higher, the lower the income. On the other hand, the lower the income, the lower the weight assigned to this income. The combination of these two contradictory effects indicates that $\eta_{FP} > \bar{\eta}$, that is, the weighted average of the growth rates of the "poor" was higher than the growth rate of the median income. Since both η_{ER} and η_{FP} were higher than $\bar{\eta}$, they influenced polarization in opposite directions and thus the change in polarization was relatively small. In the following periods (2001 and afterward), η_{ER} was lower than $\bar{\eta}$ and η_{FP} was higher or similar to $\bar{\eta}$, leading to a decrease in polarization.

The weighted average of the growth rates of total household income for the rich, η_{ER}, was higher than $\bar{\eta}$ at the beginning of the period (1995–2005), but by the end of the period (2005–2011) it was significantly lower. However, changes in income tax policies, in favor of middle and high incomes, offset much of these differences: η_{ER} for net household income was smaller than $\bar{\eta}$ at the beginning of the period (1995–2001) and higher than $\bar{\eta}$ the following period (2001–2007). Between 2007 and 2011, η_{ER} was slightly lower than $\bar{\eta}$, but this gap between η_{ER} and $\bar{\eta}$ was much

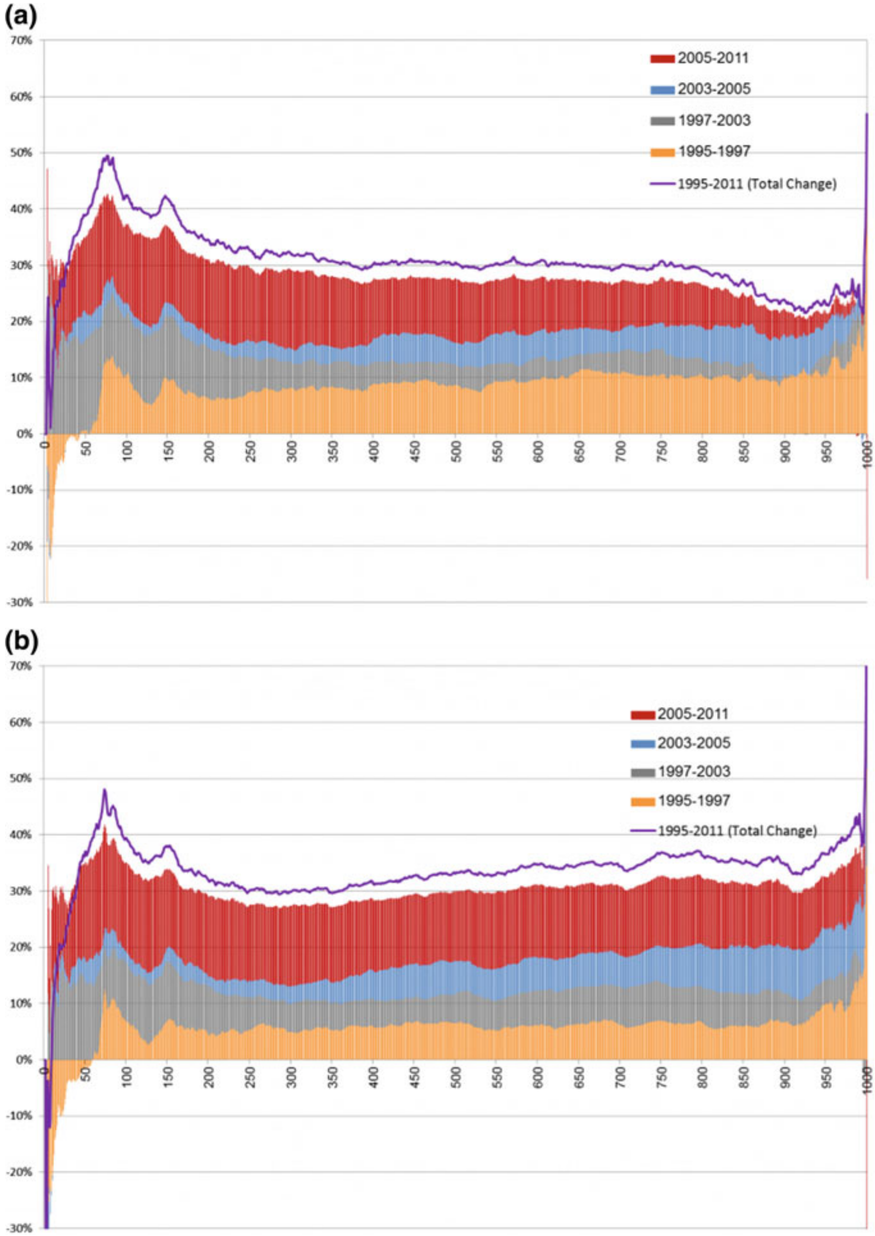


Fig. 3 **a** Change in total household income, **b** Change in net household income

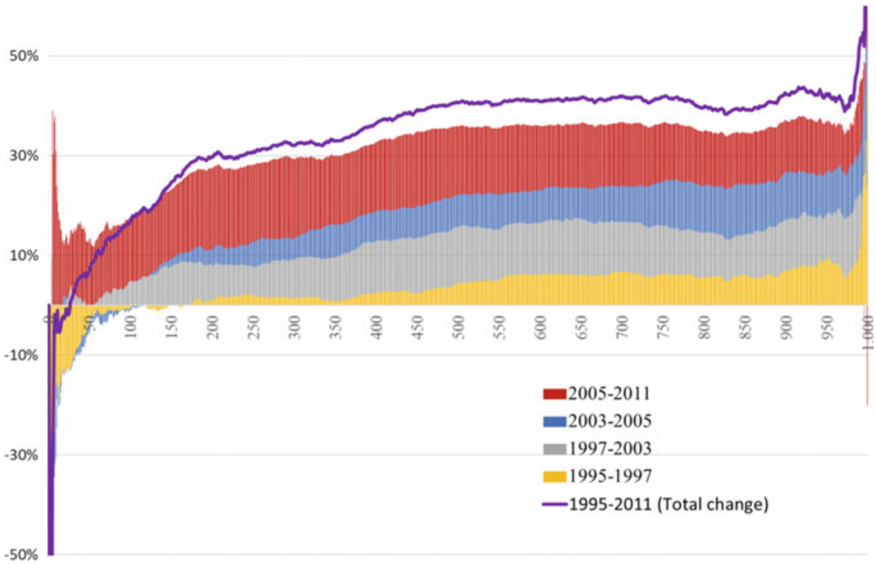


Fig. 4 Change in equivalized net income

smaller than the gap between η_{ER} and $\bar{\eta}$ for total household income during this same period. As far as the "poor" are concerned, η_{FP} was higher than $\bar{\eta}$ for total household income during the period 1997–2001, and similar to $\bar{\eta}$ in all the other periods. Conversely, for net household income, η_{FP} was significantly lower than $\bar{\eta}$ at the beginning of the period (1995–1997 and 2001–2005) and at the end of the period (2007–2011). The periods in which η_{FP} was higher than $\bar{\eta}$ were the years 1997–2001 (mostly as a consequence of an increase in the minimum wage) and the years 2005–2007 (because of an increased participation in the labor force). Thus, over the whole period, the weighted growth rate of the poor's income was lower than the growth rate of the median income. As a result of these different developments, polarization in total household income followed a downward trend starting in 1997, while the polarization of net household income continually increased.

Focusing on specific population subgroups, we should remember that in Israel, larger households are more common in the lower social strata of the population because two population subgroups, the Muslim Arabs and the ultraorthodox Jews, tend to have larger households. Moreover, these two population subgroups have also lower labor force participation rates. These two factors clearly lead to lower incomes. In addition, larger households were more affected by policy changes, since a bigger share of their income came from various allowances. Moreover, the cut in allowances pushed some of the larger households into lower quantiles of the income distribution. Thus, the increase in equivalized income (both net and total income) was significantly lower in the strata located at the bottom of the income distribution (see, Fig. 4).

Table 3 Distributional changes during the period 1995–2011 by selected income types

Income type	$\bar{\eta}$	η_{ER}	$\frac{(\eta_{ER}-\bar{\eta})}{1+\bar{\eta}}$	η_{ER}	$\frac{(\eta_{ER}-\bar{\eta})}{1+\bar{\eta}}$	ΔP_{FW}
	Percent change					Percentage points difference
Individual income from salaried work (wage)	12.5	13.1	0.5	16.2	−3.2	−0.4
Individual wage per hour worked	15.0	15.1	0.1	19.6	−3.9	−0.7
Household wage income	5.8	12.0	5.8	11.9	−5.7	1.9
Household economic income (wages, capital income, and pensions)	35.9	30.8	−3.7	47.0	−8.2	−2.7
Total household income	30.4	28.8	−1.2	31.5	−0.9	−0.6
Total equivalized household income (Israeli equivalence scale)	37.1	35.2	−1.4	33.6	2.6	−0.0
Total equivalized household income (OECD equivalence scale)	36.3	33.7	−2.0	33.2	2.3	−0.3
Net household income	33.2	34.8	1.2	31.8	1.1	0.6
Net equivalized household income (Israeli equivalence scale)	40.6	40.9	0.2	34.4	4.4	0.9
Net equivalized household income (OECD equivalence scale)	39.1	39.6	0.3	33.7	3.9	0.8

Table 3 summarizes now the distributional change of different types of incomes for the whole period 1995–2011.

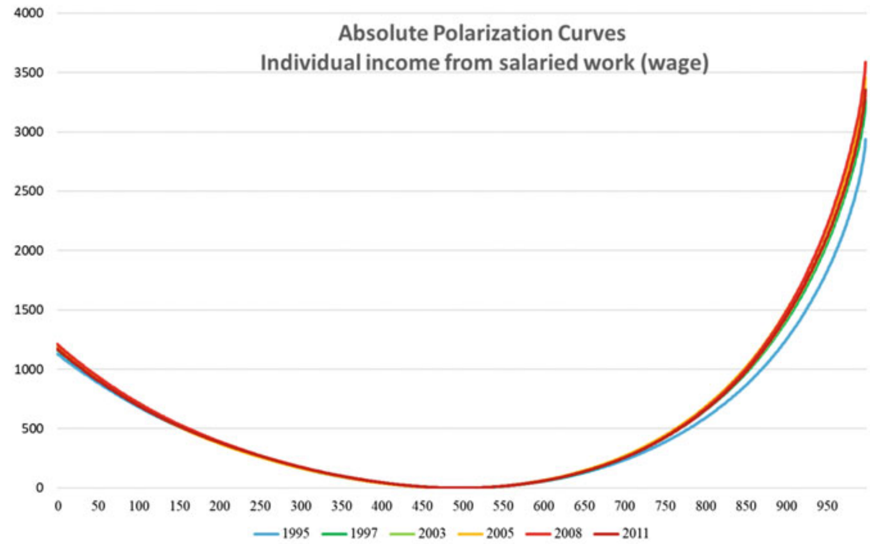
Over the whole period, the degree of bipolarization of net income increased while the bipolarization of market incomes (household economic income) decreased. This result reflects the rapid growth of wage income at the lower part of the distribution, due to increased labor force participation, and to some extent to the effect of an increase in the minimum wage ($\eta_{FP} > \bar{\eta}$ for individual income from salaried work, individual wage per hour worked, household wage income, household economic income, and total household income). However, this relatively rapid increase in market income was offset by the decrease in allowances so that the growth rate of the total income at the lower end of the distribution was similar or even lower than the growth of median income ($\eta_{FP} < \bar{\eta}$ for net household income and net equivalized household income).

For the “rich” (households whose income is higher than the median income), we observe that $\eta_{ER} > \bar{\eta}$ for individual income from salaried work, individual wage per hour worked and household wage income, net household income, and net equivalized income. This in itself would lead, ceteris paribus, to an increase in bipolarization of incomes, but $\eta_{ER} < \bar{\eta}$ for household economic income, total household income, and total equivalized income. The net result of these changes in the income of the “rich” and the “poor” is that the overall degree of bipolarization decreased for individual income from salaried work, individual wage per hour worked, household economic income (the biggest decrease in percentage terms), total household income, and total equivalized income, but increased for net income and net equivalized income. As mentioned before, the increase in the bipolarization of net income reflects the combined effect of developments in labor force participation and welfare policy, which affected mainly the lower strata, together with the effect of a major change in income tax that affected mainly the upper–middle incomes.

Absolute bipolarization curves (APCs) are presented in Fig. 5 and emphasize the differences in the evolution over time of the distribution of market and net income. The APC of wage income increased slightly below the median. In other words, the absolute change in the wage income of the “poor” (in real terms) was almost the same as the absolute change in the median income. Hence, the left part of the APC did not change significantly. The increase of the APC above the median was more pronounced, but it occurred at the beginning of the period and did not change much afterward. The APC of net equivalized income gives a completely different picture, with a continuous increase of the APC. Net income on both sides of the median moved further apart from the median, in real absolute terms.

4 Concluding Comments

Following a vast literature on pro-poor growth, this chapter proposed a definition of pro-middle-class growth that was derived from the concepts of absolute polarization indices and curves introduced by Chakravarty et al. (2007). It appears that a sufficient condition for growth to be pro-middle class is for the growth rate of the median income of the whole population, to be higher than that of the weighted average of the growth rates of the “rich,” and smaller than the weighted average growth rate of the “poor,” the “rich,” and the “poor” being, respectively, those with an income higher and lower than the median income. An empirical illustration based on Israeli data for the period 1995–2011 showed that growth was pro-middle class when looking at individual income from salaried work, individual wage per hour worked, household economic income, total household income, and total equivalized income. But bipolarization increased for net income and net equivalized income, as a consequence of the combined effect of developments in labor force participation, welfare policy, and major changes in income tax rates.



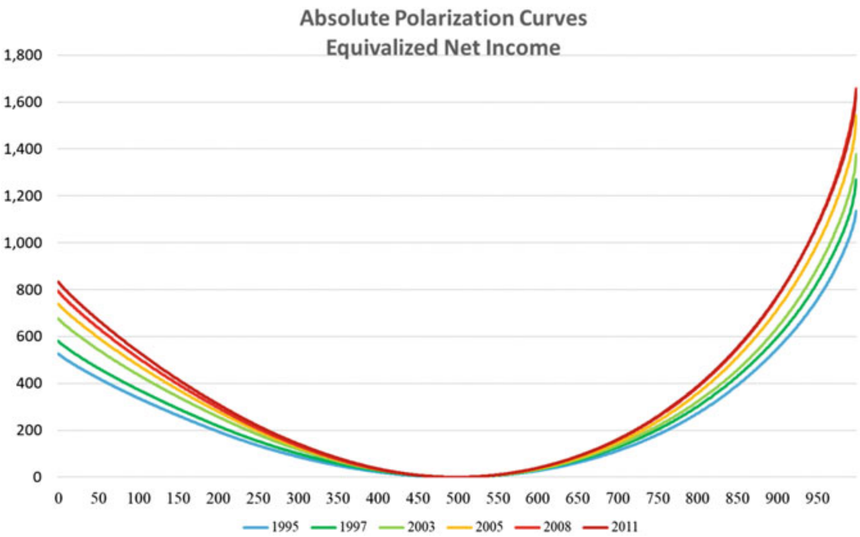


Fig. 5 Absolute bipolarization curves

References

Ahluwalia MS (1976a) income distribution and development: some stylized facts. Am Econ Rev 66(2):128–35

Ahluwalia MS (1976b) Inequality, poverty and development. J Dev Econ 3(4):307–42

Atkinson AB, Brandolini A (2013) On the identification of the middle class. In: Gornick J, Jäntti M (eds) Income inequality: economic disparities and the middle class in affluent countries. Stanford University Press, Stanford, CA, pp 77–100

Berrebi ZM, Silber J (1989) Deprivation, the Gini index of inequality and the flatness of an income distribution. Math Soc Sci 18:229–237

Bhalla S, Kharas H (2013) Middle class angst spills over. In: Poverty in focus, number 26, international policy centre for inclusive growth, poverty practice. Bureau for Development Policy, UNDP, Brasilia, Brazil, pp 6–7

Birdsall N (2007b) Reflections on the macro foundations of the middle class in the developing world. Working Paper 130, Center for Global Development

Birdsall N (2013) The middle class in developing countries–who they are and why they matter. In: Poverty in focus, number 26, international policy centre for inclusive growth, poverty practice. Bureau for Development Policy, UNDP, Brasilia, Brazil, pp 10–12

Boushey H, Hersh A (2012) The american middle class, income inequality, and the strength of our economy. Center for American Progress

Bussolo M, Maliszewska M, Murard E (2014) The long-awaited rise of the middle class in Latin America is finally happening. Policy Research Working Paper 6912, The World Bank

Chakravarty SR, Majumder A (2001) Inequality, polarization and welfare. Aust Econ Pap 40:1–13

Chakravarty SR, Majumder A, Roy S (2007) A treatment of absolute indices of polarization. Jpn Econ Rev 58(2):273–293

Chakravarty SR (2009) Inequality, polarization and poverty. advances. In: Distributional analysis. Economic studies in inequality. Social Exclusion and Well-Being, Springer, New York

Chakravarty SR, D'Ambrosio C (2010) Polarization orderings of income distributions. Rev Income Wealth 56(1):47–64

Chakravarty SR, Chattopadhyay N, Maharaj B (2010) Inequality and polarization: an axiomatic approach. In: Deutsch J, Silber J (eds) The measurement of individual well-being and group inequalities: essays in memory of ZM Berrebi. Routledge Economics, Taylor and Francis Group, pp 88–109

Chakravarty SR, Maharaj B (2011) Subgroup decomposable inequality indices and reduced form indices of polarization. Keio Econ Stud 47

Chakravarty SR, Maharaj B (2012) Ethnic polarization orderings and indices. J Econ Interac Coord 7:99–123

Chakravarty SR, Maharaj B (2015) Generalized Gini polarization indices for an ordinal dimension of human well-being. Int J Econ Theory 11:231–246

Chakravarty SR (2015) Inequality, polarization and conflict: an analytical study. Economic Studies in Inequality, Social Exclusion and Well-Being, Springer India

Deutsch J, Silber J (2011) On various ways of measuring pro-poor growth. In: The measurement of inequality and well-being: new perspectives, special issue of economics—the open-access, open-assessment e-journal, vol 5, no 13. Kiel Institute for the World Economy, pp 1–57

Dollar D, Kraay A (2002) Growth is good for the poor, 7(3):195–225

Easterly W (2001) The middle class consensus and economic development. J Econ Growth 6(4):317–335

Ferreira FHG, Messina J, Rigolini J, López-Calva L-F, Lugo MA, Vakis R (2013) Economic mobility and the rise of the Latin American middle class. The World Bank, Washington, D.C.

Foster JE, Székely M (2008) Is economic growth good for the poor? Tracking low incomes using general means. Int Econ Rev 49(4):1143–1172

Foster JE, Wolfson MC (2010) Polarization and the decline of the middle class: Canada and the US. J Econ Inequal 8(2):247–273

Kakwani N, Pernia E (2000) What is pro-poor growth. Asian Dev Rev 18:1–16
Kakwani N, Son HH (2008) Poverty equivalent growth rate. Rev Income Wealth 54(4):643–655
Kakwani N, Son HH (2018) Essay on economic growth and poverty reduction. In: Kakwani N, Son HH (eds) Economic Growth and Poverty. Edward Elgar
Kraay A (2006) When is growth pro-poor? Evidence from a panel of countries. J Dev Econ 80(1):198–227
Landes D (1998) The wealth and poverty of nations: why some are so rich and some so poor. Norton (New York, NY)
Loayza N, Rigolini J, Llorente G (2012) Do middle classes bring institutional reforms? Policy Research Working Paper No. 6015, The World Bank
Lopez-Calva LF, Ortiz-Juarez E (2012) A vulnerability approach to the definition of the middle class. J Econ Inequal 12(1):23–47
Lopez-Calva LF (2013) A new economic framework to analyse the middle classes in Latin America. In: Poverty in focus, number 26, international policy centre for inclusive growth, poverty practice. Bureau for Development Policy, UNDP, Brasilia, Brazil, pp 14–17
Massari R, Pittau GM, Zelli R (2009) A dwindling middle class? Italian evidence in the 2000s. J Econ Inequal 7:333–350
McLachlan G, Peel D (2000) Finite mixture models. Wiley, New York
Nissanov Z (2017) Economic growth and the middle class in an economy in transition. The case of Russia. Economic studies in inequality, social exclusion and well-being. Springer International Publishing AG, Cham, Switzerland
Nissanov Z, Pittau G (2016) Measuring changes in the Russian middle class between 1992 and 2008: a nonparametric distributional analysis. Empir Econ 50:503–530
Nissanov Z, Poggi A, Silber J (2010) Measuring bi-polarization and polarization: a survey. In: Deutsch J, Silber J (eds) The measurement of individual well-being and group inequalities: essays in memory of ZM Berrebi. Routledge Economics, Taylor and Francis Group, pp 49–87
OECD and the World Bank (2016) The squeezed middle class in OECD and emerging countries: myth and reality. Issues paper, OECD, Centre for Opportunity and Equality (COPE), Paris
Pittau MG, Zelli R, Johnson PA (2010) Mixture models, convergence clubs, and polarization. Rev Income Wealth 56(1):102–122
Pressman S (2007) The decline of the middle class: an international perspective. J Econ Issues 41:181–201
Ravallion M, Chen S (2003) Measuring pro-poor growth. Econ Lett 78(1):93–99
Son HH (2004) A note on pro-poor growth. Econ Lett 82(3):307–314
Son HH, Kakwani N (2008) Global estimates of pro-poor growth. World Dev 36(6):1048–1066
Thewissen S, Kenworthy L, Nolan B, Roser M, Smeeding T (2015) Rising income inequality and living standards in oecd countries. How does the middle fare? Institute for new economic thinking, employment, equity and growth programme. INET Oxford Working Paper No. 2015-01, Oxford, UK

Inequalities and Identities

Arjun Jayadev and Sanjay G. Reddy

Abstract We introduce concepts and measures relating to inequality between identity groups. We define and discuss the concepts of *Representational Inequality*, *Sequence Inequality,* and *Group Inequality Comparison*. *Representational Inequality* captures the extent to which an attribute is shared between members of distinct groups. *Sequence Inequality* captures the extent to which groups are ordered hierarchically. *Group Inequality Comparison* captures the extent of differences between groups. The concepts can be used to interpret segregation, clustering, and polarization in societies.

> Civil paths to peace also demand the removal of gross economic inequalities, social humiliations and political disenfranchisement, which can contribute to generating confrontation and hostility. Purely economic measures of inequality do not bring out the social dimension of the inequality involved. For example, when the people in the bottom groups in terms of income have different non-economic characteristics, in terms of race (such as being black rather than white), or immigration status (such as being recent arrivals rather than older residents), then the significance of the economic inequality is substantially magnified by its "coupling" with other divisions, linked with non-economic identity groups.
>
> Amartya Sen, *The Guardian*, Friday November 9, 2007

> A highly ethnically diversified society may generate tensions in the society which ultimately may lead to conflicts…. [An] objective of the society should, therefore, be to make ethnic diversity (hence ethnic polarization) as low as possible.
>
> Satya Chakravarty and Bhargav Maharaj, 2009

This is a modified version of various versions of a working paper by the same title uploaded on the ISERP website, https://academiccommons.columbia.edu/doi/10.7916/D86979DR and on the SSRN website: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1162275. We are the copyright owners on these.

A. Jayadev (✉)
Azim Premji University, Bengaluru, India
e-mail: arjun.jayadev@apu.edu.in

S. G. Reddy
New School for Social Research, New York, USA
e-mail: reddysanjayg@gmail.com

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*,
Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_7

1 Introduction

For Satya Chakravarty, whose contributions to scholarship and justice we honor here, the salience of group-based membership for evaluating social outcomes is clear. His remarks, along with those of Amartya Sen above, suggest that the interpersonal differences in advantages can be of greater concern when associated systematically with membership of groups. This can be true for two reasons. First, the avoidance of systematically arising intergroup differences may be of intrinsic importance from the perspectives of justice and fairness in the distribution of goods and opportunities. Second, the fact that there exist distinct groups in society and that these groups exhibit intergroup differences may have instrumental significance from the standpoint of their impact on social goods such as peace, stability, or economic growth

The concern with the intrinsic significance of intergroup differences has centered on the degree to which "morally irrelevant" characteristics of a person (such as belonging to a given race, sex, caste, or other groups as a result of birth) should be permitted to determine her or his life chances.[1] Such a motivation is distinct from one based on the idea that social goods or "bads" may be generated by intergroup differences in economic and social achievements, and that intergroup differences may be relevant for that reason. A body of literature in economics and other social sciences has explored this instrumental concern.[2] Both concerns have led to the development of a growing literature that has identified and empirically examined such concepts as "horizontal inequality", segregation, polarization, and related ideas about differences between groups

That a multitude of concepts concerning intergroup difference has been proposed is not entirely surprising because such differences can be understood as arising in more than one way. For example, studies on segregation focus on the degree to which members of different groups share a location, occupation, or other attribute while studies on horizontal inequality focus on the extent of difference in the income or other achievements of separate groups. In both cases, however, the subject of interest is the degree of unevenness or inequality in the possession of attributes between groups. The goal of our paper is to elucidate some distinct ways in which intergroup differences can be conceived, which encompass but are not restricted to the concerns of these existing approaches.

A common underlying concern in analyses of intergroup differences is the degree to which distinct groups are systematically over- or under-represented in their possession of various attributes (levels of income or health, club membership, political office, etc.). In this paper, we introduce the concept of *Representational Inequality*

---

[1]For a review of these debates, see, e.g., Roemer (1996), and Sen (1992). Arguments that societies should be organized so as to limit the consequences of being born into a particular position as such include those of "luck egalitarians" such as Arneson (1989), Cohen (1989), Dworkin (2000), Rawls (1971) and Roemer (1996). Egalitarians of other kinds (e.g., those concerned with relational equality, such as Anderson 1999) may come to similar conclusions for different reasons.

[2]For some examples, see, e.g., Stewart (2001), Alesina et al (2003), Alesina and La Ferrara (2000, 2002), Montalvo and Reynal Querol (2005), Miguel and Gugerty (2005), and Østby (2008).

(*RI*) as a way to capture this concern. This concept describes the extent to which a given attribute (for instance, a level of income or health, or right or left-handedness) is shared by members of distinct groups. It can be used to measure the degree of "segregation" of distinct identity groups in the attribute space.[3]

When individuals can be ordinally ranked in relation to an attribute (such as income or health but not right- or left-handedness), we may be interested not only in how segregated or separated each identity group is in terms of their achievements, but in some measure of their relative positions in the ranking. *Sequence Inequality (SI)*, understood as the degree to which members of one group are placed higher in a given hierarchy than those from another, captures this concern. Such a concept provides an intuitive framework for understanding the degree of "clustering" of various identity groups in distinct sections of a hierarchy.[4]

When individuals' level of achievement can also be cardinally identified for an attribute (as for income but not for right- or left-handedness), the *distance* between groups' attribute levels may be of interest. We may identify a distinction between two different concepts, which we term, respectively, *Group Inequality Comparison (I)* and *Group Inequality Comparison (II)* and abbreviate as *GIC (I)* and *GIC (II).* The concept of *Group Inequality Comparison (I)* involves a comparison of counter-factuals. Specifically, it is derived by comparing the inequality arising in a society in which all of the members of a group are assigned a representative income for that group and the total interpersonal inequality in a society. This concept is concerned with identifying the extent to which between-group inequality "accounts for" overall inequality in society. *Group Inequality Comparison (II)* by contrast measures only the inequality arising in the first situation, i.e., that in a society in which all the members of a group are assigned a representative income for that group. This latter concept is concerned with the absolute magnitude of the inequality generated by between-group inequality.

Our purpose in this paper is twofold. We seek not only to clarify the concepts described above and thus to recognize the complexities of intergroup differences but also to show that combining these concepts can be helpful in characterizing intergroup differences taken as a whole. "Polarization",[5] understood to involve the collection of like elements and the separation of such collections of like elements from one another, can be fruitfully described as involving the simultaneous presence of between-group differences of different kinds. The combination of Representational Inequality with Sequence Inequality alone provides a measure of what might be termed "Ordinal Polarization." Combining Group Inequality Comparison (of either type I or type II)

---

[3]Segregation is defined by the Oxford English Dictionary, *inter alia,* as "The separation of a portion of portions of a collective or complex unity from the rest; the isolation of particular constituents of a compound or mixture."

[4]A cluster is defined by the Oxford English Dictionary, *inter alia*, as "A collection of things of the same kind…growing closely together; a bunch… a number of persons, animals, or things gathered or situated close together; an assemblage, group, swarm, crowd."

[5]The Oxford English Dictionary defines the verb "polarize" as "To accentuate a division within (a group, system, etc.); to separate into two (or occas. several) opposing groups, extremes of opinion, etc."

with these other two indices can provide a richer index of Polarization applicable to the case in which the attribute is cardinally measurable as well. *Our purpose is not to provide a unique characterization of a single measure of polarization, but rather to show that a broad class of measures of polarization can be derived from a simple set of unexceptionable axioms concerning different types of between-group differences and their combination.*

The concept of polarization that we employ here is distinct from that developed in the preponderance of the existing literature in that it draws on information about the identity groups to which those who possess distinct attributes belong. In contrast, the existing frameworks generally employ a "collapsed" framework in which the level of the attribute (typically income) defines the identity group (Esteban and Ray 1994; Duclos et al. 2004). In these frameworks, polarization of an income distribution is understood to involve "identification" between individuals possessing a certain level of income and "alienation" between those individuals and others possessing different incomes. In our framework, in contrast, polarization of an income distribution is understood to involve segregation of individuals belonging to distinct identity groups at certain levels of income and the separation of these groupings of individuals in the income space from other groupings of individuals possessing distinct identities.

2 Part I: Concepts of Group Inequality

One approach to evaluating intergroup differences is to construct a measure of overall group advantage or disadvantage for each group prior to assessing the differences in these overall measures.[6] Although there can be advantages to such an approach, it can obscure the diverse aspects of intergroup difference (by reducing intergroup differences to inequalities in a single dimension). We accordingly explicitly identify here three distinct concepts of intergroup difference, and a fourth which builds upon them.

2.1 Representational Inequality

We define a situation of representational inequality as occurring when, for some attribute and some identity group, the proportion of the group possessing the attribute is either greater or less than the proportion of the group in the overall population. To provide some graphical intuition for this idea, consider the distribution of income among different groups in a society that consists of 50 percent whites and 50 percent blacks. Figure 1 depicts the situation in which there is no representational inequality. The location of each bar on the horizontal axis represents an income level ordered from lowest to highest and the proportion of persons possessing that income of either

[6] See Jayaraj and Subramanian (2006) for an example of such an approach.

group is represented through shading. At all levels of income, blacks and whites are represented in equal proportion to their share of the population as a whole (i.e., one half each). Any deviation from such equiproportionality leads to a situation of representational inequality. Such a situation is depicted in Fig. 2, in which at certain levels of income blacks or whites comprise a larger or smaller proportion of the individuals possessing that level of income than they do in the population.

While the situation depicted in Fig. 2 is one of the representational inequalities, both groups are represented at all the incomes. In contrast, Fig. 3 depicts a situation in which at each level of income there is *complete* segregation, in the sense that at each level of income there is one and only one identity group represented. It may be noted that although this is a situation of complete segregation, the incomes at which whites and blacks appear are evenly interspersed. We depict this example to make sharp the distinction between segregation and clustering as we use the terms. The former refers to a situation in which those possessing a specific attribute (in this case an income level) belong disproportionately to a group. The latter refers to a situation in which the attributes disproportionately possessed by members of a group are sited together in a certain part of an attribute hierarchy (in this case the income spectrum).

The concept of representational inequality clearly need not be restricted to a scenario in which the attribute is cardinally orderable. Thus, for example, we can apply the principle in an equally straightforward manner to unordered attributes such as location of residence, or membership in distinct clubs or legislatures. If instead of income brackets, each bar referred to a distinct legislature in a federal country,

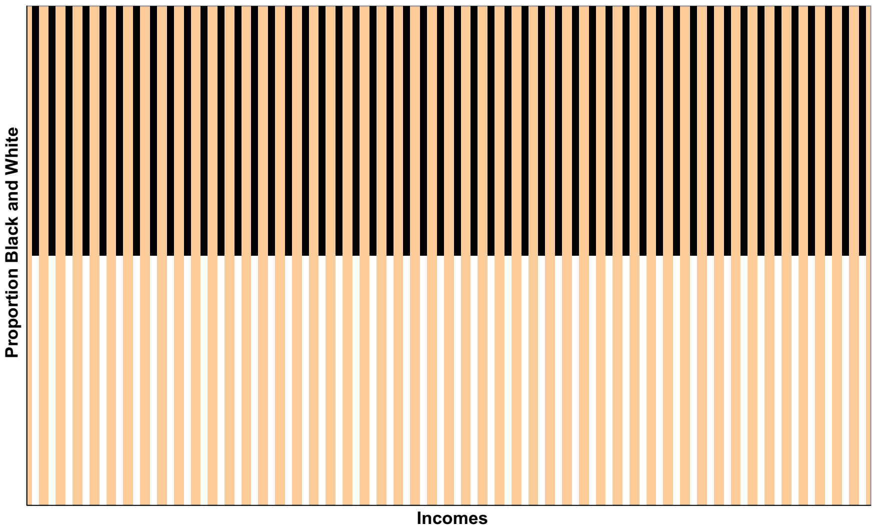


Fig. 1 Zero representational inequality

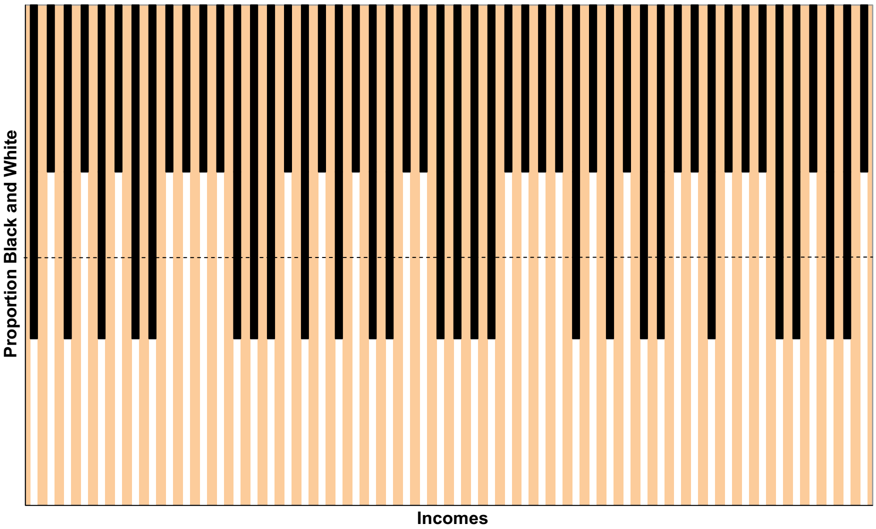


Fig. 2 Nonzero representational inequality

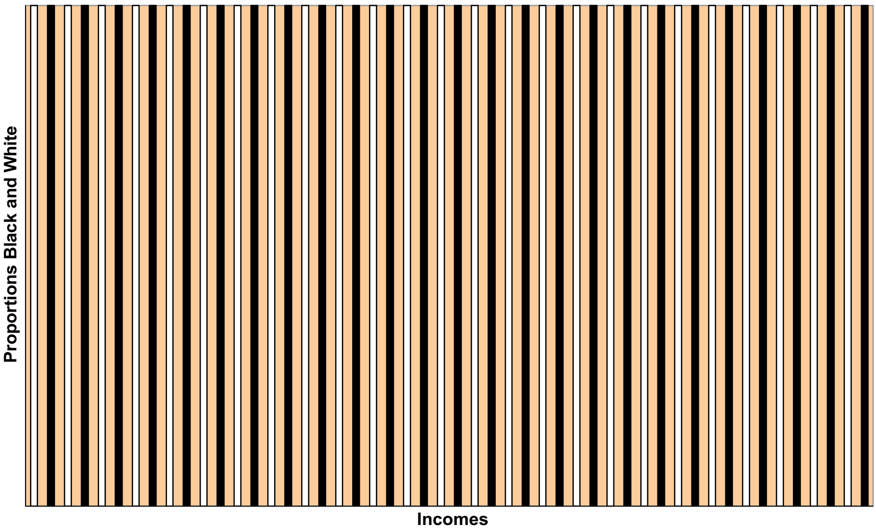


Fig. 3 “Complete” segregation

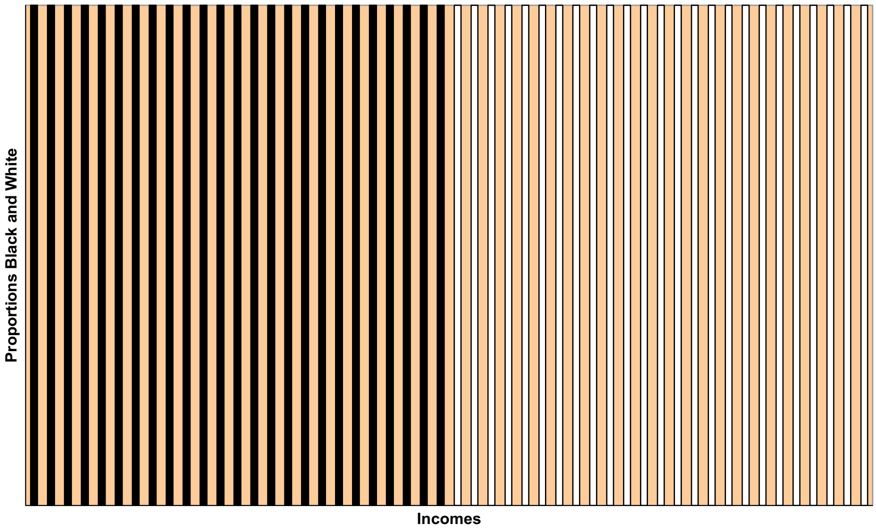


Fig. 4 Polarization

the figures we have discussed here would depict the degree of inequality in political representation.

2.2 *Sequence Inequality*

The distinction between “complete segregation” and “complete clustering” can be seen by comparing Figs. 3 and 4. Figure 4 depicts the situation that results from a transfer of incomes such that all the whites move to the richer half of society while all the blacks move to the poorer half of society. This situation is one in which each subgroup is concentrated in a different part of the income distribution. Such a situation can plausibly be described as one of “complete clustering” of groups.[7] In both cases, there is complete segregation and thus maximal representational inequality. However, in Fig. 3, whether an individual is black or white provides very little information on his or her rank in society. By contrast, in Fig. 4, whether an individual is black or white provides a great deal of information. One simple way to capture the distinction between Figs. 3and 4 is through the concept of sequence inequality, which together with representational inequality captures the clustering of the income distribution. This concept is linked to the position in the overall societal ranking possessed by individuals belonging to distinct groups in the hierarchy.

