Research Paper No. 2004/10

A Re-scaled Version of the Foster-Greer-Thorbecke Poverty Indices based on an Association with the Minkowski Distance Function

S. Subramanian

February 2004

Abstract

This note advances a family of poverty measures, \( \Pi_\alpha \), which are derived as simple, normalized Minkowski distance functions. The \( \Pi_\alpha \) indices turn out to be the \( \alpha \)th roots of the corresponding Foster, Greer and Thorbecke \( P_\alpha \) indices. The re-calibration of \( P_\alpha \) in terms of \( \Pi_\alpha \) could have certain possible advantages, which are reviewed in the note. While the \( \Pi_\alpha \) indices are not decomposable in the ordinarily understood sense of that term, they are amenable to the completely general decomposition procedure advanced by Shorrocks (‘Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value’) and discussed, here, as an application in the poverty context.

Keywords: poverty measures, decomposition

JEL classification: I13
Acknowledgements

A first draft of this paper was written in the course of a Visiting Fellowship at UNU-WIDER over April-June 2003. I would like to thank, without implicating, James Foster, Tony Shorrocks, and D. Jayaraj for helpful discussions of the subject of this note, and Adam Swallow for seeing it through.
1 Introduction

This is a simple note which addresses, by way of a possibly useful curiosum, a few simple issues in the measurement of poverty. First, a family of poverty indices, $\Pi_\alpha$ ($\alpha \geq 1$), is derived in a very elementary way as a set of normalized (Minkowski) distance functions. The indices in this family turn out, simply but interestingly, to be strictly concave transforms of the corresponding—and well-known—Foster, Greer and Thorbecke (hereafter FGT 1984) class of poverty measures $P_\alpha$. Typically, the magnitudes of the $P_\alpha$ indices—especially for ‘larger’ values of $\alpha$—are somewhat ‘inconveniently’ small; and the effect of the re-scaling, or ‘re-cardinalization’, of the $P_\alpha$ measures in terms of the $\Pi_\alpha$ measures is to provide a set of indices of more tractably larger magnitude. The $\Pi_\alpha$ indices inherit most of the important properties of the corresponding ‘parental’ $P_\alpha$ indices, except that the former are not decomposable in the FGT sense. However, the $\Pi_\alpha$ measures are amenable to a completely general decomposition procedure proposed by Shorrocks (1999), and this—a second issue—is also discussed.

2 Poverty measures and their properties

For every positive integer $n$, let $X_n$ be the set of all non-negative $n$-vectors of income $x = (x_1, x_2, \ldots, x_n)$, where $x_i$ is the income of the $i$th individual ($i = 1, \ldots, n$). $X$ will be taken to be the set $\bigcup_{n=1}^{\infty} X_n$. The poverty line, which separates the poor from the non-poor, is a strictly positive level of income denoted by $z$. $R$ is the set of reals and $S$ the set of positive reals. A poverty index is a mapping $P: X \times S \rightarrow R$ such that, for every combination of $x$ and $z$ in its domain, $P$ specifies a unique real number describing the extent of poverty in the regime $(x, z)$. For every $(x, z) \in X \times S$, $q(x, z)$ will stand for the number of poor persons, that is, for the cardinality of the set $\{i \mid x_i < z\}$.

