

Frank Cowell

Inequality and poverty measures

Book section

Original citation:

Cowell, Frank (2016) Inequality and poverty measures. In: Adler, Matthew D. and Fleurbaey, Marc, (eds.) The Oxford Handbook of Well-Being and Public Policy. Oxford University Press, Oxford, UK. ISBN 9780199325818 (In Press)

© 2014 [Oxford University Press](#)

This version available at: <http://eprints.lse.ac.uk/66022/>

Available in LSE Research Online: April 2016

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's submitted version of the book section. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

Inequality and Poverty Measures

by

Frank A. Cowell

STICERD

London School of Economics

Houghton Street

London, WC2A 2AE, UK

f.cowell@lse.ac.uk

Revised December 2014

Prepared for the *Oxford Handbook of Well-Being And Public Policy*, edited
by Matthew D. Adler and Marc Fleurbaey.

Contents

1	Introduction	3
2	A framework	4
2.1	Income distributions	5
2.2	The axiomatic approach	7
2.3	Measurement tools	9
3	Inequality measurement: principles	10
3.1	Inequality measures and their properties	11
3.1.1	Scale-invariant inequality measures	18
3.1.2	Translation-invariant inequality measures	20
3.1.3	Inequality and income levels	22
3.2	Decomposition	23
3.2.1	Arbitrary partition	24
3.2.2	Restricted partition	26
3.3	Alternative approach: reference points	29
4	Inequality: welfare and values	34
4.1	Social welfare and inequality	35
4.2	Welfare-based inequality measures	38
4.3	Welfare and individual values	47
5	Poverty measurement: principles	49

5.1	Intuitive approaches	51
5.2	Axiomatic approach	53
5.3	Other approaches	62
6	Inequality and poverty: ranking	63
6.1	A meta-analysis	63
6.2	First and second order dominance	66
6.3	Significance for welfare, inequality and poverty	71
7	Conclusion	74

1 Introduction

Inequality and poverty measurement share a common ancestry. Many economists and other social scientists are generally aware of this but, if pressed, are not too sure about what the exact dynastic connections are. The aim of this chapter is to explain a little of the “who is related to whom” and to distinguish some of the main family traits.

However, we will be thinking about close family only. We will not be going into some of the interesting further relations with deprivation, affluence, polarisation concentration and so on; the chapter will not attempt to provide a comprehensive survey;¹ nor will we be going beyond income-related inequality and poverty (although the definition of “income” can be stretched a bit). Furthermore the chapter confines itself to the problems of measurement only. So we will not be taking a detour to examine the interesting recent evidence on inequality and poverty trends in particular countries nor will we be looking at the many important questions that arise in the practical implementation of the measures – the statistical issues of estimation and inference.

As can be seen by scanning the headings below we shall devote rather more time to inequality than to poverty. There are two good reasons for this: many of the important abstract concepts were worked out first in the inequality context and then extended to the formal analysis of poverty; furthermore

¹For a broad overview of inequality measurement see Cowell (2000, 2011), Lambert (2001) and for surveys of poverty measurement see Zheng (1997, 2000).

many of these abstract concepts are, in my opinion, more appropriate in the inequality context. We begin with some explanation of terms.

2 A framework

For most of the time we will consider a society consisting of a fixed population. So, we assume that there are n persons, each with an identifiable and observable income. Each person i 's income is represented by a real number x_i that tells us all that we need to know about the individual in the inequality-measurement or poverty-measurement problem. To be more precise we will suppose that for any i , $x_i \in X$ where X is a subset of the real line. The exact assumption that we make about X is a reflection of our assumption about the nature of "income": for example, it is common to assume that X consists only of non-negative numbers (appropriate if "income" actually means consumption expenditure; but for some inequality problems (such as the inequality of net worth) negative values make perfectly good sense.

The number x_i may or may not completely represent person i 's well-being. Whether or not it does so depends on the use to which we want to put the measurement tools. Although a comprehensive definition of income would be required for welfare-based interpretations of inequality (see section 4) much of the formal analysis applies equally to broadly based and narrowly based definitions of income. The principal requirement is that the equalisand,

“income”, however it is defined, be representable on a cardinal scale.²

2.1 Income distributions

In our framework an “income distribution” is simply an n -vector of people’s incomes $\mathbf{x} := (x_1, x_2, \dots, x_n)$. It is useful to express mean income in shorthand form as

$$\mu(\mathbf{x}) = \mu(x_1, x_2, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1)$$

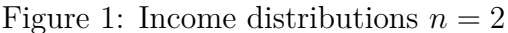
In analysing inequality it is sometimes useful to merge populations and then to analyse the income distribution in the merged society. To handle this we use the following notation: if \mathbf{x} and \mathbf{x}' are vectors with n and n' components respectively, then $(\mathbf{x}, \mathbf{x}')$ is an $(n + n')$ -vector formed by stringing the components of \mathbf{x} and \mathbf{x}' together. Given a precise specification of “income”, incorporated in the assumption about X the set of all possible income distributions is given by X^n ; but sometimes we want to refer only to the subset of this consisting of distributions with a specified mean μ : we will write this as $X^n(\mu)$. We will use $\mathbf{1}$ to denote the vector $(1, 1, \dots, 1)$: so, for example, $\mu\mathbf{1}$ is a perfectly equal distribution with mean μ .

Clearly we could talk about inequality within a very simple economy consisting of just two persons. But this would not be very interesting. In Figure 1 the horizontal axis measures Irene’s income and the vertical axis

²Inequality measurement when the equalisand is purely ordinal rather than cardinal presents a different class of problems – see Cowell and Flachaire (2014).

Janet's income; the 45° ray through the origin represents distributions that are perfectly equal, the point marked \mathbf{x} shows the supposed current income distribution (with Irene richer than Janet) and the solid line through \mathbf{x} (at right-angles to the equality ray) represents all the distributions that could be reached from \mathbf{x} by simple income transfers. So the point $\mu\mathbf{1}$ – where $\mu = \mu(\mathbf{x})$ is mean income – represents the distribution that would emerge if incomes were to be equalised between Irene and Janet. If Irene and Janet's incomes were swapped to give the income distribution \mathbf{x}' presumably inequality would remain unchanged; if the income distribution \mathbf{x} were to be moved closer to $\mu\mathbf{1}$ then presumably inequality would be reduced. But we do not need a formal inequality measure to tell us this, nor do we need a formal statement of distributional principles in order to make sensible comparisons of distribution.

By contrast Figure 2 shows that distributional comparisons could be much more interesting in a 3-person world. The axes measure the incomes of Irene, Janet and Karen; as before \mathbf{x} shows the supposed current income distribution (with Irene richer than Janet who is richer than Karen); the triangular shaped area (formally known as a simplex) is $X^3(\mu)$, representing all the possible distributions with the same mean as \mathbf{x} ; and, again as before, the point $\mu\mathbf{1}$ in this triangle represents the distribution that would emerge if incomes were to be equalised among all three persons (the ray through $\mathbf{0}$ and $\mu\mathbf{1}$ represents all such perfectly equal distributions. Straightaway we can see that this three-person case represents us with a much richer set of alternatives for



2.2 The axiomatic approach

7

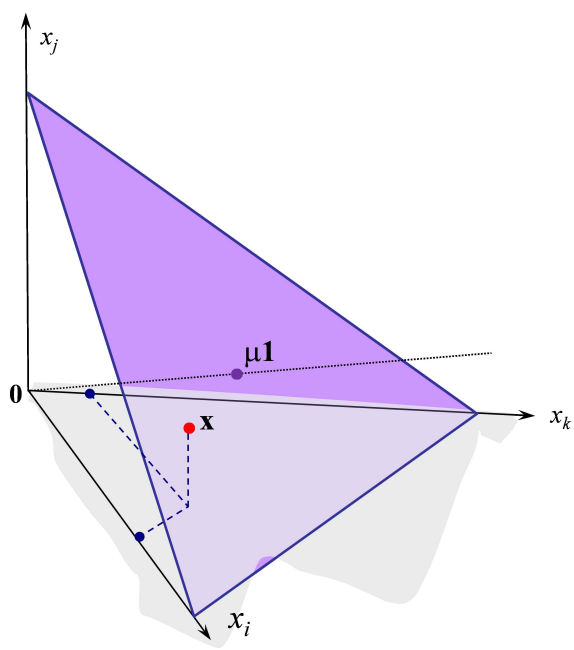


Figure 2: Income distributions $n = 3$

method, which can be briefly explained as follows:

- set out and defend on *a priori* grounds a minimal set of principles or axioms to which inequality comparisons or poverty comparisons ought to conform;
- follow mathematical logic to characterise the class of measures that exactly satisfy these principles;
- if necessary add further axioms to narrow down the class of measures.

Clearly the method is only as good as the reasonableness of the individual axioms that we choose to introduce. “Reasonable” here could mean that the axiom accords with economic intuition, or that it accords with people’s views on distributional comparisons, or that it simplifies an otherwise intractable mathematical problem, or that it is helpful in empirical implementation.

2.3 Measurement tools

There is a variety of types of measurement tool that might be considered for the analysis of inequality and poverty. We will focus on three types:

- 1 *A distributional ordering.* Here we want to be able to say something like this: “When we compare distributions \mathbf{x} and \mathbf{x}' either (i) inequality in \mathbf{x} is higher than in \mathbf{x}' , or (ii) inequality in \mathbf{x}' is higher than in \mathbf{x} , or (iii) inequality in \mathbf{x} and \mathbf{x}' is the same.” To do this it is sufficient to have an inequality index that is defined up to a monotonic transformation. The

same thing applies also to distributional orderings in terms of poverty – just substitute the word “poverty” for “inequality” in the foregoing.

2 *A distributional ranking.*³ This is similar to type 1 but we allow for one further possibility in addition to (i)-(iii) above: “(iv) \mathbf{x} and \mathbf{x}' cannot be unambiguously compared in terms of inequality (or poverty).”

3 *A cardinal index.* Inequality or poverty can be represented by a cardinal index that is defined up to a change of scale. Comparing this with type 1 we can see that it is similar but, whereas in type 1 the index could be subjected to any order-preserving transformation and still be regarded as operationally the same index, here we allow only scale changes.⁴

3 Inequality measurement: principles

Let us begin with a standard measurement tool, an inequality measure: this is a function I that assigns a numerical value to any distribution in X^n . We would write the inequality of income distribution \mathbf{x} as

$$I(\mathbf{x}) = I(x_1, x_2, \dots, x_n). \quad (2)$$

³Sometimes also known as a quasi-ordering (Weymark 2015).

⁴Example. It is arguable that the variance – equation (6) below – could be used as a satisfactory inequality measure; if so then the standard deviation would do just as well if we are merely concerned with a distributional ordering; but if we want a cardinal index then the standard deviation will give different results from the variance. Note also that, strictly speaking we should allow for scale and origin changes for a cardinal index. However, if we assume that the minimum value of the inequality or poverty index is set at zero this automatically fixes the origin.

For the moment we are only concerned with making simple inequality *comparisons* of any two members of X^n , so a particular function I could be replaced by any increasing transform (for example $\log(I), I^2, \exp(I), \dots$) and leave the comparisons unchanged; we are talking about a “type 1” measurement tool in the language of section 2.3.

3.1 Inequality measures and their properties

Following the outline in section 2.2 our first step is to set out a set of axioms (assumptions) each of which can be defended on its own merits. In fact a small group of core assumptions – Axioms 1-4 below – is sufficient to characterise completely a widely used class of inequality measures. Some of these core axioms are also relevant to poverty measurement. We will discuss their reasonableness as we go along.

Axiom 1 (*Anonymity*). Suppose \mathbf{x}' is formed from \mathbf{x} by a permutation of the components; then $I(\mathbf{x}) = I(\mathbf{x}')$.

Axiom 2 (*Population Principle*). $I(\mathbf{x}) = I(\mathbf{x}, \mathbf{x}, \dots, \mathbf{x})$.

Axiom 3 (*Transfer Principle*). Suppose \mathbf{x}' is formed from \mathbf{x} thus: $x'_i = x_i + \delta$, $x'_j = x_j - \delta$, $x'_k = x_k$ for all $k \neq i, j$ where $\delta > 0$ is such that $x'_i, x'_j \in X$. If $x_i \geq x_j$ then $I(\mathbf{x}) < I(\mathbf{x}')$.

Axioms 1 and 2 are easily expressed in plain language: for any income distribution, relabelling the persons in the distribution or simply replicating

the distribution leaves inequality unaffected. However, we should not let these two apparently innocuous assumptions pass without some comment. The anonymity principle seems fine as long as we really believe that income captures all that is important about a person's current economic status (we have sorted out any difficulties concerning differences in needs, for example) and that history is unimportant (if it were known that Irene had had been heavily disadvantaged and Janet excessively privileged in the past then some might not think that swapping the current incomes x_i and x_j is neutral in terms of inequality). Axiom 3 says that a mean-preserving redistribution from anyone to someone who is richer must increase inequality – and *vice versa*, of course (Dalton 1920). Figure 3 shows a closeup (taken from Figure 2) of the income distribution \mathbf{x} and the other distributions attainable from \mathbf{x} through simple transfers.

To find your way around this triangle notice first that the point in the centre represents complete equality, where everyone receives the mean μ . The corner labelled [Irene] represents complete inequality, where Irene receives all the income; any point on the side of the triangle *opposite* this corner represents a distribution where Irene gets nothing and total income is divided between Janet and Karen alone. So starting from any such point on this line and heading directly towards the [Irene] corner we find distributions where Irene gets more and more income (and the ratio between Janet's income and

Karen's stays constant).⁵ So point \mathbf{x} , being closest to the [Irene] corner, must represent a distribution where Irene gets the most income; and, because \mathbf{x} is closer to the [Janet] corner than to [Karen], Janet must get more than Karen in this distribution.

Let us label the point representing distribution \mathbf{x} as (i): then simply swapping incomes between pairs of people yields points (ii)-(vi) and Axiom 1 means that all the points (i),..., (vi) are equally unequal. Axiom 3 (the Transfer Principle) implies that inequality must be lower anywhere in the interior of the line connecting adjacent pairs of these points (this would involve a partial equalisation between two of the people, leaving the third person's income unchanged); from further reasoning on the Transfer Principle we can see that the inequality associated with any point in the interior of the hexagon associated with \mathbf{x} is less than $I(\mathbf{x})$ (Champernowne and Cowell 1998, Chapter 5).

This interpretation immediately reveals an apparent difficulty, illustrated in Figure 4. Suppose we try to compare distributions \mathbf{x} and \mathbf{x}' in terms of inequality: it is clear that distribution \mathbf{x}' does not lie inside the \mathbf{x} -hexagon (copied from Figure 3); nor does \mathbf{x} does not lie inside the \mathbf{x}' -hexagon. The Transfer Principle is insufficient to produce a clear inequality ranking of all the points in $X^n(\mu)$. There are two ways forward from here:

⁵Of course the same interpretation applies to the [Janet] or [Karen] corner and the line opposite.