An individual (weakly) rank dominates another if that individual is ranked equal to or higher than the other in the possession of the attribute. For any population

[7]Massey and Denton (1988) refer to equivalent concepts.

partitioned into given identity groups, there are a fixed number of between-group pair-wise comparisons between individuals from different identity groups. The share of the total number of such between-group pair-wise comparisons involving a given group in which a member of the group rank dominates a member of some other group is called its level of group rank dominance. Group rank dominance is an indicator of the position the group occupies in the ordinal hierarchy of attribute levels. Another way to understand the difference between Figs. 3 and 4 is simply that the average rank of the whites and the blacks is different. This is clearly a necessary condition for distinct groups to be clustered in different parts of the attribute space. We establish in Appendix One that a monotonic relationship exists between the concepts of group rank dominance and of average rank. Both could be seen to be indicators of the placement of groups in the attribute hierarchy (in the extreme complete clustering of groups) and will thus be referred to as indicators of a group's rank sequence position.

The level of inequality in different groups' rank sequence position (whether as measured by group rank dominance or by average rank) indicates the extent to which a population is clustered. We refer to this concept of inequality as Sequence Inequality (*SI*). Some reflection will suffice to show that this is an unambiguous criterion even when group sizes differ. In any situation sequence, inequality is minimal when the groups are evenly interspersed or symmetrically placed around the median member(s).

It is clear from this discussion that while Figs. 3and 4 depict two groups with equal representational inequality, the two groups possess different levels of group rank dominance and average rank. In Fig. 4, whites have 100% of the available instances of rank domination and higher average rank.

While sequence inequality and representational inequality are related, they are also distinct concepts. A simple example which makes this distinction transparent is provided in Figs. 5 and 6. In Fig. 5, both groups possess the same level of group rank dominance and average rank. The black group has two of the possible four instances of rank domination as does the white group, and their average rank is the same. Thus, there is no sequence inequality between the groups. In the second, both groups again share equally in levels of group rank domination (both have two of the potential four instances once again) and have the same average rank. The situation once again is one in which there is no sequence inequality. However, in the first case, there is complete representational inequality and in the second case there is zero representational inequality. In neither case is group membership always associated with higher rank, yet the cases differ in the degree to which income levels are shared by members of distinct groups.

2.3 Group Inequality Comparison

Figure 4depicts a situation of maximal representational inequality and maximal sequence inequality. It could perhaps be thought of as a situation of polarization in the sense that each group is concentrated at a given pole of the income distribution.

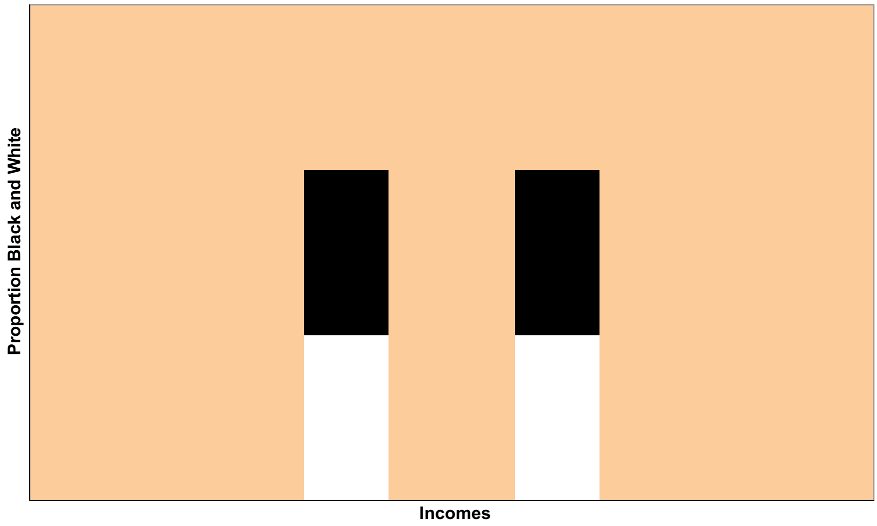


Fig. 5 Perfect sequence equality with perfect representational equality

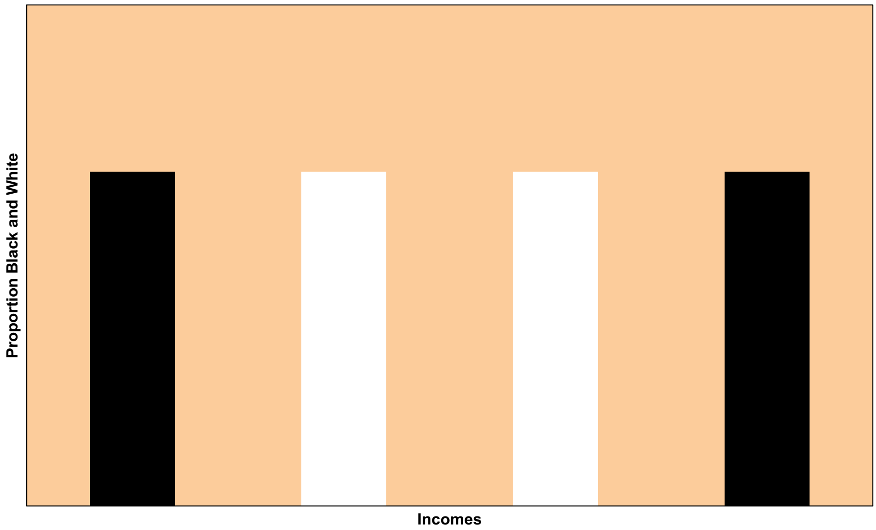


Fig. 6 Perfect sequence equality with complete segregation

However, this is true only in an ordinal sense. Both the situations depicted in Figs. 4 and 7 are *identical* from the standpoints of representational inequality and sequence inequality since neither concept takes note of cardinal information, which alone accounts for the difference between the two situations described. To take account of cardinal information (for instance, concerning the distance between distinct clusters), it is necessary to introduce an additional concept.

A common way to account for such information is to take note of the distance between the means of distinct subpopulations, for example, by using measures of inequality between group means. This indeed is the conception behind *Group Inequality Comparison (II).*However, such an approach ignores relevant information on *within*-group inequality. Consider a two-group society in which all members of each group originally, respectively, possess the mean incomes of their groups. Suppose that both groups experience within-group transfers leading to intragroup inequality. The extent of inequality in the society must be judged to have increased if the measure of inequality employed obeys the Pigou–Dalton Transfer Principle (ensuring that such transfers between persons are deemed to increase overall inequality). However, between-group inequality (understood in terms of inequality between mean incomes of groups) is unchanged. Between-group inequality must be deemed to have become relatively less substantial in comparison with total interpersonal inequality.

An approach to intergroup inequalities which is based on between-group inequalities in isolation rather than on the contribution of between-group inequality to overall interpersonal inequality (i.e., *Group Inequality Comparison (II)*) will fail to contrast situations that might be distinguished. Consider Fig. 8 which depicts a two-group

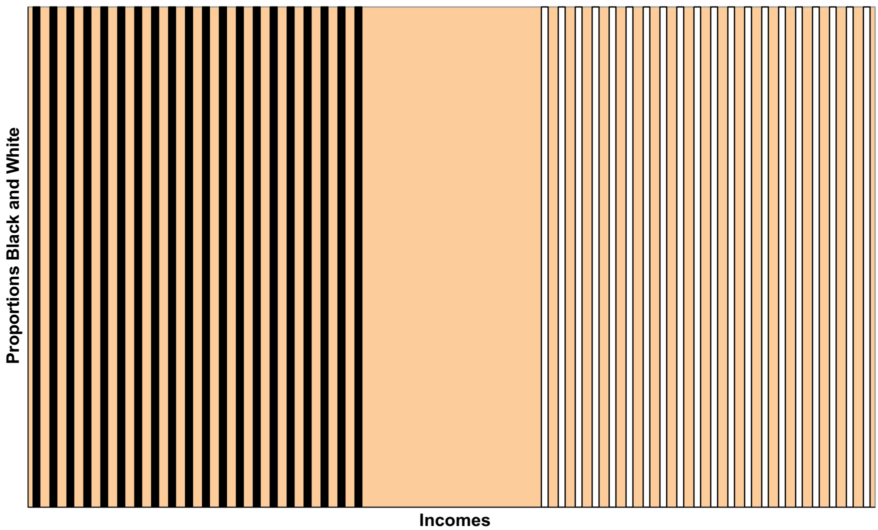


Fig. 7 Polarization

society in which all members of each group originally possess mean income A and B, respectively. Both groups now experience within-group transfers which increase inequality and their distributions are now depicted by densities A' and B', respectively. Assume further that the transfers are such that the span between the means is *D* and the span between the richest and poorest members of each group is also *D*. We might plausibly consider intergroup differences to have become less significant after the transfer since no member of the richer group is further away from some member of the poorer group than before the transfer, and all but the very richest member of the richer group is closer to some member of the poorer group.

On the other hand, Group *Inequality Comparison (I)* can have the disadvantage of ignoring information relevant for understanding the extent to which intergroup differences generate overall inequality. To see this, consider what would happen if in Fig. 8, the original populations A and B were made arbitrarily closer to each other while maintaining their separation. According to *Group Inequality Comparison (I)*,there would be no difference between the two situations. If we employed instead the concept of *Group Inequality Comparison (I)* the degree to which between-group differences generate inequality will have fallen. There are potentially good reasons to choose either approach.

Group Inequality Comparison need not be measured, of course, in terms of differences in means and could potentially be understood in other ways—for instance, in terms of differences in medians, generalized means, or other measures of central tendency. Indeed, still other ways of viewing group differences can be envisioned, for example, involving comparison of higher moments of the group-specific distributions of incomes, examination of the extent of "non-overlap" between distributions,

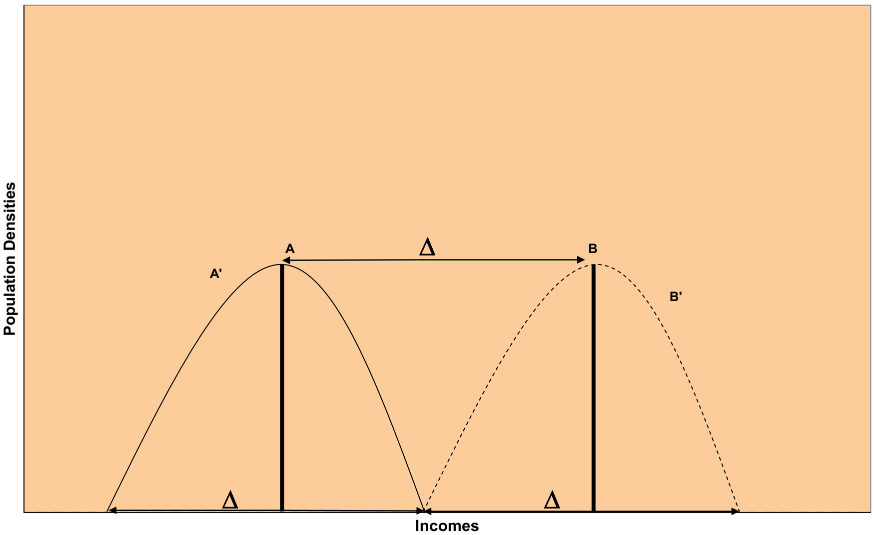


Fig. 8 Group inequality contribution versus inequality between means

etc. For a wide-ranging discussion of methods of defining group separation, see Anderson (2004, 2005). We limit our further discussions of the concept, however, to the case where it is measuring mean differences, for expositional simplicity.

2.4 Combining Concepts: Polarization

We have introduced above three concepts relating to intergroup inequalities: representational inequality, sequence inequality, and group inequality comparison. How are these concepts related to polarization? Polarization is a concept which has been used in many ways in the literature, for example, to mean the absence of "middleness" in a distribution (Wolfson 1994), the distance between the average achievements of groups (Østby, 2008) and the presence of distinct sizable groupings in the income distribution (Esteban and Ray, 1994). Many of these approaches do not explicitly rely on the identification of individuals by identity groups (understood as being distinct from attributes). A contrasting approach understands the level of polarization of a distribution in terms of the extent of intergroup differences in the possession of an attribute. If polarization is defined in this way, it becomes clear that each one of the concepts of intergroup inequality defined above is *itself* a measure of polarization. However, taken individually each may prove to be an unsatisfactory measure of polarization, because of the information to which each is individually indifferent. Thus, the relative ranking of the situations depicted in Figs. 3, 4 and 7 according to the extent of polarization depends on the expansiveness of the approach used. All the figures depict maximal polarization as judged according to *RI*, whereas Figs. 4 and 7 depict maximal polarization according to both *RI* and *SI*, and Fig. 7 depicts more polarization than does Fig. 4 according to *GIC* (taking the figures to possess the same income scale on the horizontal axis).

The fact that our judgments regarding the polarization of society may depend on more than one concept suggests the value of combining measures of intergroup differences to construct orderings of social situations according to the extent of their polarization. Such orderings can be partial and based on dominance of the vectors (two-tuples or three-tuples) defined by the individual measures of intergroup differences, or can be complete if based on some method of aggregation of these measures.

This said, orderings based on combining only a pair of the concepts we have defined (and not all three) will be indifferent to some important considerations that may be deemed relevant in any assessment of polarization. We have already seen that in the two-group case, combining representational inequality and sequence inequality will be sufficient to give us a measure of ordinal polarization. Such a combination however will be indifferent to cardinality and will be unable to distinguish, for example, between the situations depicted in Figs. 4 and 8, respectively.

A measure combining sequence inequality and group inequality comparison is not indifferent to cardinal information on the achievements of individuals but it is indifferent to the degree of clustering of identity groups in any specific income bracket. To see this, consider Figs. 5 and 6 again. Let us assume that, by construction, the mean income of both blacks and whites is the same in both groups in both situations. If this is the case, the index of group inequality comparison is the same in both figures (i.e., zero) and sequence inequality is the same, but representational inequality is different. We may argue that in Fig. 5 there is no clustering of identity groups in distinct parts of the income spectrum, as there is no representational inequality. In Fig. 6, however, blacks are clustered at the top and bottom ends of the income spectrum, and indeed there is complete segregation between the two groups. Note further that we could increase the distance between the blacks at the ends and the whites in the middle, keeping the means of both groups the same (so that the blacks at each end are very distant from the whites at the center) and yet record the same level of polarization defined according to such a measure.

Finally, combining representational inequality and group inequality comparison (I) alone leads to an approach that is indifferent to the sequencing of individuals from distinct identity groups in the income spectrum. Consider the distinction between Fig. 9a and b. Both depict cases of complete segregation. However, in Fig. 9b, some population of blacks has been moved to a higher income than all the whites, thereby increasing within-group inequality for the blacks and total interpersonal inequality. We can further imagine that every white has been given a higher income in such a way that within-group inequality among whites is unchanged and the ratio of between-group inequality to total inequality (which would otherwise have fallen) is restored to its level prior to the initial movement of blacks. In other words, the index of group inequality comparison (I) remains the same by construction, as does representational inequality. However, the sequencing of blacks and whites in the income distribution (and thus sequence inequality) is different. An analogous argument can be made for group inequality comparison (II) by moving the blacks and whites to keep mean incomes of the groups the same.

Any approach to polarization based on a pair of the group inequality concepts we have defined will capture certain judgments about social situations and neglect others. Only by combining all three concepts can an approach to polarization which takes account of the considerations reflected in each of the concepts be constructed.

A variant of group inequality comparison (I) has been proposed as a stand-alone measure of polarization (Zhang and Kanbur, 2001). However, such a measure, while attractive in its simplicity can violate some intuitions. Consider Fig. 10 in which two completely segregated and clustered groups A and B experience within-group progressive transfers which reduce within-group inequality. Further, suppose that they also experience a reduction of between-group inequality through progressive transfers between the members of the two groups in such a way that the ratio of between-group inequality to overall inequality remains unchanged and the groups (whose densities are now depicted by A' and B') overlap. If we utilize group inequality comparison (I) alone as our measure of polarization, a social configuration with A and B is viewed as being exactly as polarized as a situation with A' and B', which seems to conflict with our intuitions. If we, however, combine it with some measure of sequence inequality and/or representational inequality (both of which are lower

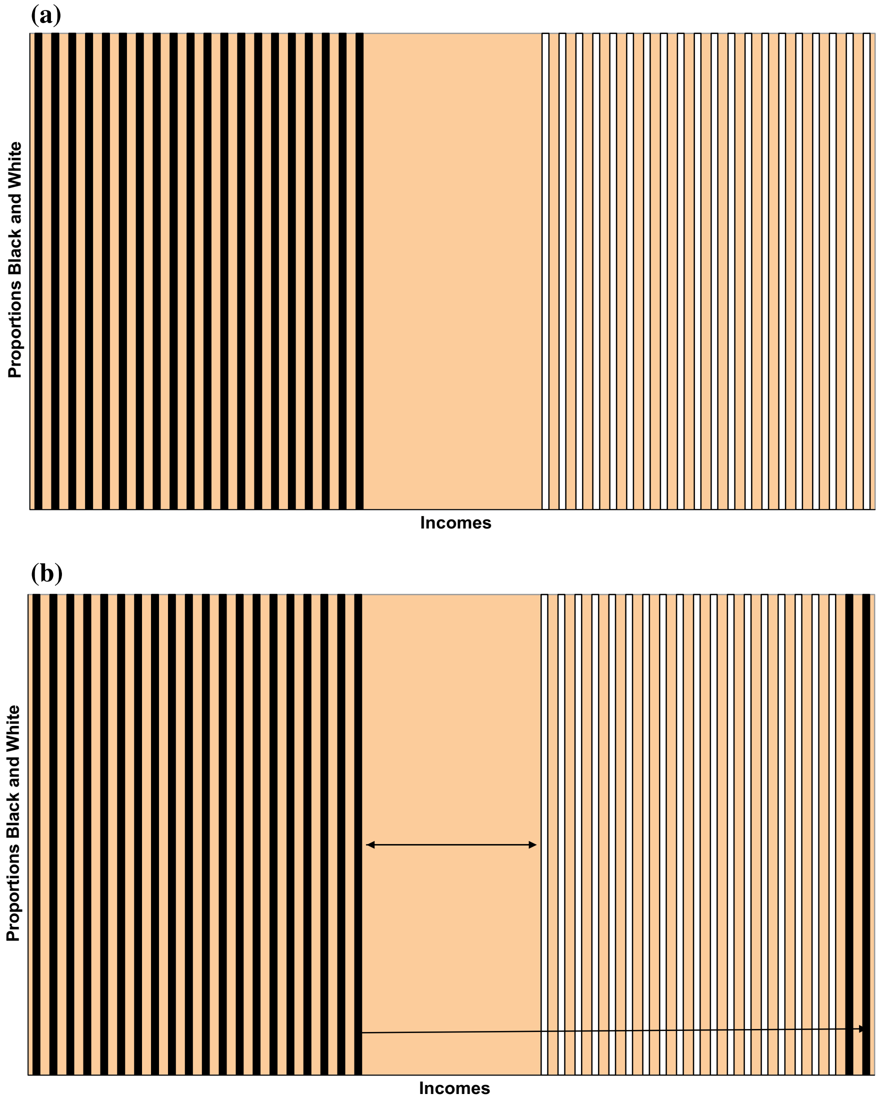


Fig. 9 **a** Maximal representational inequality, maximal sequential inequality with a fixed group inequality contribution, **b** maximal representational inequality, reduced sequential inequality with a fixed group inequality contribution

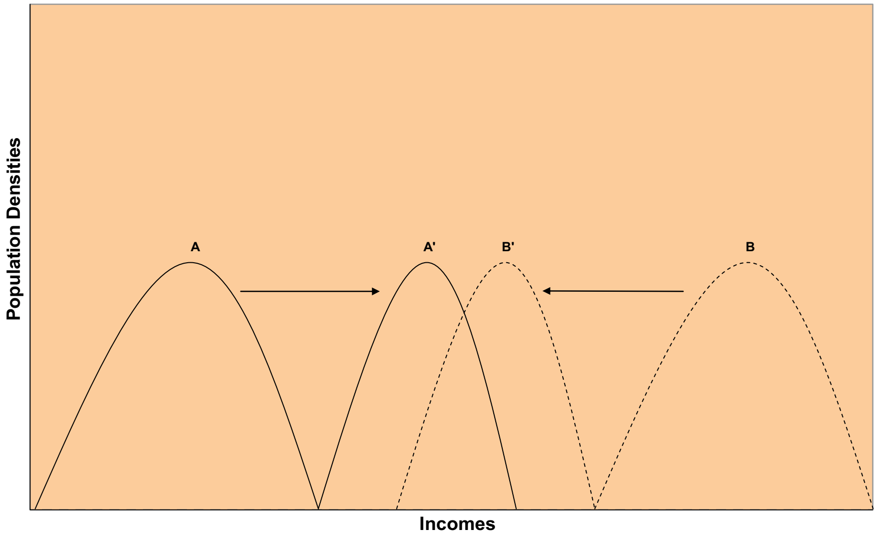


Fig. 10 Group inequality contribution alone is an incomplete measure of polarization

when the groups overlap), the first situation is unambiguously more polarized than the second.

It should be noted that the regressive transfers considered above led to a decrease in the index of group inequality comparison (I), and therefore their impact was in the *opposite* direction from that which would normally be expected of an inequality measure (i.e., to obey the Pigou–Dalton principle of responding to a regressive transfer with an increase in measured inequality). It follows that any measure of polarization which increases when the index of group inequality comparison (I) increases would similarly potentially violate the Pigou–Dalton principle.[8]

3 Part II: From Concepts to Measures

3.1 Formalizing Concepts

Our purpose in this section is to formalize the concepts relating to group differences which we have introduced above and develop measures of them.[9]

[8]This view corresponds to the findings of Esteban and Ray (1994) among others that polarization and inequality are distinct concepts and that measures of polarization need not therefore be expected to obey the Pigou–Dalton principle.

[9]These measures can be readily implemented using a Stata module that we have developed. For an example involving actual data, see Reddy and Jayadev (2011).

We begin by supposing a "social configuration" (ζ) in which there is a population, S_0, of individuals $\{i\}$ of size N partitioned[10] into K distinct identity groups (S_1, $S_2 \ldots . S_K$). The individuals possess an attribute (let us say y), drawn from an attribute set, Y. The attributes are not necessarily ordered. For example, the attribute may be a level of income (ordered and cardinally measured), a quality of health (ordered but not cardinally measured), or a club to which a person may belong (distinguished from one another, but not ordered). We employ a superscript to distinguish the information associated with distinct social configurations. For simplicity, we assume (although nothing depends on this other than notation) that the number of elements, l, in the set Y is finite.

More specifically, the individuals $\{i\}$ each belong to a distinct identity group $S_J \subseteq S_0 \quad (J \neq 0)$ so that

$\forall i \in S_0, i \in S_J$ for some $J (J \neq 0)$, with

$$S_J \bigcap S_M = \phi \, \forall J, M \in (1, \ldots K) \text{s.t } J \neq M \text{ and } \bigcup_{J=1}^{K} S_J = S_0.$$

Our assumptions imply there are at least two identity groups which are each smaller than the population as a whole and non-empty. Let the number of persons in group J be denoted by n_J. The proportion of persons of a group J in the society is defined by

$$\theta_J = \frac{n_J}{N} \text{for} \, [J \in (0, \ldots K)].$$

Each individual i has attribute y_i. The same attribute may be shared by more than one individual.

Define the membership function for group $J$by $M_J(y) = \#\{i \in S_J | y_i = y\}$, for $[J \in (0, \ldots K)]$. Moreover, define the complementary membership function for group J by $M_J^-(y) = \#\{i \notin S_J | y_i = y\}$. In other words, the membership function specifies the number of persons in group J who possess attribute y while the complementary membership function specifies the number of persons not in group J who possess attribute y.

[10]We do not consider currently the case of societies in which individuals belong to more than one identity group simultaneously and in which the identity groups do not form a partition of the society into mutually exclusive categories. Generally, a "maximal" partition of a society, generating a mutually exclusive and exhaustive set of groups, can be constructed by generating the Cartesian product of all of the identity groups in the society. This approach may not be deemed appropriate, however, in every situation. For example, a mixed-race group in a society otherwise divided into two races may be deemed to belong to *both* of the races rather than to neither, generating a different characterization of intergroup differences.

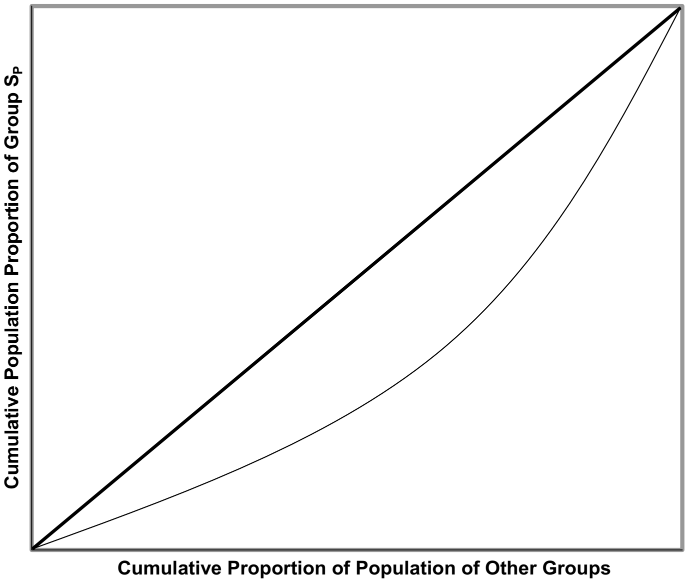


Fig. 11 The representational inequality Lorenz curve

3.2 Representational Inequality

A simple way to capture the degree to which each identity group is disproportionately represented among those who share a given attribute would be to describe the ratio of the number of the persons possessing a given attribute who belong to each group, J, to their overall number in society for any given attribute (y): $\frac{M_J(y)}{M_0(y)} = F_J(y)$. In other words, $F_J(y)$ refers to the proportion of persons who possess a given attribute who belong to group J. This information can be captured in what we call the Representational Inequality (RI) Lorenz curve (Fig. 11). As we shall see, this framework allows for a simple way of presenting information concerning these proportions and for analyzing this information using familiar tools.[11]

To construct the *RI* Lorenz Curve for each group, J, we first create a rank ordering, R_J, such that

$$F_J(y_{1_J}) \leq F_J(y_{2_J}) \leq \ldots F_J(y_{l_J}),$$

[11] In spirit, this approach is similar to that adopted by Duncan and Duncan (1955) and later, *inter alia*, by Silber (1989, 1991, 1992), and Hutchens (1991, 2004). Other references include Flückiger and Silber (1994). Boisso et al., (1994), and Reardon and Firebaugh, (2002). Silber notes that various information structures (for example, involving the frequencies with which distinct groups possess an attribute such as membership in an occupation) can be analyzed using "measures of dissimilarity" which are analogous to measures of inequality. Our approach builds upon this insight but differs from all of the authors above in explicitly going beyond the two-group case and aggregating information derived from the concentration curves of different groups.

where $F_J(y_{1_J}) \leq F_J(y_{2_J}) \leq \ldots F_J(y_{l_J})$ reflects the ordering of the attributes according to the proportion of the population in the attributes belonging to group J.The ordering starts from the attribute for which the proportion of the population consisting of members of group J is the lowest and proceeds to the attribute for which the proportion of the population consisting of members of group is the highest.

Clearly, in the case in which the attribute can itself be ordered (e.g., income), the sequence in which the y_{i_J} appear in the ordering R_J will not necessarily be from lowest to highest.

Define:

$\alpha_J(t) = \frac{\sum_{i=1}^{t} M_J^-(y_{i_J})}{N-n_J}$ and $\beta_J(t) = \frac{\sum_{i=1}^{t} M_J(y_{i_J})}{n_J}$, where $[t \in (0, \ldots l)]$ and $\alpha_J(0) \equiv 0$ and $\beta_J(0) \equiv 0$.

The RI Lorenz Curve for group J,$\widehat{L_J}$, can be defined by the following rule, which creates a piecewise linear curve:

When $x = \alpha_J(t)$, for integer values $[t \in (0, \ldots l)]$, then $\hat{L}_J(x) = \beta_J(t)$ and, when x is such that $\alpha_J(t) < x < \alpha_J(t+1)$, $[t \leq (l-1)]$, then

$$\widehat{L_J}(x) = \widehat{L_J}(\alpha_J(t))\lambda + \widehat{L_J}(\alpha_J(t+1))(1-\lambda), \text{ where } \lambda = \frac{x-\alpha_J(t))}{\alpha_J(t+1)-\alpha_J(t)}.$$

In using this definition, we follow the procedure described by Shorrocks (1983), p. 5.

This gives rise to a curve as shown in Fig. 11. By construction, the *RI* Lorenz curve must, in the familiar way, begin at (0,0) and end at (1,1), as well as slope upward, with the slope increasing as one moves to the right, since each addition to the total cumulative population of others is associated with an addition of a larger proportion of group *J*. Note that the 45-degree line here has the interpretation of being the line of equiproportionate representation (analogous to the line of perfect equality in the case of an ordinary Lorenz curve). That is, all along this line, the members of identity group *J* are represented at every attribute in the same proportion as they are represented in the population.

Any deviation from the line of equiproportionality represents a situation in which members of the group are disproportionately represented in the possession of certain attributes, leading them to be “over-represented” in the possession of certain attributes and “under-represented” in the possession of others. The *RI* Lorenz curve therefore contains information on the extent of segregation of a population in relation to the attributes possessed. Having defined it, we can draw on the analogy between the *RI* Lorenz curve and the ordinary Lorenz curve to suggest further useful concepts.

Consider for instance what might correspond to the familiar idea of a progressive transfer. Just as a progressive transfer in an income distribution involves a transfer from a person with higher income to a person with lower income, in the context of representational inequality, a progressive transfer could be defined as a transfer of a person from the set of persons who possess an attribute in which his or her identity group is represented more to one in which it is represented less. However, since we are dealing with proportions of identity groups possessing different attributes, a transfer of a single person will change the overall population that possesses each attribute involved, affecting the “denominator” used to assess population proportions for the groups possessing these attributes. We overcome this problem and maintain an unchanged denominator by instead employing the concept of a “balanced bilateral population transfer”[12]:

Definition Balanced Bilateral Population Transfers
Suppose $\exists(y_i, y_j) \in Y$ and S_P, S_Q such that

$$F_P(y_i) > F_P(y_j) \text{ and } F_Q(y_i) < F_Q(y_j)$$

with $P \neq Q$ *and* $i \neq j$.

Then, a progressive (regressive) balanced bilateral population transfer is one in which population mass Δ (i.e., some number of persons; we abstract from integer problems here) of group P is shifted from y_i to y_j and equal population mass Δ of group Q is shifted from y_j to y_i, thereby lowering (raising) $F_P(y_i)$ and $F_Q(y_j)$ while raising (lowering) $F_P(y_j)$ and $F_Q(y_i)$.

A balanced bilateral progressive population transfer results in two upward shifts in the *RI* Lorenz curves for the identity groups (and corresponding downward shifts for regressive transfers). An example of the latter is provided in Fig. 12a and b. The *RI* Lorenz curve that results from a progressive (regressive) balanced population transfer dominates (is dominated by) the *RI* Lorenz curve that preceded the transfer.[13] We note further that:

Lemma 1 *There exists a pair of identity groups and a pair of attributes* (y_i,y_j) *for which a progressive balanced bilateral population transfer can take place if all groups are not equiproportionately represented in the possession of every attribute.*

Proof: See Appendix Three

An *RI* Lorenz curve $\widehat{L}(x)$ weakly dominates an *RI* Lorenz curve $\widehat{L'}(x)$ if and only if $\widehat{L}(x) \geq \widehat{L'}(x)$ for all $x \in [0, 1]$. An implication of this framework is that any Lorenz consistent measure of inequality, for which inequality never decreases when $\widehat{L}(x)$ is replaced by $L(x)$, i.e., all income inequality measures used in practice can also be applied to measure representational inequality. It is also well known in the literature on income distribution that it is possible to shift from an income distribution that possesses a Lorenz curve $L(x)$ to another that possesses the Lorenz curve $L'(x)$ where $L(x) \leq L'(x)$ if and only if there exists a corresponding sequence of

---

[12]This concept of a balanced bilateral population transfer is related to that of a “disequalizing movement” between groups used by Hutchens (2004) in his discussion of a two-group case. However, the latter concept is insufficient in a multigroup case and necessitates the use of the alternative concept which we develop and employ. The concept is also intimately related to the idea of a “marginal preserving swap” which has appeared in the statistical literature (see, for example, Tchen, 1980, Schweizer and Wolff, 1981, and Bartolucci et al., 2001). However, to the best of our knowledge, no one has shown that in the absence of perfect representational equality, there always exists the possibility of achieving a balanced bilateral transfer (as we do in Appendix two) and that this can be given a natural interpretation in terms of Lorenz curves.

[13]For the relevant reasoning, see Shorrocks (1983).

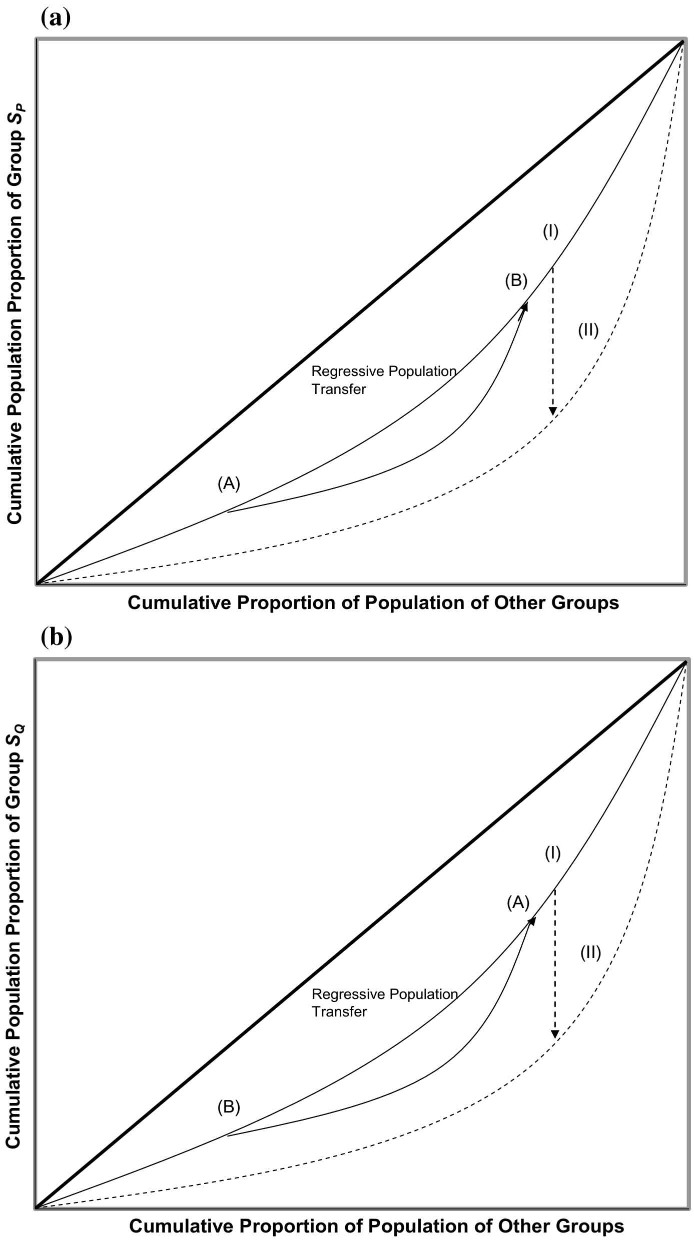


Fig. 12 **a** Balanced bilateral transfer, **b** balanced bilateral transfer

Table 1 Correspondences between conventional inequality and representational inequality concepts

Conventional inequality concept	Representational inequality concept
Inequality	Over- or under-representation
Pigou–Dalton transfers	Balanced bilateral population transfers
(First-order) Lorenz dominance	(First-order) *RI* Lorenz dominance

progressive transfers. Equivalently, in our case, it is possible to shift from a situation for which each group possesses a Lorenz curve $L_J(x)$ to another in which each group possesses a Lorenz curve $L'_J(x)$, where $L_J(x) \leq L'_J(x)$ if and only if there exists a corresponding sequence of balanced bilateral progressive population transfers. For this reason, a balanced bilateral progressive population transfer can be deemed to decrease overall representational inequality.

The consequence is a striking parallel between inequality measures in the income space and inequality measures in the representation space. Table 1 provides a map of the isomorphism between corresponding concepts introduced so far.[14]

Suppose that we apply Lorenz consistent inequality measure $\hat{I}(\hat{L}_J(x)$ to assess representational inequality for group J and denote the resulting vector of measured inequality for all groups in the society by $\bar{\hat{I}}$ and its individual components by $\bar{\hat{I}}_J = \hat{I}(\hat{L}_J(x))$, $J \in (1, \ldots K)$. Then, an overall measure of representational inequality in the society is given by $RI = f(\bar{\hat{I}}, \ldots)$,where $f(\bar{0}) = 0$, $f(\bar{1}) = 1$, and $\frac{\partial f}{\partial \hat{I}_J} \geq 0$ for all $J \in (1, \ldots K)$. One simple version of such an aggregation function, *f*, is the mean of the group-specific representational inequality measures. It may seem attractive for a measure of overall representational inequality to take into account subgroup sizes and respond to unequally sized groups differently. Indeed, it will be argued below that there can be sound reason for such weighting. We may define a population-weighted overall representational inequality measure of the following form:

$$RI = \frac{1}{K} \sum_{J=1}^{K} (\theta_J) \Big(\hat{I}(\hat{L}_J(x)) \Big),$$

where θ_J refers to the population weight of subgroup *J*.

Such a measure can be offered some justification through axiomatic underpinnings which we consider in the next section.

[14]The concepts of the generalized Lorenz curve and dominance of generalized Lorenz curves do not possess straightforward and interpretatively useful analogs in the area of representational inequality since the concept of an income mean does not possess a straightforward analog in this realm.