A number of axioms for poverty measurement have been advanced in the literature. The following is a quick and informal description of some salient axioms. Focus requires the poverty index to be invariant with respect to increases in non-poor incomes; symmetry requires that the personal identities of individuals should not matter in assessing the extent of poverty; normalization requires that the poverty index be bounded from below by zero (corresponding to the case of “no poverty”); continuity requires the poverty index to be continuous on the sub-vector of poor incomes; scale-invariance requires that for all $(x, z) \in X \times S$, $P(\beta x, \beta z) = P(x, z)$ for any positive scalar $\beta$; replication-invariance requires that for all $(x, z) \in X \times S$ and any positive integer $k$, $P(kx, z) = P(x, z)$; monotonicity requires that, other things equal, an increase in any poor person’s income should cause poverty to decline; transfer requires that, ceteris paribus, any rank-preserving transfer of income from a poor person to a poorer person should reduce poverty; transfer-sensitivity requires that, everything else remaining the same, a progressive rank-preserving transfer between two poor persons a fixed income apart should cause poverty to decline by more the poorer the pair of persons involved in the transfer (Foster 1984); subgroup consistency (Foster and Shorrocks 1991) requires that, ceteris paribus, an increase in any subgroup’s poverty should increase aggregate poverty; and decomposability, which is a strengthened version of subgroup consistency, requires the poverty index to be capable of being expressed as a population-share weighted sum of subgroup poverty levels (FGT 1984).
3  Measuring poverty as a normalized distance function

We shall here consider a class of poverty indices called the \( \Pi_\alpha \) class. Each member of the \( \Pi_\alpha \) family can be straightforwardly derived as a specific type of normalized distance function (on the construction of distance functions between income distributions, the reader is referred to Shorrock 1982, Ebert 1984, and Chakravarty and Dutta 1987). To this end, consider a non-decreasingly ordered non-negative n-vector of incomes \( x = (x_1,\ldots,x_q,x_{q+1},\ldots,x_n) \), where—to recall—\( x_i \) is the income of the \( i \)th poorest person \((i = 1,\ldots,n)\), \( x_i \leq x_{i+1} \) \((i = 1,\ldots,n-1)\), and \( q \) is the number of poor persons \((\text{so } x_q < z \text{ and } x_{q+1} \geq z)\). Given \( x \) and \( z \), we define the following three n-vectors: \( z = (z,\ldots,z) \), \( x^c = (x_1,\ldots,x_q,z,\ldots,z) \), and \( \mathbf{0} = (0,\ldots,0) \). The vector \( z \) is one in which every income level is just the poverty line income: it is the distribution with the smallest mean income that is compatible with a complete absence of poverty. The vector \( x^c \) is what Takayama (1979) has called a ‘censored’ income vector, and is obtained from the distribution \( x \) by replacing all the non-poor incomes in \( x \) by the poverty line income \( z \). The vector \( \mathbf{0} \) is one in which every person receives zero income: it represents the case of ‘total poverty’. A fairly natural way of measuring the extent of poverty in the income vector \( x \), given that the poverty line is \( z \), would be to identify it with the ‘shortfall’ or ‘distance’ of \( x^c \) from \( z \), that is to say, with the distance of the ‘actual’ situation from one that corresponds to the most efficiently achieved ‘ideal’ of ‘no poverty’. (Implicit in this interpretation is respect for the focus axiom of poverty measurement which treats the magnitude of non-poor incomes as being irrelevant for the assessment of the extent of poverty associated with an income vector). To obtain a normalized picture of the extent of poverty, all one has to do is to divide the distance of \( x^c \) from \( z \) by the distance of \( \mathbf{0} \) from \( z \): the latter distance represents the case of ‘maximal deficit’, that is to say, the distance of the ‘total poverty’ situation from the ‘no poverty’ situation. What remains to be done is to specify the form of the relevant distance function. A very natural set of candidates, in this context, is the well-worn Minkowski class of \( \alpha \)-distance functions, written as \( D_\alpha^M \), of which the ‘city-block’ function and the Euclidean norm are special cases, realized for \( \alpha = 1 \) and \( 2 \) respectively (see Wilson and Martinez 1997). Given any two vectors \( a \) and \( b \) belonging to n-dimensional Euclidean space, the distance between the two, in terms of the Minkowski \( \alpha \)-distance functions, is given by

\[
D_\alpha^M(a,b) = \left( \sum_{i=1}^n |a_i-b_i|^\alpha \right)^{1/\alpha}, \quad \alpha \geq 1
\]