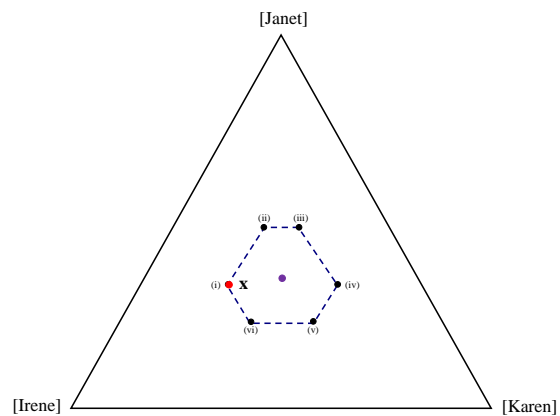


Figure 3: Income distributions $n = 3$ (close-up)

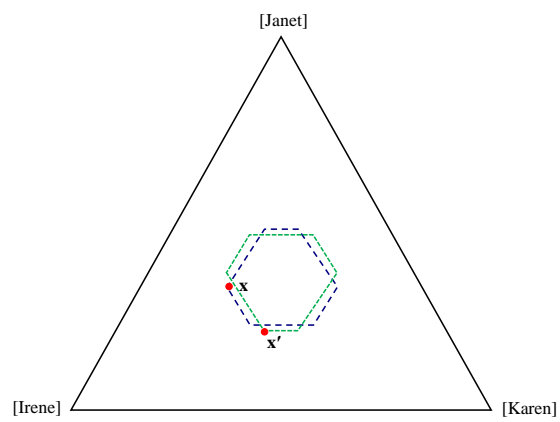


Figure 4: Distributions \mathbf{x} and \mathbf{x}' cannot be ranked by the Transfer Principle

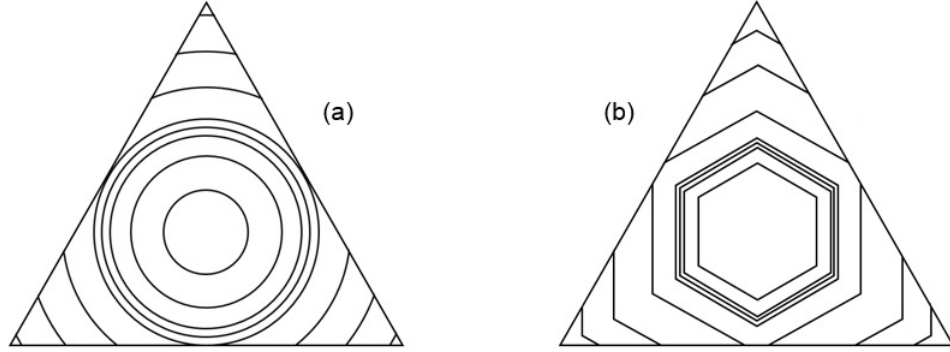


Figure 5: Two contour maps

1. One could force a resolution of the ambiguity by a contour map on the diagram. Two possible systems of contours are illustrated in Figure 5, each of which satisfies Axioms 1-3: contours further away from the centre correspond to higher inequality levels. Ideally such a contour map should be supported by further axioms with a clear rationale in terms of economic intuition.
2. One could live with the ambiguity and obtain a richer insight into inequality comparisons. We give up on finding a type-1 inequality measurement tool and consider type-2 comparisons – this is pursued in section 6.

Let us follow up the first of these routes now by introducing two further apparently reasonable axioms.

Axiom 4 (*Decomposability*). Let $\mathbf{x}, \mathbf{x}' \in X^n(\mu)$ and $\mathbf{x}'' \in X^{n'}(\mu)$. If $I(\mathbf{x}) \geq I(\mathbf{x}')$ then $I(\mathbf{x}, \mathbf{x}'') \geq I(\mathbf{x}', \mathbf{x}'')$.

Axiom 5 (*Scale Invariance*). Let $\mathbf{x}, \mathbf{x}' \in X^n(\mu)$ and $\lambda > 0$. If $I(\mathbf{x}) \geq I(\mathbf{x}')$ then $I(\lambda\mathbf{x}) \geq I(\lambda\mathbf{x}')$.

Decomposability (Axiom 4) means this: take two distributions \mathbf{x} and \mathbf{x}' with the same population size n and the same mean μ ; merge each of them with any third distribution \mathbf{x}'' that has the same mean μ (but not necessarily the same population size); then, if \mathbf{x} is more unequal than \mathbf{x}' , \mathbf{x} -merged-with- \mathbf{x}'' is also more unequal than \mathbf{x}' -merged-with- \mathbf{x}'' .

Scale Invariance (Axiom 5), also known as “homotheticity”, is a property that applies to the shape of the contour maps at different levels of income. If \mathbf{x} and \mathbf{x}' register the same amount of inequality in $X^n(\mu)$ then the “scaled-up” or “scaled-down” versions of these distributions, where each person’s income is rescaled by the same factor λ , are also regarded as equally unequal in $X^n(\lambda\mu)$. The property is illustrated in Figure 6 where the contours at the higher income level can be regarded as a “blow-up” of the inequality contours at the lower income level. Notice that Axiom 5 does *not* say that inequality remains constant under a scale change of all incomes; of course it may make sense to replace this axiom with the stronger requirement of *scale independence*, namely that $I(\lambda\mathbf{x}) = I(\mathbf{x})$; but this is not necessary for the basic results. Also, using scale invariance rather than scale independence leaves open an interesting possibility, discussed in subsection 3.1.3 below.

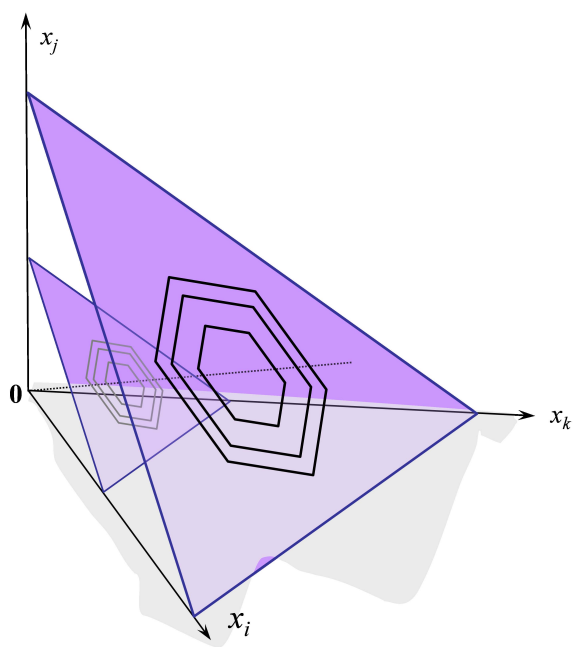


Figure 6: Scale invariance

3.1.1 Scale-invariant inequality measures

Equipped with these two extra axioms we have the following result (Bourguignon 1979; Cowell 1980; Shorrocks 1980, 1984; Russell 1985; Zagier 1982):

Theorem 1 *Axioms 1-5 imply that a continuous inequality index must be ordinally equivalent⁶ to*

$$I_{\text{GE}}(\mathbf{x}) = \frac{1}{\alpha[\alpha - 1]} \left[\frac{1}{n} \sum_{i=1}^n \left[\frac{x_i}{\mu(\mathbf{x})} \right]^\alpha - 1 \right], \quad (3)$$

where α is a sensitivity parameter that can be assigned any real value.⁷

From Theorem 1 emerges not a single inequality measure but a broad class or family of measures commonly known as the *generalised entropy* (GE) measures (Cowell 1977, Cowell and Kuga 1981a, 1981b; Toyoda 1975); the class forms an important example of so-called relative inequality indices (Blackorby and Donaldson 1978). Any one member of the family in (3), picked out by a particular value of α will do the job of ordering all the distributions in $X^n(\mu)$; so the important question is, how to pick α ? Figure 7 illustrates the iso-inequality contours for various values of α . In the extreme case $\alpha = -\infty$

⁶Continuity of the index is a technical requirement which ensures that measured inequality does not “jump” by a substantial amount when there is an infinitesimal change in the income distribution: for an interesting exception to this see note 26 on poverty indices. Two measures I and I' are said to be ordinally equivalent over $X^n(\mu)$ if there is a function f , increasing in its first argument, such that $I' = f(I; n, \mu)$.

⁷Using L'Hôpital's rule one can show that the limiting form for the case $\alpha = 0$ is given by $I(\mathbf{x}) = -\sum_{i=1}^n \log\left(\frac{x_i}{\mu(\mathbf{x})}\right)$, the so-called Mean Logarithmic Deviation and the limiting form for the case $\alpha = 1$ is given by $I(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\mu(\mathbf{x})} \log\left(\frac{x_i}{\mu(\mathbf{x})}\right) - 1$, the Theil index (in fact both indices were developed by Theil 1967).

we can see that redistribution between the richest and the next richest leaves inequality unaltered (each straight line here represents the distributions that keep the poorest person's income constant); only redistribution involving the poorest affects measured inequality. By contrast in the other extreme case $\alpha = +\infty$ we can see that redistribution between the poorest and the next poorest leaves inequality unaltered (each straight line here represents the distributions that keep the richest person's income constant) – only redistribution involving the richest affects measured inequality; by similar reasoning we can see that for positive values of α the GE index is more sensitive to changes at the top of distribution and for negative values of α the GE index is more sensitive to changes at the bottom of distribution. The cases $\alpha = +2$ (contours illustrated on the left-hand side of Figure 5) and $\alpha = -1$ represent the practical limits on the sensitivity parameter for most empirical work.

However, reasonable alternatives to Axioms 4 and 5 are available. First, let us consider modifying the statement of Axiom 4, so that there is no “overlap” between the incomes in \mathbf{x}'' and the incomes in \mathbf{x} or \mathbf{x}' – either the income of the poorest person in \mathbf{x}'' is at least as high as that of the richest person in \mathbf{x} or \mathbf{x}' , or the income of the richest person in \mathbf{x}'' is no greater than that of the poorest person in \mathbf{x} or \mathbf{x}' . The relative inequality measures in (3) also work for this case and in addition we find that now a further inequality index is also available (Ebert 1988b):

$$I_{\text{Gini}}(\mathbf{x}) := \frac{1}{2n^2\mu(\mathbf{x})} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j| \quad (4)$$

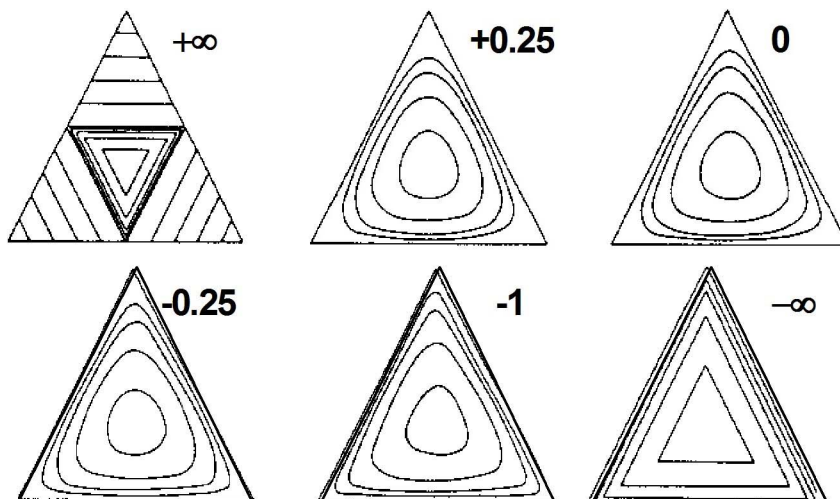


Figure 7: Generalised entropy contours for different values of α

This is the *Gini index* and its contours are depicted in Figures 5(b) and 6.

3.1.2 Translation-invariant inequality measures

Now consider a second alternative to the standard axioms. In place of scale invariance it is sometimes argued that the following structural assumption is appropriate (Kolm 1976a, 1976b):

Axiom 6 (*Translation Invariance*). Let $\mathbf{x}, \mathbf{x}' \in X^n(\mu)$. If $I(\mathbf{x}) \geq I(\mathbf{x}')$ then $I(\mathbf{x} + \delta \mathbf{1}) \geq I(\mathbf{x}' + \delta \mathbf{1})$.

Again, notice that this is just a requirement on the pattern of the contours and it is weaker than requiring translation independence, namely $I(\mathbf{x}) = I(\mathbf{x} + \delta \mathbf{1})$. If we replace the scale-invariance property used in Theorem 1

with translation invariance then we have the following result (Bosmans and Cowell 2010):

Theorem 2 *Axioms 1-4 and 6 imply that a continuous inequality index must either be ordinally equivalent to*

$$\frac{1}{n} \sum_{i=1}^n e^{\beta[x_i - \mu(\mathbf{x})]} - 1, \quad (5)$$

where $\beta \neq 0$ is a sensitivity parameter, or to

$$\frac{1}{n} \sum_{i=1}^n [x_i - \mu(\mathbf{x})]^2 \quad (6)$$

(replacing the case $\beta = 0$).

Equation (5) gives us a class of absolute measures (Blackorby and Donaldson 1980b) of inequality; members of this class for which $\beta > 0$ are ordinally equivalent to the Kolm (1976a) class of indices:

$$I_K^\beta(\mathbf{x}) := \log \left(\frac{1}{n} \sum_{i=1}^n e^{\beta[x_i - \mu(\mathbf{x})]} \right). \quad (7)$$

Equation (6) is just the variance. Again if we use the non-overlapping version of decomposability we find that there is another index that is both decomposable and translation-invariant:

$$I_{AG}(\mathbf{x}) := \frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|, \quad (8)$$

the so-called “absolute Gini” – the connection with the regular Gini in (4) is obvious. In contrast to scale invariance – where the inequality-contour map remains invariant under scalar transformations of income – translation invariance means that the contour map remains invariant under uniform additions to or subtractions from income.⁸

3.1.3 Inequality and income levels

If we review the contour maps for the inequality measures in sections 3.1.1 and 3.1.2 we will find exactly two maps that are consistent with both scale- *and* translation-invariance: these are the ones illustrated in Figure 5 above – the variance (6) and for the Gini coefficient (8). For each of these two cases Figure 8 illustrates how one can connect the contours at different levels of income either along rays through $\mathbf{0}$ (scale invariance) or along lines parallel to the ray through $\mathbf{0}$ and $\mu\mathbf{1}$ (translation invariance). However, we may be interested in a stronger requirement: scale- or translation-*independence*. Here scale independence means that inequality stays the same as all incomes are increased proportionately – $I(\mathbf{x}) = I(\lambda\mathbf{x})$ – and translation independence means that inequality stays the same as all incomes are increased by the same absolute amount – $I(\mathbf{x}) = I(\mathbf{x} + \delta\mathbf{1})$. Clearly both cannot be true at the same time unless one confines attention to trivial cases of perfectly equal distributions.