3.3 Sequence Inequality

As noted in the discussion of the previous section, representational inequality is a measure of group differences which is indifferent to the ordering of attributes as well as to their cardinal properties. To operationalize our concept of sequence inequality, therefore, we now assume that the attributes can be ordered.[15]

Considering first the concept of group rank dominance, we define a pair-wise individual rank domination function,δ_{ij}, for a given pair of individuals i and j as follows: $\delta_{ij} = 0$ if $y_i < y_j$ and $\delta_{ij} = 1$ if $y_i \geq y_j$. We can now define the group rank domination quotient for group J as follows:$\tau_J = \frac{\sum\limits_{i \in K} \sum\limits_{j \in (J \neq K)} \delta_{ij}}{n_J(N - n_J)}$. It can be seen that τ_J possesses the interpretation of the proportion of possible instances of pair-wise domination involving members of group J and members of other groups in which such domination actually occurs. It is evident that this quotient varies between a minimum of 0 and a maximum of 1 for any group. The size of the group plays no direct role in determining the value of the group rank domination quotient. Rather, it is the placement of members of the group relative to members of other groups that determine the quotient. Sequence inequality could be treated simply as the measured inequality in τ_J across groups. It is common for individuals from a given group to express pride or shame at the achievements or failures of other members from that group. Such a psychological interpretation can provide justification for treating the group rank domination quotient as defining the experience of everyone in that group and measure inequality across all individuals in possession of that experience.[16]

A seemingly puzzling asymmetry is implied by our approach to sequence inequality. Consider two populations, consisting of one white individual and ten black individuals each. In the first population, the individuals are ordered in the income space from lowest to highest as (w, b, b, b,....b), and in the second, the individuals are ordered from lowest to highest as (b, b, b, b.....w). In the first instance, all 10 black members possess a domination quotient of 1, while the white individual possesses a domination quotient of zero. The inequality in domination quotient is therefore inequality in a population having scores (0, 1,1,1,1....1). In the second case, all 10 black members possess a domination quotient of 0, while the white individual possesses a domination quotient of 1. The inequality in domination quotients is therefore the inequality in a population having scores (0,0,0,0,0....1). More sequence inequality will be recorded in the first case than in the second, even though all that has been done is to change the placement of the white from being at the bottom to being at the top of the income spectrum. While this may initially appear puzzling, it is perhaps appropriate to treat these cases asymmetrically. By the psychological interpretation,

---

[15]There is a small nascent literature on the measurement of ordinal inequality. Some key references include Allison and Foster (2004), Reardon (2008) and Abul Naga and Yalcin (2009).

[16]One way to interpret sequence inequalities is in terms of an analogy to a society wherein each group practices radical egalitarianism. In such a society, an even distribution of each group's share of the social assets, in this case instances of rank domination, results among the individuals belonging to the group.

in the first instance, most people in society do not experience a relative deprivation. By contrast, in the second, most do.

As we noted above, the average rank of a group (call it ω_J, $J \in (1, \ldots K)$) is also an indicator of group rank sequence position. In fact, it is linked in a direct and monotonic fashion to group rank dominance. It is easily shown that the relation between them, for a perfectly segregated population is

$$\tau_J = \frac{\omega_J - (n_J + 1)/2}{N - n_J}.$$

In the case of populations which are not perfectly segregated, appropriate changes to the definition of a rank maintain this relationship (see Appendix Two):

The inequality in group rank sequence position across groups can be assessed either in terms of the inequality of group rank dominance quotients or that of average group ranks.[17] In either case, if a member of a group (the "beneficiary") exchanges his or her attribute with another person in a different group who has a higher level of the attribute, then the indicator of group rank sequence position is increased for the group to which the beneficiary belongs and is decreased for the other group. We assume henceforth in this section that we are specializing to the case of group rank dominance quotients, although the concepts we present can equally be applied to average ranks.

The group rank dominance quotients achieved by members of distinct groups can be captured by what we call the Group Rank Dominance (GRD) Lorenz curve. The GRD Lorenz curve relates the cumulative proportion of the total of the group rank domination quotients to the cumulative population of groups, when the identity groups are ordered from lowest group rank domination quotient to highest. It captures the degree of inequality in group rank domination quotients. Any symmetric arrangement of identity groups in the attribute space (i.e., one in which for any instance in which a member of a given group rank dominates a member of another group, a distinct pair can be found in which the opposite is true) is one of the perfect equalities in group rank domination quotients, and will give rise to a GRD Lorenz curve which is on the 45-degree line.

We can now define coordinates of the GRD Lorenz curve associated with each group added as follows:

$\alpha(t) = \frac{\sum_{J=1}^{t} n_J}{N}$ and $\beta(t) = \frac{\sum_{J=1}^{t} \tau_J}{\sum_{J=1}^{K} \tau_J}$, where $[t \in (0, \ldots K)]$ and $\alpha(0) \equiv 0$ and $\beta(0) \equiv 0$.

We can now define the GRD Lorenz curve as a whole, $\widehat{LGRD}$, as follows.

[17] For a given group, although the ordinal ranking of social configurations according to the group's rank sequence position does not depend on the choice between these indicators (or indeed any other monotonic transformation thereof) the cardinal level of the indicator does depend on it. As a result, the choice of indicator can be consequential for determining the measured sequence inequality.

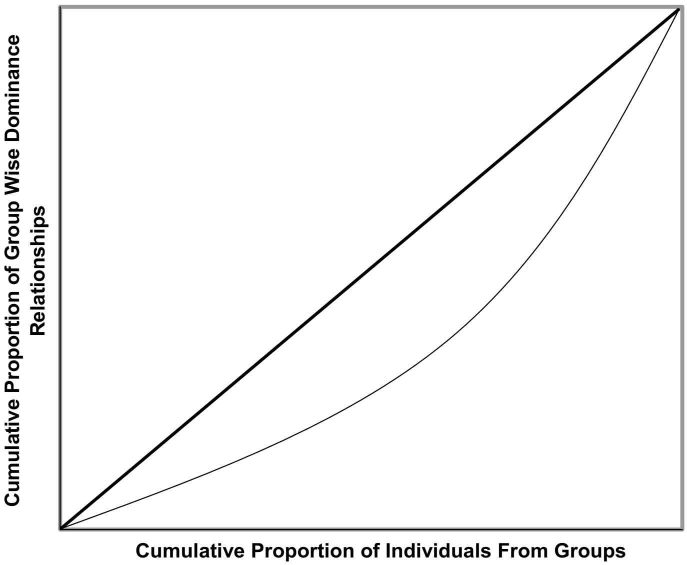


Fig. 13 The rank dominance Lorenz curve

When $x = \alpha(t)$, for integer values $[t \in (0, \ldots K)]$, then $\widehat{LGRD}(x) = \beta(t)$and, when x is such that $\alpha(t) < x < \alpha(t+1), t \leq (K-1)$, then $\widehat{LGRD}(x) = \widehat{LGRD}(\alpha(t))\lambda + \widehat{LGRD}(\alpha(t+1))(1-\lambda)$, where $\lambda = \frac{x-\alpha(t))}{\alpha(t+1)-\alpha(t)}$.

An example of such a curve is shown in Fig. 13. Since the Lorenz curve is defined for sequence inequality analogously to income inequality, with income corresponding to the group rank domination quotient of the groups to which individuals belong, the properties of the GRD Lorenz curve are analogous to those of the ordinary Lorenz curves. Once again, therefore, any Lorenz consistent measure of inequality will suffice to capture the level of sequence inequality.

3.4 Group Inequality Comparison

Group Inequality Comparison (I) refers to the degree to which between-group inequalities contribute to overall inequality. Typically, measures which are "additively separable" (such as members of the generalized entropy class) have been utilized for this purpose (see, for example, Shorrocks 1980, Foster and Shneyerov 1999 and Zhang and Kanbur 2001), although such a restriction is not required. In particular, if the between-group inequality is defined as the inequality that arises when every member of the population is assigned a representative level of an attribute (mean, generalized mean, median, or other measures of central tendency) of the group to which they belong, then the ratio of between-group inequality to total interpersonal inequality can serve as an index of Group Inequality Comparison (I). This measure has the advantage of always lying between zero and one and responding in an appropriate way to intragroup transfers. More generally, any indicator that the distributions associated with different groups are different can potentially serve as a measure of Group Inequality Comparison.

3.5 Polarization

Polarization as we have defined it above aggregates the three concepts concerning group differences which we have defined. The range of polarization measures which could be used is very wide indeed since any such measure could involve any form of aggregation of a three-tuple (*RI*,*SI*, and *GIC*), and in turn each element of this three-tuple could be defined in various ways. Further, any measure of polarization which is positively responsive to all three will only be maximized in a situation where all three are maximized.

An empirical examination which involves these four concepts can, as we have noted, be achieved using almost any common measure of inequality. The choice will naturally bring in additional implications and properties. Given this flexibility, an analyst can choose which measure to utilize to satisfy the additional properties thought important. Thus, for example, a researcher who wishes to treat sequence inequality as being decreased more in a situation where an exchange of ranks happens between members of different groups, each of whom has lower ranks to begin with, can choose an inequality measure which shows the required form of transfer sensitivity (e.g., a generalized entropy index with appropriately chosen parameters). Whether the measure of polarization can be normalized in a specific way will also depend on the choice of the underlying measures of inequality.

4 Part III: Axiomatic Framework

We define below some requirements that may reasonably be imposed on measures of each of the concepts defined above, considering each of them in turn. We also identify some classes of measures which satisfy these requirements.

4.1 Axioms (Representational Inequality)

We begin by suggesting some requirements which may be imposed on an overall representational inequality measure *RI* when it is viewed as a function of the information in a social configuration ζz. We write $RI = RI(\zeta)$ to reflect this dependence.

Axiom (RI1): *Lorenz Consistency*

Let *(ζ^1, ζ^2)* refer to two different social configurations and *(I,J)* refer to two different identity groups. If *(ζ^1, ζ^2)* are such that $\hat{L}_I^1 \geq \hat{L}_I^2$, $\hat{L}_J^1 \geq \hat{L}_J^2$, and $\hat{L}_H^1 = \hat{L}_H^2$, $\forall H \neq I, J, H \in (1, \ldots K)$, then *RI($\zeta^1$) ≤ RI(ζ^2).*

In other words, all else remaining equal, a social configuration which is at least as segregated according to the criterion of Lorenz dominance of representational inequality Lorenz curves is one which is at least as representational unequal. It may be noted that just as there is an equivalence between Lorenz consistency of an inequality measure and that measure's respect for the Pigou–Dalton Transfer Principle, there is an equivalence between Lorenz consistency of a representational inequality measure as defined here and the requirement that the representational inequality measure respond to a progressive balanced bilateral transfer by registering a decrease.

Axiom (RI2): Within-Group Anonymity

If y_{ij} represents the attribute of person i ($i \in (1, \ldots n_J)$) belonging to group J $((J \in (1, \ldots K))_J$ and
if *(ζ^1, ζ^2)* are such that $y_{iJ}^1 = y_{\pi_J(i)J}^2 \forall(i)$, where π_J is a permutation operator applied to $(1, \ldots n_J))$, then *RI(ζ^1) = RI(ζ^2).*

In other words, a measure of overall representational inequality is invariant to permutations of the attributes assigned to individuals within an identity group.

Axiom (RI3): Group Identity Anonymity

If y_{ij} represents the attribute of person i ($i \in (1, \ldots n_J)$) belonging to group J $((J \in (1, \ldots K))$ and if *(ζ^1, ζ^2)* are such that $y_{iJ}^1 = y_{i\pi(J)}^2$ and $n_J^1 = n_{i\pi(J)}^2$, $\forall(i)$,where π is a permutation operator applied to $(1, \ldots K)$, then *RI(ζ^1) = RI(ζ^2).*

In other words, a measure of overall representational inequality is invariant to permutations of the group identities with which distinct sets of individual attributes are associated. This axiom incorporates the idea that all the information relevant to assessing representational inequality is taken into account by noting the partition of the society into groups and the attributes of the members of these groups. The axiom embodies the idea that there is no need to take independent account of any other features of groups. This approach disallows the incorporation of judgments that group identities are *additionally* relevant (e.g., because of past histories or present injustices not already reflected in the information described by the social configuration).[18]

Axiom (RI4): Minimal Representational Inequality

Let $\hat{L}_E$ be the RI Lorenz curve corresponding to even representation (i.e., the line of equiproportionate representation). If $\hat{L}_J = \hat{L}_E \; \forall J \in (1, \ldots K)$, then *RI* =0.

In other words, minimal overall representational inequality is achieved when all identity groups are represented in the same proportion as their share of the population for all attributes, and has measure zero.

[18]See Loury, (2004) for an extensive discussion on the merits of the anonymity axiom as applied to groups.

Axiom (RI5): Maximal Representational Inequality

The maximum level of Representational Inequality is 1.

This is a normalization axiom which may be imposed for interpretative convenience. It may be dispensed with if it is desired to employ an unbounded inequality measure (such as a measure of the additively decomposable generalized entropy class).

Axiom (RI6): Positive Population Share Responsiveness of Overall Representational Inequality

Suppose that a measure of overall representational inequality is a function of the vector of measures of representational inequality of groups, $\bar{\hat{I}}$. Suppose further that the population share for group J is increased and that for group H is decreased, and the set of measures of representational inequality of groups remains unchanged as do the population shares for any remaining groups. Suppose further that $\bar{\hat{I}}_J > \bar{\hat{I}}_H$, i.e., that the group-specific representational inequality of group J is greater than that of group H. Then, the measure of overall representational inequality must increase.

This axiom can be motivated in different ways. We might, for example, believe that a group which is very small in the population but which is highly unequally represented simply because it is a small group in a society where there is unequal representation should not affect overall representational inequality in the same manner as a group which is much larger.

We may note that the measure of overall representational inequality defined above, $RI = \frac{1}{K} \sum_{J=1}^{K} (\theta_J)\left(\hat{I}(\hat{L}_J(x))\right)$, satisfies these axioms if the measure used to assess representational inequality for each group, $\hat{I}$, is Lorenz consistent, which will be the case if it has the form of any standard inequality measure, for example, the Gini coefficient.

From another perspective, it may not be appropriate disproportionately to disvalue the unequal representation of smaller groups. If one is interested in the experience of groups as opposed to the experience of individuals within groups, it should make no difference whether the group is small or large. Following this intuition, there is no reason to promote a population-weighted overall measure and one should instead adopt a measure which weights every group equally. This alternative may seem especially compelling if one views polarization as an attribute of the society as opposed to the individuals who belong to it. Such a measure can satisfy all the other axioms.

4.2 Axioms (Sequence Inequality)

In what follows, we shall use γ_J to refer to the indicator of group rank sequence position (which may be either the group rank domination quotient or the average rank) of group J. Let *SI* refer to the measure of overall sequence inequality. Some reasonable axioms are as follows:

Axiom (SI1): Lorenz Consistency

Let (ζ^1, ζ^2) refer to two different social configurations and *(I,J)* refer to two different identity groups. Further, let $\hat{L}$ refers to the Lorenz curve describing inequality across groups in the indicator of group rank sequence position, γJ. If (ζ^1, ζ^2) are such that $\hat{L}^1 \geq \hat{L}^2$, then $SI(\zeta^1) \leq SI(\zeta^2)$.

Axiom (SI2): Within-Group Anonymity

If y_{ij} represents the attribute of person i $(i \in (1, \ldots n_J))$ belonging to group J $((J \in (1, \ldots K))$ and
If (ζ^1, ζ^2) are such that $y^1_{iJ} = y^2_{\pi_J(i)J} \forall(i)$, where π_J is a permutation operator applied to $(1, \ldots n_J))$, then $SI(\zeta^1) = SI(\zeta^2)$.

Axiom (SI3): Group Identity Anonymity

If y_{ij} represents the attribute of person i $(i \in (1, \ldots n_J))$ belonging to group J $((J \in (1, \ldots K))$ and if (ζ^1, ζ^2) are such that $y^1_{iJ} = y^2_{i\pi(J)}$ and $n^1_J = n^2_{i\pi(J)}$, $\forall(i)$,where π is a permutation operator applied to $(1, \ldots K)$, then $SI(\zeta^1) = SI(\zeta^2)$.

Axiom (SI4): Sequence Inequality Limits

Let $\hat{L}_E$ be the Lorenz curve (describing inequality in the indicator of group rank sequence position, γ_J) that corresponds to even group rank sequence position (i.e., the case in which γ_J is the same for all groups). If $\hat{L} = \hat{L}_E$, then $SI = 0$.

Axiom (SI5): Maximal Sequence Inequality
The maximum level of sequence inequality is 1. As with Axiom *RI5* above, this is a normalization axiom which may be imposed for interpretative convenience. It may be dispensed with if it is desired to employ an unbounded inequality measure (such as a measure of the additively decomposable generalized entropy class).

4.3 Axioms (Group Inequality Comparison)

Some reasonable axioms may be as follows, if members of each group, *j*, are assigned a representative income, μ_j, and possesses an individual income, y_{ij}.

Axiom (GIC1): Between-Group Synthetic Population Lorenz Consistency

Let (ζ^1, ζ^2) refer to two different social configurations. Assume that a synthetic population is constituted in which every member of a group, *j*,is assigned the same representative income for its group, μ_j. Consider the Lorenz curve, $\tilde{L}^1$, $\tilde{L}^2$, for the resulting synthetic population in each social configuration. If (ζ^1, ζ^2) are such that $\tilde{L}^1 \geq \tilde{L}^2$ and $L^1 = L^2$(i.e., the overall Lorenz curves for the actual population remain unchanged), then GIC$(\zeta^1) \leq$ GIC(ζ^2).

This axiom states that between-group regressive transfers which do not change the overall interpersonal distribution must have an appropriate directional effect (nondecreasing) on the measure of GIC. Thus, for example, an exchange of incomes between individuals of different incomes belonging to two different groups that results in an increase in inequality in the synthetic population must increase the measure of GIC.

Axiom (GIC2): Within-Group Anonymity

If y_{ij} represents the attribute of person i ($i \in (1, \ldots n_J)$) belonging to group J (($J \in (1, \ldots K)$) and
if (ζ^1, ζ^2) are such that $y^1_{iJ} = y^2_{\pi_J(i)J} \forall(i)$, where π_J is a permutation operator applied to $(1, \ldots n_J)$), then $GIC(\zeta^1) = GIC(\zeta^2)$.

Axiom (GIC3): Group Identity Anonymity

If y_{ij} represents the attribute of person i ($i \in (1, \ldots n_J)$) belonging to group J (($J \in (1, \ldots K)$) and if (ζ^1, ζ^2) are such that $y^1_{iJ} = y^2_{i\pi(J)}$ and $n^1_J = n^2_{\pi(J)}, \forall(i, J)$, where π is a permutation operator applied to $(1, \ldots K)$, then $GIC(\zeta^1) = GIC(\zeta^2)$.

Axiom (GIC4): Within-Group Lorenz Consistency

Let (ζ^1, ζ^2) refer to two different social configurations. Further, let $\hat{L}^1_i$ and $\hat{L}^2_i$ refer to the Lorenz curves describing inequality within each group, i, in the respective social configurations. If (ζ^1, ζ^2) are such that $\hat{L}^1_i \geq \hat{L}^2_i$, but $\mu^1_i = \mu^2_i$, then GIC((ζ^1) $\geq$ GIC((ζ^2).

This axiom states that within-group (weakly) regressive transfers of income must have an appropriate directional effect (nonincreasing) on the measure of GIC, holding the representative incomes of groups constant. Clearly, since *Group Inequality Comparison (II)*does not rely on any information about within-group inequality, imposing this axiom will exclude its use.

It may be readily checked that a measure of *GIC* of the form *B/T*, where *B* represents the inequality measure for the synthetic population in which each member of the society is assigned the representative income of its group and *T* represents the total interpersonal inequality of the society, which satisfies all the axioms above. Such a measure would capture the concept of Group Inequality Comparison (I). In contrast, employing *B* alone as the measure of *GIC* would capture the concept of Group Inequality Comparison (II). Such a measure would satisfy Axioms (*GIC1*) − (*GIC3*) alone.

4.4 Axioms (Polarization)

We have proposed above to define polarization as a function of the other concepts of group difference we have defined. In a working version of this paper (Reddy and Jayadev, 2011), we provide conditions that yield a simple example of a polarization measure that permits the underlying inequality measure used to calculate *RI* and *SI* to be chosen flexibly as long as it is bounded and normalized to vary between zero and one[19]

$$P = (RI)(SI)(GIC).$$

It can be shown that it is the unique measure which satisfies the required conditions and is of the CES functional form. The measure is used to characterize group-based differences in various societies in Jayadev and Reddy (2011).

It is interesting to note that the circumstances in which this measure of polarization is maximized are different from those identified in Duclos, Esteban, and Ray (2004), in which this happens when there are two equal sized groups. The measure of polarization identified here can be maximized regardless of the number of groups, and to approach its maximum it is required that the poorer group be as large as possible relative to the richer groups, that there is complete segregation and that there is no within-group inequality (Fig. 14).

5 Part IV: Conclusion

This paper has sought to clarify how one may assess social situations according to the extent to which attributes are disproportionately possessed by different social groups. The measures we have developed capture the various ways in which experiences of members of distinct groups may differ. Thus, social situations can differ in the extent to which members of a group share experiences with members of other groups (representational inequality), experience the same or different relative positions (sequence inequality) and experience differences in the extent to which interpersonal inequalities are accounted for by intergroup differences (group inequality comparison). These concepts are distinct but complexly interrelated. They each integrate empirical observations and evaluative judgments. Judgments concerning the relative importance to be attached to different aspects of intergroup difference are also involved when they are combined (for example, to form a measure of polarization).[20] These measures have an intuitive appeal and can have widespread application in social science.

There appear to deep-seated tendencies for societies to exhibit segregation, clustering, and polarization of identity groups. This observation has important implications for both empirical investigations of societies and for social evaluation. We hope that, given Satya Chakravarty's lifelong concern with the assessment of social inequalities, he and others may find the concepts and measures that we have discussed to be useful.

[19]The requirement that *RI* and *SI* are bounded and normalized to vary between zero and one excludes certain inequality measures, such as the additively decomposable members of the generalized entropy class.

[20]The concepts we have discussed can be understood as "thick ethical concepts," on which see, e.g., Putnam (2004).

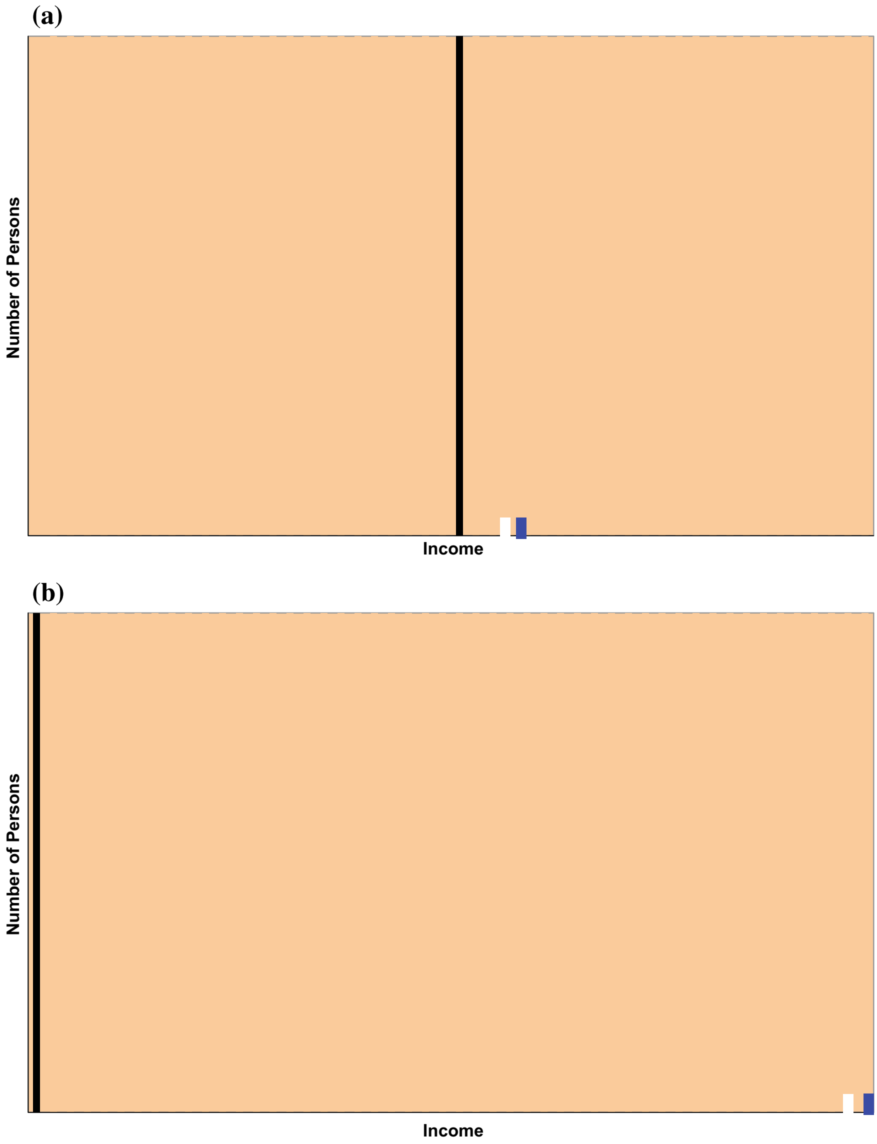


Fig. 14 **a** Very high polarization with GIC (l), **b** Very high polarization with GIC (lI)

Acknowledgements We would like to thank for their useful comments or suggestions (without implicating them in errors and imperfections) James Boyce, Indraneel Dasgupta, Rahul Lahoti, Hwok-Aun Lee, Glenn Loury, Yona Rubinstein, Peter Skott, Rajiv Sethi, Joseph Stiglitz, S. Subramanian, Roberto Veneziani, and other participants at seminars in the Dept. of Economics at Brown University, the University of Massachusetts at Amherst, the Jerome Levy Institute at Bard College, Queen Mary, University of London and the Brooks World Poverty Institute at the University of Manchester, and an anonymous referee for the current volume.

Appendix 1 (Rank Domination Quotient and Average Rank)

As we noted above, the average rank of a group (call it ω_J, $J \in (1, \ldots K)$) is also an indicator of group rank sequence position. In fact, it is linked in a direct and monotonic fashion to group rank dominance. As before, we understand rank as referring to the position in which an individual appears when incomes are sequenced from lowest to highest (the ascending order of values). When individuals from the same group share an income, we shall assign them a rank equal to the average position in which an individual appears when incomes are sequenced from lowest to highest. We shall consider subsequently the rule to be applied in assigning ranks when individuals from different groups share an income.

Consider at the outset, for simplicity, a perfectly segregated population in which there is no more than one individual in each income bracket. In such a population, the total number of instances of pair-wise rank domination that members of group J enjoy vis-à-vis others can be understood as a function of the ranks of members of group J in the population. The lowest ranked member of group J, having rank r_1 dominates $(r_1 - 1)$ persons belonging to other groups. The second lowest ranked member of group J, having rank r_2 dominates $(r_2 - 2)$ persons belonging to other groups (i.e., $(r_2 - 1)$ persons belonging to all groups -1 person belonging to the same group)). Extending this logic, the total number of instances of pair-wise domination by members of group J is

$$\sum_{i=1}^{n_J} (ri - i).$$

The rank domination quotient correspondingly is

$$\tau_J = \frac{\sum_{i=1}^{n_J} (r_i - i)}{N - n_J}$$

from which it follows that:

$$\tau_J = \frac{\omega_J - (n_J + 1)/2}{N - n_J}.$$

It is easy to see that this formula also applies in the case in which there may be more than one person in an income bracket but all persons who share an income bracket are always from the same group. In contrast, in the most general case of populations in which there may exist some income brackets which contain members of distinct groups there can be ties in the income ranks assigned to members of different groups, which will imply that this formula will no longer hold exactly unless the ranks are assigned appropriately to individuals in the same income bracket. Specifically, if strict domination is the concept that is employed then this relationship will hold exactly if individuals in the same income bracket are assigned a rank equal to the lowest of their positions in the ascending order of values. Correspondingly, if weak domination is the concept that is employed, then this relationship will hold exactly if each individual in the income bracket is assigned a value equal to the sum of the lowest of the positions of the individuals sharing the income bracket (in the ascending order of values) and the number of individuals from other groups with whom they share the income bracket.

The correspondence we have derived between τ_J and ω_J holds also in the case of continuous distributions, as can be shown through limit properties. In this case, the average rank of members of a group, *J*, is defined by

$$\omega_J = N \int F(x) g_J(x) dx$$

and the rank domination quotient for the group is defined by

$$\tau_J = \int (F(x) - \theta_J g_J(x)) dx,$$

where $F(x)$ is the cumulative distribution function for incomes of the entire population, $g_J(x)$ is the density function for incomes of members of the group, *J*, and the integrals are calculated over the domain of all possible incomes.

Appendix 2 (Proof of Lemma 1)

Without loss of generality, we shall assume that the attributes can be understood as income levels. Let $\boldsymbol{A}$ refer to a matrix of size *K* by *n* with *K* identity groups and *n* income levels. Each element in the matrix $a_{ij} \in \{0, 1, 2\}, \forall (i, j)$. We say that the *ith* identity group is “under-represented” at the *jth* income level if the proportion of persons from group *i* at income *j* is less than the proportion of persons of group *i* in the population as a whole. We denote the statement that the *ith* identity group is “under-represented” at the *jth* income level by $a_{ij} = 0$. We say further that the *ith* identity group is “over-represented” if the proportion of persons from group *i* at income *j* is greater than the proportion of persons of group *i* in the population as a whole. We denote the statement that the *ith* identity group is “over-represented” at the *jth* income level by $a_{ij} = 1$. If the *ith* identity group is represented at the *jth* income level in the same proportion as it is represented in the population as a whole, then we say that it is "equiproportionally represented" and we denote this by $a_{ij} = 2$.

Thus, $\boldsymbol{A}$ is a matrix in which every element is 0,1, or 2. We may further note that if any row or any column contains a zero, then it must contain a one and vice versa. This requirement captures the necessity that if an identity group is over-represented at an income level, it must be under-represented at another income level and that if a group is over-represented at an income level, then another group is under-represented at that same income level.

A balanced bilateral transfer is always possible if an identity group is represented to a greater extent at one income level (call it y_1) than it is at another (call it y_2) and another identity group is represented to a lesser extent at y_1 than it is at y_2. This condition is satisfied as long as it is possible to identify two rows (i and j) and two columns (l and m) of the matrix A such that they form a matrix $A^{\sim} = \begin{pmatrix} a_{il} & a_{im} \\ a_{jl} & a_{jm} \end{pmatrix}$ which is of one of the following forms, or which can be constructed from one of the following forms by permuting either their rows or their columns:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

The lemma is therefore equivalent to the statement that there exists a matrix $\boldsymbol{A}^{\sim}$ for any matrix $\boldsymbol{A}$ which contains at least a single one or zero. Suppose that the lemma is false. Then, it is possible to construct an $\boldsymbol{A}$ such that there is no $\boldsymbol{A}^{\sim}$ associated with it.

We now try to construct such a matrix $\boldsymbol{A}$. Without loss of generality, consider the case in which **A** contains at least one zero (i.e., an identity group is under-represented at a particular level of income). We can present this as occurring at the top left corner (a_{11}) of the matrix, without loss of generality, as given below:

$$\boldsymbol{A} = \begin{pmatrix} 0 & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kn} \end{pmatrix}.$$

This however means that there must be at least one level of income in the first column and in the first row in which there is over-representation of an identity group. Without loss of generality, let us say that this occurs at a_{12} and a_{21}, respectively, so that

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & \dots a_{1n} \\ 1 & a_{22} & \cdots a_{2n} \\ \vdots & \dots & \ddots \\ & & ak1 \end{pmatrix}.$$

Now, if $a_{22} = 0$ or 2, then $\boldsymbol{A}^{\sim}$ exists. If $a_{22} \neq 0$ or 2, then $a_{22} = 1$. That is

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & \ldots a_{1n} \\ 1 & 1 & \cdots a_{2n} \\ \vdots \cdots & & \ddots \\ & & ak1 \end{pmatrix}.$$

Consider row 2 and column 2 now. Since for the already fixed elements, there is over-representation, there must be elements in row 2 and in column 2, respectively, that have value zero (reflecting under-representation). Without loss of generality, let these occur at a_{23} and a_{32}, respectively, so that

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & a_{13}\ldots \\ 1 & 1 & 0\ldots\ldots \\ a31 & 0 & a33\ldots \\ & & \vdots \end{pmatrix} = \begin{pmatrix} 0\,1\,0\ldots\ldots \\ 1\,1\,0\ldots\ldots \\ 0\,0\,0\ldots\ldots \\ \vdots \end{pmatrix}.$$

But this in turn fixes a_{13}, a_{31}, and a_{33} to be 0 since if any of these are 1 or 2, we can construct matrix $\boldsymbol{A}^{\sim}$. This in turn implies that there exist elements elsewhere in row 3 and column 3 with value 1 (indicating over-representation), which we can place without loss of generality at a_{34} and a_{43}, respectively. It can readily be seen that this in turn fixes a_{41}, a_{42}, a_{44}, a_{24}, and a_{14} to be 1 since if any of these are 0 or 2, we can construct matrix $\boldsymbol{A}^{\sim}$. Thus we may construct a matrix $\boldsymbol{A}$ such that $a_{ij} = a_{ji} = 0$, if i is odd and $j \leq i$ and $a_{ij} = 1$ otherwise.

Let us now consider the matrix where the row $(k - 1)$ is odd. This means that $\boldsymbol{A}$ has the following form:

$$\boldsymbol{A} = \left(\begin{array}{ccccc|c} \vdots & \vdots & \vdots & \vdots & \vdots & a1n \\ \cdots 1 & 1 & 0 & 1 & 0 & \vdots \\ \cdots 0 & 0 & 0 & 1 & 0 & \vdots \\ \cdots 1 & 1 & 1 & 1 & 0 & \vdots \\ \cdots 0 & 0 & 0 & 0 & 0 & \vdots \\ ak1 & \cdots & \cdots & \cdots & \cdots & akn \end{array}\right)$$

This in turn implies that $a_{k-1,n} = 1$ and $a_{k,n-1} = 1$. It may be verified that for $\boldsymbol{A}^{\sim}$ not to exist all elements in row k and in column n must equal 1. However, this violates the requirements on a matrix $\boldsymbol{A}$.

Consider now the matrix where the row $(k - 1)$ is even. This means that $\boldsymbol{A}$ has the following form:

$$\boldsymbol{A} = \left(\begin{array}{ccccc|c} \vdots & \vdots & \vdots & \vdots & \vdots & a1n \\ \cdots 0 & 0 & 1 & 0 & 1 & \vdots \\ \cdots 1 & 1 & 1 & 0 & 1 & \vdots \\ \cdots 0 & 0 & 0 & 0 & 1 & \vdots \\ \cdots 1 & 1 & 1 & 1 & 1 & \vdots \\ ak1 & \cdots & \cdots & \cdots & \cdots & akn \end{array} \right)$$

This in turn implies that $a_{\mathrm{k}-1,\mathrm{n}} = 0$ and $a_{\mathrm{k,n}-1} = 0$. It may be verified that for $\boldsymbol{A}^{\sim}$ not to exist all elements in row k and in column n must equal 0. However, this violates the requirements on a matrix $\boldsymbol{A}$.

Thus, it is not possible to construct a matrix $\boldsymbol{A}$ such that $\boldsymbol{A}^{\sim}$ does not exist.

$A^{\sim}$ must exist, thereby proving the lemma.

QED

References

Abul Naga RH, Yalcin T (2009) Inequality Measurement for Ordered Response Health Data. J Health Econ 27(6):1614–1625

Alesina A, La Ferrara E (2000) Participation in heterogeneous communities. Q J Econ 115, 847–904

Alesina A, La Ferrara E (2002) Who trusts others? J Public Econ 85, 207–34

Alesina A, Devleeschauwer A, Easterly W, Kurlat S, Wacziarg R (2003) Fractionalization. J Econ Growth 8, 155–194

Allison RA, Foster J (2004) Measuring health inequalities using qualitative data. J Health Econ 23:505–524

Anderson E (1999) What is the point of equality? Ethics 1999:287–337

Anderson G (2004) Toward an Empirical Analysis of Polarization. Journal of Econometrics 122(2004):1–26

Anderson G (2005) Polarization. Working Paper, University of Toronto. Available at http://www.chass.utoronto.ca/~anderson/Polarization.pdf

Arneson R (1989) Equality and equal opportunity for welfare. Philos Stud 1989:77–93

Bartolucci F, Forcina A, Dardanoni V (2001) Positive quadrant dependence and marginal modelling in two-way tables with ordered margins. J Am Stat Assoc 96(456), 1497–1505

Boisso D, Hayes K, Hirschberg J, Silber J (1994) Occupational segregation in the multidimensional case: decomposition and tests of statistical significance. J Econ 61:161–171

Chakravarty S, Maharaj B (2009) A study on the RQ index of ethnic polarization. ECINEQ WP 134

Cohen GA (1989) On the currency of egalitarian justice. Ethics 1989:906–944

Duclos J-Y, Esteban J, Ray D (2004) Polarization: concepts, measurement, estimation. Econometrica 1737–1772

Duncan OD, Duncan B (1955) A methodological analysis of segregation indexes. Am Sociol Rev 20(2):210–217
Dworkin R (2000) Sovereign virtue: the theory and practice of equality. Harvard University Press, Cambridge, MA
Esteban J-M, Ray D (1994) On the measurement of polarization. Econometrica 62:819–851
Foster J, Shneyerov A (1999) A general class of additively decomposable inequality measures. Econ Theory 14(1)
Flückiger Y, Silber J (1994) The gini index and the measurement of multidimensional inequality. Oxf Bull Econ Stat 56:225–228
Hutchens R (1991) Segregation curves, lorenz curves, and inequality in the distribution of people across occupations. Math Soc Sci 21:31–51
Hutchens R (2004) One measure of segregation. Int Econ Rev 45(2):555–578
Jayaraj D, Subramanian S (2006) Horizontal and vertical inequality: some interconnections and indicators. Soc Indic Res 75(1):123–139
Jayadev A, Reddy S (2011) Inequalities between groups: theory and empirics. World Dev 39(2), 159–173
Loury GC (2004) The anatomy of racial inequality. Harvard University Press, Cambridge, MA
Miguel E, Gugerty M (2005) Ethnic diversity, social sanctions, and public goods in Kenya. J Public Econ 89(11–12):2325–2368
Montalvo JG, Reynal-Querol M (2005) Ethnic polarization, potential conflict and civil war. Am Econ Rev 95 (3), 796–816
Østby Gudrun (2008) Polarization, horizontal inequalities and violent civil conflict. J Peace Res 45(2):143–162
Putnam H (2004) The collapse of the fact/value dichotomy and other essays. Harvard University Press, Cambridge MA
Rawls J (1971) A theory of justice. Harvard University Press, Cambridge, MA
Reardon SF, Firebaugh G (2002) Measures of multigroup segregation. Sociol Methodol 32:33–67
Reardon S (2008) Measures of ordinal segregation. Working paper 2008–2011, Institute for Research on Education Policy and Practice, Stanford University. Forthcoming in Research on Economic Inequality, vol 17
Reddy S, Jayadev A (2011) Inequalities and identities. Available at SSRN: https://ssrn.com/abstract=1162275 or http://dx.doi.org/10.2139/ssrn.1162275
Roemer J (1996) Theories of distributive justice. Harvard University Press, Cambridge, MA
Schweizer B, Wolff EF (1981) On nonparametric measures of dependence for random variables. Ann Stat 9(4):879–885
Sen A (2007) We can best stop terror by civil, not military, means. The Guardian, Friday November 9th, 2007
Sen A (1992) Inequality reexamined. Harvard University Press, Cambridge, MA
Shorrocks AF (1980) The class of additively decomposable inequality measures. Econometrica 48(3):613–625
Shorrocks AF (1983) Ranking income distributions. Economica 50:3–17
Silber JG (1989) On the measurement of employment segregation. Econ Lett 30(3):237–243. Elsevier
Silber JG (1991) Inequality indices as measures of dissimilarity: a generalization. Stat Pap 1991(32):223–231
Silber JG (1992) Occupational segregation indices in the multidimensional case: a note. Econ Rec 68:276–277
Stewart F (2001) Horizontal inequalities: a neglected dimension of development. WIDER Annual Lectures 5, UNU WIDER
Tchen AH (1980) Inequalities for distributions with given marginals. Ann Probab 8(4):814–827
Wolfson MC (1994) When inequalities diverge. Am Econ Rev Pap Proc 94:353–358
Zhang X, Kanbur R (2001) What difference do polarisation measures make? an application to china. J Dev Stud 37:85–98

A Generalization of the Theil Measure of Inequality

Swami Tyagarupananda and Nachiketa Chattopadhyay

Abstract A general measure of inequality which, in the limit, converges to the Theil measure is developed, based on the Generalized Entropy of Degree α (Aczel and Daroczy 1975). The properties of the measure are discussed with an empirical illustration using Indian Consumer Expenditure Survey data.