In view of the preceding discussion, and given \( x \) and \( z \), a fairly straightforward expression for a class of normalized distance measures of poverty—\( \Pi_\alpha \)—would be given by

\[
\Pi_\alpha(x;z) = D_\alpha^M(z,x^c)/D_\alpha^M(z,\mathbf{0}), \quad \alpha \geq 1
\]

Given (2), (3) can be written, with a little bit of manipulation, as

\[
\Pi_\alpha(x;z) = \left( \frac{1}{n} \sum_{i=1}^q \left( \frac{z-x_i}{z} \right)^\alpha \right)^{1/\alpha}, \quad \alpha \geq 1
\]

The FGT (1984) family of poverty indices \( P_\alpha \) \((\alpha \geq 0)\) is given by

\[
P_\alpha(x;z) = \left( \frac{1}{n} \sum_{i=1}^q \left( \frac{z-x_i}{z} \right) \right)^\alpha
\]
Comparing (3) and (4), it is immediate that, for all \((x;z)\in X\times S\)

\[
\Pi_\alpha(x;z) = [P_\alpha(x;z)]^{1/\alpha}, \quad \alpha \geq 1
\]

It thus turns out that each member of the family of indices we have called \(\Pi_\alpha\) is just the \(\alpha\)th root of the corresponding member of the \(P_\alpha\) family of indices. This is in itself a possibly small, but nevertheless, interesting result. Each member of the \(\Pi_\alpha\) class is just a continuous, increasing and concave (strictly concave for \(\alpha > 1\)) transform of the counterpart member of the \(P_\alpha\) class: it should not be surprising that the two classes of indices share their respective properties. (Indeed, \(\Pi_1\) and \(P_1\) are the same index).

Specifically (see FGT 1984), the \(\Pi_\alpha\) indices like the \(P_\alpha\) indices satisfy focus, symmetry, normalization, continuity, scale- and replication-invariance, monotonicity, and subgroup consistency for all \(\alpha \geq 1\); transfer for all \(\alpha > 1\); and transfer-sensitivity for all \(\alpha > 2\).

Where the two sets of indices part company is in the matter of satisfying the decomposability property: every member of the \(P_\alpha\) family is decomposable, while no member of the \(\Pi_\alpha\) family is. We shall return to this issue later.

Setting aside for the moment the question of decomposability, are there any reasons for favouring the \(\Pi_\alpha\) class over the \(P_\alpha\) class? One can think of two reasons, each of which has something to commend it, though neither may be unqualifiedly compelling. The first has to do with a matter of interpretation: motivationally, the \(\Pi_\alpha\) family of indices makes a direct appeal to intuition by presenting a poverty measure as a normalized distance measure, while the \(P_\alpha\) family presents the poverty measure as a weighted sum of poverty-gaps, with the weights assuming the form of the poverty gaps themselves raised to a non-negative power, the basis for which is open to the (mild) criticism of a degree of arbitrariness. The second reason has to do with what, very informally, one may call ‘the psychology of numbers’. Take an index like \(P_3\), for instance. Inasmuch as it satisfies a number of desirable properties, including and in particular that of transfer-sensitivity, one would imagine that it would be widely employed in empirical work. The fact that it is not, one strongly suspects, is because, typically, \(P_3\) yields ‘very small’ values when computed for actual empirical distributions of income. Consider, for instance, the vector \(u = (7,8,9,10,10,10,10,10,10,10)\) in a situation where \(z = 10\), so that \(q = 3\) and \(n = 10\). It can be verified that \(P_3(u;z) = 0.0036\): this is an ‘inconveniently small’ number, unlike \(\Pi_3(u;z)\) which yields a ‘healthier-looking’ value of 0.1561. Indeed, this could be a matter of some practical (computational) import, as reflected in the following numerical example. Assuming that \(z = 10\) and \(n = 10\), let \(r\), \(s\) and \(t\) be three income vectors given, respectively, by \(r = (6,7,8,9,10,10,10,10,10,10)\), \(s = (6,7,1,7,9,9,10,10,10,10,10)\), and \(t = (6,7,8,1,8,9,10,10,10,10,10)\). It can be seen that each of \(s\) and \(t\) is derived from \(r\) through a progressive transfer, and transfer sensitivity would require that \(P(s;z) < P(t;z)\). Suppose we measure poverty correct to four decimal places. Then, it can be verified that \(P_3(s;z) = P_3(t;z) = 0.0099\): \(P_3\) is unable to differentiate between \(s\) and \(t\), and one has to proceed to the fifth decimal place in order to secure the required discrimination \((P_3(s;z) = 0.00987 < P_3(t;z) = 0.00992)\); on the other hand, \(\Pi_3\) does the job when poverty is measured correct to four decimal places \((\Pi_3(s;z) = 0.2145 < \Pi_3(t;z) = 0.2149)\).