⁸There are also “intermediate” versions of invariance can be specified (Bossert 1988; Bossert and Pfingsten 1990; Kolm 1969, 1976a, 1976b).

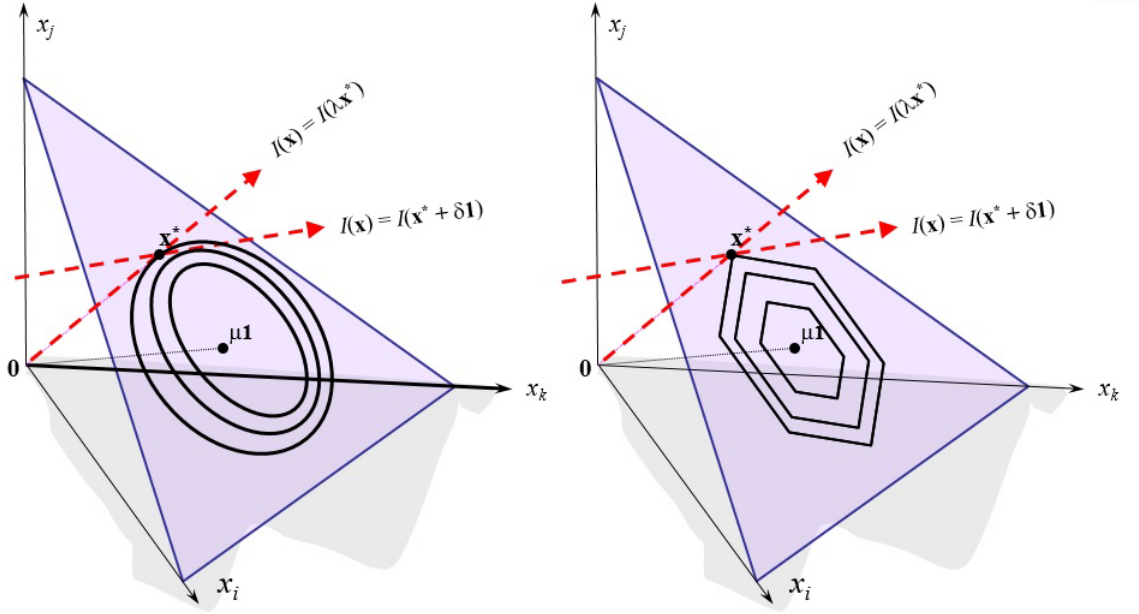


Figure 8: Scale or translation independence?

3.2 Decomposition

The decomposability property (Axiom 4) enables us to examine the breakdown of inequality by population subgroups. This property is particularly useful if we want to get some insight on the apparent “causes” of overall income inequality. There need to be heavy quotation marks around the word “causes”: it is interesting to know, for example, whether changes in world inequality appear to be associated with changes in the income distribution within individual countries or with changes in the overall income levels in the various countries; but obviously we would be doing no more than a kind

of sophisticated distributional accounting procedure. Nonetheless, this kind of accounting procedure can be informative if properly carried out.⁹

3.2.1 Arbitrary partition

To fix ideas, suppose we partition the population into m arbitrary subgroups that are mutually exclusive and exhaustive; for example if the partition were by sex we might have $m = 2$ groups (“male”, “female”) or possibly $m = 3$ groups if we allow for the value “unknown” alongside “male” and “female”. The decomposability property means that we can write overall inequality as a function of inequality in each subgroup $\ell = 1, \dots, m$ which may also depend on the population shares of each subgroup, π_ℓ (defined as n_ℓ/n , where $\sum_{\ell=1}^m n_\ell = n$) and the income shares of each subgroup, s_ℓ (defined as $\mu_\ell n_\ell / \mu n$, where $\sum_{\ell=1}^m s_\ell = 1$). In its most general form we would have:

$$I(\mathbf{x}) = \Phi(I(\mathbf{x}_1), I(\mathbf{x}_2), \dots, I(\mathbf{x}_m); \pi_1, \pi_2, \dots, \pi_m; s_1, s_2, \dots, s_m) \quad (9)$$

where \mathbf{x}_ℓ means the income distribution in subgroup ℓ . If we also require that the inequality measure be scale invariant (Axiom 5) then, as we know, the inequality measure has to take the form (3), or a transform of it. So in

⁹For an account of the issues involved and the relationship between decomposition by subgroups and decomposition by factor components see (Cowell and Fiorio 2011).

this case we could write¹⁰

$$I_{\text{GE}}(\mathbf{x}) = \sum_{\ell=1}^m w_{\ell} I_{\text{GE}}(\mathbf{x}_{\ell}) + I_{\text{GE}}(\mathbf{x}_{\text{Betw}}), \quad (10)$$

where

$$w_{\ell} := \pi_{\ell}^{1-\alpha} s_{\ell}^{\alpha} \quad (11)$$

and \mathbf{x}_{Betw} is the distribution that would emerge if, for every group ℓ , all n_{ℓ} persons were to receive exactly μ_{ℓ} , the mean of group ℓ , so that

$$I_{\text{GE}}(\mathbf{x}_{\text{Betw}}) = \frac{1}{\alpha [\alpha - 1]} \left[\sum_{\ell=1}^m \pi_{\ell} \left[\frac{\mu_{\ell}}{\mu(\mathbf{x})} \right]^{\alpha} - 1 \right]. \quad (12)$$

Equation (10) gives us an additive decomposition formula and is easy to interpret. The first term on the right-hand side represents aggregate within-group inequality and consists of the weighted sum of inequality in each subgroup where the weights are given by (11); the second term is the between-group component of inequality. So, in principle we can make a clear distinction between the contribution to overall inequality of the inequality in any one particular group, between the inequality within one group and the weight accorded to that group in the aggregation, and between the aggregate within-group component and the between-group component (for example, on the one hand, the inequality among males and among females weighted using (11) and, on the other hand inequality between males and female)s.

¹⁰Clearly we could do a similar exercise for translation-invariant (Axiom 6) indices using equation (7).

One might wonder whether (11) yields numbers that represent “weights” as conventionally understood: after all it is clear that although they will be non-negative they may not sum to 100 percent. However, the formula for the weights (11) immediately reveals that there are exactly two cases where the adding-up property will hold:

- where $\alpha = 0$, which yields pure population weights (we aggregate using the relative numbers of men and women); this is the case of the Mean Logarithmic Deviation index;
- where $\alpha = 1$, which yields pure income weights (we aggregate using the relative income levels rich men and women are); this is the case of the Theil index.

Clearly it is a matter of judgment whether one requires that the weights sum to 100 percent; it might be considered an extra that is of secondary importance compared with the flexibility of being able to select from the full range of values of the sensitivity parameter α .

3.2.2 Restricted partition

There is a further dimension of choice by the user of inequality measures. In the previous subsection we implicitly assumed that any and every possible partition of the population might be considered in the decomposition exercise. However we know from subsection 3.1.1 that if we restrict attention to cases where there is no “overlap” of incomes then the Gini coefficient (4) is also

available as a scale-independent inequality measure alongside (3). What this means in the present context is that we choose the subgroups $\ell = 1, \dots, m$ such that every income in group ℓ is less than or equal to the minimum income in group ℓ' (or every income in group ℓ is greater than or equal to the maximum income in group ℓ') then we could also use the Gini coefficient in a decomposition like (9). To see how this works, let us first note that it is always possible to rewrite the Gini formula (4) in a weighted-income form as follows

$$I_{\text{Gini}}(\mathbf{x}) = \sum_{i=1}^n \kappa_i x_{(i)}, \quad (13)$$

where (i) stands for “ i th smallest”; the κ_i act as the weights and are given by

$$\kappa_i(\mathbf{x}) := \frac{1}{n\mu(\mathbf{x})} \left[2\frac{i}{n} - \frac{1}{n} - 1 \right]. \quad (14)$$

Note that the κ_i take into account each person’s position in the distribution (i/n) . This has a nice interpretation. Assume that Irene is richer than Janet who is richer than Karen. Suppose Irene’s income increases by \$1 and Janet’s income decreases by \$1 – checking (13) and (14) this transfer by itself means that inequality must have increased by an amount $\frac{2}{n\mu} [i - j] > 0$; suppose also that Karen’s income increases by \$1 and Janet’s income decreases by \$1 – this transfer by itself means that inequality must have decreased by an amount $\frac{2}{n\mu} [k - j] < 0$; whether the combined is positive or negative depends on whether the difference in position $i - j$ is larger or smaller than $j - k$.

Suppose we break down the population into two subgroups F and M where no-one in the F group has an income higher than anyone in the M group. Clearly inequality in the F group could be written

$$I_{\text{Gini}}(\mathbf{x}_F) = \sum_{i=1}^{n_F} \kappa_i(\mathbf{x}_F) x_{(i)}, \quad (15)$$

where $\kappa_i(\mathbf{x}_F)$ is person i 's positional weight *evaluated within the F group alone*:

$$\kappa_i(\mathbf{x}_F) = \frac{1}{n_F \mu(\mathbf{x}_F)} \left[2 \frac{i}{n_F} - \frac{1}{n_F} - 1 \right] \quad (16)$$

From (14) and (16) we have

$$\kappa_i(\mathbf{x}) = \pi_F s_F \kappa_i(\mathbf{x}_F) + \frac{\pi_F - 1}{n \mu(\mathbf{x})} \quad (17)$$

and we also find:

$$I_{\text{Gini}}(\mathbf{x}) = \pi_F s_F I_{\text{Gini}}(\mathbf{x}_F) + \pi_M s_M I_{\text{Gini}}(\mathbf{x}_M) + I_{\text{Gini}}(\mathbf{x}_{\text{Betw}}) \quad (18)$$

which gives an exact formula for decomposing the Gini coefficient into non-overlapping subgroups F and M.

The importance of non-overlapping incomes in the decomposition can be easily seen if we think about the joint Janet→Irene and Janet→Karen transfer mentioned above. The effects of these two transfers on inequality within the F group are just $\pi_F s_F$ times the effects of these two transfers on inequality in the whole population: again they just depend on the relative

size of $i - j$ and $j - k$. So, if inequality goes up in the F group, then it also goes up in the population as a whole, an example of the property known as subgroup consistency. Now suppose instead that two people from the M group, Gordon and Harry let us say, have incomes that overlap with those in the F group; specifically Gordon and Harry are richer than Janet but poorer than Irene. Clearly the effect of the joint Janet→Irene and Janet→Karen transfer on inequality within the F group is just what it was in the earlier example; but the effect in the population as a whole now depends on the relative size of $i - j + 2$ and $j - k$ (remember the two guys between Irene and Janet). So it would be possible to find that inequality goes *down* in the F group and *up* in the distribution as a whole even though inequality in the M group, inequality between the groups and the weights on the groups stay the same!

3.3 Alternative approach: reference points

The driving assumption for characterising inequality in the majority of the theoretical literature is some form of the transfer principle (see Axiom 3). But in its classic form (Dalton 1920) it is often rejected by people when invited to compare income distributions (Amiel and Cowell 1992, 1999) and, while elegant, it is obviously restrictive. An alternative approach to inequality is to think of it as an aggregation of distance from a relevant reference point. The reference point could, for example, be an income that is associated with a particular group. This is the essence of the Temkin approach of characterising

inequality in terms of *complaints* about income distribution (Temkin 1986, Temkin 1993, Devooght 2003, Cowell and Ebert 2004): these “complaints” are effectively the distances just mentioned and are typically upward-looking comparisons.

This alternative approach to the assessment of income distribution is individualistic – as with the standard approach in section 3.1 – but is based on a concept of differences rather than on income levels. To make this operational we first need to specify $r(\mathbf{x})$, a reference income that usually depends on the incomes in the current distribution \mathbf{x} : we will discuss below three different specifications for $r(\mathbf{x})$ that lead to interesting forms for inequality measures. The second thing to specify is a set of axioms to characterise the implicit notion of distance from the reference point (the “complaint”) and the function which aggregates the individual distances

Cowell and Ebert (2004) showed that, under standard assumptions, distance must be of the form

$$d_i = \left| r(\mathbf{x}) - x_{(i)} \right|, \quad (19)$$

with the reference-point based inequality index given by

$$I_{\text{Temk}}(\mathbf{x}) = \left[\sum_{d_i > 0} w_i d_i^\theta \right]^{\frac{1}{\theta}}, \quad (20)$$

where the w_i are positive numbers (weights) and θ is a sensitivity param-

eter.¹¹ Both the sensitivity parameter θ and the weights may incorporate distributional values. Consider an income transfer from poorer Janet to richer Irene that leaves the reference point unchanged. This increases Janet's distance and decreases Irene's distance from the unchanged reference point. From (20) it is clear that the effect of this transfer is proportional to $w_j d_j^{\theta-1} - w_i d_i^{\theta-1}$; So if Irene is richer than Janet and if the reference point is left unchanged, this transfer will increase inequality if $w_j \geq w_i$ and $\theta > 1$ (or $w_j > w_i$ and $\theta = 1$).

Three versions of the reference point, highlighted by Temkin's seminal work, are of particular interest.

(1) The Best-off Person (BoP)

Here everyone but the richest person has a complaint about inequality. The reference point for everyone is:

$$r(\mathbf{x}) = x_{(n)}.$$

Of the three types of reference point suggested by Temkin (1986) this is, perhaps, the most intriguing. For a start in this case there is a *second* type of "transfer principle" the effect of transferring income to the richest person (everyone's reference point) from any of the other $n - 1$ persons. As we have noted, the regular transfer principle (Axiom 3 whereby inequality increases for any poorer to richer transfer) holds for only for a specific range of pa-

¹¹In the case where $\theta = 0$ (20) is replaced by $\exp(\sum_{d_i > 0} w_i \log(d_i))$.

parameter values in (20); but for all θ , the inequality measure (20) satisfies the principle of transfers-to-the-richest. So there could be some tension between the two types of transfer principle: if $\theta \leq 0$ absolute priority is placed on the salience of the best-off person: so if Irene is the richest and Janet is richer than Karen then Karen \rightarrow Janet transfers may reduce inequality (such a transfer reduces the distance between Janet and Irene) but if an Irene \rightarrow Janet transfer always reduces inequality (distance is reduced everywhere).

Let us look at the counterparts to Figures 5 and 7 in the case of reference-point inequality. Some typical contours of I_{Temk} are depicted in Figure 9 for cases where the regular transfer principle is satisfied and in Figure 10 for the case where the regular transfer principle is not satisfied. Notice that if the weights are equal each contour breaks down into three segments; if we have differential weights when aggregating the distance in (20) then we have six segments in each contour (as in the case of I_{Gini}). In Figure 10 where the segments bend the “wrong” way we can see that equalising Janet \rightarrow Karen transfers increase inequality.