Keywords Information · Inequality · Entropy

1 Introduction

Information measures have been used in economics to measure inequality and concentration. The Shannon measure of information has been widely used in inequality literature. This has generated the Theil measure of inequality (see Theil 1967 and Sen 1973). In this paper, we provide a general measure of inequality which, in the limit, converges to the Theil measure. This is based on the Generalized Entropy of degree α (see Aczel and Daroczy 1975).

We discuss the information measures and some of their properties in Sect. 2. In Sect. 3 the corresponding inequality measures are derived, while their properties are analyzed in Sect. 4. An empirical illustration using NSS data is presented in Sect. 5. Section 6 concludes the paper.

The authors would like to thank Hitesh Tripathi for the computational assistance.

S. Tyagarupananda
Ramakrishna Math and Mission, Malda, West Bengal, India
e-mail: malda@rkmm.org

N. Chattopadhyay (✉)
Sampling and Official Statistics Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700 108, India
e-mail: nachiketa@isical.ac.in

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*, Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_8

2 Information Theory Results

The results presented in this section are taken from Aczel and Daroczy (1975). The Shannon Entropy of an experiment with outcomes $A_1, A_2, \ldots, A_m$ of probabilities $p_1, p_2, \ldots, p_m$ is

$$S(P) = S(p_1, p_2, \ldots, p_m) = -\sum_{i=1}^{m} p_i \, log \, p_i, \ p \in T^m, \tag{1}$$

where T^m is the set such that $p \in T^m$ implies $\sum_{i=1}^{m} p_i = 1$, $p_i \geq 0$, for all i.

The Shannon Entropy measure satisfies a number of algebraic properties of which we note the following.

Strong Additivity:

$$S(p_1q_{11}, p_1q_{12}, \ldots, p_mq_{m1}, \ldots, p_mq_{mn}) = S(p_1, p_2 \ldots, p_m) + \sum_{i=1}^{m} p_i S(q_{i1}, q_{i2} \ldots, q_{in}), \tag{2}$$

where, $p \in T^m, q_i \in T^n$

The vector $q_i = (q_{i1}, q_{i2} \ldots, q_{in})$ gives the conditional probability of n mutually exclusive and exhaustive events, given that the event A_i has occurred. The probability of occurrence of the event A_i is P_i. The property (2) thus has a simple interpretation, the combined entropy of mn compound events is equal to the entropy of the m events $A_1, A_2, \ldots, A_m$ added to a weighted sum of the entropy of the conditional probabilities $(q_{i1}, q_{i2} \ldots, q_{in})$, the weight being equal to the probability of event A_i.

Maximality:
The maximum value of the Shannon Entropy measure is obtained when $p_1 = p_2 = \cdots = p_n = 1/n$. In this case,

$$max S = Max(-\sum_{k=1}^{n} p_k \, log \, p_k) = log \, n. \tag{3}$$

The Shannon Entropy function has been generalized to a much wider class, the Generalized Entropy of Degree α. This class, too, has a number of properties corresponding to the properties of the Shannon Entropy function. The Generalized Entropy of Degree α is given by

$$\begin{aligned} G_\alpha &= \frac{1}{2^{1-\alpha} - 1}(\sum_{k=1}^{n} p_k^\alpha - 1), \ \alpha \neq 1, \ \alpha \geq 0 \\ &= -\sum_{k=1}^{n} p_k \, log \, p_k = S, \ \alpha = 1. \end{aligned} \tag{4}$$

Thus, in the limit as $\alpha \to 1$, the Generalized Entropy class converges to the Shannon Measure of Information. The measure in (4) satisfies the property of Strong Additivity of Degree α. It can be verified that

$$\begin{aligned} & G_\alpha(p_1q_{11}, \ldots, p_1q_{1n}, p_2q_{21}, \ldots, p_mq_{m1}, \ldots, p_mq_{mn}) \\ & = G_\alpha(p_1, p_2, \ldots, p_m) + \sum_{i=1}^{m} p_i^\alpha G_\alpha(q_{i1}, q_{i2} \ldots, q_{in}). \end{aligned} \tag{5}$$

This condition is a generalization of the condition of Strong Additivity [Eq. (2)] which the Shannon Entropy Measure satisfies. The vector $(q_{i1}, q_{i2} \ldots, q_{in})$ again gives the conditional probability of n mutually exclusive and exhaustive events, given that the event A_i has occurred. As earlier, the probability of occurrence of the event A_i is p_i. The combined entropy of mn compound events is equal to the entropy of the m events $A_1, A_2, \ldots, A_m$ added to a weighted sum of the entropy of the conditional probabilities $(q_{i1}, q_{i2} \ldots, q_{in})$, the weight being equal to p_i^α. As $\alpha \to 1$ and the Generalized Entropy class tends to the Shannon Entropy Function which satisfies Eq. (2), the condition of Strong Additivity. For Maximal, note that the index G_α, again, is maximum when $p_i = 1/n$ for all i. Thus,

$$\begin{aligned} max G_\alpha &= \frac{n^{1-\alpha} - 1}{2^{1-\alpha} - 1}, \ \alpha \neq 1, \ \alpha > 0 \\ &= log\, n, \quad \alpha = 1. \end{aligned} \tag{6}$$

As $\alpha \to 1$, the maximal value of G_α coincides with that of S.

3 Inequality Measures

3.1 Theil Measure

Let s_i be the share of the i th person in the income profile $Y = (y_1, y_2, \ldots, y_n)$. So $s_i = \frac{Y_i}{n\mu}$, where μ is the mean income of Y. If we substitute the probabilities in the Shannon Entropy Function by the shares of incomes (both the probabilities and shares add up to unity), we have the Shannon Entropy for the profile Y

$$\begin{aligned} S &= -\sum_{i=1}^{n} \frac{y_i}{n\mu} log \frac{y_i}{n\mu} \\ &= log\, n - \sum_{i=1}^{n} \frac{y_i}{n\mu} log \frac{y_i}{\mu} \\ &= log\, n - \sum_{i=1}^{n} s_i \, log\, ns_i \\ &= log\, n - T, \end{aligned} \tag{7}$$

where

$$T = \sum_{i=1}^{n} s_i \, log \, ns_i$$
$$= \sum_{i=1}^{n} \frac{y_i}{n\mu} \, log \, \frac{y_i}{\mu} \tag{8}$$

is the Theil Measure of Inequality. From (7), it is seen that the Theil Measure is equal to the maximal value of the Shannon Measure less the value of the Shannon Measure for the income profile Y. We approximate the probabilities p_i by the income shares $s_i = \frac{y_i}{n\mu}$. Theil (1967) and Sen (1973) argue that the value of S can be regarded as a measure of equality. By subtracting the actual equality from the maximum equality, we get the measure of inequality, T, introduced by Theil. The Theil Measure T achieves its maximum value when all the income is monopolized by a single person whose income becomes $n\mu$. The rest have 0 incomes. In this case, $T = log\, n$ (assuming $0\, log\, 0 = 0$). The measure T achieves its maximum when the measure of equality S becomes 0 in value.

3.2 *The Inequality Measure I_G*

Now, we derive the class of inequality measures obtainable from the Generalized Entropy of Degree α. Putting $s_i = \frac{y_i}{n\mu}$, the share of the i th person in place of the probabilities in (4), we have for the income vector Y

$$\begin{aligned} G_\alpha(Y) &= \frac{1}{2^{1-\alpha}-1}\Big(\sum_{i=1}^{n}\Big(\frac{y_i}{n\mu}\Big)^\alpha - 1\Big) \\ &= \frac{1}{2^{1-\alpha}-1}\Big(\frac{1}{n^\alpha}\sum_{i=1}^{n}\Big(\Big(\frac{y_i}{\mu}\Big)^\alpha - 1\Big) + n^{1-\alpha} - 1\Big) \\ &= \frac{n^{1-\alpha}-1}{2^{1-\alpha}-1} - \frac{1}{n^\alpha(1-2^{1-\alpha})}\sum_{i=1}^{n}\Big(\Big(\frac{y_i}{\mu}\Big)^\alpha - 1\Big),\ \alpha \neq 1,\ \alpha \geq 0 \\ &= -\sum_{i=1}^{n}\frac{y_i}{n\mu}\, log\, \frac{y_i}{n\mu},\ \alpha = 1. \end{aligned} \tag{9}$$

The inequality measures derived from the Generalized Entropy of Degree α, in line with the derivation of the Theil Measure in (8) is given by

$$\begin{aligned} I_G &= \frac{1}{n^\alpha(1-2^{1-\alpha})}\Big(\sum_{i=1}^{n}\Big(\frac{y_i}{\mu}\Big)^\alpha - 1\Big) \\ &= \frac{\sum_{i=1}^{n} s_i^\alpha - n^{1-\alpha}}{1-2^{1-\alpha}}. \end{aligned} \tag{10}$$

From (9), we have

$$I_G = \frac{n^{1-\alpha} - 1}{2^{1-\alpha} - 1} - G_\alpha.$$

If G_α is again interpreted as a measure of equality, I_G is obtained by subtracting the actual equality from the maximum equality. The value of I_G is $\frac{n^{1-\alpha}-1}{2^{1-\alpha}-1}$ when all incomes accrue to a single person. In this case, the measure of equality G_α has the value 0. At the other extreme, when the incomes of all persons are equal to the mean of the profile the inequality measure I_G has the value 0. Again,

$$lim_{\alpha\to 1}\, max\, I_G = lim_{\alpha\to 1} \frac{n^{1-\alpha} - 1}{2^{1-\alpha} - 1} = log\, n$$

$$\text{and } lim_{\alpha\to 1}\, I_G = lim_{\alpha\to 1} \frac{1}{n^\alpha(1 - 2^{1-\alpha})} \sum_{i=1}^{n} ((\frac{y_i}{\mu})^\alpha - 1) \tag{11}$$

$$= \sum_{i=1}^{n} \frac{y_i}{n\mu}\, log\, \frac{y_i}{\mu} = T.$$

(We use L Hospital s rule in evaluating both the limits) Thus, I_G contains the Theil index as a particular case and obviously the maximum value of the inequality measure I_G tends to the maximum of the Theil Measure as $\alpha \to 1$.

4 Decomposability Properties

Theil noted that his index satisfies a useful decomposability property. To explain this, let there be m disjoint groups of incomes where the number of members in the ith group is n_i. So, if Y stands for the vector of incomes of all the m groups taken together, we have

$$Y = (Y^1, Y^2, \ldots, Y^m) = [(y_{11}, y_{12}, \ldots, y_{1n_1}), (y_{21}, y_{22}, \ldots, y_{2n_2}), \ldots, (y_{m1}, y_{m2}, \ldots, y_{mn_m})].$$

The Theil Measure of Inequality for Y is given by

$$T = \sum_{i=1}^{m} \sum_{j=1}^{n_i} s_{ij}\, log\, (ns_{ij}), \tag{12}$$

where $n = \sum_{i=1}^{m} n_i$. Let $y_{i0} = \sum_{j=1}^{n_i} y_{ij}$ and $y_0 = \sum_{k=1}^{m} y_{i0} = \sum_{i=1}^{m} \sum_{j=1}^{n_i} y_{ij}$. Then, $s_{ij} = \frac{y_{ij}}{y_0} = \frac{y_{ij}}{y_{i0}}.\frac{y_{i0}}{y_0} = s^*_{ij}.s_{i0}$ (say). So, from (12), we have

$$\begin{aligned} T &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(s^*_{ij}s_{i0})\, log\,(ns_{ij}) = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(s^*_{ij}s_{i0}) log\,(n_i\frac{m}{n_i}s_{i0}.s^*_{ij}) \\ &= \sum_{i=1}^{m}\frac{y_{i0}}{y_0}\,T(Y^i) + \sum_{i=1}^{m}\frac{n_i\mu_i}{n\mu}\,log\,\frac{\mu_i}{\mu} \\ &= \sum_{i=1}^{m}\frac{y_{i0}}{y_0}\,T(Y^i) + T(\mu_1 1^{n_1}, \mu_2 1^{n_2}, \ldots, \mu_m 1^{n_m}). \end{aligned} \tag{13}$$

This additive decomposability property of the Theil index (with same number of members in each group) corresponds to the strong Additivity property of the Shannon Entropy Index of Information [Eq. (2)].

Turning now to the decomposability of the Generalized-Entropy-based index of inequality I_G, we have, from (10)

$$\begin{aligned} I_G(Y) &= I_G[(y_{11}, y_{12}, \ldots, y_{1n_1}), (y_{21}, y_{22}, \ldots, y_{2n_2}), \ldots, (y_{m1}, y_{m2}, \ldots, y_{mn_m})] \\ &= \frac{1}{1-2^{1-\alpha}}[\sum_{i=1}^{m}\sum_{j=1}^{n_i} s^{\alpha}_{ij} - n^{1-\alpha}], \text{ where } n = \sum_{i=1}^{m} n_i \\ &= \frac{1}{1-2^{1-\alpha}}[\sum_{i=1}^{m} s^{\alpha}_{i0}\sum_{j=1}^{n_i} s^{*\alpha}_{ij} - n^{1-\alpha}] \text{ as } \; s_{ij} = s^*_{ij}s_{i0} \\ &= \frac{1}{1-2^{1-\alpha}}[\sum_{i=1}^{m} s^{\alpha}_{i0}(\sum_{j=1}^{n_i} s^{*\alpha}_{ij} - n_i^{1-\alpha}) - n^{1-\alpha} + \sum_{i=1}^{m} s^{\alpha}_{i0} n_i^{1-\alpha}]. \end{aligned} \tag{14}$$

On simplification, this gives

$$I_G(Y) = \sum_{i=1}^{m}(\frac{y_{i0}}{y_0})^{\alpha} I_G(Y^i) \,+ I_G(\mu_1 1^{n_1}, \mu_2 1^{n_2}, \ldots, \mu_m 1^{n_m}). \tag{15}$$

The weights to $I_G(Y^i)$, in general, do not add up to 1 (except in the limiting case). Just as the Strong Additivity property of the Shannon Entropy Index generated the additive decomposability property of the Theil index as given in (14), the property of Strong Additivity of Degree α of the Entropy Index class G_{α} generates the additive decomposability property (15) of I_G with same number of members in each group.

4.1 Social Welfare Interpretation of I_G

We now define the social welfare function associated with I_G as

$$W_G = \mu(max\ I_G - I_G) = \mu G_\alpha, \tag{16}$$

where G_α is the Generalized Entropy of Degree α. Clearly, W_G corresponds to I_G in a negative monotonic way. That is, for a given population size n and given mean income μ, W_G will rank income distributions in exactly the same way as the negative of I_G. Since G_α is homogeneous of degree $0, W_G$ is linearly homogeneous. W_G also satisfies continuity and strict S-concavity (hence symmetry). When efficiency considerations are absent (that is, μ is fixed), an increase in I_G is equivalent to a reduction in social welfare. On the other hand, an equiproportionate increase in all incomes does not alter inequality but increases the mean income μ leading to an increase in W_G. Thus, the society now moves to a Pareto superior state. The parameter α reflects different perceptions of inequality (social welfare). A transfer of income from a poor person i to a richer person j increases (decreases) I_G (W_G) by a larger amount the higher is α.

5 Empirical Illustration

We use National Sample Survey data of Consumer Expenditure to analyze changes in inequality in Monthly Per Capita Expenditure (MPCE) across states of India for the time period 1993–94 to 2014–15. The specific rounds of survey considered are 50, 55, 61, 66, 68, and 72 with respective years as 1993–94, 1999–2000, 2004–2005, 2009–10, 2011–12, and 2014–15. The computations are carried for the sectors rural, urban, and rural–urban combined. It may be noted that MPCE is taken as an approximation of Monthly Per Capita Income (MPCI) in absence of reliable income data. Obviously, the inequality in the savings pattern and its correlations with consumption expenditure have definitive influence on inequality of income. Tables 1, 2, and 3 present the inequality of MPCE for the three sectors, rural–urban combined, rural, and urban in terms of the Theil index ($\alpha = 1$). The values are sorted by the level of inequality of the most current round (72) in the year 2014–15. Note that for $\alpha = 1$, the index is population replication invariant, and hence is comparable across state having different population sizes. An useful comparison of inequality can also be based on the ranks of the states across rounds which are presented in Tables 4, 5, and 6. We have also observed that the inequalities as well as ranks generally differ for the choice of α. Further, for $\alpha \neq 1$, the indexes are not comparable across populations differing in sizes being not satisfying population principle. Hence, they are not reported here. However, our computation shows that the functional form of the index and the corresponding perspective on inequality are crucial in judging inequality across groups within a population.

Table 1 All India and state inequality (Theil index) in MPCE (rural–urban combined across NSS rounds)

NSS round	50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
All India	0.247	0.239	0.278	0.329	0.296	0.202
Lakshadweep	0.149	0.150	0.120	0.268	0.181	0.067
Nagaland	0.066	0.099	0.118	0.081	0.085	0.077
Goa	0.167	0.167	0.197	0.115	0.179	0.094
Meghalaya	0.244	0.102	0.132	0.101	0.109	0.096
Bihar	0.145	0.139	0.128	0.134	0.105	0.102
Assam	0.110	0.136	0.137	0.169	0.173	0.107
Uttaranchal			0.190	0.552	0.213	0.109
Manipur	0.047	0.103	0.051	0.068	0.080	0.111
Sikkim	0.137	0.139	0.186	0.196	0.124	0.113
Pondicheri	0.180	0.164	0.260	0.801	0.142	0.114
Himachal Pradesh	0.401	0.169	0.201	0.233	0.223	0.117
Tripura	0.126	0.113	0.172	0.134	0.138	0.119
Kerala	0.259	0.195	0.287	0.340	0.353	0.134
Jammu and Kashmir	0.179	0.098	0.132	0.157	0.165	0.148
Andaman and Nicober	0.274	0.140	0.304	0.342	0.272	0.151
Chandigarh	0.404	0.199	0.226	0.265	0.266	0.154
Gujrat	0.165	0.173	0.241	0.239	0.189	0.157
Dadra and Nagar Haveli	0.184	0.232	0.321	0.159	0.283	0.157
Daman and Diu	0.121	0.098	0.244	0.135	0.054	0.158
Jharkhand			0.207	0.194	0.235	0.167
Delhi	0.319	0.187	0.194	0.228	0.248	0.167
Punjab	0.171	0.143	0.227	0.242	0.203	0.168
Tamil Nadu	0.283	0.412	0.296	0.244	0.226	0.170
Rajasthan	0.187	0.149	0.187	0.186	0.196	0.171
Andhra Pradesh	0.207	0.192	0.272	0.266	0.198	0.172
Uttar Pradesh	0.196	0.184	0.204	0.410	0.281	0.173
Orissa	0.185	0.166	0.220	0.266	0.203	0.175
Chhattisgarh			0.273	0.210	0.277	0.186
Madhya Pradesh	0.272	0.204	0.267	0.315	0.293	0.187
West Bengal	0.237	0.241	0.261	0.259	0.291	0.188
Mizoram	0.080	0.117	0.108	0.118	0.151	0.189
Haryana	0.209	0.137	0.278	0.204	0.238	0.194
Karnataka	0.206	0.220	0.314	0.372	0.368	0.197
Telangana						0.213
Arunachal Pradesh	0.209	0.238	0.128	0.182	0.226	0.224
Maharastra	0.293	0.271	0.299	0.332	0.317	0.246

Table 2 All India and state inequality (Theil index) in MPCE (rural) across NSS rounds)

NSS round	50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
All India	0.184	0.147	0.182	0.211	0.186	0.149
Daman and Diu	0.134	0.094	0.247	0.138	0.04	0.067
Nagaland	0.047	0.072	0.09	0.067	0.066	0.07
Meghalaya	0.256	0.053	0.054	0.07	0.063	0.075
Lakshadweep	0.124	0.115	0.12	0.351	0.161	0.085
Bihar	0.103	0.089	0.071	0.088	0.08	0.085
Assam	0.058	0.076	0.066	0.088	0.097	0.09
Goa	0.157	0.152	0.16	0.106	0.151	0.092
Tripura	0.106	0.065	0.091	0.078	0.091	0.096
Sikkim	0.118	0.122	0.169	0.205	0.086	0.097
Uttaranchal			0.127	0.586	0.138	0.1
Delhi	0.135	0.061	0.151	0.073	0.087	0.106
Himachal Pradesh	0.193	0.134	0.192	0.205	0.172	0.108
Gujrat	0.119	0.111	0.164	0.191	0.138	0.11
Kerala	0.199	0.186	0.259	0.3	0.355	0.111
Jharkhand			0.087	0.079	0.109	0.112
Telangana						0.112
Dadra and Nagar Haveli	0.155	0.214	0.291	0.094	0.23	0.113
Manipur	0.045	0.067	0.047	0.057	0.081	0.114
Pondicheri	0.168	0.121	0.258	0.122	0.136	0.116
Jammu and Kashmir	0.16	0.081	0.118	0.103	0.14	0.117
Karnataka	0.146	0.123	0.192	0.113	0.153	0.12
Andaman and Nicober	0.138	0.126	0.219	0.439	0.191	0.122
Andhra Pradesh	0.189	0.127	0.153	0.207	0.125	0.126
Chandigarh	0.105	0.117	0.082	0.167	0.121	0.128
Punjab	0.177	0.125	0.16	0.23	0.155	0.13
West Bengal	0.195	0.126	0.154	0.104	0.111	0.132
Uttar Pradesh	0.164	0.133	0.132	0.12	0.146	0.133
Tamil Nadu	0.202	0.175	0.184	0.157	0.163	0.138
Chhattisgarh			0.15	0.115	0.118	0.14
Haryana	0.226	0.112	0.291	0.138	0.122	0.144
Orissa	0.131	0.118	0.144	0.136	0.115	0.145
Madhya Pradesh	0.196	0.122	0.139	0.151	0.169	0.146
Rajasthan	0.175	0.098	0.097	0.093	0.111	0.147
Mizoram	0.066	0.098	0.076	0.074	0.105	0.153
Maharastra	0.206	0.15	0.168	0.117	0.184	0.174
Arunachal Pradesh	0.204	0.248	0.129	0.179	0.221	0.242

Table 3 All India and state inequality (Theil index) in MPCE (urban) across NSS rounds

NSS round	50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
All India	0.257	0.262	0.265	0.320	0.282	0.196
Lakshadweep	0.175	0.168	0.111	0.181	0.201	0.061
Nagaland	0.089	0.099	0.110	0.090	0.087	0.086
Goa	0.168	0.162	0.236	0.108	0.193	0.094
Manipur	0.048	0.161	0.053	0.087	0.077	0.101
Uttaranchal			0.206	0.186	0.255	0.102
Pondicheri	0.177	0.165	0.237	0.871	0.136	0.109
Dadra and Nagar Haveli	0.180	0.112	0.162	0.114	0.193	0.113
Sikkim	0.121	0.115	0.160	0.090	0.085	0.120
Meghalaya	0.115	0.094	0.173	0.104	0.092	0.128
Gujrat	0.170	0.173	0.217	0.193	0.153	0.143
Bihar	0.188	0.211	0.221	0.219	0.166	0.146
Tripura	0.156	0.153	0.207	0.154	0.148	0.149
Tamil Nadu	0.323	0.530	0.269	0.231	0.215	0.150
Kerala	0.360	0.192	0.315	0.392	0.323	0.154
Chandigarh	0.402	0.201	0.216	0.266	0.273	0.154
Himachal Pradesh	0.772	0.176	0.140	0.258	0.225	0.156
Mizoram	0.067	0.106	0.097	0.098	0.116	0.157
Arunachal Pradesh	0.151	0.148	0.113	0.172	0.189	0.157
Assam	0.153	0.156	0.180	0.249	0.230	0.158
Orissa	0.191	0.183	0.256	0.337	0.219	0.159
Daman and Diu	0.102	0.102	0.236	0.131	0.093	0.160
Andaman and Nicober	0.346	0.132	0.307	0.206	0.272	0.161
Delhi	0.339	0.196	0.192	0.231	0.257	0.167
Rajasthan	0.174	0.192	0.247	0.249	0.239	0.169
Jammu and Kashmir	0.154	0.104	0.120	0.238	0.170	0.171
Chhattisgarh			0.304	0.235	0.348	0.178
Andhra Pradesh	0.197	0.196	0.317	0.234	0.205	0.179
Karnataka	0.206	0.214	0.275	0.348	0.366	0.183
Telangana						0.184
Uttar Pradesh	0.218	0.226	0.249	0.582	0.358	0.195
West Bengal	0.194	0.278	0.249	0.278	0.302	0.198
Madhya Pradesh	0.307	0.240	0.305	0.345	0.335	0.202
Maharastra	0.243	0.242	0.256	0.307	0.278	0.210
Punjab	0.145	0.151	0.249	0.233	0.233	0.212
Jharkhand			0.194	0.229	0.230	0.218
Haryana	0.156	0.147	0.230	0.246	0.275	0.237

Table 4 Rural–urban combined Theil inequality for states by ranks

NSS round	50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
Lakshadweep	9	15	4	27	13	1
Nagaland	2	3	3	2	3	2
Goa	11	18	15	4	12	3
Meghalaya	24	4	8	3	5	4
Bihar	8	10	6	7	4	5
Assam	4	8	9	11	11	6
Uttaranchal			13	34	19	7
Manipur	1	5	1	1	2	8
Sikkim	7	11	11	15	6	9
Pondicheri	14	16	24	35	8	10
Himachal Pradesh	31	19	16	19	20	11
Tripura	6	6	10	6	7	12
Kerala	25	24	30	30	34	13
Jammu and Kashmir	13	1	7	9	10	14
Andaman and Nicober	27	12	33	31	27	15
Chandigarh	32	25	20	24	26	16
Gujrat	10	20	22	20	14	17
Dadra and Nagar Haveli	15	28	35	10	30	18
Daman and Diu	5	2	23	8	1	19
Jharkhand			18	14	23	20
Delhi	30	22	14	18	25	21
Punjab	12	13	21	21	18	22
Tamil Nadu	28	32	31	22	22	23
Rajasthan	17	14	12	13	15	24
Andhra Pradesh	20	23	27	26	16	25
Uttar Pradesh	18	21	17	33	29	26
Orissa	16	17	19	25	17	27
Chhattisgarh			28	17	28	28
Madhya Pradesh	26	26	26	28	32	29
West Bengal	23	30	25	23	31	30
Mizoram	3	7	2	5	9	31
Haryana	21	9	29	16	24	32
Karnataka	19	27	34	32	35	33
Telangana						34
Arunachal Pradesh	22	29	5	12	21	35
Maharastra	29	31	32	29	33	36

Table 5 Rural Theil inequality for states by ranks

NSS round	50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
Daman and Diu	12	9	31	21	1	1
Nagaland	2	5	8	2	3	2
Meghalaya	32	1	2	3	2	3
Lakshadweep	10	14	12	33	27	4
Bihar	5	8	4	8	4	5
Assam	3	6	3	9	9	6
Goa	17	28	23	14	24	7
Tripura	7	3	9	6	8	8
Sikkim	8	18	26	29	6	9
Uttaranchal			13	35	20	10
Delhi	13	2	19	4	7	11
Himachal Pradesh	24	26	28	28	30	12
Gujrat	9	12	24	27	21	13
Kerala	27	30	33	32	35	14
Jharkhand			7	7	11	15
Telangana						16
Dadra and Nagar Haveli	16	31	35	11	34	17
Manipur	1	4	1	1	5	18
Pondicheri	20	17	32	19	19	19
Jammu and Kashmir	18	7	11	12	22	20
Karnataka	15	20	29	15	25	21
Andaman and Nicober	14	22	30	34	32	22
Andhra Pradesh	23	24	20	30	18	23
Chandigarh	6	15	6	25	16	24
Punjab	22	21	22	31	26	25
West Bengal	25	23	21	13	12	26
Uttar Pradesh	19	25	15	18	23	27
Tamil Nadu	28	29	27	24	28	28
Chhattisgarh			18	16	15	29
Haryana	31	13	34	22	17	30
Orissa	11	16	17	20	14	31
Madhya Pradesh	26	19	16	23	29	32
Rajasthan	21	10	10	10	13	33
Mizoram	4	11	5	5	10	34
Maharastra	30	27	25	17	31	35
Arunachal Pradesh	29	32	14	26	33	36

Table 6 Urban Theil inequality for states by ranks

NSS round	50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
Lakshadweep	16	17	4	11	15	1
Nagaland	3	2	3	3	3	2
Goa	13	15	20	6	14	3
Manipur	1	14	1	1	1	4
Uttaranchal			14	12	24	5
Pondicheri	17	16	22	35	7	6
Dadra and Nagar Haveli	18	6	9	7	13	7
Sikkim	6	7	8	2	2	8
Meghalaya	5	1	10	5	4	9
Gujrat	14	18	17	13	9	10
Bihar	19	26	18	15	10	11
Tripura	12	12	15	9	8	12
Tamil Nadu	27	32	29	17	17	13
Kerala	30	22	34	33	31	14
Chandigarh	31	25	16	27	27	15
Himachal Pradesh	32	19	7	26	19	16
Mizoram	2	5	2	4	6	17
Arunachal Pradesh	8	10	5	10	12	18
Assam	9	13	11	24	20	19
Orissa	20	20	28	30	18	20
Daman and Diu	4	3	21	8	5	21
Andaman and Nicober	29	8	33	14	26	22
Delhi	28	23	12	18	25	23
Rajasthan	15	21	23	25	23	24
Jammu and Kashmir	10	4	6	22	11	25
Chhattisgarh			31	21	33	26
Andhra Pradesh	22	24	35	20	16	27
Karnataka	23	27	30	32	35	28
Telangana						29
Uttar Pradesh	24	28	25	34	34	30
West Bengal	21	31	26	28	30	31
Madhya Pradesh	26	29	32	31	32	32
Maharastra	25	30	27	29	29	33
Punjab	7	11	24	19	22	34
Jharkhand			13	16	21	35
Haryana	11	9	19	23	28	36

6 Conclusion

We have introduced a generalization of the Theil index of inequality, the index I_G. No characterization, however, has been attempted. We have simply explored the relationships between the inequality and information measures and studied some of their properties. Information-theory-based characterization of inequality measures are rare, Fosters (1983) article being one exception. Lasso de la Vega et. al. (2013) characterized the Theil inequality ordering, which in our case falls under $\alpha = 1$. We have already noted that for $\alpha \neq 1$, I_G does not satisfy the principle of population (PP), though in the particular case of $\alpha = 1$ it does meet this property. Consequently, $\alpha \neq 1$, I_G is not a member of the (Shorrocks 1980) class of indices since the latter fulfills the Principle of Population (PP). This may explain the absence of characterization of the general class of indexes introduced in this article.

References

Aczel J, Daroczy Z (1975) On measures of information and their characterizations. Academic Press, New York

Foster JE (1983) An axiomatic characterization of the Theil measure of income inequality. J Econ Theory 31:105–121

Lasso de la Vega C, Urrutia A, Volij O (2013) An axiomatic characterization of the Theil inequality ordering. Econ Theory. https://doi.org/10.1007/s00199-012-0739-6

Sen AK (1973) On economic inequality. Clarendon, Oxford

Shorrocks AF (1980) The class of additively decomposable inequality measures. Econometrica 48:613–625

Theil H (1967) Economics and information theory, North Holland, Amsterdam

Correlation and Inequality in Weighted Majority Voting Games

Sanjay Bhattacherjee and Palash Sarkar

Abstract In a weighted majority voting game, the weights of the players are determined based on some socioeconomic parameter. A number of measures have been proposed to measure the voting powers of the different players. A basic question in this area is to what extent does the variation in the voting powers reflect the variation in the weights? The voting powers depend on the winning threshold. So, a second question is what is the appropriate value of the winning threshold? In this work, we propose two simple ideas to address these and related questions in a quantifiable manner. The first idea is to use Pearson's Correlation Coefficient between the weight vector and the power profile to measure the similarity between weight and power. The second idea is to use standard inequality measures to quantify the inequality in the weight vector as well as in the power profile. These two ideas answer the first question. Both the weight–power similarity and inequality scores of voting power profiles depend on the value of the winning threshold. For situations of practical interest, it turns out that it is possible to choose a value of the winning threshold which maximizes the similarity score and also minimizes the difference in the inequality scores of the weight vector and the power profile. This provides an answer to the second question. Using the above formalization, we are able to quantitatively argue that it is sufficient to consider only the vector of swings for the players as the power measure. We apply our methodology to the voting games arising in the decision-making processes of the International Monetary Fund (IMF) and the European Union (EU). In the case of IMF, we provide quantitative evidence that the actual winning threshold that is currently used is suboptimal and instead proposes a winning threshold which has a firm analytical backing. On the other hand, in the case of EU, we provide quantitative evidence that the presently used threshold is very close to the optimal.

An earlier version of this paper was posted as a technical report at https://mpra.ub.uni-muenchen.de/83168/1/MPRA_paper_83168.pdf.

S. Bhattacherjee (✉) · P. Sarkar
Applied Statistics Unit, Indian Statistical Institute, 203, B.T. Road, Kolkata 700108, India
e-mail: sanjay.bhattacherjee@gmail.com

P. Sarkar
e-mail: palash@isical.ac.in

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*,
Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_9

Keywords Voting games · Weighted majority · Power measure · Correlation · Inequality · Gini index · Coefficient of variation

1 Introduction

Voting is arguably the most important aspect of decision-making in a democratic setup. A committee settles an issue by accepting or rejecting some resolution related to the issue. While unanimity or consensus is desirable, this may not always be possible due to the conflicting interests of the different committee members. In such a situation, a voting procedure among the members is used to either accept or reject a resolution. A resolution is accepted or passed, if a certain number of persons vote in its favor, else it fails and is rejected.

In its basic form, each committee member has a single vote. Many scenarios of practical interest, on the other hand, assign weights to the committee members. These weights need not be the same for all the members. In the context of weighted voting, a resolution is accepted, if the sum total of the weights of the members who vote in its favor cross a previously decided upon threshold. A common example of weighted voting is a company boardroom, where the members have weights in proportion to the shares that they hold in the company. Important examples of weighted voting in the context of public policy are the European Union (EU) and the International Monetary Fund (IMF).

Voting procedures have been formally studied in the game theory literature under the name of voting games. Due to its real-life importance, weighted majority voting games have received a lot of attention. In the literature on voting games, the members are called players. One of the basic questions is how much influence does a player have in determining the outcome of a voting procedure? In other words, what is the power of a player in a voting game? In quantitative terms, it is desirable to measure the power of a player in a voting game by assigning a nonnegative real number to the player. A power measure assigns such a number to each player in the game. This leads to the basic question of what constitutes a good measure of power of a player in a voting game. The literature contains a number of power measures. Each one of these measures aims to capture certain aspects of the informal notion of power in a voting game. We refer to Felsenthal and Machover (1998) for a comprehensive discussion to voting games and the various power measures. An introduction to the area can be found in Chakravarty et al. (2015).

Consider the setting of weighted majority voting games. For any such game, the players are assigned weights based on socioeconomic parameters. As a result, there is a variation in the weights of the players. Further, given any power measure, we obtain a variation in the powers of the different players. It is well known that the variation in the voting powers does not necessarily reflect the variation in the weights. In this context, the following three questions can be formulated:

1. To what extent does the variation in the voting powers reflect the variation in the weights?
2. Is the inequality present in the weights preserved in the voting powers?
3. How does the value of the winning threshold (i.e., the threshold which is required to be crossed for a motion to be passed) affect the above two questions?

This work addresses the above questions. The questions posed are not merely theoretical. Similar questions have been posed in Leech (2002b) in the context of measurement of voting power in IMF. For example, the following text fragments appear in Leech (2002b):

> ... weighted voting raises the important question of whether the resulting inequality of power over actual decisions is precisely what was intended for the relationship between power and contribution.
>
> How does the voting power of individual countries compare with their nominal votes? To what extent is the degree of inequality in the distribution of votes reflected in the distribution of voting power?
>
> Different types of decisions use different decision rules, some requiring a special supermajority. What effect do different decision rules have on the distribution of power and also on the power of the voting body itself to act?

The work Leech (2002b) makes a qualitative analysis of the above issues. Our work allows a quantitative analysis of these issues. In more detail, our work makes the following contributions.

Measurement of Similarity Between Weights and Voting Powers. We propose the use of Pearson's correlation coefficient as a measure of similarity between the weight vector of the players and the vector of voting powers of the players.

Measurement of Inequality in Weights and Voting Powers. There is a large literature on the measurement of inequality in a vector of values obtained from measurement of various social parameters. A survey on measurement of inequality appears in Cowell (2016). The variation of the values in such a vector is captured by an inequality index. A number of inequality indices have been proposed in the literature. We propose the use of such inequality indices to measure the inequality in the weights and also in the voting powers. This allows the comparison of the inequality present in the weights to that present in the voting powers.