(2) All Those Better Off

As with BoP, everyone but the best off has a “complaint” about income distribution. However in this case people at different positions in the income

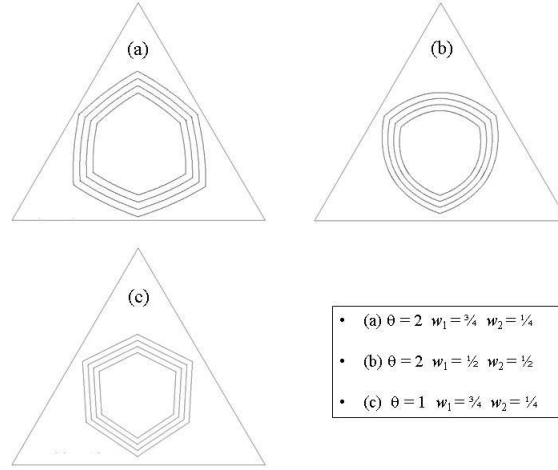


Figure 9: BoP-reference inequality (Axiom 3 is satisfied)

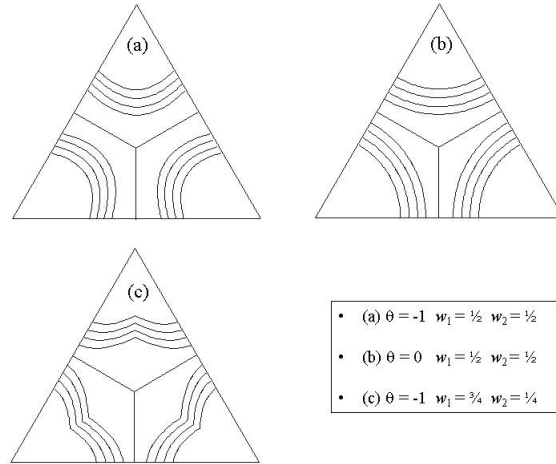


Figure 10: BoP-reference inequality (Axiom 3 is not satisfied)

distribution have different reference points:

$$r_i(\mathbf{x}) = \frac{1}{n-i} \sum_{x_k > x_{(i)}} x_k .$$

If we replace $r(\mathbf{x})$ by $r_i(\mathbf{x})$ in (19) then the resulting inequality measure in (20) is essentially the same as the deprivation index suggested by Chakravarty and Mukherjee (1999).¹²

(3) Average income

If the reference point is just mean income

$$r(\mathbf{x}) = \mu(\mathbf{x}),$$

then substituting in (19) yields a class of inequality measures (20) that is closely related to the family of “compromise” inequality measures in Ebert (1988a).

4 Inequality: welfare and values

Inequality measurement is not, or should not be, just about formal propositions in mathematical language. Inequality is something about which some individuals feel passionately and which is, arguably, a proper concern of public policy. So, it is important to examine the ways in which ethical principles

¹²Of course the general idea of a connection between inequality and deprivation had been developed much earlier – see Yitzhaki (1979).

may be incorporated in the analysis of inequality measurement.

4.1 Social welfare and inequality

For a discussion of the role of social values or ethical principles we need an additional basic tool, the *social-welfare function*, written as

$$W(\mathbf{x}) = W(x_1, x_2, \dots, x_n), \quad (21)$$

which gives a numerical score for every income distribution \mathbf{x} in X^n ; in this context each x_i is assumed to completely represent person i 's well-being.¹³ The question immediately arises – what values or principles should be embodied in W ? If social values broadly consider that income increases are on the whole a good thing and inequality increases are a bad thing we would want this to be built into the structure of W . Clearly it is necessary to connect the formula for social welfare (21) to the formulas for mean income (1) and inequality measure (2).

A simple representation representation of this connection is as the reduced-form social-welfare function:

$$W(\mathbf{x}) = \Omega(\mu(\mathbf{x}), I(\mathbf{x})) \quad (22)$$

¹³To achieve this in a heterogeneous society it may be necessary to adjust incomes by a factor the appropriately reflects, say, the needs of individuals living in different types of household (Fleurbaey 2015). For more on the nature of individual well-being in connection with social welfare functions see Weymark (2015).

where the function Ω is increasing in its first argument (mean income) and decreasing in its second argument (inequality). A simple example of this is the social-welfare function given by

$$\mu(\mathbf{x}) [1 - I_{\text{Gini}}(\mathbf{x})]; \quad (23)$$

from (13) it is clear that this is just the weighted sum of the ordered incomes $x_{(i)}$ where, in view of (14), the weights must be strictly decreasing as i goes from 1 to n .

Clearly the properties of the function W and the function I are intimately connected: for any given value of μ they will have the same contours in X^n (but of course the contours will be numbered in opposite directions – inequality increases imply welfare decreases. However, the relationship with income levels needs further comment. It is common to assume that W is strictly increasing in each income x_i (the so-called *monotonicity* property) – this would mean that, if the status quo is distribution \mathbf{x} then social welfare would be increasing in the direction of *any* of the arrows in Figure 11 (a). But it could be argued that this is unacceptably strong; as an alternative we might suggest either the weaker requirement that welfare should increase if all incomes were to be increased simultaneously by the same proportional amount (if $\lambda > 1$ we would have $W(\lambda\mathbf{x}) > W(\mathbf{x})$ so that welfare increases along the steeper of the arrows in Figure 11 b) or the requirement that welfare should increase if all incomes were to be increased simultaneously by

the same absolute amount (if $\delta > 0$ we would have $W(\mathbf{x} + \delta \mathbf{1}) > W(\mathbf{x})$ so that welfare increases along the flatter of the arrows in Figure 11 b)¹⁴

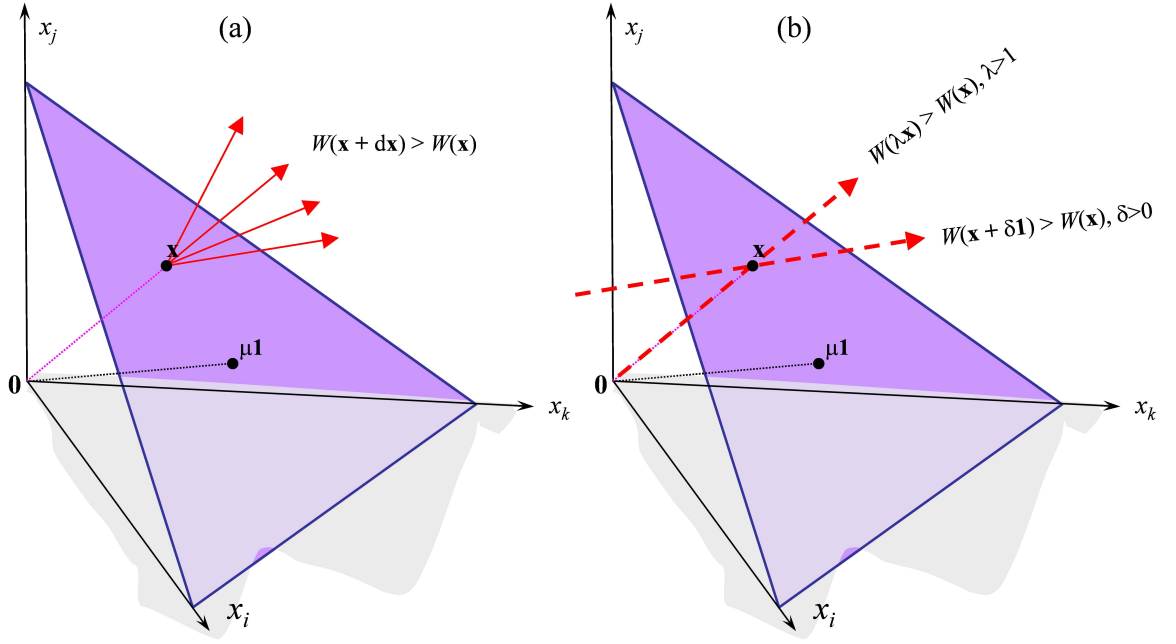


Figure 11: Social welfare and income growth

If we now rethink the contour maps in Figures 5 to 8 as contours of the SWF for given mean income, it is clear that one could read this as though the function W in (21) and (22) “inherits” the properties of the inequality measure I in (2). But it may be more appropriate to read this in the reverse direction: inequality is something that is considered socially undesirable or

¹⁴Known as the “principle of uniform income growth” (Champernowne and Cowell 1998) or the “incremental improvement condition” (Chakravarty 1990).

even wrong; this principle is built into the specification of the SWF; this specification determines the contours of the SWF and therefore the inequality contours that appear in the diagrams of section 3. But what values? They can be essentially characterised as two types of tradeoff:

1. Tradeoffs between equality and efficiency. Inequality represents a loss of welfare – as one moves within the simplex away from the centre one moves to lower welfare contours. How much should one give weight to a movement towards the centre of the simplex $\mu\mathbf{1}$, relative to movement from one μ -simplex to a “higher” one?
2. Tradeoffs between income differences in one part of the distribution relative to those in another: if Irene is richer than Janet who is richer than Karen should we prioritise the Irene-Janet income differences or the Janet-Karen differences?

4.2 Welfare-based inequality measures

Once one is persuaded of the idea that the social wrong of inequality can be expressed in terms of a loss of social welfare and that inequality measures inherit their properties from SWFs or vice versa there emerges a natural way of characterising inequality in welfare terms. It might at first be thought that there is an obstacle to using the SWF to specify an inequality index since there are no “natural” units for the measurement of welfare. But we could measure welfare – and therefore gains or losses of welfare – in terms of income.

To do this, introduce the concept of the *equally-distributed equivalent* (EDE) income. For any distribution \mathbf{x} this can be represented as a number ξ such that

$$W(\mathbf{x}) = W(\xi \mathbf{1}) = W(\xi, \xi, \dots, \xi). \quad (24)$$

The EDE-income is a remarkably robust concept since little more than that W be a continuous function is needed in order to be sure that any distribution \mathbf{x} has its corresponding EDE-income ξ . We could, if we wish, invert the relationship (24) to make the dependence of ξ on the income distribution explicit

$$\xi(\mathbf{x}) = \xi(x_1, x_2, \dots, x_n). \quad (25)$$

It is important to recognise that (21) and (25) do the same job: in using ξ rather than W we have just chosen a cardinalisation of the social-welfare function that is convenient to work with.

If the SWF has been specified so that unequal distributions register a loss of social welfare then we may deduce that:

$$\xi(\mathbf{x}) \leq \mu(\mathbf{x}) \quad (26)$$

where μ is the mean, defined in (1).¹⁵ A simple twist on (26) gives us the following generic welfare-based index of inequality

¹⁵The reasoning is this. In principle you could take the total income in a distribution $n\mu$ and allocate it equally among the n persons to give you the distribution $\mu \mathbf{1} = (\mu, \mu, \dots, \mu)$; the EDE-income for this very special case is of course $\xi = \mu$. Any perturbation that produces a different distribution with the same mean must introduce some inequality which reduces social welfare. Therefore (26) must be true.

$$I(\mathbf{x}) = 1 - \frac{\xi(\mathbf{x})}{\mu(\mathbf{x})}, \quad (27)$$

which is unit-free, non-negative and assumes the value 0 when $\mathbf{x} = \mu\mathbf{1}$. Once again, similarly to (23) it is clear that, by taking the EDE-income cardinalisation of welfare, we can write

$$\text{welfare} = \xi(\mathbf{x}) = \mu(\mathbf{x}) [1 - I(\mathbf{x})]$$

– compare this with equation (23).

The idea is illustrated in Figure 12 which shows the contours of W in the space of Irene and Janet's incomes. Point \mathbf{x} represents the current income distribution and the line through \mathbf{x} and $\mu\mathbf{1}$ gives all the distributions with the same mean μ . The dotted 45-degree ray represents equal distributions and it is clear that the contours are symmetric about this ray (swapping Irene's and Janet's identities leaves social welfare unchanged). The contour through the point \mathbf{x} intersects the equality ray at point $\xi\mathbf{1}$ where ξ is the EDE income. The gap between $\xi(\mathbf{x})$ and $\mu(\mathbf{x})$ is the loss of welfare associated with the inequality implicit in \mathbf{x} and can be used to calculate the inequality index (27). Obviously to go further we would need to know more about the shape of W .

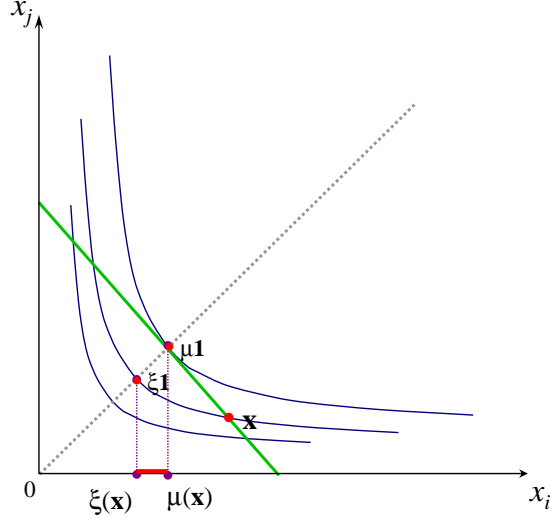


Figure 12: Social welfare and inequality

So, to make these ideas clearer, let us take a specific type of SWF: it is useful to consider all the functions W that can be written in the *additively separable* form:

$$W(\mathbf{x}) = \sum_{i=1}^n \zeta(x_i), \quad (28)$$

where ζ is a *social-evaluation function* that is the same for all individuals; in this specification we evaluate social welfare by going through the population, evaluating each person's income x_i and then summing the individual evaluations. In particular we may need two further restrictions on the family of SWFs in (28):

1. functions W where $\zeta(x)$ is increasing in x : see Figure 13 (a) and (b);

2. functions W where $\zeta(x)$ is increasing in x and also the slope of ζ is decreasing in x : see Figure 13 (b).

This specification is appealing and intuitive; but it is also quite restrictive because it effectively allows only one type of relationship between mean income and welfare: (the monotonicity property described above, implied by restriction 1) and it collapses the two types of tradeoff mentioned at the end of section 4.1 into a single tradeoff (implicit in the curvature of ζ in Figure 13 b).

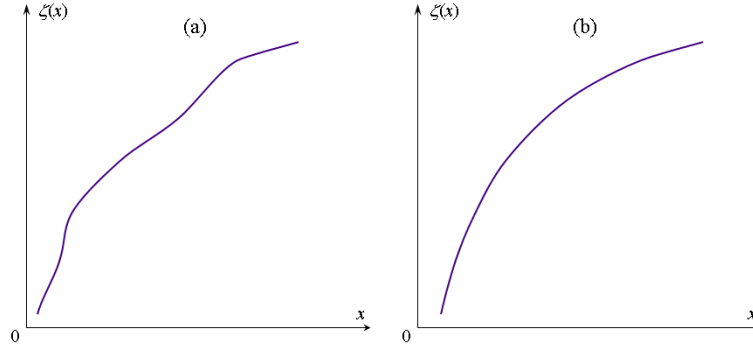


Figure 13: Evaluation functions ζ

In this set-up social attitudes towards income distribution and inequality are embodied in the social-evaluation function ζ . If this function is differentiable let us write the slope of ζ at any point x as $\zeta'(x)$: given restriction 1, $\zeta'(x)$ is positive for any x ; and, given restriction 2, the slope of the function ζ' must be negative ($\zeta''(x) < 0$). The importance of this can be seen if

we use equation (28) to express the change in social welfare dW that would arise if, for example, some policy or event changed individual incomes by the amounts $(dx_1, dx_2, \dots, dx_n)$. Clearly we have

$$dW(\mathbf{x}) = \sum_{i=1}^n \zeta'(x_i) dx_i, \quad (29)$$

which could be thought of as being the weighted sum of the income changes, where the weight on the change in Irene's income equals the marginal social evaluation of Irene's income. The weight is positive and decreases with income; how sharply the weights decrease depends on how sharply curved is the function ζ .

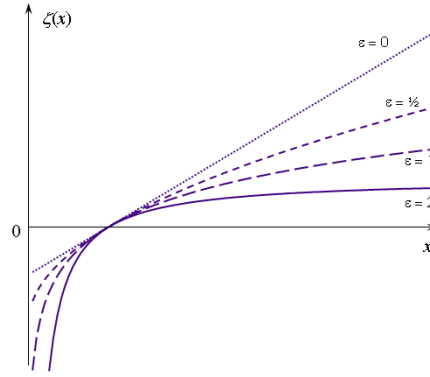


Figure 14: Isoelastic ζ for different values of ε

We can characterise the curvature of ζ by using $\varepsilon(x)$, the *elasticity of marginal social evaluation* at x , namely the percentage by which the weight

falls given a one percent increase in income:

$$\varepsilon(x) := -x \frac{\zeta''(x)}{\zeta'(x)}. \quad (30)$$

To fix ideas consider the family of *isoelastic* functions ζ , functions for which ε is a given constant; some of these are illustrated in Figure 14. The general formula is¹⁶

$$\zeta(x) = \frac{x^{1-\varepsilon} - 1}{1-\varepsilon}, \quad \varepsilon \geq 0, \quad (31)$$

so that the formula for the weights in this case is just

$$\zeta'(x) = x^{-\varepsilon}. \quad (32)$$

Let us compute the effect on social welfare of a change in the incomes of Irene and Janet; from (29) and (32) we have:

$$dW = x_i^{-\varepsilon} dx_i + x_j^{-\varepsilon} dx_j. \quad (33)$$

Let us suppose that Irene has t times the income of Janet, with $t > 1$. In this case we can easily find the tradeoff between Irene and Janet that would keep social-welfare constant. By putting $x_i = tx_j$ and $dW = 0$ in (33) we get

$$-\frac{dx_i}{dx_j} = t^\varepsilon, \quad (34)$$

¹⁶Using L'Hôpital's rule we find that in the case where $\varepsilon = 1$ equation (31) becomes $\zeta(x) = \log x$.

What this tells us is the tradeoff between equality and efficiency implicit in W . If one were to raise poor Janet's income by \$1 then a sacrifice of up to $\$t^\varepsilon$ by rich Irene would maintain or improve the current level of social welfare. For example, suppose $t = 5$: then if $\varepsilon = \frac{1}{2}$ a sacrifice of up to \$2.24 by Irene would be warranted to increase Irene's income by \$1; but if $\varepsilon = 1$ the allowable sacrifice by Irene would rise to \$5 ... and so on. Clearly this amount increases with the value of the elasticity ε and for this reason ε is usually known as the *inequality aversion parameter*.

An alternative way of interpreting ε is in terms of simple transfers in different parts of the income distribution. Suppose that Irene's income is t times that of Janet's and that Janet's income is t times that of Karen ($t > 1$). Then, using (33) we can calculate the (welfare-increasing) effect of a straight Irene-Janet transfer and that of a Janet-Karen transfer of the same amount:

$$dW_{ij} = [-t^{-\varepsilon} + 1] x_j^{-\varepsilon} dx > 0, \quad (35)$$

$$dW_{jk} = [-1 + t^\varepsilon] x_j^{-\varepsilon} dx > 0, \quad (36)$$

Dividing the one by the other we have

$$\frac{dW_{jk}}{dW_{ij}} = \frac{[-1 + t^\varepsilon] x_j^{-\varepsilon}}{[-t^{-\varepsilon} + 1] x_j^{-\varepsilon}} = t^\varepsilon. \quad (37)$$

Since $t > 1$ by assumption, this is obviously greater than one if $\varepsilon > 0$, so that $dW_{jk} > dW_{ij}$ (the transfer is more effective lower down the income scale).

But we can see more: the relative size of the lower-income transfer effect gets larger, the higher is ε . This second interpretation of ε in equation (37) is logically separate from the equality-efficiency interpretation in (34): but, in the special case where social welfare is additive (28), the two concepts happen to have the same value.

If we use the additive, isoelastic form of the SWF defined by (28) and (31) then the EDE income in (24) takes the form of a generalised mean:

$$\xi(\mathbf{x}) = \left[\frac{1}{n} \sum_{i=1}^n x_i^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \quad (38)$$

and the associated inequality measure (27) takes the form of the *Atkinson class* of indices (Atkinson 1970):¹⁷

$$I_{\text{Atk}}(\mathbf{x}) = 1 - \left[\frac{1}{n} \sum_{i=1}^n \left[\frac{x_i}{\mu(\mathbf{x})} \right]^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}. \quad (39)$$

For a given unequal distribution \mathbf{x} the higher is ε the lower is $\xi(\mathbf{x})$ and the higher is $I_{\text{Atk}}(\mathbf{x})$. Comparing $I_{\text{GE}}(\mathbf{x})$ with $I_{\text{Atk}}(\mathbf{x})$ – equations (3) and (39) respectively – several points are immediately striking:

- In both cases we have a class of indices rather than a single inequality measure; individual members of the class are characterised by a sensitivity parameter (α in the case of I_{GE} and ε in the case of I_{Atk}); the user of the inequality measure brings some personal or social judgment

¹⁷In the case where $\varepsilon = 1$ equation (39) becomes $1 - \exp\left(\frac{1}{n} \sum_{i=1}^n \log\left(\frac{x_i}{\mu(\mathbf{x})}\right)\right)$.

to the measurement problem by the choice of the parameter.

- The ordering of income distributions produced by a member of the Atkinson class with parameter ε is the same as the corresponding GE index with parameter $\alpha = 1 - \varepsilon$: they produce the same inequality contours in the distribution simplex of Figure 2.¹⁸
- In both cases the inequality contours are induced by a structural assumption – scale invariance (Axiom 5) in the case of the GE indices, isoelasticity of the social-evaluation function in the case of the Atkinson indices.

4.3 Welfare and individual values

Clearly we can use the welfare-function approach to determine the shape of the inequality contours in section 3.1. But a more important issue is, how the characteristics of the welfare function are to be determined. There are several routes through which we might imagine this to happen, which can be grouped into two broad categories:

- Representation of innate desire for redistribution. Individual members of society might have a preference for equality that is expressed through the political process, through charitable giving or through survey responses.¹⁹ Society's aversion to inequality is based on an externality

¹⁸There is no Atkinson index corresponding to a member of the GE family with $\alpha \geq 1$.

¹⁹On direct approach using questionnaires with student subjects see Amiel, Creedy, and

involving other people's incomes (Hochman and Rodgers 1969; Kolm 1964, 1969; Thurow 1971).

- Imputation of a social ranking of income distributions from individual rankings of probability distributions. One might take the position that in some sense social values are a representation of individual preferences or views behind a “veil of ignorance.” Here no-one need be averse to inequality per se. However, everyone may be averse to risk and in particular everyone may be averse to the risk associated with the lottery of life (Harsanyi, 1953, 1955 1978; Rawls, 1971, 1999). Because individuals experience a loss of utility through risk and are willing to sacrifice income in order to insure against it, inequality – the social counterpart of risk – leads to social welfare loss.

In practice aversion to inequality may come from both of these routes and people may not view risk and inequality in exactly the same way (Cowell and Schokkaert 2001, Kroll and Davidovitz 2003, Carlsson et al. 2005).

From the discussion of equation (34) above it is clear that the degree of inequality aversion can be seen in terms of the calibration of a tradeoff – either (a) between greater equality and higher overall income or (b) between transfers in different parts of the distribution. For the Atkinson indices (39) the precise tradeoff can be characterised by the value of ε , the elasticity of

Hurn (1999), Gevers et al. (1979), Glejser et al. (1977); see also Gaertner and Schokkaert (2012), Van Praag (1977, 1978), Herwaarden et al. (1977), Van Batenburg and Van Praag (1980).

marginal social evaluation. Clearly then, an important further question for policy-makers is how to determine reasonable values for ε . If we want to base this on the views of everyday people we could appeal to experimental or questionnaire evidence. However, caution is necessary in interpreting such evidence: preferences reversals can arise in the context of social choice similar to those that arise in the context of individual choice (Amiel et al. 2008) and estimates of ε from leaky-bucket exercises may be affected by the way the issue is put (Pirttilä and Uusitalo 2010). Evidence from happiness studies suggest a value of 1.0 to 1.5 for ε (Layard et al. 2008); inferring ε from tax schedules we find a value of around 1.2 to 1.4 (Cowell and Gardiner 2000).

5 Poverty measurement: principles

The fundamental difference between the inequality-measurement problem and the poverty-measurement problem is, of course, the poverty line. The poverty line performs several roles in the poverty-measurement problem: it partitions the population into two groups that we want to treat differently in analysing income distribution; it forms a reference point; it can be a policy parameter in its own right. But, of course, reading through this list of roles we can spot several points where there appear to be important links with inequality analysis: the decomposition into poor and non-poor subgroups and the concept of a reference income, for example. So we may expect –

and indeed we will find – some carry-over of the measurement analysis from Section 3.

A lot could be written on the income concept that is appropriate for the analysis of poverty and the determination of the poverty line, but such arguments are usually about the implementation of the principles of poverty measurement rather than the principles themselves. Once again we will suppose that an income concept has been specified (this “income” could, for example, be defined as consumption expenditure or as assets) and so an income distribution is just a vector of real numbers, as we discussed in section 2.1. The poverty line could be determined with reference to the living standards in the society represented by the income distribution; it could be fixed by international standards or by the observer’s own personal introspection; it could be determined from the distribution itself (for example many empirical studies use half the median income as a poverty line). For our purposes it is convenient for the moment to suppose that the poverty line is an exogenously given number z : any person i for whom $x_i \leq z$ is deemed to be poor.²⁰

What do we mean by a poverty measure? The language is important here because the intuitive approaches discussed in section 5.1 focus on simple mea-

²⁰There is a point of detail to note here: the majority of the formal literature use this so-called “strong” version of poverty where $x_i \leq z$ means “ i is poor” rather than its “weak” counterpart where the criterion is $x_i < z$ (an interesting exception is Watts 1968), but in practice governmental agencies and others often adopt the weak version. Clearly this is important for poverty measures that rely on a head count of the poor (are people with incomes exactly equal to z poor or not?) such as (40) below, but not for the other poverty measures and results discussed below (Zheng 1997, Donaldson and Weymark 1986).

asures with cardinal significance,²¹ but for formal approaches the discussion is sometimes of specific indices, sometimes of orderings: in section 5.2 we will focus on poverty orderings, as in section 3 on inequality. We will consider more general poverty rankings separately, in section 6.

5.1 Intuitive approaches

There are several poverty indices that spring to mind naturally from no more than a brief scan of the data and the simplest calculations. These are indices that are used by policy makers and that provide numbers with a compelling, common-sense interpretation. They also serve as reference points for the more sophisticated analysis in section 5.2. The leading indices are as follows.

- *Headcount ratio*. For any z we just work out n_z , the number in the population with incomes less than or equal to z , and divide this by the population:

$$\frac{n_z}{n} \text{ where } n_z := \sum_{x_i \leq z} 1. \quad (40)$$

- *Average income gap ratio (1)*. For each person i define the individual income gap is the income shortfall (if any) below the poverty line:

$$g_i = \max \{z - x_i, 0\}. \quad (41)$$

²¹See Section 2.3 for a discussion of the terminology.

Then the average income gap ratio (relative to the poverty line) is

$$\frac{1}{n_z} \sum_{i=1}^n \frac{g_i}{z}. \quad (42)$$

- *Average income gap ratio (2).* An alternative normalisation for the average income gap is

$$\frac{1}{n} \sum_{i=1}^n \frac{g_i}{\mu(\mathbf{x})}. \quad (43)$$

Notice that the average in (42) is taken over the number who are poor, whereas the average in (43) is over the whole population.

Which of the poverty measures (40), (42) and (43) is “right”? In a sense, all of them are. We can regard the headcount ratio and the average income-gap ratio as picking up two different, complementary aspects of poverty: the head-count ratio focuses on the labelling effect of poverty, the proportion of the population that carries the disfiguring mark of being poor; the average gap concept gives us some idea of the resources necessary to eliminate poverty. Version 1 (equation 42) expresses the resource-requirement interpretation relative to the poverty line; version 2 (equation 43) expresses the same idea as a proportion of resources available in the economy. It is clear that any two of these intuitive measures could contradict each other in terms of answering the question whether poverty has increased or decreased, even with a fixed poverty line.²² But this does not necessarily matter: the differ-

²²Example. Suppose the poverty line is fixed at 100. Then, if the income distribution changes from (10,50,90,1000) to (50,50,110,500), the headcount ratio falls from 0.75 to 0.5,

ent measures just reveal different things about the concept of poverty and our understanding is enriched by having these different perspectives.

5.2 Axiomatic approach

Why should we want to go beyond these simple, intuitive approaches? A good answer to this question is that it may be appropriate to consider whether poverty measurement should take into account more information about the income distribution. The simple summary statistics in (40)-(43) remain unchanged if there is some redistribution among the poor that leaves unchanged the numbers and average income of the poor. However, this would imply that we consider the following two distributions to represent the same level of aggregate poverty: (a) a distribution where all poor people have exactly the same income and (b) a distribution in which poor people are polarised into two groups, those close to destitution and those just below the poverty line. If the distribution amongst the poor is of concern then we may need a more sophisticated poverty measure (Watts 1968, Sen 1976).