Winning Threshold as a Controllable Parameter. Our formalizations of both the similarity between the weights and the voting powers as well as the measurement of inequality in the voting powers have the winning threshold as a parameter. By varying this parameter, both the weight–power similarity and the voting power inequality can be controlled. So, given a vector of weights, the winning threshold can be set to a certain value to maximize the weight–power similarity or to minimize the difference between the inequality in the weights and the inequality in the voting powers.

In this context, we would like to discuss the broader issue of designing games to achieve certain desirable power profiles. This is often called the inverse problem for voting games. Usually, the goal is to determine a set of weights which result in the target powers. For example, in the context of the IMF voting game, an iterative algorithm to determine weights has been proposed in Leech (2002b). There is one major drawback of this approach. As mentioned earlier, in a weighted majority voting game, the weights often represent a socioeconomic parameter. When the weights are artificially obtained (say, using an iterative algorithm), their interpretation in the socioeconomic context is lost. It then becomes hard to provide a natural justification of the weights.

Our approach of having the winning threshold as a controllable parameter provides an alternative method of designing games. For the complete specification of a game, both the weights and the winning threshold need to be specified. In our approach, the weights do not change, and hence they retain their original interpretation arising from the background socioeconomic application. We only suggest tuning the winning threshold so that the resulting power profile is "imbued" with the intuitive natural justification of the weights. Games designed using such an approach can be much better explained to the general public than games where the weights are artificially obtained.

Detailed Study. We consider seven different voting power measures and two different inequality indices. We show that the scaling invariance property of an inequality measure as well as that of the Pearson's correlation coefficient divides the voting power measures into three groups. The non-normalized Banzhaf measure, the normalized Banzhaf index, and the two Coleman measures fall into one group; the public good measure and public good index defined by Holler fall into a second group, and the Deegan–Packel measure is in the third group. We show that any two power measures in the same group have the same behavior with respect to both the similarity index and the inequality index. This brings down the complexity of the analysis.

There has been a lot of discussion in the literature on the comparative suitabilities of the Banzhaf and the Coleman indices Banzhaf (1965), Brink and Laan (1998), Coleman (1971), Dubey and Shapley (1979), Laruelle and Valenciano (2001), Lehrer (1998), Laruelle and Valenciano (2011), Barua et al. (2009). This discussion has both been qualitative and also formal in the sense of axiomatically deriving the indices Brink and Laan (1998), Lehrer (1998). Our work provides a new perspective to this discussion. The stand-alone values of the powers of the players as measured by any power measure are perhaps not of much interest by themselves. It is only in a relative sense that they acquire relevance. There are two ways to consider this relative sense, in comparison to the weights and in comparison among themselves. We propose to quantify the relative notion in comparison to the weights by the correlation between the weight vector and the power profile and to quantify the relative values of the powers among themselves by an appropriate inequality score. Under both of these quantifications, we prove that the Banzhaf and the Coleman power measures turn out to be the same. Based on this result, we put forward the suggestion that there is perhaps no *essential* difference between these power measures. It is sufficient to consider only the swings for the different players as was originally proposed by Banzhaf (and sometimes called the raw Banzhaf measure). While this may sound a bit radical, our analysis based on correlation and inequality does not leave scope for any other considerations. It is of course possible that there is some other quantifiable ways of distinguishing between the relative spreads of the Banzhaf and the Coleman measures. This can be a possible future research question.

The literature contains a number of voting power measures which have been proposed as fundamentally different from the swing-based Banzhaf measure. Intuitive arguments have been forwarded as to why these measures are appropriate for certain applications. In our opinion, a basic requirement for any power measure is to reflect the "content" of weights. In addition to the Banzhaf measure, we have also considered the Holler measures and the Deegan–Packel measure. Our simulation experiments as well as computations with real-life data show that the "content" of the weights is best captured by the Banzhaf measure and neither the Holler measure nor the Deegan–Packel measure is good indicators of this "content". Based on this evidence, we put forward the suggestion that it is sufficient to consider the swings as the only measure of power in voting games.

Applications. IMF decision-making procedures have been modeled as voting games Leech (2002b). Decision-making in the EU has also been discussed in the context of voting games Leech (2002a).

The notions of similarity between the variations in the weights and the voting powers as well as the relation between the inequality in the weights and that in the voting powers have been informally discussed. Our proposals for measuring weight–power similarity and the voting power inequality formalizes this intuition. We compute the various measures for the IMF game and (a simplified version of) the EU voting game and suggest that the winning threshold can be used as a parameter in achieving target values of similarity or inequality.

In both the IMF and the EU voting games, there is a "natural" justification for assigning weights to the different players. In the context of IMF, the weights reflect the proportion of financial contribution made by the different countries, while in the case of EU, the weights reflect the population of the different countries. This is reasonable, since the IMF is a financial organization, while the EU is essentially a political organization. In both cases, however, the choice of the winning threshold is not backed by any quantifiable parameter.

Our work provides methods for choosing a winning threshold which has a *quantifiable* justification. There are two options. In the first option, one should choose a value of the winning threshold which maximizes the correlation between the weight vector and an appropriate power profile. In the second option, one should choose a value of the winning threshold which yields an inequality score for an appropriate power profile which is closest to the inequality score of the weight vector. In both the cases of IMF and EU, both the options lead to similar values of the winning threshold. Based on this analysis, we put forward the suggestion that the voting rule for IMF should be modified to reflect the optimal value of the winning threshold. In the case of EU, our results provide evidence that the presently used winning threshold is close to the optimal value.

Previous and Related Works

The Shapley–Shubik power index was introduced in Shapley (1953), Shapley and Shubik (1954), Banzhaf index was introduced in Banzhaf (1965) while Coleman indices were introduced in Coleman (1971). Later work by Holler (1982) and Holler and Packel (1983) introduced the public good measure/index. Deegan and Packel introduced another power measure in Deegan and Packel (1978). There are other known measures/indices and we refer to Felsenthal and Machover (1998), Chakravarty et al. (2015) for further details.

The first work to address the problem of inequality in voting games is Einy and Peleg (1991). They provided an axiomatic deduction of an inequality index for the Shapley–Shubik power measure. A more general axiomatic treatment of inequality for power measures appears in a paper by Laruelle and Valenciano (2004). This work postulates axioms and deduces an inequality measure for a class of power indices which includes the Banzhaf index. A more recent work by Weber (2016) suggests the use of the Coefficient of Variation as an inequality index for measuring inequality arising from the Banzhaf index. Later, we provide a more detailed discussion of the relationship of these prior works to our contribution.

2 Preliminaries

2.1 Voting Games

We provide some standard definitions arising in the context of voting games. For details, the reader may consult Felsenthal and Machover (1998), Chakravarty et al. (2015). In the following, the cardinality of a finite set S will be denoted by $\#S$ and the absolute value of a real number x will be denoted by $|x|$.

Let $N = \{A_1, A_2, \ldots, A_n\}$ be a set of n players. A subset of N is called a *voting coalition*. The set of all voting coalitions is denoted by 2^N. A voting game G is given by its characteristic function $\widehat{G} : 2^N \to \{0, 1\}$, where a winning coalition is assigned the value 1 and a losing coalition is assigned the value 0. Below, we recall some basic notions about voting games:

1. For any $S \subseteq N$ and player $A_i \in N$, A_i is said to be a *swing* in S if $A_i \in S$, $\widehat{G}(S) = 1$ but $\widehat{G}(S \setminus \{A_i\}) = 0$.
2. For a voting game G, the number of swings for A_i will be denoted by $m_G(A_i)$.
3. A player $A_i \in N$ is called a *dummy player* if A_i is not a swing in any coalition, i.e., if $m_G(A_i) = 0$.
4. For a voting game G, the set of all winning coalitions will be denoted by $W(G)$ and the set of all losing coalitions will be denoted by $L(G)$.
5. A coalition $S \subseteq N$ is called a *minimal winning coalition* if $\widehat{G}(S) = 1$ and there is no $T \subset S$ for which $\widehat{G}(T) = 1$.
6. The set of all minimal winning coalitions in G will be denoted by $\mathsf{MW}(G)$ and the set of minimal winning coalitions containing the player A_i will be denoted as $\mathsf{MW}_G(A_i)$.
7. A voting game G is said to be *proper* if for any coalition $S \subseteq N$, $\widehat{G}(S) = 1$ implies that $\widehat{G}(N \setminus S) = 0$. In other words, in a proper game, it is not allowed for both S and its complement to be winning.

Definition 1 Consider a triplet $(N, \mathbf{w}, q)$, where $N = \{A_1, \ldots, A_n\}$ is a set of players; $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ is a vector of nonnegative weights with w_i being the weight of A_i; and q is a real number in $(0, 1)$. Let $\omega = \sum_{i=1}^{n} w_i$. The triplet $(N, \mathbf{w}, q)$ defines a *weighted majority voting game* G given by its characteristic function $\widehat{G} : 2^N \rightarrow \{0, 1\}$ in the following manner. Let $w_S = \sum_{A_i \in S} w_i$ denote the sum of the weights of all the players in the coalition $S \subseteq N$. Then

$$\widehat{G}(S) = \begin{cases} 1 & \text{if } w_S/\omega \geq q, \\ 0 & \text{otherwise.} \end{cases}$$

We will write $G = (N, \mathbf{w}, q)$ to denote the weighted majority voting game arising from the triplet $(N, \mathbf{w}, q)$.

For a weighted majority voting game $G = (N, \mathbf{w}, q)$ to be proper, it is necessary that $q > 0.5$. For the technical analysis of weighted majority voting games, we do not restrict to proper games. When considering applications, as is conventional, one should consider only proper games.

2.2 *Voting Power*

The notion of power is an important concept in a voting system. A *power measure* captures the capability of a player to influence the outcome of a vote.

Given a game G on a set of players N and a player A_i in N, a power measure $\mathcal{P}$ associates a nonnegative real number $v_i = \mathcal{P}_G(A_i)$ to the player A_i. The number v_i captures the power that A_i has in the game G. If $\sum_{A_i \in G} \mathcal{P}_G(A_i) = 1$ for all games G, then $\mathcal{P}$ is called a *power index*. In other words, for a power index, the powers of the individual players sum to 1.

A widely studied index of voting power is the Shapley–Shubik index. This index, however, is defined for a voting game where the order in which the players cast their votes is important. In our application of voting power to the voting games arising in the IMF and EU decision-making processes, the order of casting votes is not important. So, we do not consider the Shapley–Shubik index in this work. Below, we provide the definitions of some of the previously proposed power measures. See Felsenthal and Machover (1998), Chakravarty et al. (2015) for further details.

Banzhaf Power Measures. The *raw Banzhaf power measure* $\mathsf{BR}_G(A_i)$ for a player A_i in the game G is defined as the number of distinct coalitions in which A_i is a swing. Hence,

$$\mathsf{BR}_G(A_i) = m_G(A_i).$$

The *non-normalized Banzhaf power measure* $\mathsf{BZN}_G(A_i)$ is defined as follows:

$$\mathsf{BZN}_G(A_i) = \frac{\mathsf{BR}_G(A_i)}{2^{n-1}} = \frac{m_G(A_i)}{2^{n-1}}.$$

The *Banzhaf normalized power index* $\mathsf{BZ}_G(A_i)$ is defined as follows:

$$\mathsf{BZ}_G(A_i) = \frac{\mathsf{BR}_G(A_i)}{\sum_{j=1}^{n} \mathsf{BR}_G(A_j)} = \frac{m_G(A_i)}{\sum_{j=1}^{n} m_G(A_j)}.$$

Coleman Power Measures. The *Coleman preventive power measure* $\mathsf{CP}_G(A_i)$ for a player A_i in the game G is a measure of its ability to stop a coalition S from achieving $w_S \geq q$. It is defined as follows:

$$\mathsf{CP}_G(A_i) = \frac{m_G(A_i)}{\#W(G)}.$$

The *Coleman initiative power measure* $\mathsf{CI}_G(A_i)$ for a player A_i in the game G is a measure of its ability to turn an otherwise losing coalition S with $w_S < q$ into a winning coalition with $w_{S\cup\{A_i\}} \geq q$. It is defined as follows:

$$\mathsf{CI}_G(A_i) = \frac{m_G(A_i)}{\#L(G)} = \frac{m_G(A_i)}{2^n - \#W(G)} = \frac{\mathsf{CP}_G(A_i)}{\frac{2^n}{\#W(G)} - 1}.$$

Holler Public Good Index. Holler proposed the *public good index* $\mathsf{PGI}_{A_i}(G)$ as follows:

$$\mathsf{PGI}_{A_i}(G) = \frac{\#\mathsf{MW}_G(A_i)}{\sum_{A_j \in N} \#\mathsf{MW}_G(A_j)}.$$

The non-normalized version of $\mathsf{PGI}_{A_i}(G)$ is called the *absolute public good measure*. It is defined as

$$\mathsf{PGM}_{A_i}(G) = \frac{\#\mathsf{MW}_G(A_i)}{\#\mathsf{MW}(G)}.$$

Deegan–Packel Power Measure. The *Deegan–Packel power measure* $\mathsf{DP}_G(A_i)$ for a player A_i in the game G is defined to be

$$\mathsf{DP}_G(A_i) = \frac{1}{\#\mathsf{MW}(G)} \sum_{S \in \mathsf{MW}_G(A_i)} \frac{1}{\#S}.$$

Power Profile. Suppose $\mathcal{P}$ is a measure of voting power. Then, $\mathcal{P}$ assigns a nonnegative real number to each of the n players in the game. So, $\mathcal{P}$ is given by a vector of nonnegative real numbers. This vector is called the $\mathcal{P}$-power profile of the game.

Computing Voting Powers. A weighted majority voting game $G = (N, \mathbf{w}, q)$ is completely specified by the set of players N, a weight vector $\mathbf{w}$, and the threshold q. Given this data, it is of interest to be able to compute the $\mathcal{P}$-power profile for any power measure $\mathcal{P}$. There are known dynamic-programming-based algorithms for computing the values of the different voting power indices. We refer to Matsui and Matsui

(2000), Chakravarty et al. (2015) for an introduction to algorithms for computing voting powers. In our work, we have implemented the algorithms for computing the $\mathcal{P}$-power profiles where $\mathcal{P}$ is any of the power measures defined above.

There is a large literature on voting powers. The various indices mentioned above have been introduced to model certain aspects of voting games which are not adequately covered by the other indices. There have been axiomatic characterizations of these indices. A detailed discussion of the relevant literature is not really within the focus of the present work. Instead, we refer to the highly respected monograph Felsenthal and Machover (1998) and the more recent textbook Chakravarty et al. (2015) for such details. *Our concern in this work is how to quantify the efficacy of any particular voting power measure that one may choose for a particular application.*

2.3 Pearson's Correlation Coefficient

Given vectors $\mathbf{w} = (w_1, \ldots, w_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$, Pearson's correlation coefficient is the standard measure of linear correlation between these two vectors. It is defined as follows:

$$\mathsf{PCC}(\mathbf{w}, \mathbf{v}) = \begin{cases} 0 & \text{if } w_1 = \cdots = w_n \text{ or } v_1 = \cdots = v_n; \\ \frac{\sum_{i=1}^{n}(w_i - \mu_{\mathbf{w}})(v_i - \mu_{\mathbf{v}})}{\sqrt{\sum_{i=1}^{n}(w_i - \mu_{\mathbf{w}})^2}\sqrt{\sum_{i=1}^{n}(v_i - \mu_{\mathbf{v}})^2}} & \text{otherwise.} \end{cases} \tag{1}$$

Here, $\mu_{\mathbf{w}}$ and $\mu_{\mathbf{v}}$ are the means of $\mathbf{w}$ and $\mathbf{v}$, respectively.

From (1), it follows that for any two positive real numbers γ and δ,

$$\mathsf{PCC}(\mathbf{w}, \mathbf{v}) = \mathsf{PCC}(\gamma\mathbf{w}, \delta\mathbf{v}). \tag{2}$$

The relation captured in (2) can be considered to be a scale invariance property of the Pearson's correlation coefficient.

2.4 Inequality Indices

The notion of inequality has been considered for social parameters including income, skills, education, health, and wealth Cowell (2016). There are several methods for measuring inequality. At a basic level, the idea of an inequality index $\mathcal{I}$ is the following. Given a vector $\mathbf{a}$ whose components are real numbers, $\mathcal{I}(\mathbf{a})$ produces a nonnegative real number r. In other words, the index $\mathcal{I}$ assigns an inequality score of r to the vector $\mathbf{a}$. There is a large literature on inequality indices including the measurement of multidimensional inequality Chakravarty and Lugo (2016), Chakravarty (2017). In this work, we will consider only basic inequality indices. Some of the most commonly used inequality indices are mentioned below.

Given a vector $\mathbf{a}$ of real numbers, let $\mu_{\mathbf{a}}$ and $\sigma_{\mathbf{a}}$ denote the mean and standard deviation of $\mathbf{a}$. *In the definition of the inequality indices below, we will assume that the entries of* $\mathbf{a}$ *are nonnegative and* $\mu_{\mathbf{a}}$ *is positive.*

Gini Index. The value of the Gini index of a vector $\mathbf{a} = (a_1, \ldots, a_n)$ is given by

$$\mathsf{GI}(\mathbf{a}) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |a_i - a_j|}{2n \sum_{i=1}^{n} a_i}. \tag{3}$$

Coefficient of Variation. For a vector $\mathbf{a} = (a_1, \ldots, a_n)$, the Coefficient of Variation is computed as the ratio of the standard deviation $\sigma_{\mathbf{a}}$ to the mean $\mu_{\mathbf{a}}$ of $\mathbf{a}$.

$$\mathsf{CoV}(\mathbf{a}) = \frac{\sigma_{\mathbf{a}}}{\mu_{\mathbf{a}}} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} a_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} a_i\right)^2}}{\frac{1}{n}\sum_{i=1}^{n} a_i}. \tag{4}$$

Generalized Entropy Index. The generalized entropy index is a measure of inequality based on information theory. For a real number α, the generalized entropy index $\mathsf{GEI}_{\alpha}(\mathbf{a})$ is defined in the following manner:

$$\mathsf{GEI}_{\alpha}(\mathbf{a}) = \begin{cases} \frac{1}{\alpha(\alpha-1)}\left(\frac{1}{n}\sum_{A_i \in N}\left(\frac{a_i}{\mu_{\mathbf{a}}}\right)^{\alpha} - 1\right) & \text{if } \alpha \neq 0, 1; \\ \frac{1}{n}\sum_{A_i \in N, a_i > 0}\left(\frac{a_i}{\mu_{\mathbf{a}}^{+}}\right)\ln\left(\frac{a_i}{\mu_{\mathbf{a}}^{+}}\right) & \text{if } \alpha = 1; \\ -\frac{1}{n}\sum_{A_i \in N, a_i > 0}\ln\left(\frac{a_i}{\mu_{\mathbf{a}}^{+}}\right) & \text{if } \alpha = 0. \end{cases} \tag{5}$$

Here, ln denotes the natural logarithm and $\mu_{\mathbf{a}}^{+}$ denotes the mean of the positive entries in $\mathbf{a}$. Also, note that for $\alpha = 0, 1$ the sum is over positive values of a_i as otherwise the ln function gets applied to 0. In other words, for $\alpha = 0$ and 1, the computation of inequality considers only the positive entries of $\mathbf{a}$. GEI_1 is called the *Theil Index* and GEI_2 is half the square of CoV.

Remark For application to the context of voting powers, a power of zero implies that the player is a dummy. If GEI_0 or GEI_1 is used to measure inequality, then such dummy players will get ignored. As a result, the inequality in the power profile will not be adequately captured by these two measures. Due to this reason, GEI_0 and GEI_1 are not suitable for measuring inequality in voting powers. GEI_2 is half the square of CoV and will essentially spread out the value of CoV. The relevance of GEI_k for $k > 2$ to the context of voting power is not clear. So, though we have computed, we do not report the values of GEI in this work.

Computing Inequality Indices. It is quite routine to implement an algorithm, which given a vector of nonnegative quantities computes the values of the various inequality indices. In our work, we have implemented algorithms to compute the Gini index and the Coefficient of Variation.

Desirable Properties of an Inequality Index. A few basic and natural properties have been postulated which any reasonable inequality measure should satisfy. Below, we mention these properties. See Cowell (2016) for more details. Let $\mathcal{I}$ be a postulated inequality index and $\mathbf{a} = (a_1, \ldots, a_n)$ be a vector of nonnegative real numbers.

Let π be a bijection from $\{1, \ldots, n\}$ to itself, i.e., π is a permutation of $\{1, \ldots, n\}$. Define $\mathbf{a}_\pi$ to be the vector $(a_{\pi(1)}, \ldots, a_{\pi(n)})$, i.e., $\mathbf{a}_\pi$ is a reordering of the components of $\mathbf{a}$.

Anonymity (ANON): $\mathcal{I}$ is said to satisfy anonymity if $\mathcal{I}(\mathbf{a}) = \mathcal{I}(\mathbf{a}_\pi)$ for all permutations π of $\{1, \ldots, n\}$. Anonymity captures the property that inequality depends only on the (multi-)set of values $\{a_1, \ldots, a_n\}$. Information related to ordering or labeling of these values using names are irrelevant for the measurement of inequality.

Egalitarian Principle (EP): $\mathcal{I}$ is said to satisfy the egalitarian principle if $\mathcal{I}(\mathbf{a}) = 0$ for all $\mathbf{a}$ such that $a_1 = \cdots = a_n$. EP captures the property that the inequality is the minimum possible when all components of the vector $\mathbf{a}$ have the same value.

Scale Invariance (ScI). $\mathcal{I}$ is said to satisfy scale invariance if $\mathcal{I}(\mathbf{a}) = \mathcal{I}(\gamma\mathbf{a})$ for all real $\gamma > 0$. The idea behind scale invariance is that if all the values are scaled by the same factor then the inequality remains unchanged.

Let $\mathbf{a}^{[k]}$ denote the vector

$$\Bigg(\underbrace{a_1, \ldots, a_1}_{k}, \underbrace{a_2, \ldots, a_2}_{k} \ldots, \underbrace{a_n, \ldots, a_n}_{k}\Bigg).$$

Population Principle (PP). $\mathcal{I}$ is said to satisfy the population principle if $\mathcal{I}(\mathbf{a}) = \mathcal{I}(\mathbf{a}^{[k]})$ for any integer $k \geq 1$. The vector $a^{[k]}$ contains k copies of each of the values $a_1, \ldots, a_n$. PP says that the inequality in such a vector remains the same as in the original vector, i.e., by replicating each of the components of the original vector the same number of times does not change the inequality.

For $1 \leq i < j \leq n$, let $\mathbf{a}_{i,j,\delta}$ be the vector

$$(a_1, \ldots, a_{i-1}, a_i + \delta, \ldots, a_j - \delta, a_{j+1}, \ldots, a_n).$$

Transfer Principle (TP). $\mathcal{I}$ is said to satisfy the transfer principle if $\mathcal{I}(\mathbf{a}) \geq \mathcal{I}(\mathbf{a}_{i,j,\delta})$ for any $1 \leq i < j \leq n$ and $\delta > 0$ such that $a_i < a_j$ and $a_i + \delta \leq a_j - \delta$. The transfer principle says that if δ units are transferred from a richer person to a poorer person without changing their relative ordering, then inequality cannot increase.

Suppose $\mathbf{a}_1, \ldots, \mathbf{a}_k$ are vectors of dimensions $n_1, \ldots, n_k$, respectively, with non-negative real entries. Let μ_i be the mean of $\mathbf{a}_i$ and define $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k)$. Let $\mathbf{a}$ be the vector formed by concatenating the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$.

Decomposability (Decom). $\mathcal{I}$ is said to satisfy decomposability if

$$\mathcal{I}(\mathbf{a}) = \sum_{i=1}^{k} n_i \mathcal{I}(\mathbf{a}_i) + \mathcal{I}(\boldsymbol{\mu}).$$

Decomposability captures the following idea. The vector **a** is divided into k groups and inequality is measured for each of the groups. Further, the mean of each group is computed and inequality is computed for the vector composing of the means. The inequality for each group is “within-group inequality’, whereas the inequality in the vector of means is some kind of “across group inequality”. The index $\mathcal{I}$ satisfies decomposability if the overall inequality in the vector can be decomposed into a sum of “within-group inequality” and “across group inequality”.

The Gini Index, the Coefficient of Variation, and the Generalized Entropy Indices satisfy ANON, EP, ScI, PP, and TP. It has been shown Shorrocks (1980) that any index which satisfies ANON, ScI, PP, TP, and Decom must necessarily have the form of a generalized entropy index for some value of α.

ANON, EP, ScI, and TP are natural properties that any inequality index should satisfy irrespective of the domain to which it is applied. PP becomes relevant in the context of variable population size. For voting games, the players constitute the population which is fixed. So, the application of PP to voting games is vacuous. On the other hand, it is not clear that Decom is necessarily a desirable property for *all* applications of inequality. In particular, it is not clear that Decom is relevant in the context of voting powers which is the focus of the present work.

3 Weight–Power Similarity

Let $\mathcal{P}$ be a measure of voting power. Suppose this is applied to a weighted majority voting game $G = (N, \mathbf{w}, q)$. Let $\mathbf{v}$ be the resulting power profile. It is of interest to know how similar the power profile vector $\mathbf{v}$ is to the weight vector $\mathbf{w}$. Note that the power profile vector $\mathbf{v}$ depends on the winning threshold q. Based on the Pearson’s correlation coefficient, we define the similarity index $\mathcal{P}\text{-}\mathsf{SI}_{\mathbf{w}}(q)$ as follows:

$$\mathcal{P}\text{-}\mathsf{SI}_{\mathbf{w}}(q) = \mathsf{PCC}(\mathbf{w}, \mathbf{v}), \tag{6}$$

where $\mathbf{v}$ is the power profile vector generated by the voting power measure $\mathcal{P}$ applied to the weighted majority voting game $G = (N, \mathbf{w}, q)$.

So, for a fixed q, $\mathcal{P}\text{-}\mathsf{SI}_{\mathbf{w}}(q)$ measures the similarity of the power profile vector to the weight vector by the correlation between these two vectors. Note that $\mathcal{P}\text{-}\mathsf{SI}_{\mathbf{w}}(q)$ is a function of q. As, q changes, the power profile vector $\mathbf{v}$ will also change, though the weight vector $\mathbf{w}$ will not change. So, with change in q, the correlation between $\mathbf{w}$ and $\mathbf{v}$ changes. By varying q, it is possible to study the change in the correlation between $\mathbf{w}$ and $\mathbf{v}$.

Theorem 1 *Let* $G = (N, \mathbf{w}, q)$ *be a weighted majority voting game such that* $0 < \#W(G) < 2^n$. *Then, for any* $q \in (0, 1)$ *the following holds:*

1. $\mathsf{BZN}\text{-}\mathsf{SI}_{\mathbf{w}}(q) = \mathsf{BZ}\text{-}\mathsf{SI}_{\mathbf{w}}(q) = \mathsf{CP}\text{-}\mathsf{SI}_{\mathbf{w}}(q) = \mathsf{CI}\text{-}\mathsf{SI}_{\mathbf{w}}(q)$.
2. $\mathsf{PGI}\text{-}\mathsf{SI}_{\mathbf{w}}(q) = \mathsf{PGM}\text{-}\mathsf{SI}_{\mathbf{w}}(q)$.

Proof Let $\mathbf{v} = (m_G(A_1), \ldots, m_G(A_n))$ be the vector of swings for the players $A_1, \ldots, A_n$ in the game G. Suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, and $\mathbf{v}_4$ are the power profiles for G corresponding to BZN, BZ, CP, and CI, respectively. Then

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 = \alpha_2 \mathbf{v}_2 = \alpha_3 \mathbf{v}_3 = \alpha_4 \mathbf{v}_4,$$

where

$$\alpha_1 = 2^{n-1},\ \alpha_2 = \sum_{j \in N} m_G(A_j),\ \alpha_3 = \#W(G) \text{ and } \alpha_4 = \#L(G).$$

In G, the values 2^n, $\sum_{j \in N} m_G(A_j)$, $\#W(G)$, and $\#L(G)$ are fixed. So, α_1, α_2, α_3, and α_4 are constants. Further, since $0 < \#W(G) < 2^n$, it follows that $0 < \#L(G) < 2^n$ and so $\sum_{j \in N} m_G(A_j) > 0$. This in particular means that $\alpha_2, \alpha_3, \alpha_4 > 0$ and clearly $\alpha_1 > 0$. So, using (2), we have

$$\mathsf{BZN\text{-}SI}_{\mathbf{w}}(q) = \mathsf{PCC}(\mathbf{w}, \mathbf{v}_1) = \mathsf{PCC}(\mathbf{w}, \mathbf{v}/\alpha_1) = \mathsf{PCC}(\mathbf{w}, \mathbf{v}).$$

In a similar manner, it follows that $\mathsf{BZ\text{-}SI}_{\mathbf{w}}(q) = \mathsf{PCC}(\mathbf{w}, \mathbf{v})$, $\mathsf{CP\text{-}SI}_{\mathbf{w}}(q) = \mathsf{PCC}(\mathbf{w}, \mathbf{v})$, and $\mathsf{CI\text{-}SI}_{\mathbf{w}}(q) = \mathsf{PCC}(\mathbf{w}, \mathbf{v})$.

The argument for $\mathsf{PGI\text{-}SI}_{\mathbf{w}}(q) = \mathsf{PGM\text{-}SI}_{\mathbf{w}}(q)$ is similar. □

From the viewpoint of weight–power similarity, Theorem 1 shows that it is sufficient to consider only $\mathsf{BZ\text{-}SI}_{\mathbf{w}}(q)$, $\mathsf{PGI\text{-}SI}_{\mathbf{w}}(q)$, and $\mathsf{DP\text{-}SI}_{\mathbf{w}}(q)$.

4 Measuring Inequality of Voting Powers

Let $G = (N, \mathbf{w}, q)$ be a weighted majority voting game. The weights of all the players are not equal. In fact, in several important practical situations, the voting game is designed in a manner such that the weights are indeed unequal. The inequality in the weights can be captured by applying an appropriate inequality measure. Suppose $\mathcal{I}$ is an inequality measure. Then, $\mathcal{I}(\mathbf{w})$ is the inequality present in the weights.

Let $\mathcal{P}$ be a measure of voting power. Suppose $\mathcal{P}$ is applied to G to obtain the power profile vector $\mathbf{v}$. Then, $\mathbf{v}$ is a vector consisting of nonnegative real numbers. The inequality in the vector $\mathbf{v}$ can be measured by the inequality index $\mathcal{I}$ as $\mathcal{I}(\mathbf{v})$. The value of $\mathcal{I}(\mathbf{v})$ depends on the winning threshold q, whereas the value of $\mathcal{I}(\mathbf{w})$ does not depend on q. So, by varying q, it is possible to vary $\mathcal{I}(\mathbf{v})$ with the goal of making it as close to $\mathcal{I}(\mathbf{w})$ as possible. Then, one can say that the inequality present in the weights is more or less reflected in the inequality that arises in the voting powers.

Given a weighted majority voting game $G = (N, \mathbf{w}, q)$, we define the *weight inequality* of G with respect to an inequality measure $\mathcal{I}$ as

$$\mathcal{I}\text{-}\mathsf{WI}(\mathbf{w}) = \mathcal{I}(\mathbf{w}). \tag{7}$$

We consider two different options for $\mathcal{I}$, namely, GI and CoV. This gives rise to two different measures of weight inequality, which are GI-WI and CoV-WI.

Given a weighted majority voting game $G = (N, \mathbf{w}, q)$, a voting power measure $\mathcal{P}$, and an inequality measure $\mathcal{I}$, the *power inequality* of $\mathcal{P}$ as determined by $\mathcal{I}$ is denoted by $(\mathcal{P}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$ and is defined in the following manner:

$$(\mathcal{P}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = \mathcal{I}(\mathbf{v}), \tag{8}$$

where $\mathbf{v}$ is the power profile vector generated by the power measure $\mathcal{P}$ applied to the weighted majority voting game $G = (N, \mathbf{w}, q)$.

We have considered seven options for $\mathcal{P}$, namely, BZN, BZ, CP, CI, PGI, PGM, and DP. The following results show that under a simple and reasonable condition on an inequality measure $\mathcal{I}$, it is sufficient to consider only three of these.

Theorem 2 *Let $\mathcal{I}$ be an inequality index satisfying scale invariance. Let $G = (N, \mathbf{w}, q)$ be a weighted majority voting game such that $0 < \#W(G) < 2^n$. Then, for any $q \in (0, 1)$, the following holds:*

1. $(\mathsf{BZ}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = (\mathsf{BZ}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = (\mathsf{CP}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = (\mathsf{CI}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$.
2. $(\mathsf{PGI}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = (\mathsf{PGM}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$.

Proof As in the proof of Theorem 1, let $\mathbf{v} = (m_G(A_1), \ldots, m_G(A_n))$ be the vector of swings and $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ be the power profiles corresponding to BZN, BZ, CP, and CI, respectively, so that

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 = \alpha_2 \mathbf{v}_2 = \alpha_3 \mathbf{v}_3 = \alpha_4 \mathbf{v}_4, \tag{9}$$

where α_1, α_2, α_3, and α_4 are the positive constants defined in the proof of Theorem 1.

Using the scale invariance of $\mathcal{I}$, we have

$$(\mathsf{BZN}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = \mathcal{I}(\mathbf{v}_1) = \mathcal{I}(\mathbf{v}/\alpha_1) = \mathcal{I}(\mathbf{v}).$$

In a similar manner, it follows that $(\mathsf{BZ}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = \mathcal{I}(\mathbf{v})$ $(\mathsf{CP}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = \mathcal{I}(\mathbf{v})$ and $(\mathsf{CI}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = \mathcal{I}(\mathbf{v})$.

The argument for $(\mathsf{PGI}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q) = (\mathsf{PGM}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$ is similar. □

The Gini Index, the Coefficient of Variation, and the Generalized Entropy Index satisfy the scale invariance property. Based on Theorem 2, if $\mathcal{I}$ is any of these indices, then from the viewpoint of power inequality, it is sufficient to consider $(\mathsf{BZ}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$, $(\mathsf{PGI}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$, and $(\mathsf{DP}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$. For $\mathcal{I}$, we will consider the Gini Index and the Coefficient of Variation. This means that we need to consider six possibilities.

Remark Theorem 2 has been stated for weighted majority voting games. The crux of the argument is based on (9). This relation does require G to be a weighted majority voting game. So, it is possible to rewrite the proof to show that for general voting games (which are not necessarily weighted majority voting games), the inequalities of the different Banzhaf power profiles and the Coleman power profiles are all equal and also the inequality of the Holler public good index is equal to that of the Holler public good measure.

Comparison to Previous Works. The work Einy and Peleg (1991) on inequality in voting system is concerned with measuring inequality arising in the Shapley–Shubik power index. Since we do not consider this index in our work, we do not comment any further on the work in Einy and Peleg (1991). Instead, we simply remark that our approach can also be applied to the Shapley–Shubik power index.

Laruelle and Valenciano (2004) axiomatically derive an inequality index for a class of power indices which includes the normalized Banzhaf index. We note the following points about the work in Laruelle and Valenciano (2004):

1. The approach works only for an index, i.e., the sum of the powers must sum to one. So, for example, it cannot be applied to measure inequality arising in either of the Coleman power measures.
2. The notion of power considered in the work is based on swings. So, the power measures given by PGI, PGM, and DP are not covered by their work.
3. Among the axioms, ANON and EP are assumed and it is shown that the obtained measure satisfies ScI. On the other hand, PP and TP are not mentioned in the paper and it is not clear whether these two properties hold for the obtained measure.
4. Justification for one of the axioms (namely, Constant Sensitivity to Null Players) is not clear. In the discussion leading up to this axiom, the authors remark: "Thus, at this point any further step is questionable, though necessary to specify an index."

In view of the above, we feel that it might be preferable to study the behavior of standard inequality indices on both power measures and power indices rather than relying on an axiomatically derived inequality index where at least one of the axioms does not necessarily have a natural justification.

Weber (2016) considered the application[1] of the Coefficient of Variation to the measurement of inequality for essentially the normalized Banzhaf index. In contrast, we consider all the standard inequality indices and a much larger class of power measures/indices. Even for the Coefficient of Variation, the result that the scale invariance property implies that all the swing-based measures have the same inequality is not present in Weber (2016).

5 Variation of Similarity and Inequality with Winning Threshold

We have conducted some experiments to understand the dependence of the similarity and inequality indices of power profiles on the winning threshold.

[1]The author remarks: "To the best of my knowledge, I am the first to propose a measure of inequality of voting systems that can be used across different constituencies." This is an oversight since the work by Laruelle and Valenciano Laruelle and Valenciano (2004) is much earlier.

In the first experiment, N was taken to be $\{A_1, \ldots, A_{30}\}$ and one hundred weight vectors were generated where the individual weights were chosen to be integers independently and uniformly in the range $[1, \ldots, 100]$. For each of the 100 weight vectors $\mathbf{w}$, the value of the winning threshold q was varied from 0.01 to 0.99 in steps of 0.01. For the game defined by the triplet $(N, \mathbf{w}, q)$, the power profiles for the different power measures were computed. From this, the similarity indices $\mathsf{BZ\text{-}SI}_{\mathbf{w}}(q)$, $\mathsf{PGI\text{-}SI}_{\mathbf{w}}(q)$, and $\mathsf{DP\text{-}SI}_{\mathbf{w}}(q)$ were computed and the inequality indices $\mathcal{I}\text{-}\mathsf{WI}(\mathbf{w})$, $(\mathsf{BZN}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$, $(\mathsf{PGI}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$, and $(\mathsf{DP}, \mathcal{I})\text{-}\mathsf{PI}_{\mathbf{w}}(q)$ were computed where $\mathcal{I}$ was taken to be GI and CoV. All the obtained results show a definite pattern.

A second experiment was conducted with $n = 30$ and nonrandom weights. In particular, two distinct values of weights were used, namely, n_1 of the weights were taken to be 100 and n_2 of the weights were taken to be 1 with $n_1 + n_2 = 30$. The value of q was varied as mentioned above and the corresponding similarity and inequality indices were computed. In this case, no definite pattern was observed and there was a rich variation in the behavior.

To further explain the above experiments, we report three particular cases with $n = 30$.

Case-I: A random weight vector. The actual value of $\mathbf{w}$ (after sorting into descending order) came out to be

$$\{93, 92, 90, 86, 86, 83, 82, 77, 72, 68, 67, 67, 63, 62, 62, 59, \\ 49, 40, 36, 35, 35, 30, 29, 27, 26, 26, 23, 21, 15, 11\}.$$

The plots of similarity, Gini Index, and the Coefficient of Variation are shown in Figs. 1, 2, and 3, respectively.

Case-II: 15 of the weights were taken to be 100 and the other 15 of the weights were taken to be 1. The plots of similarity, Gini Index and the Coefficient of Variation are shown in Figs. 4, 5 and 6 respectively.

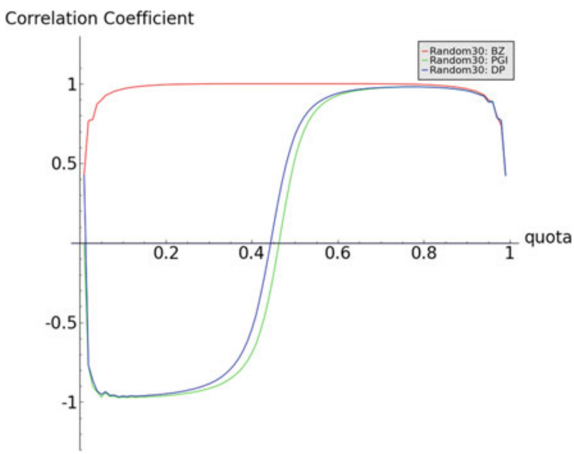


Fig. 1 Similarity indices for Case-I

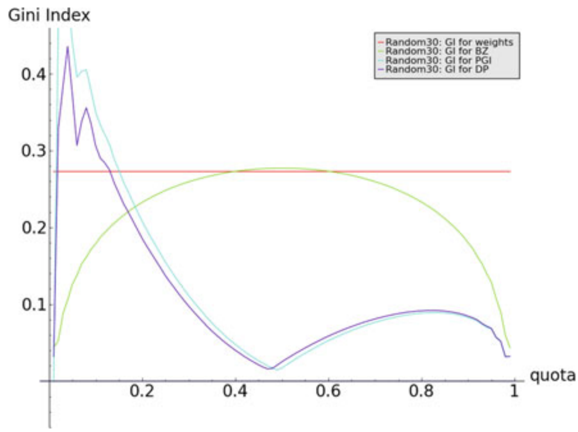


Fig. 2 Gini index for Case-I

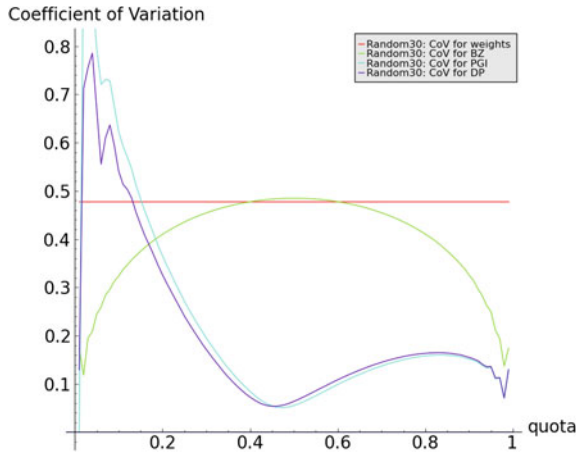


Fig. 3 Coefficient of variation for Case-I

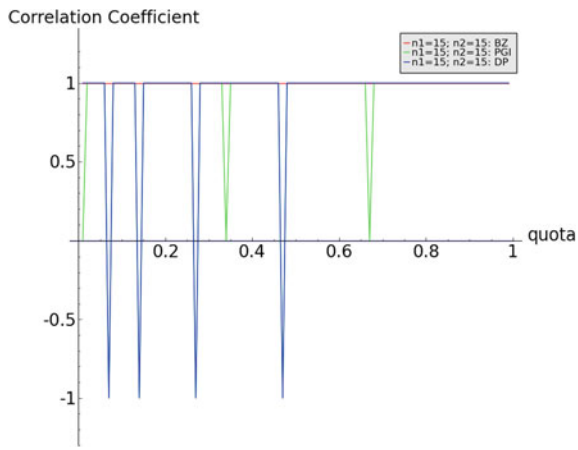


Fig. 4 Similarity indices for Case-II

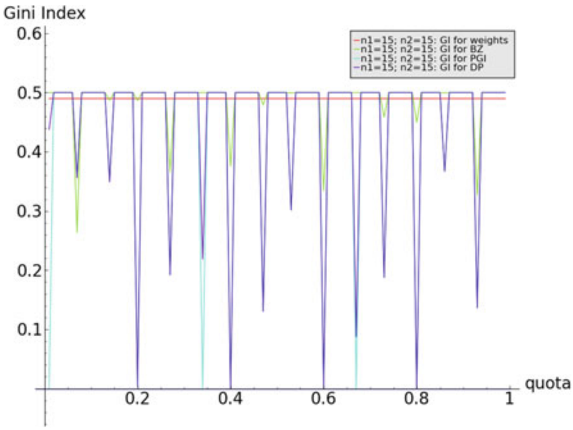


Fig. 5 Gini index for Case-II

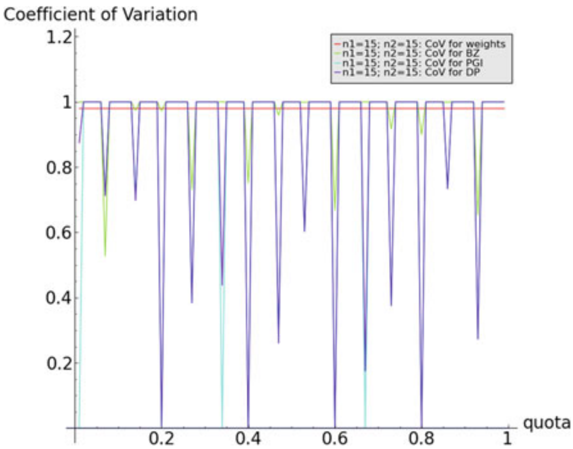


Fig. 6 Coefficient of variation for Case-II

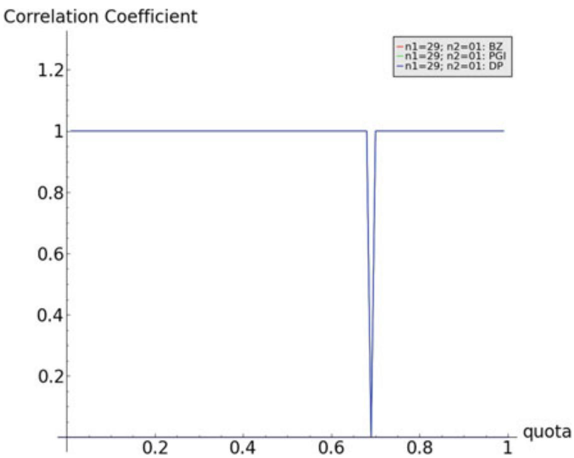


Fig. 7 Similarity indices for Case-III

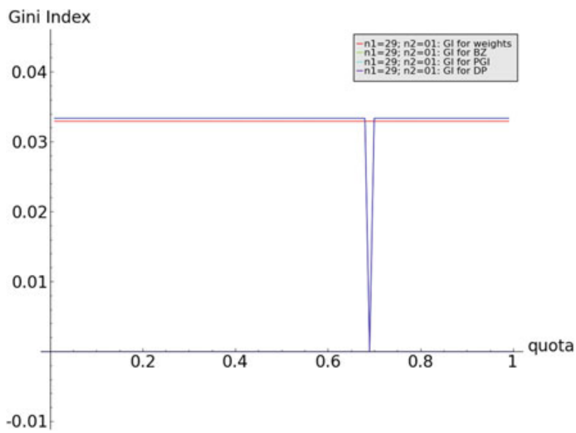


Fig. 8 Gini index for Case-III

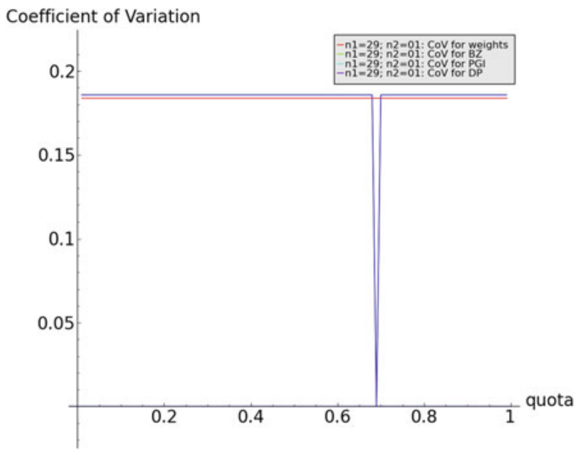


Fig. 9 Coefficient of variation for Case-III

Case-III: 29 of the weights were taken to be 100 and the other weight was taken to be 1. The plots of similarity, Gini Index, and the Coefficient of Variation are shown in Figs. 7, 8, and 9, respectively.

Based on the plots, we make the following observations:

1. For the random case, compared to the Holler index and the Deegan–Packel measure, the Banzhaf index is a much better marker of similarity to the weight vector and it is also much better at capturing the inequality present in the weights. For the two nonrandom cases, there is not much difference between the three power measures.
2. For Case-II, there are sharp spikes in the similarity and inequality plots. These correspond to choices of q for which the low weight players achieve power similar to the high weight players.

3. For Case-III, there is one player with low weight. Apart from a small range of q where this player gains significant power, for other values of q this player becomes a dummy. The inequality indices, however, do not reflect this. The inequality scores are generally quite low indicating that there is small inequality in the system. This is a feature of the inequality indices which are not sensitive to low scores of a small number of players.

6 Applications

6.1 IMF Voting Games

The IMF has two decision-making bodies, namely, the Board of Governors (BoG) and the Executive Board (EB).

A total of 189 member countries make up the BoG. Each country has a specified voting share. The voting share or weight of a country is calculated as the sum of a basic weight plus an amount which is proportional to the special drawing rights of the country. The EB consists of 24 Executive Directors (EDs) representing all the 189 member countries. Eight of these directors are nominated by eight member countries while each of the other directors is elected by a group of countries. Each ED has a voting weight which is the sum of the voting weights of the countries that he or she represents.

The BoG is the highest decision-making body of the IMF and is officially responsible for all major decisions. In practice, however, the BoG has delegated most of its powers to the EB.[2] Accordingly, in this work, we will consider only the voting game arising from the EB weights.

Actual weights of the members of the EB are available from the IMF website.[3] These weights range from the minimum of 80157 to the maximum of 831407. Since these values are rather large, for the purposes of computation of voting powers, we have divided these voting weights by 1000 and then rounded to the nearest integer. While this is an approximation, it does not significantly affect the voting powers. In particular, we have checked that no two members with originally unequal weights get the same weight after this rounding off process. The actual weight vector that has been used to compute the voting powers is the following:

$$\mathbf{w}_{\mathsf{imf}} = \{831, 310, 306, 273, 268, 267, 219, 208, 203, 203, 196, 170, 165, 162, 155, 154, 150, 149, 138, 131, 111, 101, 82, 80\}.$$

The rules specify several winning thresholds.[4] We mention these below:

[2] https://en.wikipedia.org/wiki/International_Monetary_Fund.

[3] https://www.imf.org/external/np/sec/memdir/eds.aspx.

[4] https://www.imf.org/external/pubs/ft/aa/index.htm.

> Except as otherwise specifically provided, all decisions of the Fund shall be made by a majority of the votes cast.
>
> The Fund, by a seventy percent majority of the total voting power, may decide at any time to distribute any part of the general reserve.
>
> The Fund may use a member's currency held in the Investment Account for investment as it may determine, in accordance with rules and regulations adopted by the Fund by a seventy percent majority of the total voting power.
>
> An eighty-five percent majority of the total voting power shall be required for any change in quotas.

So, three possible values of q are used: $q = 0.5$, $q = 0.7$, and $q = 0.85$.

We have computed the similarity and inequality indices for the IMF EB voting game with q varying from 0.01 to 0.99 in steps of 0.01. The plots of $\mathsf{BZ\text{-}SI}_{\mathbf{w}_{\mathrm{imf}}}(q)$, $\mathsf{PGI\text{-}SI}_{\mathbf{w}_{\mathrm{imf}}}(q)$, and $\mathsf{DP\text{-}SI}_{\mathbf{w}_{\mathrm{imf}}}(q)$ are shown in Fig. 10; the plots of $(\mathsf{BZ}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathrm{imf}}}(q)$, $(\mathsf{PGI}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathrm{imf}}}(q)$, and $(\mathsf{DP}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathrm{imf}}}(q)$ are shown in Fig. 11; and the plots of

Fig. 10 Similarity indices for IMF-EB weights

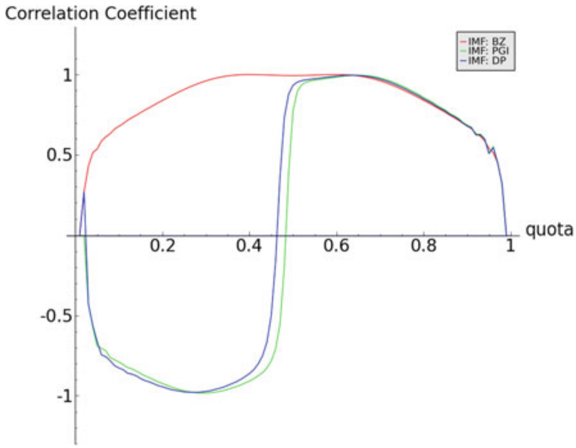
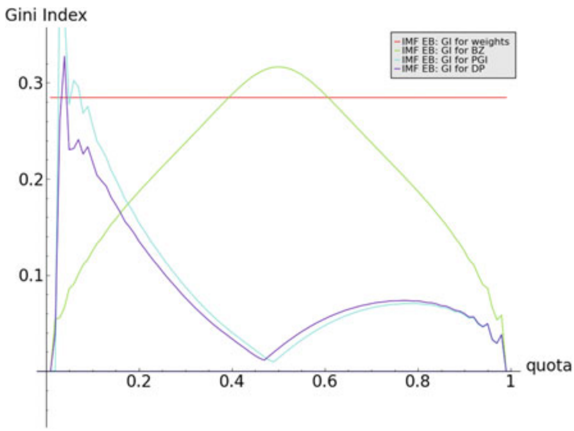


Fig. 11 Gini index for IMF-EB weights



(BZ, CoV)-$\mathrm{PI}_{\mathbf{w}_{\mathrm{imf}}}(q)$, (PGI, CoV)-$\mathrm{PI}_{\mathbf{w}_{\mathrm{imf}}}(q)$, and (DP, CoV)-$\mathrm{PI}_{\mathbf{w}_{\mathrm{imf}}}(q)$ are shown in Fig. 12. The actual values of these indices for the range [0.5, 0.65] along with the values for $q = 0.70$ and $q = 0.85$ are shown in Tables 1 and 2. Based on these data, we have the following observations:

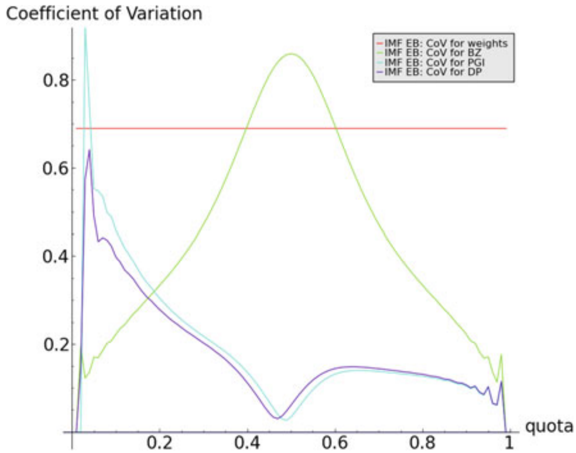


Fig. 12 Coefficient of variation for IMF-EB weights

Table 1 IMF-EB similarity indices

q	BZ-$\mathrm{SI}_{\mathbf{w}_{\mathrm{imf}}}(q)$	PGI-$\mathrm{SI}_{\mathbf{w}_{\mathrm{imf}}}(q)$	DP-$\mathrm{SI}_{\mathbf{w}_{\mathrm{imf}}}(q)$
0.50	0.9938690	0.7743330	0.9330398
0.51	0.9939952	0.8965981	0.9531445
0.52	0.9943706	0.9363455	0.9618091
0.53	0.9949599	0.9522328	0.9669979
0.54	0.9957522	0.9600696	0.9708993
0.55	0.9966513	0.9649342	0.9740920
0.56	0.9975953	0.9688469	0.9771681
0.57	0.9985210	0.9730949	0.9806975
0.58	0.9992881	0.9769905	0.9839294
0.59	0.9998168	0.9806887	0.9869004
0.60	**0.9999911**	0.9846956	0.9899614
0.61	0.9996918	0.9882609	0.9924378
0.62	0.9988216	0.9912144	0.9942195
0.63	0.9972257	0.9936152	0.9953372
0.64	0.9948817	0.9951439	0.9954396
0.65	0.9916875	0.9956169	0.9945365
0.70	0.9609858	0.9788570	0.9715726
0.85	0.7634234	0.7787653	0.7721551

Table 2 Inequality as measured by GI and CoV in the IMF Game. In the table, $\mathfrak{f}_1(q) = (\mathsf{BZ}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{imf}}}(q)$, $\mathfrak{f}_2(q) = (\mathsf{PGI}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{imf}}}(q)$, $\mathfrak{f}_3(q) = (\mathsf{DP}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{imf}}}(q)$, $\mathfrak{g}_1(q) = (\mathsf{BZ}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{imf}}}(q)$, $\mathfrak{g}_2(q) = (\mathsf{PGI}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{imf}}}(q)$, $\mathfrak{g}_3(q) = (\mathsf{DP}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{imf}}}(q)$. Note $\mathsf{GI}\text{-}\mathsf{WI}(\mathbf{w}_{\mathsf{imf}}) = \mathbf{0.285042}$ and $\mathsf{CoV}\text{-}\mathsf{WI}(\mathbf{w}_{\mathsf{imf}}) = \mathbf{0.689527}$

q	$\mathfrak{f}_1(q)$	$\mathfrak{f}_2(q)$	$\mathfrak{f}_3(q)$	$\mathfrak{g}_1(q)$	$\mathfrak{g}_2(q)$	$\mathfrak{g}_3(q)$
0.50	0.316535	0.013684	0.023564	0.859515	0.033982	0.061049
0.51	0.316212	0.017718	0.027366	0.857502	0.046205	0.074206
0.52	0.315214	0.021654	0.031084	0.851387	0.059712	0.086957
0.53	0.313531	0.025291	0.034457	0.841354	0.072727	0.098447
0.54	0.311215	0.028989	0.037852	0.827460	0.085147	0.108920
0.55	0.308401	0.032545	0.041122	0.810689	0.096331	0.118073
0.56	0.305067	0.035747	0.044020	0.791192	0.105924	0.125614
0.57	0.301236	0.038982	0.046930	0.769085	0.114385	0.132059
0.58	0.297133	0.042029	0.049670	0.745825	0.121321	0.137161
0.59	0.292743	0.044794	0.052123	0.721395	0.126885	0.141047
0.60	0.288019	0.047504	0.054517	**0.695812**	0.131366	0.144021
0.61	**0.283235**	0.050075	0.056795	0.670527	0.134712	0.131366
0.62	0.278285	0.052419	0.058857	0.645342	0.137090	0.147386
0.63	0.273201	0.054555	0.060691	0.620092	0.138635	0.147978
0.64	0.268155	0.056714	0.062592	0.596040	0.139694	0.148302
0.65	0.263088	0.058572	0.064200	0.572826	0.140068	0.148049
0.66	0.257927	0.060254	0.065613	0.550134	0.139974	0.147422
0.67	0.252875	0.061895	0.067025	0.528952	0.139751	0.146796
0.68	0.247854	0.063419	0.068331	0.508843	0.139248	0.145961
0.69	0.242726	0.064596	0.069269	0.489429	0.138348	0.144745
0.70	0.237753	0.065816	0.070276	0.471442	0.137616	0.143776
0.85	0.158497	0.066284	0.068153	0.278829	0.116985	0.120198

1. The Holler and the Deegan–Packel indices are not good indicators of either the similarity to or the inequality present in the weights. So, we focus only on the Banzhaf index.
2. The following holds for the Banzhaf index:
   - The plots of the two inequality indices have bell curve shapes. To a lesser extent, the same is also true of the similarity index.
   - The maximum similarity is achieved for $q = 0.60$.
   - For the Gini Index, the inequality in the power profile is closest to the inequality in the weights for $q = 0.61$.
   - For the Coefficient of Variation, the inequality in the power profile is closest to the inequality in the weights for $q = 0.60$.

From Tables 1 and 2, we note that in comparison to $q = 0.60$, the choices $q = 0.50$, $q = 0.70$, and $q = 0.85$ are suboptimal.

Table 3 Voting powers of the players under various power measures for the IMF-EB game with $q = 0.60$. Instead of the players, the voting powers are shown against the weights of the players

Wts	BZN	BZ	CP	CI	PGI	PGM	DP
831	0.4033509	0.1661335	0.8949349	0.2603447	0.0662802	0.8888586	0.0682873
310	0.1508728	0.0621420	0.3347491	0.0973816	0.0435712	0.5843174	0.0442821
306	0.1488262	0.0612990	0.3302082	0.0960606	0.0434846	0.5831563	0.0441693
273	0.1322311	0.0544638	0.2933878	0.0853492	0.0427837	0.5737562	0.0432420
268	0.1297327	0.0534347	0.2878445	0.0837366	0.0426849	0.5724312	0.0431133
267	0.1292485	0.0532353	0.2867701	0.0834240	0.0426597	0.5720929	0.0430820
219	0.1055294	0.0434658	0.2341434	0.0681144	0.0416246	0.5582125	0.0417465
208	0.1001409	0.0412464	0.2221877	0.0646364	0.0414027	0.5552361	0.0414634
203	0.0976979	0.0402401	0.2167671	0.0630595	0.0413069	0.5539514	0.0413394
203	0.0976979	0.0402401	0.2167671	0.0630595	0.0413069	0.5539514	0.0413394
196	0.0942787	0.0388318	0.2091809	0.0608526	0.0411606	0.5519899	0.0411522
170	0.0816418	0.0336269	0.1811427	0.0526961	0.0406361	0.5449551	0.0404690
165	0.0792183	0.0326287	0.1757655	0.0511318	0.0405488	0.5437851	0.0403568
162	0.0777606	0.0320283	0.1725313	0.0501909	0.0404977	0.5430995	0.0402881
155	0.0743695	0.0306316	0.1650074	0.0480022	0.0403450	0.5410517	0.0400980
154	0.0738934	0.0304355	0.1639510	0.0476948	0.0403227	0.5407521	0.0400702
150	0.0719494	0.0296347	0.1596377	0.0464400	0.0402490	0.5397639	0.0399749
149	0.0714759	0.0294397	0.1585871	0.0461344	0.0402247	0.5394390	0.0399449
138	0.0661672	0.0272532	0.1468086	0.0427079	0.0399681	0.5359976	0.0396272
131	0.0627846	0.0258599	0.1393032	0.0405246	0.0397938	0.5336591	0.0394117
111	0.0531555	0.0218939	0.1179389	0.0343095	0.0389966	0.5229683	0.0384651
101	0.0483502	0.0199146	0.1072770	0.0312079	0.0383484	0.5142761	0.0377462
82	0.0392314	0.0161588	0.0870447	0.0253221	0.0361261	0.4844734	0.0353936
80	0.0382680	0.0157619	0.0849071	0.0247002	0.0356770	0.4784506	0.0349375

Based on the above analysis, we put forward the suggestion that the winning threshold of $q = 0.60$ be seriously considered for any future possible change in voting rule. For $q = 0.60$, the actual values of the different power measures are shown in Table 3.

6.2 EU Voting Games

Until Brexit is effective, the European Union Council has 28 members. It votes on different types of matters in three different ways.[5] The first is the unanimity voting where all members have to vote in favor or against for the motion to be passed or

[5] http://www.consilium.europa.eu/en/council-eu/voting-system/.

Table 4 Population percentages of the countries in the European Union

Country	Pop (%)	w_{eu}
Germany	16.06	1606
France	13.05	1305
United Kingdom	12.79	1279
Italy	12.00	1200
Spain	9.09	909
Poland	7.43	743
Romania	3.87	387
Netherlands	3.37	337
Belgium	2.21	221
Greece	2.11	211
Czech Republic	2.04	204
Portugal	2.02	202
Sweden	1.96	196
Hungary	1.92	192
Austria	1.71	172
Bulgaria	1.40	140
Denmark	1.12	112
Finland	1.07	107
Slovakia	1.06	106
Ireland	0.91	91
Croatia	0.82	82
Lithuania	0.57	57
Slovenia	0.40	40
Latvia	0.39	39
Estonia	0.26	26
Cyprus	0.17	17
Luxembourg	0.11	11
Malta	0.09	9

dismissed. In nonlegislative issues, a simple majority voting is done where at least 15 out of the 28 members have to vote in favor. For most (80%) of the issues that are voted upon in the EU Council, the "qualified majority" method is used. This is stated as follows[6]:

> A qualified majority needs 55% of member states, representing at least 65% of the EU population.

It is the qualified majority voting rule that we consider in the context of weighted majority voting games. The population percentages of the individual countries are

[6]http://www.consilium.europa.eu/en/council-eu/voting-system/qualified-majority/.

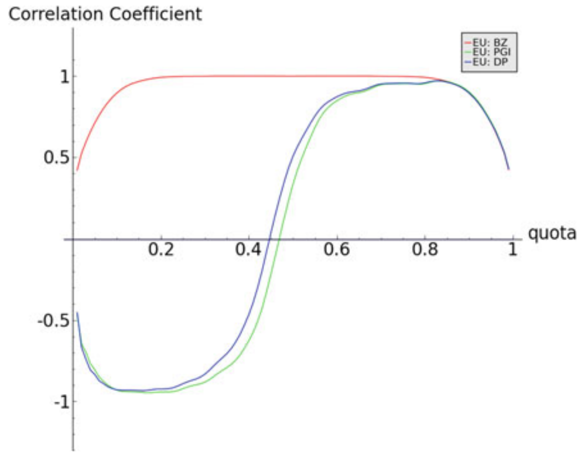


Fig. 13 Similarity indices for EU weights

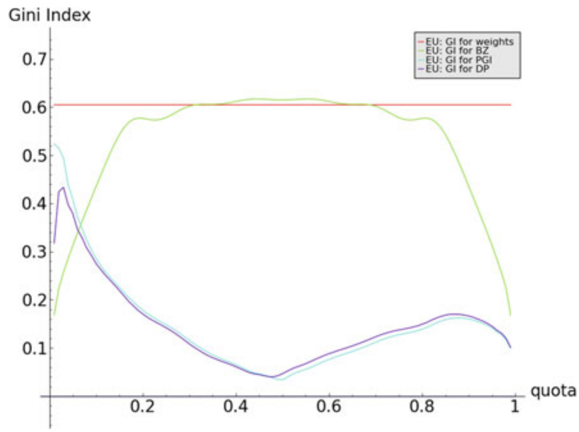


Fig. 14 Gini index for EU weights

available[7] and are reproduced in Table 4. These percentages are the weights of the individual countries. For computation of the voting powers, the percentage values are multiplied by 100 to convert these into integers which are then used as the weights. This scaling does not affect the decision-making process.

In the qualified majority voting, the passing rule is a joint condition, one on the number of member states which are involved and the other on the population percentage. While analyzing the joint condition would be more accurate, for the purpose of this work, we have worked with the simpler setting where only the winning condition on the population percentage is considered. This leads to a weighted majority voting game where the weight vector $\mathbf{w}_{\text{eu}}$ is specified in Table 4 and the winning threshold is $q = 0.65$.

[7]http://www.consilium.europa.eu/en/council-eu/voting-system/voting-calculator/.

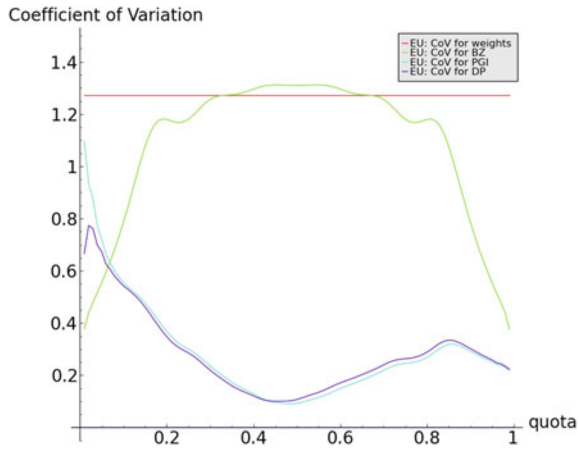


Fig. 15 Coefficient of variation for EU weights

Table 5 EU game similarity indices

q	BZ-SI$_{\mathbf{w}_{\mathrm{eu}}}(q)$	PGI-SI$_{\mathbf{w}_{\mathrm{eu}}}(q)$	DP-SI$_{\mathbf{w}_{\mathrm{eu}}}(q)$
0.50	0.9988826	0.3320853	0.5051382
0.51	0.9989581	0.4181647	0.5640333
0.52	0.9991501	0.4929785	0.6151950
0.53	0.9993725	0.5596608	0.6631804
0.54	0.9995411	0.6230357	0.7099130
0.55	0.9996231	0.6830571	0.7540455
0.56	0.9996430	0.7350513	0.7910700
0.57	0.9996501	0.7765455	0.8199512
0.58	0.9996763	0.8063575	0.8409968
0.59	0.9997123	0.8289370	0.8576475
0.60	0.9997249	0.8477416	0.8722319
0.61	0.9996959	0.8642583	0.8846472
0.62	0.9996486	0.8777664	0.8944232
0.63	0.9996354	0.8870017	0.9006101
0.64	0.9996914	0.8932824	0.9048869
0.65	0.9997886	0.8976327	0.9086210
0.66	**0.9998332**	0.9039397	0.9150038
0.67	0.9997214	0.9135211	0.9242301
0.68	0.9994175	0.9258465	0.9351369
0.69	0.9989968	0.9381400	0.9448867
0.70	0.9986065	0.9465521	0.9507084

Table 6 Inequality as measured by GI and CoV in the EU Game. In the table $\mathfrak{h}_1(q) = (\mathsf{BZ}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $\mathfrak{h}_2(q) = (\mathsf{PGI}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $\mathfrak{h}_3(q) = (\mathsf{DP}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $\mathfrak{k}_1(q) = (\mathsf{BZ}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $\mathfrak{k}_2(q) = (\mathsf{PGI}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $\mathfrak{k}_3(q) = (\mathsf{DP}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$. Note $\mathsf{GI}\text{-}\mathsf{WI}(\mathbf{w}_{\mathsf{eu}}) = \mathbf{0.605671}$ and $\mathsf{CoV}\text{-}\mathsf{WI}(\mathbf{w}_{\mathsf{eu}}) = \mathbf{1.272825}$

q	$\mathfrak{h}_1(q)$	$\mathfrak{h}_2(q)$	$\mathfrak{h}_3(q)$	$\mathfrak{k}_1(q)$	$\mathfrak{k}_2(q)$	$\mathfrak{k}_3(q)$
0.50	0.614849	0.033960	0.046614	1.309745	0.091907	0.107286
0.51	0.614993	0.039425	0.051749	1.309901	0.095374	0.112346
0.52	0.615391	0.044681	0.056573	1.310338	0.100199	0.118139
0.53	0.615945	0.049083	0.060556	1.310962	0.105176	0.123691
0.54	0.616483	0.052608	0.063782	1.311545	0.109910	0.128951
0.55	0.616834	0.055576	0.067481	1.311731	0.114808	0.134668
0.56	0.616822	0.059365	0.071789	1.311070	0.120243	0.140936
0.57	0.616302	0.063896	0.076268	1.309147	0.126922	0.148409
0.58	0.615243	0.068081	0.080370	1.305796	0.134171	0.156036
0.59	0.613694	0.072437	0.084545	1.301119	0.142004	0.163861
0.60	0.611813	0.076339	0.088256	1.295536	0.149118	0.170797
0.61	0.609851	0.079853	0.091521	1.289678	0.155445	0.176837
0.62	0.608078	0.083431	0.094969	1.284198	0.161705	0.183060
0.63	0.606708	0.086467	0.097851	1.279626	0.167582	0.188834
0.64	0.605893	0.089760	0.101116	1.276296	0.174214	0.195451
0.65	0.605616	0.093045	0.104334	1.274202	0.181046	0.202085
0.66	**0.605689**	0.096528	0.107787	**1.272960**	0.188107	0.209023
0.67	0.605797	0.100171	0.111427	1.271778	0.195381	0.216243
0.68	0.605497	0.103949	0.115322	1.269457	0.202594	0.223577
0.69	0.604333	0.107898	0.119249	1.264638	0.210268	0.231500
0.70	0.601994	0.111935	0.123198	1.256239	0.218490	0.239777

For the weight vector $\mathbf{w}_{\mathsf{eu}}$ given in Table 4, we have computed the similarity and inequality indices with q varying from 0.01 to 0.99 in steps of 0.01. The plots of $\mathsf{BZ}\text{-}\mathsf{SI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $\mathsf{PGI}\text{-}\mathsf{SI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, and $\mathsf{DP}\text{-}\mathsf{SI}_{\mathbf{w}_{\mathsf{eu}}}(q)$ are shown in Fig. 13; the plots of $(\mathsf{BZ}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $(\mathsf{PGI}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, and $(\mathsf{DP}, \mathsf{GI})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$ are shown in Fig. 14; and the plots of $(\mathsf{BZ}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, $(\mathsf{PGI}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$, and $(\mathsf{DP}, \mathsf{CoV})\text{-}\mathsf{PI}_{\mathbf{w}_{\mathsf{eu}}}(q)$ are shown in Fig. 15. The actual values of these indices for the range $[0.5, 0.7]$ are shown in Tables 5 and 6.

Based on these data, we have the following observations:

1. As in the case of the IMF-EB voting game, the Holler and the Deegan–Packel indices are not good indicators of either the similarity to or the inequality present in the weights. So, again we focus only on the Banzhaf index.
2. The following holds for the Banzhaf index:
   - The plots of the two inequality indices as well as the similarity index have a somewhat bell-shaped nature.

Table 7 Voting powers of the players under various power measures for the EU game with $q = 0.66$. Instead of the players, the voting powers are shown against the weights of the players

Wts	BZN	BZ	CP	CI	PGI	PGM	DP
1606	0.2459032	0.1582166	0.7950917	0.1454426	0.0537148	0.8173260	0.0554496
1305	0.2044021	0.1315144	0.6609039	0.1208962	0.0479912	0.7302363	0.0496147
1279	0.2003657	0.1289173	0.6478528	0.1185089	0.0475521	0.7235548	0.0491300
1200	0.1881286	0.1210439	0.6082861	0.1112711	0.0463214	0.7048273	0.0477734
909	0.1444251	0.0929245	0.4669771	0.0854221	0.0410962	0.6253212	0.0423199
743	0.1115583	0.0717777	0.3607073	0.0659826	0.0368236	0.5603092	0.0372597
387	0.0600651	0.0386465	0.1942116	0.0355263	0.0379704	0.5777581	0.0380942
337	0.0523199	0.0336631	0.1691686	0.0309453	0.0371804	0.5657380	0.0372361
221	0.0343284	0.0220872	0.1109957	0.0203039	0.0355657	0.5411688	0.0354550
211	0.0327751	0.0210878	0.1059733	0.0193852	0.0354097	0.5387953	0.0352811
204	0.0316896	0.0203894	0.1024638	0.0187432	0.0353493	0.5378758	0.0352135
202	0.0313774	0.0201886	0.1014543	0.0185586	0.0353118	0.5373050	0.0351703
196	0.0304473	0.0195901	0.0984469	0.0180085	0.0352300	0.5360610	0.0350762
192	0.0298266	0.0191907	0.0964398	0.0176413	0.0352006	0.5356133	0.0350399
171	0.0265634	0.0170911	0.0858887	0.0157112	0.0349022	0.5310729	0.0347002
140	0.0217525	0.0139958	0.0703336	0.0128658	0.0345472	0.5256718	0.0343204
112	0.0173997	0.0111951	0.0562592	0.0102913	0.0341762	0.5200262	0.0338917
107	0.0166246	0.0106964	0.0537533	0.0098328	0.0341039	0.5189262	0.0338115
106	0.0164687	0.0105961	0.0532489	0.0097406	0.0341017	0.5188932	0.0338069
91	0.0141386	0.0090969	0.0457151	0.0083625	0.0338729	0.5154111	0.0335409
82	0.0127383	0.0081959	0.0411873	0.0075342	0.0336632	0.5122204	0.0332955
57	0.0088505	0.0056945	0.0286168	0.0052347	0.0329459	0.5013056	0.0324475
40	0.0062171	0.0040002	0.0201021	0.0036772	0.0321913	0.4898234	0.0315918
39	0.0060615	0.0039000	0.0195989	0.0035851	0.0321094	0.4885777	0.0314986
26	0.0040399	0.0025993	0.0130624	0.0023894	0.0300802	0.4577017	0.0292896
17	0.0026410	0.0016992	0.0085393	0.0015621	0.0273734	0.4165147	0.0264409
11	0.0017116	0.0011013	0.0055342	0.0010123	0.0238031	0.3621886	0.0228048
9	0.0013991	0.0009002	0.0045238	0.0008275	0.0214120	0.3258058	0.0204462

– The maximum similarity is achieved for $q = 0.66$.
– For the Gini Index, the inequality in the power profile is closest to the inequality in the weights for $q = 0.66$.
– For the Coefficient of Variation, the inequality in the power profile is closest to the inequality in the weights for $q = 0.66$.

So, the value of $q = 0.66$ is the best in terms of similarity and also for inequality measured by either the Gini Index or the Coefficient of Variation.

The value of q actually used in the EU voting games is $q = 0.65$. The corresponding similarity value and the values for Gini Index and Coefficient of Variation are shown in Tables 1 and 2, respectively. These values show that in comparison to $q = 0.66$, the choice $q = 0.65$ is suboptimal but, very close. For $q = 0.66$, the actual values of the different power measures are shown in Table 7.

Unlike the case of the IMF voting game, our analysis shows that for the weighted majority game arising in the context of EU, the winning threshold is very close to the optimal value. So, our analysis provides some quantitative backing to the actual winning threshold used in the EU game.

7 Conclusion

In this paper, we have addressed the problem of quantifying whether a power profile adequately captures the natural variation in the weights of a weighted majority voting game. Ideas based on Pearson's correlation coefficient and standard inequality measures such as the Gini Index and the Coefficient of Variation have been used in the formalization. These ideas have been applied to the voting games arising in the context of the IMF and the EU. We provide concrete quantitative evidence that the actual winning threshold used in the IMF games is suboptimal and instead propose a new value of the winning threshold which has firm theoretical justification. In the case of the EU game, the actual value of the winning threshold used is close to the optimal value. So, in this case, our analysis provides some quantitative backing to the value of the winning threshold that is actually used.