Why could an axiomatic approach be useful? A good answer to this question is that, as with inequality measurement, there may be specific aspects of the comparison of income distributions in terms of poverty that are best considered as abstract principles; expressing these principles in terms of formal axioms may help to narrow down the class of poverty measures that

the average income gap (1) stays the same (at 0.5) and the average income gap (2) rises from 0.13 to 0.14.

may be worth considering in addition to – or even instead of – the intuitive approaches outlined in section 5.1. However, in the case of poverty, it may be that some of the additional formality that has been introduced has tended to obscure rather than illuminate the central issues. Some of the axioms²³ that have been suggested are less persuasive than their counterparts in inequality analysis and some may actually obscure the insights available from the simpler approaches in section 5.1. Perhaps poverty is in danger of being over-axiomed.

Methodologically the approach is similar to what we did in the case of inequality: the principal difference is that rather than just comparing distributions n -vectors we are now working with $n + 1$ incomes $(x_1, x_2, \dots, x_n, z)$ – or, written more compactly, (\mathbf{x}, z) . Some of the axioms from section 3 can be carried over with just minor notational modification: where this works we will just note how the axiom needs to be adapted (this is usually little more than cosmetic) rather than writing everything out explicitly in the new notation.

Basic form

The first key result needs the anonymity property (Axioms 1), adapted by replacing I with P and three new axioms:

²³Among the suggested axioms for poverty measurement are Transfer axiom (a reworking of axiom 3 for the population of the poor), upward transfers (a combination of the transfer axiom and focus), population growth amongst the poor, population growth amongst the non-poor. Unsurprisingly perhaps, it is easy to find fundamental conflict among some of the proposed axioms (Kundu and Smith 1983). See also Donaldson and Weymark (1986).

Axiom 7 (*Focus*). Suppose \mathbf{x}' is formed from \mathbf{x} thus: $x'_i = x_i + \delta$ where $x_i > z$ (and $x'_j = x_j$ for all $j \neq i$); then $P(\mathbf{x}', z) = P(\mathbf{x}, z)$.

Axiom 8 (*Monotonicity*). Suppose \mathbf{x}' is formed from \mathbf{x} thus: $x_i < x'_i \leq z$ and $x'_j = x_j$ for all $j \neq i$; then $P(\mathbf{x}', z) < P(\mathbf{x}, z)$.

Axiom 9 (*Independence*). Consider $\mathbf{x}, \mathbf{y} \in X^n$ such that $P(\mathbf{x}, z) = P(\mathbf{y}, z)$ and, for some person i , $x_i = y_i < z$. Suppose \mathbf{x}', \mathbf{y}' satisfy $x'_i = y'_i \leq z$ and $x'_j = x_j, y'_j = y_j$ for all $j \neq i$; then $P(\mathbf{x}', z) = P(\mathbf{y}', z)$.

We can interpret Axiom 7 as requiring that the incomes of people who are not poor are irrelevant to poverty comparisons. This might be considered the essential axiom of poverty measurement, although in fact the second version of the Average income gap ratio violates it. Axiom 8 means that if the income of any poor person is increased then poverty must decrease²⁴ (but note that the headcount ratio violates this property). Axiom 9 says the following: take two distributions that are exactly equal in terms of aggregate poverty; suppose there is a particular income x_i that is common to both distributions; then varying that common value will change the measured poverty in the two distributions by the same amount, whatever the other incomes in each distribution may be. This property is exhibited by each of the intuitive measures in section 5.1.

²⁴There is an important difference in the precise definition of monotonicity as defined here in Axiom 8 and the monotonicity property discussed in connection with the SWF in section 4.1: the monotonicity principle for the function W in (21) means that increase in *anybody's* income (rich or poor) increases welfare; but the monotonicity axiom here applies only to incomes below the poverty line.

Theorem 3 *The poverty version of Axiom 1 and Axioms 7-9 jointly imply that a continuous poverty ordering must be representable by the measure*

$$P(\mathbf{x}, z) = \frac{1}{n} \sum_{x_i \leq z} p(x_i, z) \quad (44)$$

where p is a continuous function that is decreasing in x_i as long as $x_i < z$.

So, invoking just a few appealing principles results in a particularly useful class of poverty measures, the *Additively Separable Poverty* (ASP) class: the crucial assumption for this is Axiom 9, which rules out concepts of poverty that explicitly take into account the person's position in the distribution. At the centre of (44) is the individual *poverty-evaluation function* p which gives a natural interpretation to the ASP measures. One goes through the population and evaluates each person's income x_i in the light of the poverty line: if person i is above the poverty line then the evaluation is zero; if i is exactly on the poverty line then the evaluation is zero; for all other cases the evaluation is higher the lower is i 's income.

Of course the form (44) allows a lot of leeway in the precise precise specification of the poverty-evaluation function. Just as we did for inequality measures in section 3.1 it is useful to see what additional principles might be usefully imposed on the structure of income distributions to narrow down the class of admissible poverty measures.

Income structure

In the analysis of inequality measurement we considered scale invariance and translation invariance (Axioms 5 and 6) as two different principles inducing what might appear to be a sensible structure on the map of iso-inequality contours (in the case of just two sets of contours both principles could be applied at the same time – see Figure 5). In the analysis of poverty measurement, we can again appeal to these principles but with two important differences: (1) we require them to apply to all $n + 1$ incomes $(\mathbf{x}, z) \in X^{n+1}$ and (2) they not only induce a set of iso-poverty contours but also imply that the resulting poverty indices incorporate a specific relationship between individual incomes and the poverty line. The key result is (Ebert and Moyes 2002):

Theorem 4 *If the ASP poverty indices also satisfy the poverty version of Axioms 5 and 6 then the individual poverty-evaluation function p must take either the form g_i^θ or the form $[g_i/z]^\theta$, where g_i is defined in (41) and θ is a positive parameter.*

An important implication of Theorem 4 is that the scale-invariance and translation-invariance properties alone are enough to restrict the ASP class to some of the most widely employed poverty measures, the so-called Foster-Greer-Thorbecke (FGT) indices:

$$P_{\text{FGT}}^\theta(\mathbf{x}, z) := \frac{1}{n} \sum_{x_i \leq z} \left[\frac{g_i}{z} \right]^\theta, \quad (45)$$

(see Foster et al. 1984).²⁵ Once again, as with the GE, Temkin and Atkinson measures in the inequality context, one has a family of measures, indexed by a sensitivity parameter θ . A glance back at section 3.3 also shows the strong similarity between FGT poverty measures and Temkin-type inequality indices, which should not come as a big surprise since they both use a reference point and they are both built on similar axioms regarding scale- and translation-invariance.

Theorem 4 implies that θ is a positive number, but it is useful to consider the extension of the class (45) to the case $\theta = 0$.²⁶ The two cases $\theta = 0, 1$ have a natural interpretation; these two cases and $\theta = 2$ are illustrated in Figure 15:

- $\theta = 0$. Letting θ go to zero in (45) we can see that $P_{\text{FGT}}^0(\mathbf{x}, z)$ is simply the headcount ratio (40). The individual poverty-evaluation function in this case is illustrated in the top-left panel of Figure 15: $p(x, z)$ jumps from 1 to 0 as x goes from less than z to more than z .
- $\theta = 1$. In this special case we find

$$P_{\text{FGT}}^1(\mathbf{x}, z) = \frac{1}{n} \sum_{i=1}^n \frac{g_i}{z}, \quad (46)$$

²⁵These correspond to the case $p(x_i, z) = [g_i/z]^\theta$; theorem 4 also allows for an “absolute” version of this class where $p(x_i, z) = g_i^\theta$.

²⁶Foster et al. (1984) originally specified (45) with $\theta \geq 0$; but in the special case $\theta = 0$ the FGT index is discontinuous (there is a jump as x_i approaches z and the Ebert and Moyes (2002) result requires the index to be continuous).

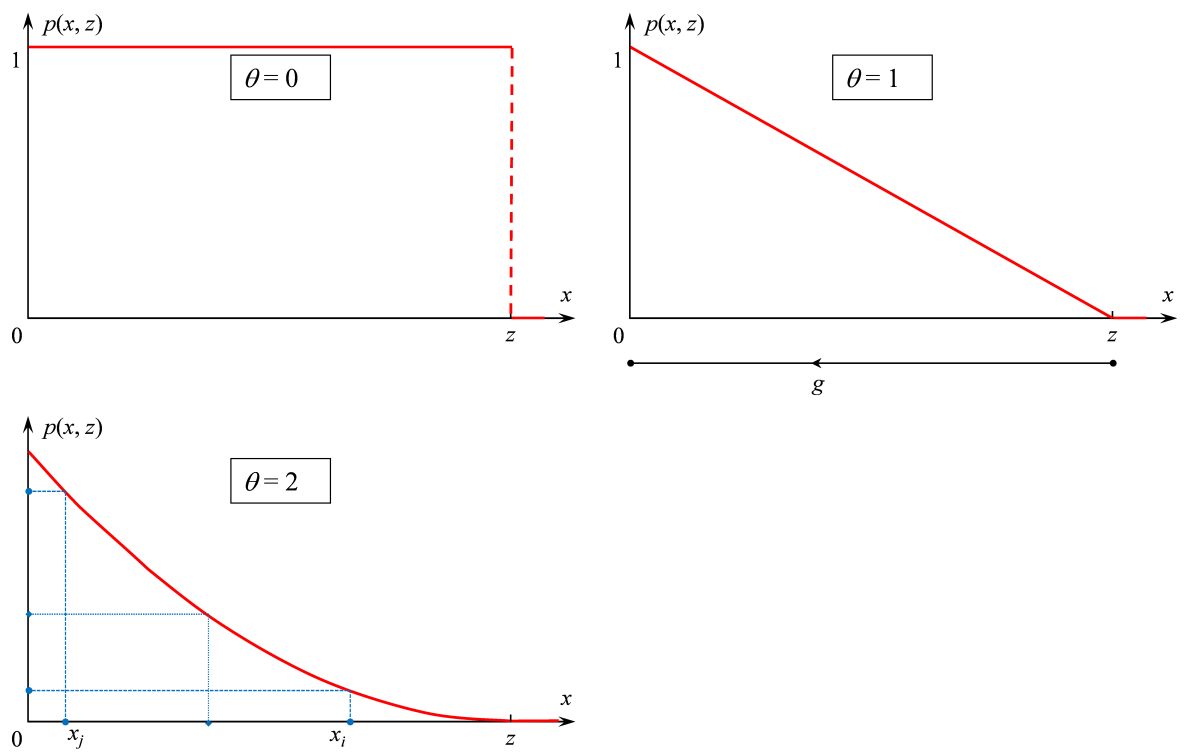


Figure 15: Poverty evaluation, FGT indices

the product of two intuitive poverty indices, the headcount ratio (40) and the average income gap ratio (42). In fact $P_{\text{FGT}}^1(\mathbf{x}, z)$ can be taken as yet another convenient way of normalising the average income gap. The individual poverty-evaluation function is illustrated in the top-right panel of Figure 15: the individual poverty gap g is measured in the opposite direction to x (and is zero for $x > z$); $p(x, z)$ is proportional to g .

- $\theta = 2$. This is an example of where the poverty measure is sensitive to inequality amongst the poor – it satisfies the transfer principle (Axiom 3). To see this check the individual poverty evaluations on the vertical axis for fairly poor Irene $p(x_i, z)$ and very poor Janet $p(x_j, z)$; now suppose Irene’s and Janet’s incomes are equalised and check the poverty evaluation of their averaged incomes on the vertical axis; clearly

$$p\left(\frac{1}{2}[x_i + x_j], z\right) < \frac{1}{2}[p(x_i, z) + p(x_j, z)]$$

and so partial equalisation amongst the poor reduces measured poverty.

Poverty contours

The contours of any member of the class (45) can of course be represented in the income simplex, just as we did with inequality contours. Figure 16 shows an example with $\theta = 1$: the central triangle is determined by the given poverty line z . Within that triangle everyone is out of poverty; outside

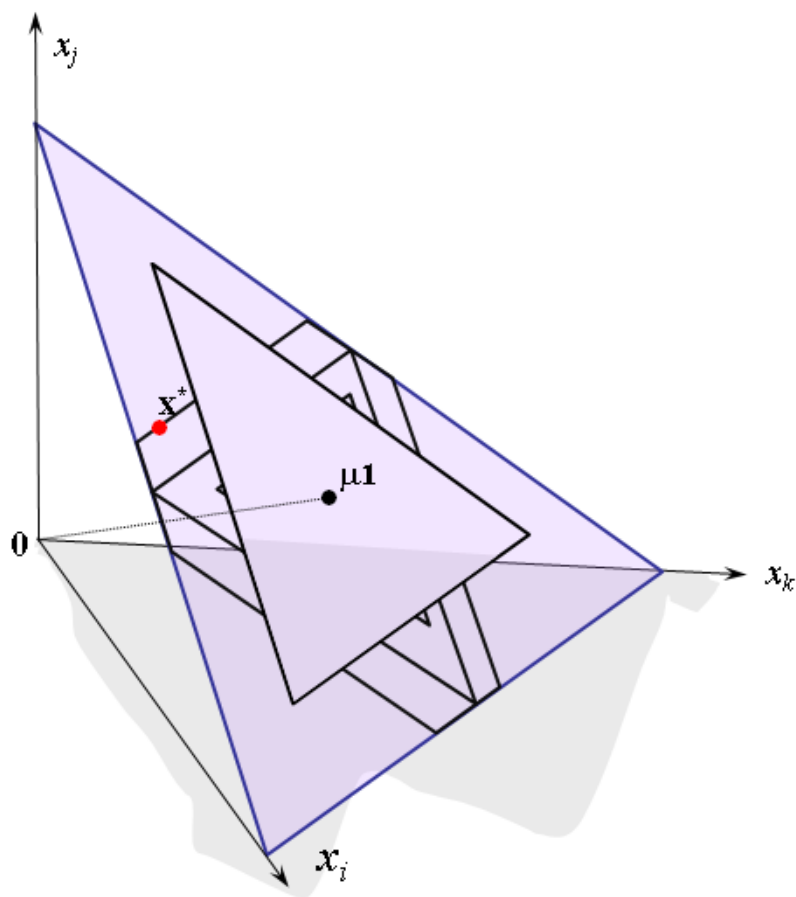


Figure 16: Poverty contours, $\theta = 1$

the triangle some are in poverty and the figure depicts the contours of the average income gap (the inverted triangle shapes). At the specific income distribution \mathbf{x}^* Irene and Janet are out of poverty, but Karen is below the poverty line.

5.3 Other approaches

Axiom 9 rules out the possibility that individual poverty evaluations depend on the individual's position in the distribution. So, clearly, if we drop the restriction imposed by this axiom other possibilities for poverty measurement are opened up. This was part of the insight of the Sen (1976) contribution which explicitly designed the transfer principle into the construction of the poverty measure. Let \mathbf{x}_P be the income distribution among the poor so that $I_{\text{Gini}}(\mathbf{x}_P)$ represents inequality among the poor. Then the Sen index can be represented as

$$P_{\text{Sen}}(\mathbf{x}, z) = P_{\text{FGT}}^0(\mathbf{x}, z) I_{\text{Gini}}(\mathbf{x}_P) + P_{\text{FGT}}^1(\mathbf{x}, z) [1 - I_{\text{Gini}}(\mathbf{x}_P)]. \quad (47)$$

The contour map can be found by replacing the inverted triangles in Figure 16 by hexagonal shapes (Cf Figures 5 (b) and 7 in the case of inequality). Of course one could go further and replace the contours in Figure 16 with contours derived from some alternative principles: for example we might wish to invoke social-welfare principles similar to those invoked in section 4.1 for inequality (Blackorby and Donaldson 1980a, Chakravarty 1983) but it is not clear how much additional insight is to be obtained over the basic FGT structure (45).