Acknowledgements We would like to thank Professor Satya Ranjan Chakravarty for his kind comments on the paper. We would also like to thank the anonymous reviewers for their comments and suggestions.

References

Banzhaf JF (1965) Weighted voting doesn't work: a mathematical analysis. Rutgers Law Rev 19:317–343

Barua R, Chakravarty SR, Sarkar P (2009) Minimal-axiom characterizations of the coleman and banzhaf indices of voting power. Math Soc Sci 58:367–375

Brink R, Laan G (1998) Axiomatizations of the normalized banzhaf value and the shapley value. Soc Choice Welf 15(4):567–582

Chakravarty SR (2017) Analyzing multidimensional well-being: a quantitative approach. Wiley, New Jersey in press

Chakravarty SR, Lugo MA (2016) Multidimensional indicators of inequality and poverty. In: Adler MD, Fleurbaey M (eds) Oxford handbook of well-being and public policy. Oxford University Press, New York, pp 246–285

Chakravarty SR, Mitra M, Sarkar P (2015) A course on cooperative game theory. Cambridge University Press

Coleman JS (1971) Control of collectives and the power of a collectivity to act. In: Lieberma B (ed) Social choice. Gordon and Breach, New York, pp 269–298

Cowell FA (2016) Inequality and poverty measures. In: Adler MD, Fleurbaey M (eds) Oxford handbook of well-being and public policy. Oxford University Press, New York, pp 82–125

Deegan J, Packel EW (1978) A new index of power for simple n-person games. Int J Game Theory 7(2):113–123

Dubey P, Shapley LS (1979) Mathematical properties of the banzhaf power index. Math Oper Res 4(2):99–131

Einy E, Peleg B (1991) Linear measures of inequality for cooperative games. J Econ Theory 53(2):328–344

Felsenthal DS, Machover M (1998) The measurement of voting power. Edward Elgar, Cheltenham

Holler MJ (1982) Forming coalitions and measuring voting power. Polit Stud 30(2):262–271

Holler MJ, Packel EW (1983) Power, luck and the right index. J Econ 43(1):21–29

Laruelle A, Valenciano F (2001) Shapley-shubik and banzhaf indices revisited. Math Oper Res 26(1):89–104

Laruelle A, Valenciano F (2004) Inequality in voting power. Soc Choice Welf 22(2):413–431

Laruelle A, Valenciano F (2011) Voting and collective decision-making. Cambridge University Press, Cambridge

Leech D (2002a) Designing the voting system for the council of the european union. Public Choice 113:437–464

Leech D (2002b) Power in the governance of the international monetary fund. Ann Oper Res 109(1–4):375–397

Lehrer E (1998) An axiomatization of the banzhaf value. Int J Game Theory 17(2):89–99

Matsui T, Matsui Y (2000) A survey of algorithms for calculating power indices of weighted voting games. J Oper Res Soc Jpn 43:71–86

Shapley LS (1953) A value for n-person games. In: Kuhn HW, Tucker AW (eds) Contributions to the theory of Games II. Annals of mathematics studies. Princeton University Press, pp 307–317

Shapley LS, Shubik MJ (1954) A method for evaluating the distribution of power in a committee system. Am Polit Sci Rev 48:787–792

Shorrocks AF (1980) The class of additively decomposable inequality measures. Econometrica 48(3):613–625

Weber M (2016) Two-tier voting: measuring inequality and specifying the inverse power problem. Math Soc Sci 79:40–45

Part III
Polarization

Reduced Form Polarization Index: Alternative Formulations and Extensions

Bhargav Maharaj and Nachiketa Chattopadhyay

Abstract In the literature on polarization, an increasing map of the between-group expression and a decreasing map of the within-group expression with respect to a population subgroup-decomposable inequality index is known as an abbreviated or reduced form monotonic polarization index. The between-group expression is said to represent the alienation characteristic of polarization and the within-group expression negatively reflects the identification characteristic. Chakravarty and Maharaj (2011) developed an ordering for ranking alternative distributions of income based on such type of polarization indices. They also characterized several polarization indices of the said type using axioms which are easily comprehensible and intuitive. In this paper, we consider some alternative formulations of the axioms used in Chakravarty and Maharaj (2011) from a more primitive standpoint. A general nature of a polarization map is derived and the setup is extended to consider pooling of two populations. An empirical illustration of polarization indices using Indian National Sample Survey data is presented.

Keywords Reduced form polarization · Ordering · Axioms · Characterization · NSS

The authors thank Hitesh Tripathi for computational assistance.

B. Maharaj
Ramakrishna Mission Vidyamandira, Belur Math, Howrah, West Bengal, India
e-mail: ekachittananda@gmail.com

N. Chattopadhyay (✉)
Sampling and Official Statistics Unit, Indian Statistical Institute, 203, B. T. Road, Kolkata 700 108, India
e-mail: nachiketa@isical.ac.in

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*, Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_10

1 Introduction

The measurement of polarization has gained considerable significance due to its usability in analysis of the evolution of income distributions, growth, conflicts, and many other contemporary issues. Generally, polarization means a clustering (subgrouping) of incomes around some local values called “poles” in the data, where the persons belonging to the same cluster possess a feeling of identification or solidarity among them and share a feeling of alienation or enmity against persons in a different cluster or subgroups (see Esteban and Ray, 1994 for details). Hence, persons in the same cluster identify themselves with the members of the cluster in terms of income but in terms of the same income variable they have a feeling of distance from persons of the other clusters. Clearly, an increase in the identification is synonymous with increase in homogeneity, or more precisely equality within a cluster or subgroup. On the other hand, an increased alienation generally gives rise to an increased heterogeneity or inequality between subgroups. Thus, both the expressions identification and alienation are positively related to polarization. Expanding on these ideas, Esteban and Ray (1994) developed an axiomatic characterization of an index of polarization in a quasi-additive framework. Following them, Zhang and Kanbur (2001) came up with an index of polarization incorporating the concept related to the identification and alienation components. Rodriguez and Salas (2003) considered a similar approach and bi-partitioned the population in terms of the median and proposed a bi-polarization index which turned out to be the difference between the between-group and within-group components of the Donaldson and Weymark (1980) S-Gini index of inequality (see also Silber et al. 2007). Such indices, now described as a reduced form polarization index in the literature, can be used to characterize the trade-off between the alienation and identification characteristics of polarization.

Chakravarty and Maharaj (2011) makes some analytical and rigorous investigation using the idea that polarization is related to Between-Group Inequality (BI) and Within-Group Inequality (WI) in increasing and decreasing ways, respectively. They define an ordering that makes one distribution more or less polarized than another unequivocally and derive the corresponding necessary and sufficient conditions. They also characterize several polarization indices, including a generalization of the Rodriguez-Salas form and the structure of a normalized ratio form index is shown to parallel that of the Zhang–Kanbur index. The axioms are shown to be independent too.

In this paper, we extend the work of Chakravarty and Maharaj (2011) in several directions. First, we have a relook at the characterization result and suggest alternative axioms which preserve the characterization and which seems more primitive in nature. This is given in Sect. 3. We derive a necessary and sufficient condition for a mapping to be considered as a reduced form polarization index in Sect. 4. Next, we present some results in Sect. 5 on polarization index based on two independent populations. The next section illustrates several polarization indices using Indian National Sample Survey data for the period 1999–2015. Finally, Sect. 7 concludes the paper.

2 Some Definition and Existing Results

As already stated, this paper is an extension of the work of Chakravarty and Maharaj (2011). Hence, in order to recognizing this aspect and ease of understanding, we retain the notation and basic definitions used in that paper as far as practicable. For a population of size n, the vector $x = (x_1, \ldots, x_n)$ represents the distribution of income, where each x_i is assumed to be drawn from the nondegenerate interval $[\nu, \infty)$ in the positive part R^1_{++} of the real line R^1. Here, x_i stands for the income of i-th person of the population, $i = 1, \ldots, n$. For any i, $x_i \in [\nu, \infty)$, and so, $x \in D^n = [\nu, \infty)^n$, the n-fold Cartesian product of $[\nu, \infty)$. The set of all possible income distributions is $D = \cup_{n \in N} D^n$, where N is the set of natural numbers. For all $n \in N$, for all $x = (x_1, \ldots, x_n) \in D^n$, 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 nonnegative orthant of the n-dimensional Euclidean space R^n is denoted by R^n_+. An inequality index is a function $I : D \to R^1_+$.

As argued by Chakravarty and Maharaj (2011), the most relevant inequality indices in the context of polarization are the Theil mean logarithmic deviation index I_{ML} and the variance I_V. These two indices share a very useful property of subgroup decomposability (see Shorrocks 1980; Foster 1985 and Chakravarty 2009a for details) for inequality indices and are invariant to relative and absolute changes in incomes, respectively. The axiomatic analysis of polarization indices by Chakravarty and Maharaj (2011) is therefore based on these two inequality indices. Formally, the indices are given by

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

and

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

As already indicated, a polarization index is taken as a real-valued map of income distributions of arbitrary number of clusters (subgroups) of a population, the clusters being formed with respect to homogeneity of some characteristic. Formally,

Definition 1 Any continuous function $P : \Omega \to R^1$, where, $\Omega = \cup_{k \in \Gamma} (\Pi_{n_i \in \Gamma, 1 \le i \le k} D^{n_i})$ and $\Gamma = N \backslash \{1\}$ is said to be defined as a polarization index.

$P(x)$ indicates the level of polarization associated with the vector x for any $x = (x^1, \ldots, x^k) \in \Omega, k \in \Gamma$.

Next, we describe a reduced form index formally.

Definition 2 A polarization index P is called abbreviated or reduced form if for all $x = (x^1, \ldots, x^k) \in \Omega, k \in \Gamma$, $P(x)$ can be expressed as $P(x) = f(BI(x), WI(x))$, where BI (WI) is the between-group inequality (within-group inequality) with respect to I, $I = I_{ML}$ or I_V is arbitrary and the real-valued map f defined on R^2_+ is continuous.

The function f considered above is known as the characteristic function. Further, the polarization index defined above will be called a relative or an absolute index according as the inequality index used is I_{ML} or I_V.

Definition 2 is consistent with the existing literature where there are economic indicators summarizing the entire distribution in terms of two or more characteristics or parameters of the distribution. For instance, a reduced form welfare function captures social welfare as an increasing map of mean income (the efficiency parameter) and a decreasing map of inequality (the equity parameter) (For details, see Ebert 1987; Amiel and Cowell 2003 and Chakravarty 2009a, b).

Note that the concepts identification and alienation are regarded as intrinsic to the characteristic of polarization. For proper reflection of this, we further assume that the mapping is monotonic, that is, it is increasing in BI and decreasing in WI. A polarization index with this additional property is said to be feasible. Formally, we have,

Definition 3 A feasible polarization index is a reduced form polarization index $P(x) = f(BI(x), WI(x))$, where $I = I_{ML}$ or I_V, where $x = (x^1, \ldots, x^k) \in \Omega, k \in \Gamma$ are arbitrary and the real-valued mapping f defined on R_+^2 is continuous, increasing in BI and decreasing in WI.

Chakravarty and Maharaj (2011) postulate some axioms to characterize a class of feasible polarization indices. In other words, a set of necessary and sufficient conditions for identifying an index uniquely is developed. Such an exercise helps in understanding the underlying polarization index in an intuitive manner. Thus, characterization of an index focuses on the underlying implicit value judgments in an explicit way. We have, therefore,

Axiom 1 (A1): Incremental Effect on BI: For all $x = (x^1, \ldots, x^k) \in \Omega, k \in \Gamma$, and for any nonnegative 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 nondecreasing in its first argument, g is increasing, $g(0) = 0$ and $I = I_{ML}$ or I_V.

Axiom 2 (A2): Incremental Effect on WI: For all $x = (x^1, \ldots, x^k) \in \Omega, k \in \Gamma$, and for any nonnegative β, $f(BI(x), WI(x) + \beta) - f(BI(x), WI(x)) = \phi(BI(x), WI(x))\, h(\beta)$ for some continuous functions $\phi : R_+^2 \to R_+^1$ and $h : R_+^1 \to R_+^1$, where ϕ is nondecreasing in its second argument, h is increasing, $h(0) = 0$ and $I = I_{ML}$ or I_V.

Axiom 3 (A3): Normalization: For arbitrary $k \in \Gamma$, if $x = (x^1, \ldots, x^k) \in \Omega$ is of the form $x^i = c\, 1^{n_i}$, where $n_i \in \Gamma$ for all $1 \le i \le k$ and $c > 0$ is a scalar, then for any $I = I_{ML}$ or I_V, $f(BI(x), WI(x)) = 0$. Since for a perfectly equal distribution x, $BI(x) = WI(x)) = 0$, we may restate axiom A3 as $f(0, 0) = 0$.

The idea behind Axiom 1 (2) is as follows.

The increase in polarization resulting from a higher value of $BI(WI)$ by the amount $\alpha(\beta)$ is assumed to be proportional to an increasing transform of $\alpha(\beta)$. This means that the increase can be decomposed into two continuous factors, one a nonnegative map of $\alpha(\beta)$ alone and the other a nonnegative-valued map of BI and WI, which is nondecreasing in $BI(WI)$. Consequently, given differentiability of the function $f(g)$, the polarization index becomes convex (concave) in $BI(WI)$. The functions f and g are interpreted, respectively, by Chakravarty and Maharaj (2011) as alienation and identification sensitivity functions (see Chakravarty and Maharaj (2011) for details).

In order to explain the alternative formulations and extensions developed in the later sections, we quote the following theorem.

Theorem 1 Chakravarty and Maharaj (2011): *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 $P : \Omega \rightarrow R^1$ 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) $P_2(x) = \frac{c_1}{\log a} a^{BI(x)-1} - c_2 WI(x), a > 1$,

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

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

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

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

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

(viii) $P_8(x) = \frac{c_1}{\log a} a^{BI(x)-1} - \frac{c_2}{\log b} b^{WI(x)-1} + \eta a^{BI(x)-1}) b^{WI(x)-1}, \quad 0 < a < 1, b > 1, -c_2 \leq \eta \log b \leq 0$,

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

where $x = (x^1, \ldots, x^k) \in \Omega$, $k \in \Gamma$ and $I = I_{ML}$ or I_V are arbitrary.

3 Alternative Formulation of Axiom 1 and Axiom 2

The essential behavior underlying the Axioms 1 and 2 is that there is a threshold level/ tolerance limit of polarization exceeding which a society becomes turbulent. 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. Also, the tolerance limit, in general, will vary from society to society. In particular, for a highly peaceful society, it is expected to be quite low. Further, it could well be the case that the net increment in polarization will not be lower for a society characterized by a higher level of conflict/ polarization. These considerations lead to the assumption that the change in polarization is nondecreasingly related to alienation and identification over the entire domain.

As an alternative formulation to A1 which is in line with the above discussion and also has some generality, we assume,

Assumption 1 *(AS1)*: The increment in polarization resulting from an increase in BI by the amount α and keeping WI constant is nondecreasing in BI and is proportional to an increasing transform of α.

To get to the difference between A1 and AS1, let $I = I_{ML}$ or I_V and let f be the characteristic function of a feasible polarization index $P : \Omega \to R^1$. Then, AS1 says that for any nonnegative α, the polarization difference $f(BI(x) + \alpha, WI(x)) - f(BI(x), WI(x))$ is proportional to $g(\alpha)$ for some increasing function $g : R^1_+ \to R^1_+$. Notice that the constant of proportionality will invariably depend on the initial distribution $x = (x^1, \ldots, x^k) \in \Omega$, where $k \in \Gamma$. Consequently, we can formally express AS1 as

AS1:

$$f(BI(x) + \alpha, WI(x)) - f(BI(x), WI(x)) = \pi(x)\, g(\alpha) \tag{3}$$

for all $x \in \Omega$ and some continuous function $\pi : D \to R^1$.

Note that while in A1 the incremental effect on polarization was explicitly depending on the levels of alienation and identification in a society captured through BI and WI, in AS1, we generalize this by requiring the effect to be dependent on the initial distribution $x = (x^1, \ldots, x^k)$ itself. Hence, it is a weaker requirement. We now show that AS1 implies A1 for a general class of nontrivial distributions. It may be noted that the alienation and identification levels are both crucial factors determining polarization, and hence the most relevant distributions for polarization measures are those where the alienation and identification, represented by BI and WI, respectively, are nonzero. We may call such distributions a nontrivial class of distributions in the context of polarization. Formally, a nontrivial class of distributions, denoted as Υ, is given by $\Upsilon = \{x = (x^1, \ldots, x^k) \in \Omega, k \in \Gamma, BI(x) \neq 0, WI(x) \neq 0\}$

Result 1: If $x \in \Upsilon$, Assumption 1 implies Axiom 1.

Proof If there is no change in BI, there will be no change in the value of the polarization index (assuming that WI remains unaltered). Putting $\alpha = 0$ in (3), we get, $\pi(x)\, g(0) = 0$ for all $x \in \Omega$. This implies that $g(0) = 0$, for otherwise, $\pi(x) = 0$, which, when substituted in (3), will lead to a violation of increasingness of f in its first argument. So, by increasingness of g, it follows that $g(\alpha) > 0$ for all $\alpha > 0$.

Next, since (3) holds for all $\alpha \geq 0$, replacing α by $BI(x)$ (where $BI(x) \neq 0$), we get, $f(2\,BI(x), WI(x)) - f(BI(x), WI(x)) = \pi(x)\, g(BI(x))$, from which it follows that $\pi(x) = (f(2\,BI(x), WI(x)) - f(BI(x), WI(x)))/g(BI(x)) = \psi(BI(x), WI(x))$ [note that $BI(x) \neq 0$ implies $g(BI(x)) \neq 0$] for some continuous function $\psi : R^2_+ \to R^1$. Substituting in (3), we get,

$$f(BI(x) + \alpha, WI(x)) - f(BI(x), WI(x)) = \psi(BI(x), WI(x))\, g(\alpha). \tag{4}$$

But f is increasing in its first argument, and so by $g(\alpha) \geq 0$, we have $\psi(BI(x), WI(x)) \geq 0$ for all $x \in \Omega$. Finally, note that the difference $f(BI(x) + \alpha, WI(x)) - f(BI(x), WI(x))$ should be nondecreasing in BI. So, we require that ψ is nondecreasing in its first argument which completes the proof.

Again, as an alternative formulation to A2 which is similar to the discussion above except that it is concerned with the changes in WI, we assume,

Assumption 2 *(AS2)*: The increment in polarization resulting from an increase in WI by the amount β and keeping BI constant is nonincreasing in WI and is proportional to an increasing transform of β. More formally, we have

AS2:

$$f(BI(x), WI(x) + \beta) - f(BI(x), WI(x)) = \chi(x)\, h(\beta) \tag{5}$$

for all $x \in \Omega$ and some continuous function $\chi : D \to R^1$.

It is easy to observe that AS2 is weaker than A2. Using a similar logic as in Result 1, we can have

Result 2: If $x \in \Upsilon$, Assumption 2 implies Axiom 2.

Thus, the Theorem 1 in Chakravarty and Maharaj (2011) can be restated in terms of the set of weaker conditions AS1, AS2, and A3 for a very general class of nontrivial distributions.

4 General Nature of a Reduced form Polarization Map

Proposition 1 *Let* $\underline{n} = (n_1, \ldots, n_k) \in N^k$, $k \in \Gamma$ *and let* $D_0 = (\Pi_{i=1}^k D^{n_i}) \cup (\cup_{i=1}^k D^{n_i})$. *A necessary and sufficient condition for an arbitrary map* $P : D_0 \to R$ *to be a reduced form polarization index with respect to some (not predetermined) subgroup-decomposable inequality index* $I : D_0 \to R_+^1$ *is that there exists a map* $f_0 : R_+^2 \to R^1$ *which is increasing in its first argument and decreasing in the second, satisfying the following relation:*

$$P(x) = f_0(P(\underline{\lambda}^{(\underline{n})}), \sum_{i=1}^{k} w_i(\underline{n}, \underline{\lambda}) g(x^i)) \quad \textit{for all } x \in D_0, \tag{6}$$

where $\underline{\lambda} = (\lambda_1, \ldots, \lambda_k)$, $\underline{\lambda}^{(\underline{n})} = (\lambda_1 1^{n_1}, \ldots, \lambda_k 1^{n_k})$, $w_i(\underline{n}, \underline{\lambda})$ *are weights and* $g : (\cup_{i=1}^k D^{n_i}) \to R_+^1$ *vanishes for the perfectly equal distribution.*

Proof Suppose P is a reduced form polarization index on D_0. Then, there is some subgroup-decomposable inequality index $I : D_0 \to R_+^1$ and a map $f : R_+^2 \to R^1$, which is increasing in its first argument and decreasing in the second, such that

$$P(x) = f(I(\underline{\lambda}^{(\underline{n})}), \sum_{i=1}^{k} w_i(\underline{n}, \underline{\lambda}) I(x^i)) \tag{7}$$

for all $x \in D_0$. In particular, $P(\underline{\lambda}^{(\underline{n})}) = f(I(\underline{\lambda}^{(\underline{n})}, 0) = f_1(I(\underline{\lambda}^{(\underline{n})})$, where $f_1 : R^1_+ \to R^1$ is continuous and strictly increasing. This implies that $I(\underline{\lambda}^{(\underline{n})}) = f_1^{-1}(P(\underline{\lambda}^{(\underline{n})})) = f_2(P(\underline{\lambda}^{(\underline{n})}))$, say with $f_1^{-1} = f_2$. Note then that f_2 is strictly increasing. Substituting in (7), we get,

$$P(x) = f(f_2(P(\underline{\lambda}^{(\underline{n})}), \sum_{i=1}^{k} w_i(\underline{n}, \underline{\lambda})I(x^i)). \tag{8}$$

Define $f_0 : R^2_+ \to R^1$ by $f_0(s,t) = f(f_2(s), t)$ for $(s,t) \in R^2_+$. Then, f_0 is increasing in its first argument and decreasing in the second.
So, (8) yields: $P(x) = f_0(P(\underline{\lambda}^{(\underline{n})}), \sum_{i=1}^{k} w_i(\underline{n}, \underline{\lambda})g(x^i))$ with $g = I$, for all $x \in D_0$. This proves our claim.

Conversely, if P satisfies (6), then define $I : D_0 \to R^1_+$ by $I(x) = P(\underline{\lambda}^{(\underline{n})}) + \sum_{i=1}^{k} w_i(\underline{n}, \underline{\lambda})g(x^i))$ if $x \in \Pi_{i=1}^{k} D^{n_i}$ and $I(x) = g(x)$ if $x \in \cup_{i=1}^{k} D^{n_i}$.

Then, $I(\underline{\lambda}^{(\underline{n})}) = P(\underline{\lambda}^{(\underline{n})})$. Consequently, $I(x) = I(\underline{\lambda}^{(\underline{n})}) + \sum_{i=1}^{k} w_i(\underline{n}, \underline{\lambda})I(x^i))$ for all $x \in \Pi_{i=1}^{k} D^{n_i}$, which means I is subgroup decomposable on D_0.

5 Polarization of Combined Populations

Given a polarization index and two independent populations, it is, in general, not easy to predict anything regarding the effect of combination of the two populations. However, in case of a reduced form polarization index $P(x) = f(BI(x), WI(x)$ (f being the characteristic function), which is, by definition, increasing in BI and decreasing in WI, if we assume further convexity in BI and concavity in WI, then it is possible to furnish bounds(s) (upper/lower) of the polarization for the combined population in some special situations, like having two populations with equal number of groups and which are identical in respective sizes. Thus, we have $x = (x^1, \ldots, x^k)$, $y = (y^1, \ldots, y^k) \in \Pi_{i=1}^{k} D^{n_i}$ two income distributions with respective group mean vectors $(\lambda_1 1^{n_1}, \ldots, \lambda_k 1^{n_k})$ and $(\mu_1 1^{n_1}, \ldots, \mu_k 1^{n_k})$. Further, let the mean of $x(y)$ be, respectively, denoted as $\lambda(\mu)$. We want to study the polarization of the population obtained by combining x and y.

We consider the following two methods of combining the population groups in the pooled population:

(a) The group-structures remain unaffected and the new population, which is a simple union of the previous ones, comprises $2k$ subgroups.

(b) Groups with equal sizes are fused together and the new population has the same number (k) of subgroups with sizes $2n_i$.

Proposition 2 *Consider a situation as in (a) above. Let $w = (x^1, \ldots, x^k; y^1, \ldots, y^k) \in \Pi_{i=1}^{k} D^{n_i} \Pi_{i=1}^{k} D^{n_i}$ be the population obtained by combining x and y, and P be a Reduced form Polarization index with respect to an inequality index $I = I_{ML}$ or I_V. Assume that P is convex in BI and concave in WI.*

(i) If $f(BI(y), WI(x)) \geq 0$, then $P(w) \geq \frac{1}{2}min\{P(x), P(y)\}$.

(ii) If the mean vectors of x and y are identical, that is, $\lambda = \mu$, then $P(w) \geq \frac{1}{2}(P(x) + P(y))$.

Proof (i) By subgroup decomposability of I used in P, we have $BI(w) = I(\lambda_1 1^{n_1}, \ldots, \lambda_k 1^{n_k}; \mu_1 1^{n_1}, \ldots, \mu_k 1^{n_k}) = I(\lambda 1^n, \mu 1^n) + \frac{1}{2}(BI(x) + BI(y)) \geq \frac{1}{2}(BI(x) + BI(y))$. Also, $WI(w) = \sum_{i=1}^{k} \frac{n_i}{2n}(I(x^i) + I(y^i)) = \frac{1}{2}(WI(x) + WI(y))$. Suppose $BI(x) \geq BI(y)$. Then, by increasingness of f in its first argument, we have, $P(w) = f(BI(w), WI(w)) \geq f(\frac{1}{2}(BI(x) + BI(y)), \frac{1}{2}(WI(x) + WI(y)) \geq f(BI(y), \frac{1}{2}(WI(x) + WI(y))) \geq \frac{1}{2}(f(BI(y), WI(x)) + f(BI(y), WI(y)))$, the inequality being a result of concavity of f in the second argument. Hence, using the assumption that, $f(BI(y), WI(x)) \geq 0$, we have, $P(w) \geq \frac{1}{2}f(BI(y), WI(y)) = \frac{1}{2}P(y)$. Similarly, if $BI(y) \geq BI(x)$, then $P(w) \geq \frac{1}{2}P(x)$. Combining the two observations, we get $P(w) \geq \frac{1}{2}min\{P(x), P(y)\}$.

(ii) Next, if x and y have the same mean vector, we have $BI(x) = BI(y) = B$, say. Then, $BI(w) = B$, by population replication invariance of I. Hence, $P(w) = f(BI(w), WI(w)) = f(B, \frac{1}{2}(WI(x) + WI(y))) \geq \frac{1}{2}(f(B, WI(x)) + f(B, WI(y))) = \frac{1}{2}(P(x) + P(y))$, by concavity of f in the second argument.

Proposition 3 *Consider a situation as in (b) above. Let P be a Reduced form Polarization index with respect to I_V. Further, assume that (i) $I(x^i) = I(y^i)$ for each i and (ii) $\sum_{i=1}^{k} n_i \lambda_i \mu_i \leq n\lambda\mu$. If $z = (z^1, \ldots, z^k) \in \Pi_{i=1}^{k} D^{2n_i}$, where $z^i = (x^i, y^i)$ for $1 \leq i \leq k$, be obtained by combining x with y and clubbing together the groups of same size, then $P(z) \leq \frac{1}{2}(P(x) + P(y))$.*

Proof Observe that mean of z is $\frac{1}{2}(\lambda + \mu)$ and $WI(x) = WI(y) = W$, say. Also, $I(z^i) = I(x^i, y^i) \geq \frac{1}{2n_i}(n_i I(x^i) + n_i I(y^i)) = \frac{1}{2}(I(x^i) + I(y^i))$ so that $WI(z) = \sum_{i=1}^{k} \frac{n_i}{n} I(z^i) \geq \frac{1}{2}(WI(x) + WI(y)) = W$. Further, $BI(x) = \sum_{i=1}^{k} n_i(\lambda_i - \lambda)^2 = V_1$ and $BI(y) = \sum_{i=1}^{k} n_i(\mu_i - \mu)^2 = V_2$, say.

Then, $BI(z) = \sum_{i=1}^{k} 2n_i(\frac{\lambda_i+\mu_i}{2} - \frac{\lambda+\mu}{2})^2 = \frac{1}{2}\sum_{i=1}^{k} n_i(\lambda_i - \lambda)^2 + \frac{1}{2}\sum_{i=1}^{k} n_i(\mu_i - \mu)^2 + \frac{1}{2}\sum_{i=1}^{k} n_i(\lambda_i - \lambda)(\mu_i - \mu) = \frac{1}{2}(V_1 + V_2) + \frac{1}{2}(\sum_{i=1}^{k} n_i\lambda_i\mu_i - n\lambda\mu) \leq \frac{1}{2}(V_1 + V_2)$, by assumption (ii). Consequently, $P(z) = f(BI(z), WI(z)) \leq f(\frac{1}{2}(V_1 + V_2), W)$, by monotonicity of f in both the arguments. So, convexity of f in the first argument yields.

$P(z) \leq \frac{1}{2}(f(V_1, W) + f(V_2, W)) = \frac{1}{2}(P(x) + P(y))$.

6 An Empirical Illustration

We use National Sample Survey data of Consumer Expenditure to analyze changes in polarization as captured through Monthly Per Capita Expenditure (MPCE) across states and union territories of India for the time period 1993–94 to 2014–15.

The specific rounds of survey considered are 50, 55, 61, 66, 68, and 72 with respective years as 1993–94, 1999–2000, 2004–2005, 2009–10, 2011–12, and 2014–15. The computations are carried out separately for the rural and urban sectors. We have considered the grouping of the population in the following three classes: the lowest 30%, the middle 40%, and the top 30%, where the observations are arranged in the increasing order. The choice of the lower 30% is in line with the current poverty headcount in India and this class may be considered as a representation of the "poor". The other two classes analogously may be termed as "middle class" and the "rich", respectively. Obviously, the polarization index will depend on the grouping. Also, the unit of observation is the households and MPCE is the representative monthly expenditure taking the household size into consideration. It may be noted that MPCE is taken as an approximation of Monthly Per Capita Income (MPCI) in absence of reliable income data. Obviously, the inequality in the savings pattern and its correlations with consumption expenditure have definitive influence on inequality of income and hence on the polarization indices based on such inequalities.

Since the polarization indices considered here are based on BI and WI which have definitive interpretations, we first present these values in Tables 1 and 2, in terms of the inequality index I_{ML} for rural and urban sectors. It is observed (Table 1) that except for some few states and territories, the between-group inequalities are higher for urban than rural implying a higher level of alienation of the three classes in urban areas. However, with a lesser degree, the same is true for the within-group inequalities (Table 2) with the urban sectors of different states and territories displaying a higher level of within-group inequality. That is, the three classes in the rural sector have a higher level of identification than the corresponding classes in the urban sector. This further implies for polarization, things may turn out differently which we explore in Table 3, where, the simple difference form polarization index is shown for rural and urban sectors. We find that here the situation more or less resembles the scenario observed in the case of between-group inequalities. All the indices are positive implying a higher level of alienation than identification across the states and the sectors.

However, in all the three cases, we get quite different states or territories which show higher values for the rural sector across various rounds. To look into this more precisely, we have a table (Table 4), where we have presented the polarization ranks for both the sectors. The ranks are sorted by the level of the rural polarization of the most current round (72) in the year 2014–15, from lower to higher value of polarization. In terms of these ranks, we find that there has been considerable changes in the polarization levels of rural and urban Indian states across time. The general feature remains the same if we use the variance inequality measure, which is, therefore, not reported here.

Table 1 Between-Group Inequality (by I_{ML}) of MPCE: Rural and Urban

NSS round	50		55		61		66		68		72	
Year	1993–94		1999–2000		2004–05		2009–10		2011–12		2014–15	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
All India	0.110	0.168	0.097	0.165	0.111	0.190	0.119	0.215	0.116	0.197	0.117	0.157
Andhra Pradesh	0.108	0.145	0.082	0.146	0.102	0.206	0.126	0.187	0.087	0.144	0.103	0.157
Arunachal Pradesh	0.143	0.114	0.144	0.105	0.095	0.086	0.134	0.130	0.173	0.151	0.197	0.130
Assam	0.043	0.120	0.055	0.128	0.049	0.144	0.066	0.188	0.062	0.175	0.069	0.125
Bihar	0.070	0.141	0.061	0.158	0.052	0.178	0.065	0.167	0.058	0.128	0.067	0.113
Goa	0.119	0.124	0.117	0.113	0.127	0.160	0.077	0.075	0.110	0.120	0.074	0.065
Gujrat	0.078	0.112	0.079	0.120	0.107	0.147	0.099	0.139	0.089	0.104	0.092	0.115
Haryana	0.134	0.122	0.083	0.114	0.158	0.164	0.106	0.183	0.089	0.208	0.110	0.173
Himachal Pradesh	0.119	0.297	0.095	0.121	0.122	0.096	0.127	0.164	0.113	0.156	0.092	0.138
Jammu & Kashmir	0.079	0.118	0.052	0.073	0.070	0.084	0.065	0.131	0.085	0.131	0.095	0.151
Karnataka	0.096	0.153	0.083	0.154	0.098	0.205	0.075	0.218	0.093	0.261	0.097	0.155
Kerala	0.120	0.178	0.117	0.148	0.164	0.215	0.175	0.240	0.188	0.224	0.093	0.111
Madhya Pradesh	0.102	0.162	0.083	0.150	0.089	0.208	0.103	0.241	0.104	0.224	0.119	0.169
Maharastra	0.126	0.178	0.100	0.167	0.112	0.190	0.081	0.212	0.101	0.189	0.136	0.167
Manipur	0.032	0.036	0.049	0.062	0.030	0.039	0.035	0.058	0.054	0.056	0.075	0.070
Meghalaya	0.081	0.091	0.034	0.072	0.035	0.124	0.046	0.084	0.045	0.072	0.061	0.106
Mizoram	0.046	0.048	0.062	0.073	0.052	0.076	0.054	0.077	0.077	0.086	0.127	0.132
Nagaland	0.036	0.067	0.054	0.076	0.059	0.077	0.045	0.073	0.050	0.072	0.058	0.068
Orissa	0.085	0.147	0.083	0.130	0.095	0.184	0.090	0.236	0.079	0.166	0.099	0.139
Punjab	0.103	0.105	0.082	0.109	0.105	0.163	0.124	0.175	0.104	0.144	0.098	0.177
Rajasthan	0.095	0.129	0.067	0.130	0.067	0.167	0.063	0.169	0.074	0.173	0.120	0.145
Sikkim	0.083	0.091	0.083	0.090	0.109	0.127	0.140	0.064	0.062	0.068	0.082	0.098
Tamil Nadu	0.121	0.190	0.103	0.191	0.108	0.196	0.100	0.163	0.108	0.156	0.109	0.121
Tripura	0.079	0.119	0.049	0.124	0.062	0.147	0.054	0.121	0.063	0.121	0.075	0.114
Uttar Pradesh	0.111	0.157	0.089	0.164	0.087	0.168	0.079	0.347	0.091	0.247	0.102	0.143
West Bengal	0.084	0.145	0.070	0.155	0.089	0.181	0.066	0.202	0.076	0.205	0.101	0.155
Andaman & Nicober	0.110	0.198	0.093	0.088	0.128	0.166	0.156	0.126	0.137	0.173	0.096	0.139
Chandigarh	0.086	0.290	0.099	0.158	0.067	0.196	0.099	0.205	0.087	0.211	0.135	0.122
Dadra & Nagar Haveli	0.112	0.146	0.132	0.096	0.265	0.127	0.069	0.081	0.184	0.165	0.090	0.097
Daman & Diu	0.103	0.067	0.076	0.086	0.186	0.158	0.125	0.110	0.035	0.084	0.051	0.173
Delhi	0.113	0.178	0.045	0.151	0.110	0.153	0.064	0.174	0.083	0.179	0.104	0.130
Lakshadweep	0.104	0.114	0.086	0.119	0.085	0.089	0.222	0.137	0.129	0.149	0.080	0.052
Pondicheri	0.114	0.124	0.100	0.123	0.158	0.165	0.078	0.306	0.095	0.094	0.105	0.097
Uttaranchal					0.084	0.146	0.539	0.146	0.091	0.176	0.086	0.087
Jharkhand					0.061	0.154	0.063	0.176	0.069	0.174	0.079	0.190
Chhattisgarh					0.090	0.211	0.075	0.167	0.080	0.237	0.113	0.145
Telangana											0.091	0.164
No of States with Higher Rural BI		3		4		3		7		3		5