6 Inequality and poverty: ranking

Section 2.3 briefly introduced the idea of ranking (in contrast to ordering) as a measurement concept. Let us now see why a tool that may yield an ambiguous answer can still be useful as a method of comparing income distributions. The key idea is that of *distributional dominance* which can take a variety of forms, discussed in section 6.2;²⁷ first let us consider the common thread of analysis.

6.1 A meta-analysis

There are several ranking approaches that are of potential interest in studying inequality and poverty. For this reason it is useful to consider a meta analysis that will enable us to see how any particular ranking argument may be used.

Suppose that for each person i in the distribution we define some concept of individual status q_i which may depend on i 's own income, on the rest of the distribution and possibly on other information. Then consider the problem of comparing two distributions with associated status vectors (q_1, q_2, \dots, q_n) and $(q'_1, q'_2, \dots, q'_n)$ for a population of size n . A natural way of doing this is to introduce the idea of dominance: given the definition of status we say that (q_1, q_2, \dots, q_n) *dominates* $(q'_1, q'_2, \dots, q'_n)$ if it is true that

²⁷These are “type-2” measures, in the terminology of section 2.3.

$$q_{(1)} \geq q'_{(1)},$$

$$q_{(2)} \geq q'_{(2)},$$

...

$$q_{(n)} \geq q'_{(n)}$$

and for at least one person $q_{(i)} > q'_{(i)}$, where the subscript (1) denotes the person with the lowest income, (2) denotes the person with the second smallest income, and so on. Notice that the each status vector is separately arranged according to income to carry out this comparison: so if one income vector were formed just by rearranging the components of the other, the two resulting status vectors would be regarded as equivalent.

A typical way of checking for dominance is to draw a graph of $q_{(k)}$ against $\frac{k}{n}$ for two distributions in a population of size n . Clearly this defines a set of n points rather than a curve; so the convention is to define an extra starting point $(0, \underline{q})$ (where \underline{q} is the smallest possible value of status, often assumed

to be zero)²⁸ and to plot the $n + 1$ points $(A_0, A_1, A_2, \dots, A_n)$ given by

$$\left. \begin{aligned} A_0 &= (0, \underline{q}), \\ A_1 &= \left(\frac{1}{n}, q_{(1)}\right), \\ A_2 &= \left(\frac{2}{n}, q_{(2)}\right), \\ &\dots \dots \\ A_{n-1} &= \left(\frac{n-1}{n}, q_{(n-1)}\right), \\ A_n &= (1, q_{(n)}) \end{aligned} \right\} \quad (48)$$

connected by straight lines. Comparing the curves for (q_1, \dots, q_n) and (q'_1, \dots, q'_n) immediately reveals whether or not there is dominance.

If the distributions being compared really do have the same population then the above graphing convention is just a convention; but if the status vectors being compared are from different-sized populations then the connected curves are often important in checking whether or not there is dominance.²⁹

Thus far the dominance definition and its associated graph is devoid of interpretation. But this immediately acquires importance once we plug into (48) different versions of status with associated meanings in terms of inequality, social welfare and poverty. Then (48) acts as a graph template for differ-

²⁸For second-order dominance (see section 6.2) $\underline{q} = 0$, but for first-order dominance \underline{q} can be taken as the lower bound to the support of the income distribution.

²⁹Example. If the first distribution has $n = 2$, then it is represented by three points $A_0 = (0, \underline{q})$, $A_1 = (\frac{1}{2}, q_{(1)})$, $A_2 = (1, q_{(2)})$; if the second distribution has $n = 4$, then it is represented by five points $A'_0 = (0, \underline{q}')$, $A'_1 = (\frac{1}{4}, q'_{(1)})$, $A'_2 = (\frac{1}{2}, q'_{(2)})$, $A'_3 = (\frac{3}{4}, q'_{(3)})$, $A'_4 = (1, q'_{(4)})$. To check for dominance you need to check whether A'_1 lies above or below a point half-way between A_0 and A_1 and whether A'_3 lies above or below a point half-way between A_1 and A_2 .

ent dominance concepts associated with important results in the literature. In section 6.2 we examine the principal dominance concepts that are useful in the inequality-measurement and poverty-measurement problems and then in section 6.3 we examine the results associated with these concepts.

6.2 First and second order dominance

Dominance criteria are widely known from the literature on risk taking where they are usually known as stochastic dominance criteria. In the analysis of income distributions no uncertainty is involved (so the “stochastic” is superfluous) and the results are extended to include applications in the context of inequality and poverty.

First-order dominance.

Perhaps the simplest application is to associate each person’s status with his/her income so that, for $i = 1, 2, \dots, n$:

$$\begin{aligned} q_{(i)} &= x_{(i)}, \\ q'_{(i)} &= x'_{(i)}. \end{aligned}$$

Here the q values are simply the quantiles of the distribution – so if $i/n = 0.25$ then $q_{(i)}$ is the lower quartile, if $i/n = 0.50$ then $q_{(i)}$ is the median and so on. Plugging this into (48) gives Pen’s (1974) Parade diagram which, because we have sorted the x -values into ascending order, must be an upward sloping

graph.

We can use the Parade to give us a simple and fairly intuitive result on welfare: if \mathbf{x} quantile-dominates \mathbf{x}' then we must have $W(\mathbf{x}) > W(\mathbf{x}')$ for any monotonic social-welfare function W (Saposnik 1981, 1983); we can also obtain from the quantiles a number of intuitive indices that have been used to characterise the spread of distributions, such as the interquartile range, the 90-10 ratio and so on. However, these first-order concepts do not conform to the key axioms outlined in section 3.1; the principal contributions to the theory of inequality measurement come from the concept of second-order dominance.

Second-order dominance.

Let us first introduce the idea of a normalised income cumulation, where one sums the first k incomes in ascending order, $x_{(1)} + x_{(2)} + \dots + x_{(k)}$, and then divides by the population size:

$$c_k := \frac{1}{n} \sum_{i=1}^k x_{(i)}, \quad k = 1, 2, \dots, n. \quad (49)$$

In this case we associate status with income cumulation. By setting $A_0 = (0, 0)$ and $q_{(k)} = c_k$ in the template graph (48) we obtain the so-called *Generalised Lorenz curves* (GLC). Because the x -values are arranged in ascending order (so that one is adding into the sum in (49) ever larger incomes as k increases) the GLC must be a convex curve. Also, by definition, $c_n = \mu(\mathbf{x})$.

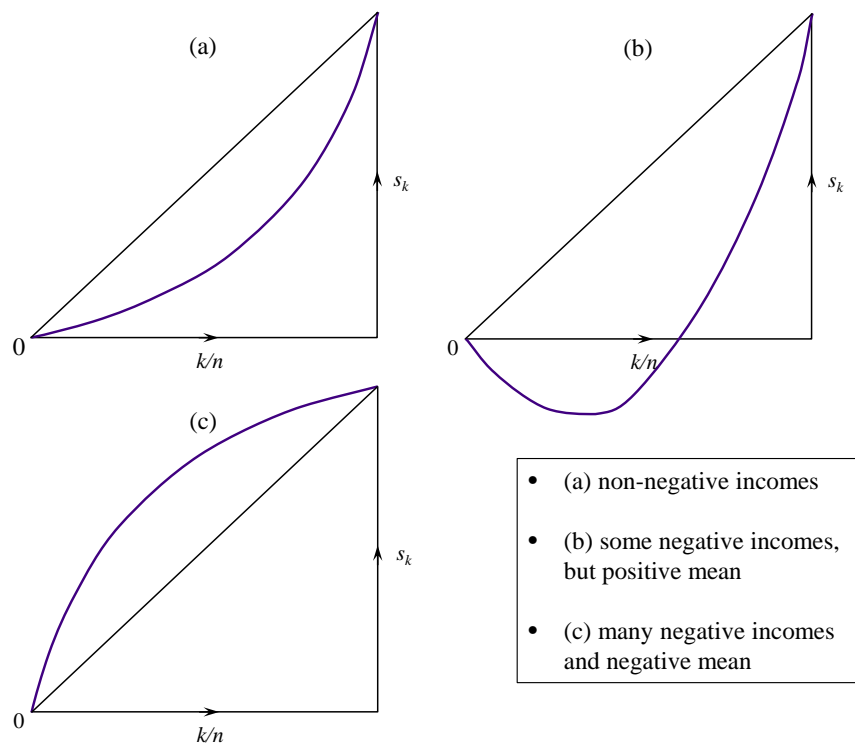


Figure 17: Lorenz curves for different types of distribution

Second-order dominance has a great significance for a number of results on social welfare, inequality comparisons and poverty comparisons which are discussed in section 6.3. Some of these results follow from simple extensions of the basic second-order dominance concept which we will review first.

Extensions of second-order dominance.

Associated with the GLC are a number of related concepts that are widely used in the income-distribution literature.

- *(Relative) Lorenz Curves.* If we divide the income cumulations by the population mean $\mu(\mathbf{x})$ we get the income shares

$$s_k := \frac{c_k}{\mu(\mathbf{x})} = \frac{1}{n\mu(\mathbf{x})} \sum_{i=1}^k x_{(i)}, \quad k = 1, 2, \dots, n. \quad (50)$$

More precisely s_k is the share in total income of the bottom $100k/n$ percent of the population. Plugging the shares (50) into the graph template (48) yields a persuasive pictorial representation of income distribution, the Lorenz curve (the “Relative” tag is sometimes added to distinguish it from Generalised and Absolute Lorenz curves). The Lorenz or shares ranking provides an intuitive method of inequality comparison.: if all incomes are non-negative the Lorenz curve will look like Figure 17 (a), an increasing, convex curve from (0,0) to (1,1); if the curve for distribution \mathbf{x}' lies everywhere above that for \mathbf{x} then it is clear that the income-share of the bottom 10 percent, the bottom 20 percent... of the population is higher in \mathbf{x}' than in \mathbf{x} , so that \mathbf{x}' appears unambiguously more equal than \mathbf{x} ; the limiting case of perfect equality is represented by the diagonal line. There is a potential problem if x can take negative values, a situation that is particularly important when dealing with the inequality of wealth: if there are just a few negative values so that we still have $\mu(\mathbf{x}) > 0$ then the Lorenz curve looks like Figure 17 (b); if there are so many negative incomes that $\mu(\mathbf{x}) < 0$ then the Lorenz curve bends “the wrong way” as in Figure 17 (c); and

if $\mu(\mathbf{x}) = 0$ the Lorenz curve is undefined (Amiel et al. 1996).

- *Absolute Lorenz curves.* In this case the counterpart to (50) is the absolute Lorenz curve (ALC). For this we let

$$q_{(k)} = c_k - \frac{k}{n} \mu(\mathbf{x}) \quad (51)$$

in the graph template. It is clear that, in contrast to the relative Lorenz curve, the resulting convex curve starts from (0,0), dips below the horizontal axis and ends at (1,0). The ALC is a particularly convenient tool for comparing distributions where a large proportion of the incomes are negative.

- *Cumulative Complaint Contours.* In the context of the reference-point approach to inequality measurement, the Cumulative Complaint Contour (CCC) can be used to provide a visual representation of distance-based “Temkin” inequality. Using the definition of complaint or distance in (19) calculate the cumulative complaints:

$$D_k := \frac{1}{n} \sum_{i=1}^k d_i, \quad k = 1, 2, \dots, n. \quad (52)$$

Putting $q_{(k)} = D_k$ in the template (48) we obtain the CCC. It is clear that this is an increasing concave curve because the ordering by income ensures that the largest distances are included first in the sum (52).

- *TIP curves.* The discussion in section 5 focused on three aspects of

poverty that can be summarised as the Three “I”s of Poverty (Jenkins and Lambert 1997): Incidence (as indicated by the headcount ratio), Intensity (as indicated by the average income gap) and Inequality among the poor. These three features are evident if we construct cumulative poverty gaps in the same way as we constructed income cumulations in (49), namely

$$G_k := \frac{1}{n} \sum_{i=1}^k g_i, \quad k = 1, 2, \dots, n. \quad (53)$$

where $g_i = \max \{z - x_{(i)}, 0\}$. Setting $q_{(k)} = G_k$ in the template (48) we obtain the TIP curve. This is a concave curve (for the same reason as given for the CCC) and is increasing up to k^* , the first value of k such that $g_k = 0$, and constant thereafter; k^*/n is the headcount ratio (the Incidence component of the TIP); G_{k^*} is the (non-normalised) average income gap (the Intensity component) and the curvature of the TIP represents the Inequality of the income gaps.

We now turn to the results that can be obtained from each of these graphs.

6.3 Significance for welfare, inequality and poverty

Welfare and distributional rankings.

To show how the second-order dominance criteria can be interpreted in terms of social welfare we need to do a small adaptation to the statement of Axiom 3 in order to obtain a welfare version of this principle – just substitute the

expression “ $W(\mathbf{x}) > W(\mathbf{x}')$ ” for “ $I(\mathbf{x}) < I(\mathbf{x}')$ ” in the statement of Axiom 3.

The first key result, originally formulated by Atkinson (1970), links (relative) Lorenz curves and SWFs that satisfy the key principle of inequality measurement:³⁰

Theorem 5 *For any $\mathbf{x}, \mathbf{x}' \in X^n(\mu)$, \mathbf{x} Lorenz-dominates \mathbf{x}' if and only if $W(\mathbf{x}) > W(\mathbf{x}')$ for all welfare functions satisfying the transfer principle.*

So the Lorenz curve is intimately connected with welfare judgments that respect the transfer principle. However, it may be regarded as a weakness that this result restricts attention solely to distributions for which $\mu(\mathbf{x}) = \mu(\mathbf{x}')$: it remains silent on cases where the distributions compared differ in terms of total income.

Our second key result deals with the income-change issue. Recall that we distinguished a number of possibly interesting cases, depicted in Figure 11. Panel (a) of that figure illustrates the case of monotonicity whereby $W(\mathbf{x} + d\mathbf{x}) > W(\mathbf{x})$ for any $d\mathbf{x} > 0$. Panel (b) illustrates both the principle of scale growth, whereby $W(\lambda\mathbf{x}) > W(\mathbf{x})$ for any $\lambda > 1$, and the principle of uniform income growth, whereby $W(\mathbf{x} + \delta\mathbf{1}) > W(\mathbf{x})$ for any $\delta > 0$. Then we may state (Shorrocks 1983):

³⁰ Atkinson (1970) established Theorem 5 for a more restrictive class of welfare function, where W can be expressed in the additive form (28). However it can easily be broadened to the version stated here – see Arnold (1987), Dasgupta et al. (1973), Fields and Fei (1978), Kolm (1966, 1968, 1969), Kurabayashi and Yatsuka (1977), Rothschild and Stiglitz (1973). These approaches are based on the classic work of Hardy et al. (1934).