Table 2 Within-Group Inequality (by I_{ML}) of MPCE: Rural and Urban

NSS round	50		55		61		66		68		72	
Year	1993–94		1999–2000		2004–05		2009–10		2011–12		2014–15	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
All India	0.039	0.051	0.031	0.047	0.039	0.049	0.045	0.059	0.038	0.053	0.037	0.048
Andhra Pradesh	0.042	0.037	0.032	0.034	0.039	0.066	0.046	0.041	0.026	0.041	0.039	0.047
Arunachal Pradesh	0.037	0.031	0.051	0.061	0.025	0.024	0.032	0.035	0.036	0.039	0.071	0.038
Assam	0.012	0.025	0.017	0.027	0.015	0.030	0.016	0.046	0.021	0.040	0.026	0.034
Bihar	0.023	0.034	0.020	0.037	0.014	0.036	0.018	0.036	0.016	0.029	0.022	0.035
Goa	0.032	0.037	0.026	0.034	0.027	0.045	0.022	0.027	0.030	0.045	0.026	0.032
Gujrat	0.025	0.035	0.024	0.034	0.035	0.042	0.041	0.038	0.029	0.033	0.028	0.046
Haryana	0.045	0.027	0.023	0.031	0.067	0.045	0.026	0.043	0.024	0.048	0.032	0.054
Himachal Pradesh	0.040	0.122	0.026	0.036	0.038	0.030	0.041	0.054	0.033	0.048	0.025	0.028
Jammu & Kashmir	0.033	0.026	0.019	0.021	0.027	0.025	0.025	0.048	0.032	0.030	0.028	0.039
Karnataka	0.030	0.038	0.027	0.041	0.047	0.051	0.024	0.068	0.033	0.062	0.033	0.045
Kerala	0.042	0.065	0.038	0.035	0.054	0.059	0.059	0.074	0.071	0.058	0.027	0.038
Madhya Pradesh	0.039	0.060	0.026	0.048	0.030	0.052	0.031	0.062	0.036	0.059	0.034	0.052
Maharastra	0.046	0.047	0.033	0.047	0.037	0.048	0.023	0.058	0.044	0.053	0.042	0.049
Manipur	0.010	0.010	0.014	0.039	0.011	0.011	0.014	0.019	0.017	0.016	0.040	0.032
Meghalaya	0.050	0.019	0.013	0.016	0.013	0.033	0.016	0.019	0.014	0.019	0.018	0.035
Mizoram	0.015	0.014	0.023	0.023	0.018	0.017	0.014	0.020	0.022	0.021	0.037	0.036
Nagaland	0.010	0.018	0.014	0.019	0.020	0.023	0.016	0.013	0.015	0.014	0.014	0.017
Orissa	0.029	0.033	0.025	0.035	0.030	0.046	0.031	0.059	0.023	0.038	0.036	0.036
Punjab	0.037	0.029	0.025	0.029	0.031	0.048	0.047	0.042	0.030	0.049	0.033	0.046
Rajasthan	0.035	0.032	0.022	0.036	0.020	0.049	0.020	0.047	0.026	0.043	0.037	0.038
Sikkim	0.023	0.023	0.025	0.025	0.033	0.031	0.036	0.024	0.016	0.016	0.026	0.025
Tamil Nadu	0.047	0.100	0.038	0.086	0.040	0.048	0.035	0.042	0.032	0.039	0.037	0.035
Tripura	0.022	0.029	0.014	0.024	0.020	0.038	0.017	0.027	0.018	0.025	0.028	0.039
Uttar Pradesh	0.032	0.040	0.027	0.040	0.027	0.047	0.025	0.109	0.032	0.061	0.033	0.048
West Bengal	0.039	0.037	0.026	0.048	0.035	0.046	0.024	0.049	0.023	0.056	0.034	0.049
Andaman & Nicober	0.023	0.074	0.022	0.029	0.046	0.057	0.106	0.044	0.033	0.048	0.025	0.028
Chandigarh	0.018	0.072	0.027	0.041	0.018	0.035	0.049	0.060	0.026	0.052	0.022	0.047
Dadra & Nagar Haveli	0.030	0.034	0.045	0.024	0.054	0.049	0.018	0.023	0.029	0.026	0.023	0.036
Daman & Diu	0.027	0.022	0.018	0.016	0.036	0.042	0.019	0.021	0.006	0.015	0.016	0.023
Delhi	0.026	0.061	0.014	0.037	0.028	0.032	0.014	0.042	0.016	0.051	0.013	0.037
Lakshadweep	0.017	0.036	0.021	0.031	0.022	0.034	0.067	0.030	0.024	0.040	0.025	0.023
Pondicheri	0.031	0.037	0.021	0.036	0.049	0.042	0.026	0.163	0.037	0.030	0.023	0.032
Uttaranchal					0.027	0.039	0.062	0.032	0.028	0.047	0.025	0.029
Jharkhand					0.018	0.036	0.015	0.041	0.024	0.042	0.028	0.041
Chhattisgarh					0.034	0.055	0.024	0.048	0.024	0.065	0.034	0.047
Telangana											0.027	0.052
No of States with Higher Rural WI		9		3		8		7		7		6

Table 3 Difference Form Polarization Index (by I_{ML}) of MPCE: Rural and Urban

NSS round	50		55		61		66		68		72	
Year	1993–94		1999–2000		2004–05		2009–10		2011–12		2014–15	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
All India	0.071	0.117	0.067	0.118	0.072	0.141	0.074	0.156	0.078	0.144	0.080	0.109
Andhra Pradesh	0.066	0.108	0.050	0.112	0.063	0.139	0.080	0.147	0.061	0.103	0.064	0.110
Arunachal Pradesh	0.107	0.082	0.093	0.044	0.069	0.062	0.102	0.094	0.136	0.113	0.126	0.091
Assam	0.030	0.094	0.038	0.101	0.034	0.115	0.051	0.142	0.041	0.135	0.044	0.091
Bihar	0.046	0.107	0.041	0.121	0.039	0.142	0.048	0.131	0.042	0.100	0.045	0.079
Goa	0.087	0.087	0.091	0.079	0.099	0.115	0.055	0.047	0.081	0.075	0.048	0.034
Gujrat	0.052	0.077	0.056	0.087	0.072	0.105	0.058	0.100	0.060	0.070	0.064	0.069
Haryana	0.089	0.095	0.060	0.083	0.091	0.119	0.080	0.140	0.066	0.160	0.078	0.119
Himachal Pradesh	0.079	0.175	0.070	0.085	0.084	0.066	0.086	0.110	0.080	0.108	0.066	0.110
Jammu & Kashmir	0.046	0.091	0.032	0.052	0.043	0.059	0.040	0.083	0.053	0.101	0.067	0.112
Karnataka	0.066	0.116	0.056	0.113	0.052	0.155	0.052	0.150	0.060	0.199	0.064	0.109
Kerala	0.079	0.114	0.079	0.113	0.110	0.156	0.116	0.166	0.117	0.166	0.065	0.073
Madhya Pradesh	0.063	0.102	0.056	0.102	0.059	0.156	0.072	0.179	0.069	0.165	0.086	0.118
Maharastra	0.080	0.131	0.067	0.120	0.075	0.142	0.058	0.154	0.057	0.136	0.093	0.118
Manipur	0.022	0.026	0.035	0.023	0.019	0.028	0.021	0.039	0.037	0.039	0.036	0.038
Meghalaya	0.031	0.072	0.022	0.056	0.022	0.090	0.029	0.065	0.031	0.053	0.043	0.071
Mizoram	0.031	0.034	0.038	0.050	0.034	0.059	0.040	0.056	0.055	0.065	0.090	0.096
Nagaland	0.026	0.049	0.039	0.058	0.039	0.055	0.029	0.059	0.035	0.058	0.044	0.051
Orissa	0.056	0.115	0.059	0.095	0.065	0.138	0.060	0.177	0.056	0.128	0.063	0.103
Punjab	0.065	0.077	0.057	0.081	0.074	0.115	0.077	0.133	0.073	0.095	0.064	0.130
Rajasthan	0.061	0.097	0.045	0.094	0.047	0.119	0.043	0.122	0.048	0.130	0.083	0.107
Sikkim	0.060	0.068	0.059	0.065	0.077	0.096	0.104	0.040	0.046	0.053	0.055	0.073
Tamil Nadu	0.074	0.090	0.065	0.105	0.068	0.147	0.065	0.122	0.076	0.117	0.072	0.086
Tripura	0.057	0.090	0.035	0.100	0.042	0.110	0.037	0.093	0.045	0.096	0.047	0.075
Uttar Pradesh	0.079	0.116	0.061	0.124	0.060	0.120	0.054	0.238	0.059	0.186	0.069	0.095
West Bengal	0.046	0.108	0.044	0.107	0.053	0.135	0.042	0.153	0.053	0.149	0.067	0.106
Andaman & Nicober	0.087	0.124	0.071	0.059	0.083	0.109	0.050	0.082	0.103	0.125	0.071	0.111
Chandigarh	0.068	0.218	0.072	0.118	0.049	0.160	0.050	0.145	0.061	0.159	0.113	0.074
Dadra & Nagar Haveli	0.081	0.112	0.087	0.072	0.211	0.079	0.051	0.058	0.155	0.139	0.067	0.061
Daman & Diu	0.076	0.046	0.059	0.070	0.150	0.116	0.106	0.089	0.029	0.069	0.035	0.151
Delhi	0.087	0.117	0.031	0.114	0.082	0.121	0.050	0.132	0.067	0.128	0.091	0.094
Lakshadweep	0.086	0.078	0.065	0.088	0.063	0.055	0.155	0.106	0.105	0.109	0.055	0.029
Pondicheri	0.083	0.088	0.079	0.086	0.110	0.123	0.052	0.143	0.058	0.064	0.082	0.065
Uttaranchal					0.057	0.107	0.478	0.114	0.063	0.129	0.062	0.059
Jharkhand					0.042	0.118	0.048	0.135	0.045	0.131	0.051	0.149
Chhattisgarh					0.055	0.156	0.051	0.118	0.056	0.173	0.079	0.099
Telangana											0.064	0.112
No of States with Higher Rural Polarization		3		5		5		2		4		7

Table 4 Polarization Ranks (by I_{ML}) of MPCE: Rural and Urban

NSS round	50	55	61	66	68	72		50	55	61	66	68	72
Year	1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15		1993–94	1999–2000	2004–05	2009–10	2011–12	2014–15
			RURAL							URBAN			
Daman & Diu	20	19	34	32	1	1		3	9	18	10	7	36
Manipur	1	5	1	1	4	2		1	1	1	1	1	3
Meghalaya	5	1	2	3	2	3		6	5	9	7	3	9
Assam	3	6	4	14	5	4		16	21	15	25	25	16
Nagaland	2	8	6	2	3	5		4	6	2	6	4	4
Bihar	7	9	5	10	6	6		20	31	29	20	12	14
Tripura	11	4	7	4	7	7		13	20	14	11	11	13
Goa	29	31	31	20	30	8		11	11	17	3	9	2
Jharkhand			8	9	8	9				19	23	24	35
Lakshadweep	27	22	19	34	32	10		9	17	3	14	16	1
Sikkim	12	17	26	31	9	11		5	8	10	2	2	11
Uttaranchal			15	35	23	12				12	16	22	5
Orissa	10	18	20	23	15	13		25	19	26	33	20	22
Andhra Pradesh	16	12	18	27	21	14		21	25	27	28	14	26
Punjab	15	16	24	26	27	15		7	12	16	22	10	34
Karnataka	17	13	12	18	19	16		26	26	31	29	35	25
Telangana						17							29
Gujrat	9	14	23	22	20	18		8	16	11	13	8	8
Kerala	21	28	32	33	33	19		24	27	33	32	32	10
Himachal Pradesh	23	25	29	29	29	20		31	14	7	15	15	27
Jammu & Kashmir	8	3	9	6	12	21		15	4	5	9	13	30
Dadra & Nagar Haveli	25	30	35	16	35	22		23	10	8	5	27	6
West Bengal	6	10	13	7	11	23		22	24	25	30	28	23
Uttar Pradesh	22	21	17	19	18	24		27	32	22	35	34	19
Andaman & Nicober	30	26	28	13	31	25		29	7	13	8	19	28
Tamil Nadu	19	23	21	24	28	26		14	23	30	18	18	15
Haryana	31	20	30	28	24	27		17	13	20	24	30	33
Chhattisgarh			14	15	14	28				32	17	33	21
Pondicheri	26	29	33	17	17	29		12	15	24	26	5	7
Rajasthan	13	11	10	8	10	30		18	18	21	19	23	24
Madhya Pradesh	14	15	16	25	26	31		19	22	34	34	31	31
Mizoram	4	7	3	5	13	32		2	3	4	4	6	20
Delhi	28	2	27	11	25	33		28	28	23	21	21	18
Maharastra	24	24	25	21	16	34		30	30	28	31	26	32
Chandigarh	18	27	11	12	22	35		32	29	35	27	29	12
Arunachal Pradesh	32	32	22	30	34	36		10	2	6	12	17	17

7 Concluding Remarks

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. Alternative sets of independent axioms have been proposed for polarization indices characterized by Chakravarty and Maharaj (2011). The approach is extended to get a necessary and sufficient condition for polarization mapping and to analyze polarization for combined population. An empirical illustration using Indian National Sample Survey data over several rounds is presented.

References

Amiel Y, Cowell FA (2003) Inequality, welfare and monotinicity. Res Econ Inequal 9:35–46
Chakravarty SR (2009a) Inequality, polarization and poverty: advances in distributional analysis. Springer, New York
Chakravarty SR (2009b) Equity and efficiency as components of a social welfare function. Int J Econ Theory 5:181–199
Chakravarty SR, Maharaj B (2011) Subgroup decomposable inequality indices and reduced-form indices of polarization. Keio Econ Stud 47:57–83
Donaldson D, Weymark JA (1980) A single parameter generalization of the Gini indices of inequality. J Econ Theory 22:67–86
Ebert U (1987) Size and distribution of income as determinants of social welfare. J Econ Theory 41:25–33
Esteban JM, Ray D (1994) On the measurement of polarization. Econometrica 62:819–851
Foster JE (1985) Inequality measurement in fair allocation by Young HP (ed) American Mathematical Society: Providence
Rodriguez J, Salas R (2003) Extended bipolarization and inequality measures. Res Econ Inequal 9:69–83
Shorrocks AF (1980) The class of additively decomposable inequality measures. Econometrica 48:613–625
Silber J, Deutsch J, Hanoka M (2007) On the link between the concepts of kurtosis and bipolarization. Econ Bull 4:1–6
Zhang X, Kanbur R (2001) What differences do polarization measures make? Appl China J Dev Stud 37:85–98

On the Measurement of "Grayness" of Cities

Sripad Motiram and Vamsi Vakulabharanam

Abstract We consider a situation where individuals belonging to multiple groups inhabit a space that can be divided into smaller distinguishable units, a feature characterizing many cities in the world. When data on an economic attribute (in our case, income) is available, we conceptualize a phenomenon that we refer to as "Grayness"—a combination of spatial integration based upon group-identity and income. Grayness is high when cities display a high degree of spatial coexistence in terms of both identity and income. We lay down some desirable properties of a measure of Grayness and develop a simple and intuitive index that satisfies them. We provide an illustration by using data from the Indian city of Hyderabad, and selected American cities.

Keywords Segregation · Inequality · Group-based disparities · Cities · Grayness

JEL Codes: D61 · D63

We are delighted to contribute to a Festschrift for Prof. Satya Chakravarty. We have learnt a lot from his work and benefitted from his comments and suggestions. We would like to thank the Indian Council of Social Science Research (ICSSR) for a generous grant (No. RESPRO/31/ICSSR/2013-14/RPS) that supported this research. We would also like to thank an anonymous referee whose comments on a previous version helped us improve the paper. For discussions on the US Census, we thank Mike Carr and Andrew Perumal.

S. Motiram
University of Massachusetts Boston, Boston, MA, USA
e-mail: sripad.motiram@umb.edu

V. Vakulabharanam
University of Massachusetts Amherst, Amherst, MA, USA
e-mail: vamsi@econs.umass.edu

I. Dasgupta and M. Mitra (eds.), *Deprivation, Inequality and Polarization*,
Economic Studies in Inequality, Social Exclusion and Well-Being,
https://doi.org/10.1007/978-981-13-7944-4_11

1 Introduction

It is being increasingly acknowledged that the world is predominantly urban and urbanization will continue into the foreseeable future (Davis 2007; UN 2015). This paper is therefore concerned with cities, in particular, group-based ("horizontal") disparities within cities.[1] Many cities in the world are characterized by severe disparities among groups. The particular facet of group-based disparity that we are interested in can be illustrated by the following examples. Blacks and Whites live in American cities which have distinct neighborhoods or administrative divisions. Similarly, different caste groups live in Indian cities which can be divided into several wards. Across the world, different ethnic groups inhabit urban areas, which are characterized by some form of spatial division. What is common to all these examples is the presence of identity groups in an urban context with spatial heterogeneity. Apart from this, we may have information available on an economic attribute of individuals and groups, e.g., incomes or wages. Cities could display various degrees of spatial coexistence (or lack of it) in terms of this attribute, e.g., the rich and poor in a city could live together in the same neighborhood or live completely apart. Combining spatial integration on income and non-income dimensions reveals certain interesting features and dynamics of cities. For example, irrespective of their ethnic identities, if the rich are separating themselves into "enclaves" or "gated-communities", whereas the poor are being pushed into "slums", then this can be understood as a process where spatial integration is low in terms of income, but not so in terms of ethnic identity. Such a phenomenon is being witnessed in India after economic reforms were initiated in the early 1990s, and such "neoliberal" cities can be distinguished from "mixed" cities that prevailed earlier.[2] Essentially, we are interested in a phenomenon that is a combination (or intersection) of spatial integration based upon identity and income (or some other economic attributes). We refer to this phenomenon as "Grayness". Grayness is high when cities display a high degree of spatial coexistence of both identity group and economic groups. When Grayness is negligibly small, cities become "stark".

Our focus on space is inspired by the recent recognition in social sciences and humanities of the importance of explicit considerations of space. Scholars have argued that such a "spatial turn" and the idea of "spatial justice" have conferred both theoretical and practical advantages, e.g., the discourse on the "right to the city" which has assumed political salience today (Soja 2009; Harvey 2013). One of the important ideas in this literature is that space and society are intricately linked ("sociospatial dialectics"). Space is not an inert given, and individuals and groups shape it, even as it influences them. Our attempt is to bring these ideas to bear on the literature in welfare economics. While the above body of knowledge has seen

[1] Following Stewart (2002), it has become customary to distinguish between interpersonal ("vertical") and group-based ("horizontal") disparities.

[2] On poverty and rising inequality in urban India since economic reforms, see Vakulabharanam and Motiram (2012).

contributions mostly from noneconomists, economists have also recognized that spatial location confers both advantages and disadvantages (see, e.g., Reardon 2017; Chetty et al. 2016) and spatial patterns influence outcomes like crime (e.g., see the Lewis Mumford lecture of Sen 2017).

In light of the above, we consider an abstract city that is comprised of multiple spatial units. The population of the city, and in each spatial unit is divided into several groups. Apart from his/her group-identity, we have information on an economic attribute (income) of an individual. We lay down the desirable properties of a "Grayness Index" and develop a simple and intuitive index which satisfies these properties. To the best of our knowledge, our paper is the first one that focuses explicitly on space and explores the interaction of spatial integration of different kinds. In doing so, it differs in spirit from the literature on spatial (residential) segregation, and adds to it. The consideration of spatial integration on different dimensions simultaneously captures an important social phenomenon that cannot be reduced to considering spatial integration on these dimensions separately and then bringing them together. The interaction among these dimensions will be lost. We will discuss this further in the next section. The paper draws upon existing literature by conceptualizing the Grayness index as a function of two components—identity group and income—these are in turn indices of spatial integration (and inverse of spatial segregation). To keep the exposition brief, we do not present a survey of the literature on segregation, but refer interested readers to Chakravarty (2009) and Silber et al. (2009). The most commonly used index of segregation is the Duncan–Duncan dissimilarity index, which is useful when there are only two groups. In our case, there can be multiple groups, so we draw upon Reardon and Firebaugh (2002) Since income is a continuous variable, we draw upon Kim and Jargowsky (2009), who demonstrate how the Gini Index can be used for segregation for both continuous and binary variables.[3]

Urbanization in recent times has been driven by growth of cities in developing countries (Davis 2007). Projecting into the future, the United Nations estimates that several of the largest cities in the world will be located in the global South.[4] We therefore implement our index on an Indian city viz. Hyderabad. We show that the Grayness index of Hyderabad is high, e.g., as compared to selected American cities. We hypothesize that this maybe an important characteristic of Indian cities vis-a-vis cities in the developed world.

The remaining portion of the paper is divided into two sections. The next section develops the index and presents an illustration from the Indian city of Hyderabad and some American cities. The third section concludes with a discussion.

---

[3]Also, see Reardon (2009) and Hutchins (2009). Reardon (2009) develops indices of segregation with multiple groups, when one of the dimensions (e.g., occupation, education) can be ordered. Hutchins (2009) develops an "augmented index" of gender-based occupational segregation where occupations can be ranked in terms of a scalar variable (e.g., average wage).

[4]The top ten urban agglomerations in 2030 are expected to be: Tokyo, Delhi, Shanghai, Beijing, Mumbai, Mexico City, Cairo, São Paulo, Osaka, and New York-Newark (UN 2015).

2 Grayness: Theory and Illustration

2.1 An Index to Measure Grayness

Consider a city which is divided into N (> 1) spatial units, which we index by m and n. $G(> 1)$ groups live in the city, and we index these groups by g. Let the shares of group g in spatial unit m, and in the city be denoted by p_g^m and p_g^c, respectively. If p_g^m is less (more) than p_g^c, then group g is considered to be under-represented (over-represented) in the spatial unit m. Let $P^m = (p_1^m, p_2^m, \ldots, p_G^m)$ denote the vector of group shares in spatial unit m. Let $P = (P^1, P^2, \ldots, P^N)$ denote a vector that captures the spatial distribution of group shares for the entire city. Let $P^c = (p_1^c, p_2^c, \ldots, p_G^c)$ denote the vector of city shares. Let Y_g^m denote the income distribution of group $g = 1, \ldots, G$ in spatial unit $m = 1, \ldots, N$. We could consider the income distribution for either individuals or households. Let $I = ((Y_1^1, Y_2^1, \ldots, Y_G^1), \ldots, (Y_1^N, Y_2^N, \ldots, Y_G^N))$ denote the vector of income distributions for the entire city. Let $S = (T, s^1, \ldots, s^N)$ where T (> 0) denotes the total population of the city, and s^m, $m = 1, \ldots, N$ denotes the fraction of the population of the city that resides in the spatial unit m. We conceptualize the Grayness Index (GI) as a function, $GI : (P, P^c, S, I) \rightarrow [0, 1]$ that combines spatial integration on identity groups and income.

Formally, we can think of GI as a function (f) of a "Group Component (GC)" and an "Income Component (IC)" where these two components measure spatial integration on identity groups and income, respectively. We will discuss the properties of GC and IC later. It suffices here to point out that since they are measures of spatial integration, they lie in [0, 1]. We propose that GI satisfies the following properties/axioms in terms of its components:

(A1) Minimum Grayness

GI is at its minimum value of zero if and only if GC and IC are both at their minimum values of zero, i.e., there is complete lack of spatial integration in terms of both group-identity and income.

(A2) Maximum Grayness

GI is at its maximum value of one if and only if GC and IC are both at their maximum values of one, i.e., there is complete spatial integration in terms of both group-identity and income.

(A3) Monotonicity: Grayness as an Increasing Function of Spatial Integration

GI increases (decreases) if spatial integration increases (decreases) either among identity groups or on the income dimension, i.e., $\frac{\partial GI}{\partial GC} > 0$ and $\frac{\partial GI}{\partial IC} > 0$.

The above axioms are straightforward. (A3) considers the impact on Grayness of spatial integration on one dimension. How do we consider the impact on Grayness of spatial integration on multiple dimensions and how does this compare with the situation depicted in (A3)? A simple example can be used to explore this question.

Let us imagine three cases: (1) $GC = 0.8, IC = 0$, (2) $GC = 0, IC = 0.8$, and (3) $GC = 0.4, IC = 0.4$. In Cases (1) and (2), there is complete lack of spatial integration on one dimension and high spatial integration on the other, whereas in Case (3), there is modest spatial integration on both dimensions. Starting from a situation where there is complete lack of spatial integration on both dimensions ($GC = 0$ and $IC = 0$), we can imagine three different processes: A, B, and C, that can result in Cases (1), (2), and (3), respectively. A is a process that increases cohesion among identity groups while preserving income-based/class-based exclusions and prejudices. B is a process similar to A, except that the roles of identity groups and income are interchanged. C is a process that promotes cohesion on both identity group and income dimensions, albeit in a modest manner. We believe that a city can be considered to be more spatially integrated in Case (3) as compared to Cases (1) and (2). In other words, process C contributes more to spatial integration and Grayness compared to processes A and B. Essentially, for a given "total spatial integration" ($GC + IC$), we consider a city to be more spatially integrated if the "mix" of spatial integration on multiple dimensions is better. This idea is analogous to the preference for variety in international trade under monopolistic competition (see, e.g., Grossman 1992). The axiom below captures this idea more formally.

(A4) Preference for Mix of Spatial Integration on Multiple Dimensions

For a given total spatial integration ($GC + IC$), consider a process that increases spatial integration on the dimension that has lower integration (say by $\delta > 0$) and decreases spatial integration on the other dimension by an equal amount (i.e., by δ). Such a process will result in a better mix of spatial integration on the two dimensions, and thereby increases GI.

Note the similarity with the ideas of "mean-preserving spread" and "Dalton–Pigou transfer principle" in the measurement of risk and inequality, respectively, (see, e.g., Chakravarty (2009) for a discussion). In a way, we are applying these ideas to GC and IC. Finally, we would like to consider the interaction of the two components explicitly. It is reasonable to argue that the phenomenon of interest to us should depend upon the interaction of the two components, and not just upon "pure" spatial integration among either identity groups or on income. This is formalized in the axiom below:

(A5) Interaction: *GI* Depends upon Interaction of *GC* and *IC*, i.e., $\frac{\partial^2 GI}{\partial GC \partial IC} \neq 0$

We can consider several functional forms for f, although some simple ones like the arithmetic mean or geometric mean of GC and IC are ruled out because they violate one or more of the above axioms. Interestingly, a "Mean-Variance" form satisfies the above axioms, and we propose it as follows:

$$GI = \alpha \frac{(GC + IC)}{2} - \beta [\frac{(GC^2 + IC^2)}{2} - (\frac{GC + IC}{2})^2]. \tag{1}$$

Note that the first term $\frac{(GC+IC)}{2}$ is the mean spatial integration (i.e., average of GC and IC) and the second term $[\frac{(GC^2+IC^2)}{2} - (\frac{GC+IC}{2})^2]$ is the variance between the two components of spatial integration (GC and IC). As we show in the proposition below, when $\alpha = 1$ and $0 < \beta < 1$, GI satisfies the axioms $(A1) - (A5)$. An interesting result concerns the decomposition properties of GI. We can show that

$$GI = f(GC, IC) = \alpha\frac{GC}{2} - \beta[\frac{GC^2}{2} - (\frac{GC}{2})^2] + \alpha\frac{IC}{2} - \beta[\frac{IC^2}{2} - (\frac{IC}{2})^2] + \beta\frac{GC * IC}{2} \quad (2)$$

$$= f(GC, 0) + f(0, IC) + \beta\frac{GC * IC}{2} \quad (3)$$

$f(GC, 0)$ and $f(0, IC)$ represent pure spatial integration, in terms of identity groups and income, respectively. Hence, we can see that GI can be decomposed into three parts, representing pure spatial integration in terms of identity groups, pure spatial integration in terms of income, and interaction between spatial integration on identity groups and income. The parameter β captures the strength of interaction between the two components (we will see this more clearly below) and the impact of interaction will vanish if $\beta = 0$.

Proposition 1 *If* $\alpha = 1$ *and* $0 < \beta < 1$, *then* GI *satisfies* $(A1) - (A5)$.

Proof It is easy to establish that GI satisfies $(A1)$. If $GC = IC = 1$, then $GI = \alpha$. Hence, if $\alpha = 1$, then GI satisfies $(A2)$. $\frac{\partial GI}{\partial GC} = \frac{1}{2} - \beta\frac{(GC - IC)}{2}$. Since the maximum value that $(GC - IC)$ can take is 1, the condition $\beta < 1$ ensures that $\frac{\partial GI}{\partial GC} > 0$. On similar lines, we can show that it ensures that $\frac{\partial GI}{\partial IC} > 0$. As long as $\beta > 0$, the process referred to in (A4) reduces the variance between GC and IC, and thereby increases GI. Hence, if $\beta > 0$, GI satisfies $(A4)$. From Eq. (4), we can see that $\frac{\partial^2 GI}{\partial GC \partial IC} = \frac{\beta}{2} \neq 0$. Hence, GI satisfies $(A5)$. ■

Note that, in general, (i.e., given that $GC \in [0, 1]$ and $IC \in [0, 1]$), β needs to be less than one. But, for particular values of GC and IC, we can work with higher values of β. For example, for modest values of GC and IC (less than 0.5), we can use values of β in excess of 1, but less than 2. Having characterized GI, we will now move to its components, GC and IC. Let $Gini_a$ denote the Gini Index of average incomes of spatial units and $Gini_t$ denote the Gini Index for the income distribution in the city. Spatial integration on income can be considered as the inverse of income-based spatial segregation. Income is a continuous variable, and we can draw upon the literature on segregation for continuous variables. In particular, Kim and Jargowsky (2009) demonstrate that the ratio $Gini_a/Gini_t$ can be considered as an index of segregation which lies in [0, 1]. Following this, we can characterize IC as

$$IC = 1 - \frac{Gini_a}{Gini_t}. \quad (4)$$

Note that IC lies in [0, 1]. It takes the maximum value of 1 when the city is completely spatially integrated in terms of income, i.e., all the spatial units have identical average incomes. It takes the minimum value of zero when the city is completely spatially segregated (or atomized) in terms of income, i.e., each spatial unit comprises just one individual or household.

Since we have characterized IC using the Gini Index, it would be advantageous to consider a Gini-based characterization for GC too. As in the case of IC, we can consider spatial integration among identity groups as the opposite of group-based spatial segregation. Since the number of groups can be greater than two, we draw upon the literature on multigroup segregation indices. Reardon and Firebaugh (2002) present a comprehensive overview of this issue, including the various notions of segregation. They demonstrate how an index of segregation based upon the Gini Index can be constructed by comparing the group proportions across all organizational (in our case, spatial) units, and for all groups. Following them, we characterize GC as

$$GC = 1 - \frac{\Sigma_{g=1}^{G} p_g^c \Sigma_{m=1}^{N} \Sigma_{n=1}^{N} s^m s^n |p_g^m / p_g^c - p_g^n / p_g^c|}{2\Sigma_{g=1}^{G} p_g^c (1 - p_g^c)}. \tag{5}$$

As in the case of IC, GC lies in [0, 1]. It takes the maximum value of 1 if the city is completely spatially integrated in terms of the identity group, i.e., for each group, its share in every spatial unit is the same as its city share. It takes the minimum value of zero if the city is completely segregated in terms of the identity group, i.e., each spatial unit comprises just one group.

The above formulation of GI attaches equal weightage to the two different kinds of spatial integration, i.e., to GC and IC. It is easy to see that this is not necessary, and we could privilege one kind of spatial integration over another. Let w_g and w_i denote the weights on GC and IC, respectively, where $(w_g + w_i) = 1$. In the analysis above, we have considered $w_g = w_i = 0.5$. A general formulation would be as follows:

$$GI = (w_g GC + w_i IC) - \beta[(w_g GC^2 + w_i IC^2) - (w_g GC + w_i IC)^2]. \tag{6}$$

Before moving on to the illustration of GI, it is worthwhile to point out that the index can be extended to spatial integration on more than two dimensions. This can be done by simply considering a mean-variance form for the multiple components. For example, if there are three dimensions (say race, religion, and income) for which the components are GC_1, GC_2, and IC, then the index can be expressed as

$$GI = \frac{GC_1 + GC_2 + IC}{3} - \beta[\frac{GC_1^2 + GC_2^2 + IC^2}{3} - (\frac{GC_1 + GC_2 + IC}{3})^2]. \tag{7}$$

2.2 An Illustration of Grayness

We will now illustrate the above analysis by considering the cases of Hyderabad city in India and some American cities. The data for Hyderabad comes from a spatially representative household survey conducted by us during 2015–17. The survey is described in detail in Motiram and Vakulabharanam (2017), and we briefly discuss the relevant features here. The survey focuses on the completely urban part of Hyderabad city (viz., the district of Hyderabad). The methodology is a multistage stratified one which draws upon the latest (2011) decennial Indian Census. The survey comprises 1000 households which are spread across 100 Enumeration Blocks (EBs)—10 households in each EB. To ensure spatial representation, the 16 subdistricts of Hyderabad are treated as strata and the EBs are spread across them. For the computation of the Grayness Index, we consider census wards as the spatial units. The Census ward is a larger area compared to the EB, but is smaller than the subdistrict. We divide the population of Hyderabad into two groups based upon their caste status: Dalits (Scheduled Castes and Tribes) and Non-Dalits. We could use household per-capita income (total monthly household income/household size) or household income. The former is defined at the individual level, whereas the latter is defined at the household level. Since the literature on US income inequality that draws upon census data has largely used the household as the unit of analysis, we focus on the former. The ranking of groups in terms of household income is as expected: Dalits—Rs. 20,613.62, and Non-Dalits—Rs. 23,400.71.[5] This difference would be much starker in other Indian cities since Hyderabad has a substantially larger proportion of Muslims, who are mostly included under Non-Dalits, and whose economic status in urban India is quite low.

We present estimates for two American cities: Chicago (Chicago–Naperville–Elgin IL-IN-WI Metropolitan Statistical Area) and New York (New York–Newark–Jersey City-NY-NJ-PA Metropolitan Statistical Area).[6] We use census tracts as the spatial units and use the American Consumer Survey 2016, 5-year estimates (i.e., 2012–16) from the Factfinder site of US Census Bureau.[7] Analogous to the analysis from Hyderabad, we consider two racial groups: Black or African American alone and White alone, Black or African American alone, and Others. On the average, the household incomes of Blacks are considerably lower than those of Whites, e.g., for New York city the average annual household incomes in 2016 inflation-adjusted dollars are \$24,103 and \$45,952 for Blacks and Whites, respectively (Table S1902, U.S Census Bureau).

In Table 1, we present the estimates of GI and its components for various values of β.[8] As we discussed above, given the particular estimates for GC and IC, we can experiment with values of β that are greater than 1. The estimates for GC are slightly higher for Hyderabad as compared to the American cities. This reflects the fact that Dalit and Non-Dalit spatial integration in Hyderabad is much better than race-based

[5] The corresponding figures for per-capita income are also on expected lines: Dalits—Rs. 4,858.43, and Non-Dalits—Rs. 5,534.04.

[6] We have also conducted analysis for several other American cities, and they turn out to have a smaller Grayness index than Hyderabad, i.e., our main finding is not altered if we include some more American cities.

[7] https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml.

[8] For the computation of GC, we use "seg", the module in Stata that computes various segregation indices with multiple groups.

Table 1 Estimates of grayness index and its components

β	City	D	GC	IC	Mean	Variance	GI
0.95	Hyderabad	0.5327	0.3154	0.3854	0.3504	0.0012	0.3492
	Chicago (B-W)	0.7256	0.1249	0.4443	0.2846	0.0255	0.2604
	Chicago (B-O)	0.7105	0.1395	0.4443	0.2919	0.0232	0.2698
	New York (B-W)	0.7059	0.1446	0.4727	0.3087	0.0269	0.2831
	New York (B-O)	0.6349	0.2023	0.4727	0.3375	0.0183	0.3201
1.0	Hyderabad	0.5327	0.3154	0.3854	0.3504	0,0012	0.3492
	Chicago (B-W)	0.7256	0.1249	0.4443	0.2846	0.0255	0.2591
	Chicago (B-O)	0.7105	0.1395	0.4443	0.2919	0.0232	0.2687
	New York (B-W)	0.7059	0.1446	0.4727	0.3087	0.0269	0.2817
	New York (B-O)	0.6349	0.2023	0.4727	0.3375	0.0183	0.3192
1.5	Hyderabad	0.5327	0.3154	0.3854	0.3504	0.0012	0.3485
	Chicago (B-W)	0.7256	0.1249	0.4443	0.2846	0.0255	0.2464
	Chicago (B-O)	0.7105	0.1395	0.4443	0.2919	0.0232	0.2571
	New York (B-W)	0.7059	0.1446	0.4727	0.3087	0.0269	0.2683
	New York (B-O)	0.6349	0.2023	0.4727	0.3375	0.0183	0.3101

Note Authors’ computations using household survey data for Hyderabad and American Community Survey (ACS) 2016, 5-year estimates. For IC, estimates of Gini are from table B19083, US Census Bureau. D: Duncan–Duncan Dissimilarity index, B-W: Black-White, B-O: Black-Others

spatial integration in American cities. To shed further light on this, we also present the Duncan–Duncan dissimilarity index, which confirms this observation. The estimates of IC for Hyderabad are slightly lower compared to the same for American cities. However, this is more than compensated by higher GC in Hyderabad and the lower variance component. Consequently, the Grayness Index GI for Hyderabad is higher. Note that as the value of β rises, the importance of the variance component increases and the value of the Grayness Index falls.

3 Discussion and Conclusions

Recent scholarship in several social sciences has emphasized the centrality of space and the need to incorporate spatial considerations explicitly. In the above analysis, we have taken this idea seriously and considered a feature that characterizes many cities. We examine the existence of identity groups in cities that are internally spatially heterogeneous by considering a phenomenon (“Grayness”) that is a combination of spatial integration based upon identity and income. We develop an index of “Grayness” that satisfies several desirable properties. We illustrate this index by applying it to the Indian city of Hyderabad and some American cities.

The spatial units or identity groups in a city could be ordered on some attribute (e.g., average income, educational opportunities) and this ordering can be explicitly incorporated in the analysis. While we have not addressed this, it can be taken up in future research. Also, we have focused on measurement issues only, but it would be quite fascinating to examine the interrelationship between Grayness and outcomes and the mechanisms through which these relationships work.

References

Chakravarty S (2009) Inequality, polarization and poverty: advances in distributional analysis. Springer, Heidelberg

Chetty R, Hendren N, Katz L (2016) The effects of exposure to better neighborhoods on children: new evidence from the moving to opportunity project. Am Econ Rev 106:855–902

Davis M (2007) The planer of slums. Verso, London

Grossman G (ed) (1992) Imperfect Competition and International Trade. MIT Press, Cambridge USA

Harvey D (2013) Rebel cities: from the right to the city to the urban revolution. Verso, London

Hutchins R (2009) Occupational segregation with economic disadvantage: an investigation of decomposable indexes. In: Silber J et al (eds) Occupational and residential segregation. Emerald Insight Publishing, Bingley, pp 99–120

Kim J, Jargowsky P (2009) The GINI coefficient and segregation on a continuous variable. In: Silber J et al (eds) Occupational and residential segregation. Emerald Insight Publishing, Bingley, pp 57–70

Motiram S, Vakulabharanam V (2017) Class, caste, and production of city space in India. University of Massachusetts Boston and Amherst (mimeo)

Reardon S (2009) Measures of ordinal segregation. In: Silber J et al (eds) Occupational and residential segregation. Emerald Insight Publishing, Bingley, pp 129–155

Reardon S (2017) Educational opportunity in early and middle childhood: variation by place and age. Working Paper No. 17-12. Center for Education Policy Analysis, Stanford University

Reardon S, Firebaugh G (2002) Measures of multigroup segregation. Sociol Methodol 32:33–67

Sen A.K (2017) The urbanity of Calcutta. Fourth annual Lewis Mumford lecture. City College of New York

Silber J, Fluckiger Y, Reardon S (eds) (2009) Occupational and residential segregation. Emerald Insight Publishing, Bingley

Soja E (2009) The city and spatial justice. Spatial justice

Stewart F (2002) Horizontal inequality: a neglected dimension of development. WIDER Annual Lecture, UNU-WIDER, Helsinki

UN (2015) World urbanization prospects: the 2014 revision. United Nations, New York

Vakulabharanam V, Motiram S (2012) Understanding Poverty and Inequality in Urban India since Reforms: Bringing Quantitative and Qualitative Approaches Together. Economic and Political Weekly XLVII: 44–52