Theorem 6 *For any $\mathbf{x}, \mathbf{x}' \in X^n$:*

(a) \mathbf{x} generalised Lorenz-dominates \mathbf{x}' if and only if $W(\mathbf{x}) > W(\mathbf{x}')$ for all SWFs satisfying monotonicity and the transfer principle.

(b) \mathbf{x} relative Lorenz-dominates \mathbf{x}' if and only if $W(\mathbf{x}) > W(\mathbf{x}')$ and $\mu(\mathbf{x}) \geq \mu(\mathbf{x}')$, for all SWFs satisfying the principle of scale growth and the transfer principle.

(c) \mathbf{x} absolute Lorenz-dominates \mathbf{x}' if and only if $W(\mathbf{x}) > W(\mathbf{x}')$ for all SWFs satisfying the principle of uniform income growth and the transfer principle.

Inequality and Poverty.

The essential inequality result follows immediately from Theorem 5. We know that a welfare-based inequality measure can be constructed from any SWF using (24) and (27) and the inequality contours and the welfare contours in $X^n(\mu)$ will be exactly the same. Theorem 5 thus implies that \mathbf{x} Lorenz-dominates \mathbf{x}' if and only if $I(\mathbf{x}) < I(\mathbf{x}')$ for *all* inequality measures satisfying the transfer principle. So, in terms of the hexagon diagrams in Figures 3 and 4, if \mathbf{x} Lorenz-dominates \mathbf{x}' then \mathbf{x} must lie inside the \mathbf{x}' -hexagon; if the Lorenz curves intersect then we have a situation similar to that in Figure 4.

The result obtainable for generalised Lorenz dominance (Theorem 6) can be adapted to yield a nice result for the case of reference-point inequality. If \mathbf{x}' CCC-dominates \mathbf{x} then $I_{\text{Temk}}(\mathbf{x}') > I_{\text{Temk}}(\mathbf{x})$ for all Temkin-type inequality measures based on the BoP reference point and that satisfy the transfer

principle (Cowell and Ebert 2004). A similar result is available for poverty using the TIP curve: if (\mathbf{x}', z) TIP-dominates (\mathbf{x}, z) then $P(\mathbf{x}', z') > P(\mathbf{x}, z')$ for all poverty measures P that satisfy the population principle (Axiom 2) and the transfer principle for poverty measures and for all $z' \leq z$ (Jenkins and Lambert 1997).³¹

7 Conclusion

Inequality and poverty measures in common use can be seen to be founded on a relatively small number of principles. In the main these principles accord well with intuition and can be represented using a relatively small number of mathematical functions. In many cases there is a natural interpretation of inequality measures in terms of welfare economics.

³¹Note that the headcount ratio and the average income gap ratio do not fall into this class of measures.

References

- Amiel, Y. and F. A. Cowell (1992). Measurement of income inequality: Experimental test by questionnaire. *Journal of Public Economics* 47, 3–26.
- Amiel, Y. and F. A. Cowell (1999). *Thinking about Inequality*. Cambridge: Cambridge University Press.
- Amiel, Y., F. A. Cowell, L. Davidovitz, and A. Polovin (2008). Preference reversals and the analysis of income distributions. *Social Choice and Welfare* 30, 305–330.
- Amiel, Y., F. A. Cowell, and A. Polovin (1996). Inequality amongst the kibbutzim. *Economica* 63, S63–S85.
- Amiel, Y., J. Creedy, and S. Hurn (1999). Attitudes towards inequality. *The Scandinavian Journal of Economics* 101, 83–96.
- Arnold, B. C. (1987). *Majorization and the Lorenz Order: A Brief Introduction*. Heidelberg: Springer-Verlag.
- Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory* 2, 244–263.
- Blackorby, C. and D. Donaldson (1978). Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory* 18, 59–80.
- Blackorby, C. and D. Donaldson (1980a). Ethical indices for the measure-

- ment of poverty. *Econometrica* 48, 1053–1060.
- Blackorby, C. and D. Donaldson (1980b). A theoretical treatment of indices of absolute inequality. *International Economic Review* 21, 107–136.
- Bosmans, K. and F. A. Cowell (2010). The class of absolute decomposable inequality measures. *Economics Letters* 109, 154–156.
- Bossert, W. (1988). A note on intermediate inequality indices which are quasilinear means. Discussion Paper 289, Karlsruhe University.
- Bossert, W. and A. Pfingsten (1990). Intermediate inequality: concepts, indices and welfare implications. *Mathematical Social Science* 19, 117–134.
- Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica* 47, 901–920.
- Carlsson, F., D. Daruvala, and O. Johansson-Stenman (2005). Are people inequality averse or just risk averse? *Economica* 72, 375–396.
- Chakravarty, S. R. (1983). Ethically flexible measures of poverty. *The Canadian Journal of Economics* 16, 74–85.
- Chakravarty, S. R. (1990). *Ethical social index numbers*. Berlin: Springer-Verlag.
- Chakravarty, S. R. and D. Mukherjee (1999). Measures of deprivation and their meaning in terms of social satisfaction. *Theory and Decision* 47, 89–100.

- Champernowne, D. G. and F. A. Cowell (1998). *Economic Inequality and Income Distribution*. Cambridge: Cambridge University Press.
- Cowell, F. A. (1977). *Measuring Inequality* (First ed.). Oxford: Phillip Allan.
- Cowell, F. A. (1980). On the structure of additive inequality measures. *Review of Economic Studies* 47, 521–531.
- Cowell, F. A. (2000). Measurement of inequality. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Chapter 2, pp. 87–166. New York: Elsevier Science B. V.
- Cowell, F. A. (2011). *Measuring Inequality* (Third ed.). Oxford: Oxford University Press.
- Cowell, F. A. and U. Ebert (2004). Complaints and inequality. *Social Choice and Welfare* 23, 71–89.
- Cowell, F. A. and C. Fiorio (2011). Inequality decompositions – a reconciliation. *Journal of Economic Inequality* 9, 509–528.
- Cowell, F. A. and E. Flachaire (2014). Inequality with ordinal data. Public Economics Programme Discussion Paper 16, London School of Economics, <http://darp.lse.ac.uk/pdf/IneqOrdinal.pdf>.
- Cowell, F. A. and K. A. Gardiner (2000). Welfare weights. OFT Economic Research Paper 202, Office of Fair Training, Salisbury Square, London.
- Cowell, F. A. and K. Kuga (1981a). Additivity and the entropy concept:

- An axiomatic approach to inequality measurement. *Journal of Economic Theory* 25, 131–143.
- Cowell, F. A. and K. Kuga (1981b). Inequality measurement: an axiomatic approach. *European Economic Review* 15, 287–305.
- Cowell, F. A. and E. Schokkaert (2001). Risk perceptions and distributional judgments. *European Economic Review* 42, 941–952.
- Dalton, H. (1920). Measurement of the inequality of incomes. *The Economic Journal* 30, 348–361.
- Dasgupta, P. S., A. K. Sen, and D. A. Starrett (1973). Notes on the measurement of inequality. *Journal of Economic Theory* 6, 180–187.
- Devooght, K. (2003). Measuring inequality by counting ‘complaints’: theory and empirics. *Economics and Philosophy* 19, 241–263.
- Donaldson, D. and J. A. Weymark (1986). Properties of fixed-population poverty indices. *International Economic Review* 27, 667–688.
- Ebert, U. (1988a). A family of aggregative compromise inequality measure. *International Economic Review* 29(5), 363–376.
- Ebert, U. (1988b). Measurement of inequality: an attempt at unification and generalization. *Social Choice and Welfare* 5, 147–169.
- Ebert, U. and P. Moyes (2002). A simple axiomatization of the Foster-Greer-Thorbecke poverty orderings. *Journal of Public Economic Theory* 4, 455–473.

- Fields, G. S. and J. C. H. Fei (1978). On inequality comparisons. *Econometrica* 46, 303–316.
- Fleurbaey (2015). Equivalent income. In M. D. Adler and M. Fleurbaey (Eds.), *Oxford Handbook of Well-Being And Public Policy*, Chapter 16. Oxford University Press.
- Foster, J. E., J. Greer, and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica* 52, 761–776.
- Gaertner, W. and E. Schokkaert (2012). *Empirical Social Choice: Questionnaire-Experimental Studies on Distributive Justice*. Cambridge University Press.
- Gevers, L., H. Glejser, and J. Rouyer (1979). Professed inequality aversion and its error component. *Scandinavian Journal of Economics* 81, 238–243.
- Glejser, H., L. Gevers, P. Lambot, and J. A. Morales (1977). An econometric study of the variables determining inequality aversion among students. *European Economic Review* 10, 173–188.
- Hardy, G., J. Littlewood, and G. Pólya (1934). *Inequalities*. London: Cambridge University Press.
- Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy* 61, 434–435.
- Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics and inter-

- personal comparisons of utility. *Journal of Political Economy* 63, 309–321.
- Harsanyi, J. C. (1978). Bayesian decision theory and utilitarian ethics. *American Economic Review* 68, 223–228.
- Herwaarden, F. V., A. Kapteyn, and B. V. Praag (1977). Twelve thousand individual welfare functions: A comparison of six samples in Belgium and the Netherlands. *European Economic Review* 9, 283–300.
- Hochman, H. and J. D. Rodgers (1969). Pareto-optimal redistribution. *American Economic Review* 59, 542–557.
- Jenkins, S. P. and P. J. Lambert (1997). Three ‘I’s of poverty curves, with an analysis of UK poverty trends. *Oxford Economic Papers* 49, 317–327.
- Kolm, S.-C. (1964). *Les Fondements de l’Économie Publique: Introduction a la théorie du rôle économique de l’état*. Paris: IFP.
- Kolm, S.-C. (1966). *Les Choix Financiers et Monétaires: Théorie et Technique Modernes*. Paris: Dunod.
- Kolm, S.-C. (1968). The optimal production of social justice. In J. Margolis and H. Guitton (Eds.), *Économie Publique*. Paris: CNRS.
- Kolm, S.-C. (1969). The optimal production of social justice. In J. Margolis and H. Guitton (Eds.), *Public Economics*, pp. 145–200. London: Macmillan.

- Kolm, S.-C. (1976a). Unequal inequalities I. *Journal of Economic Theory* 12, 416–442.
- Kolm, S.-C. (1976b). Unequal inequalities II. *Journal of Economic Theory* 13, 82–111.
- Kroll, Y. and L. Davidovitz (2003). Inequality aversion versus risk aversion. *Economica* 70, 19 – 29.
- Kundu, A. and T. R. Smith (1983). An impossibility theorem on poverty indices. *International Economic Review* 24, 423–434.
- Kurabayashi, Y. and A. Yatsuka (1977). Redistribution of income and measures of income inequality. In T. Fuji and R. Sato (Eds.), *Resource Allocation and Division of Space*. New York: Springer-Verlag.
- Lambert, P. J. (2001). *The Distribution and Redistribution of Income* (Third ed.). Manchester: Manchester University Press.
- Layard, R., S. Nickell, and G. Mayraz (2008). The marginal utility of income. *Journal of Public Economics* 92, 1846–1857.
- Pen, J. (1974). *Income Distribution* (second ed.), Chapter 3, pp. 48–59. London: Allen Lane, The Penguin Press.
- Pirttilä, J. and R. Uusitalo (2010). A ‘leaky bucket’ in the real world: Estimating inequality aversion using survey data. *Economica* 77, 60–76.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, Massachusetts: Harvard

University Press.

Rawls, J. (1999). *A Theory of Justice* (Revised ed.). Cambridge, Massachusetts: Harvard University Press.

Rothschild, M. and J. E. Stiglitz (1973). Some further results on the measurement of inequality. *Journal of Economic Theory* 6, 188–203.

Russell, R. R. (1985). A note on decomposable inequality measures. *Review of Economic Studies* 52(4), 347–352.

Saposnik, R. (1981). Rank-dominance in income distribution. *Public Choice* 36, 147–151.

Saposnik, R. (1983). On evaluating income distributions: Rank dominance, the Suppes-Sen grading principle of justice and Pareto optimality. *Public Choice* 40, 329–336.

Sen, A. K. (1976). Poverty: An ordinal approach to measurement. *Econometrica* 44, 219–231.

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

Shorrocks, A. F. (1983). Ranking income distributions. *Economica* 50, 3–17.

Shorrocks, A. F. (1984). Inequality decomposition by population subgroups. *Econometrica* 52, 1369–1385.

Temkin, L. S. (1986). Inequality. *Philosophy and Public Affairs* 15, 99–121.

- Temkin, L. S. (1993). *Inequality*. Oxford: Oxford University Press.
- Theil, H. (1967). *Economics and Information Theory*, Chapter 4, pp. 91–134. Amsterdam: North Holland.
- Thurow, L. C. (1971). The income distribution as a pure public good. *Quarterly Journal of Economics* 85, 327–36.
- Toyoda, T. (1975). Inequalities of income distributions: their comparisons and inequality measures. *Kokumin Keizai* 134, 15–41. In Japanese.
- Van Batenburg, P. C. and B. M. S. Van Praag (1980). The perception of welfare inequality: a correction note. *European Economic Review* 13(3), 259–261.
- Van Praag, B. M. S. (1977). The perception of welfare inequality. *European Economic Review* 10(11), 189–207.
- Van Praag, B. M. S. (1978). The perception of income inequality. In W. Krelle and A. F. Shorrocks (Eds.), *Personal Income Distribution*. North Holland, Amsterdam.
- Watts, H. W. (1968). An economic definition of poverty. In D. P. Moynihan (Ed.), *Understanding Poverty*, Chapter 11, pp. 316–329. New York: Basic Books.
- Weymark, J. (2015). Social welfare functions. In M. D. Adler and M. Fleurbaey (Eds.), *Oxford Handbook of Well-Being And Public Policy*, Chapter 5. Oxford University Press.

- Yitzhaki, S. (1979). Relative deprivation and the Gini coefficient. *Quarterly Journal of Economics* 93, 321–324.
- Zagier, D. (1982). On the decomposability of the Gini coefficient and other indices of inequality. Discussion paper, Institut für Gesellschafts-und Wirtschaftswissenschaften, Universität Bonn.
- Zheng, B. (1997). Aggregate poverty measures. *Journal of Economic Surveys* 11(2), 123–162.
- Zheng, B. (2000). Poverty orderings. *Journal of Economic Surveys* 14, 427–466.