The question remains whether the instrumental advantages discussed above in favour of \(\Pi_\alpha\) outweigh the fact that no member of this class is decomposable, as that term is ordinarily understood. It may be as well here to formally define ‘decomposability’ after
the FGT (1984) fashion. Let \( M = \{1, \ldots, j, \ldots, m\} \) be a set of \( m \) mutually exclusive and exhaustive subgroups into which a population of \( n \) individuals is partitioned. Let \( P \) be a measure of aggregate poverty for the population, and let \( P_j \) (respectively, \( t_j \)) be the poverty level (respectively, population share) of subgroup \( j \), for every \( j \) belonging to \( M \). The poverty index \( P \) may be written as a function \( f \) of the subgroup poverty levels

\[
P = f(P_1, \ldots, P_j, \ldots, P_m)
\]

Decomposability (or Axiom D), as FGT (1984) define it, requires the function \( f \) in (6) to take the form of a population share weighted sum of subgroup poverty levels, namely \( P = \sum_{j \in M} t_j P_j \). Axiom D is obviously of considerable value in facilitating qualitative and quantitative assessments of the contributions of different subgroups to aggregate poverty, and, as such, holds a very useful key to policy analysis. For instance, it is common practice to interpret the decomposability of the poverty index \( P \) as suggesting that, for every \( j \in M \), the contribution of subgroup \( j \) to total poverty is just \( t_j P_j \). This is clearly useful information, and in particular not the sort of information one might wish to lose in the cause of a poverty index belonging to the \( \Pi_\alpha \) family, despite the two reasons mentioned earlier in favour of this class of indices. This issue is investigated further in the next section.

4 Shorrocks-Shapley decomposition

The necessity of having to choose between indices which belong, respectively, to the \( \Pi_\alpha \) and \( P_\alpha \) families, would not arise if there were available a workable and intuitively acceptable notion of ‘decomposability’ which is not simply defined to coincide with the demand made by Axiom D. Fortunately, such a notion of decomposability does exist, and can be found in Shorrocks’ regretfully unpublished paper (Shorrocks 1999). In this work Shorrocks presents a completely general technique of decomposition, based on Shapley’s (1953) formulation of the ‘Shapley value’ solution to the problem of allocation of output/costs among contributors/beneficiaries within the setting of an n-person cooperative game. Shorrocks’ adaptation of the Shapley value approach leads to a formulation of a general decomposition procedure which will here be called Shorrocks-Shapley decomposition, or SS-decomposition, for short. As applied specifically to the decomposition of a poverty index into subgroup poverty contributions, SS-decomposition can be described along the following lines (this description is heavily dependent on Shorrocks 1999).

We begin by recalling equation (6), in which aggregate poverty \( P \) is written as a function \( f \) of the set of \( m \) subgroup poverty levels \( \{P_j\}_{j \in M} \). Let \( S \) be any subset of \( M \), and define \( F(S) \) to be the value which \( P \) assumes when all subgroup poverty levels \( P_j \), \( j \in S \), have been eliminated from consideration, viz. \( F(S) = f(\{P_j\}_{j \in S}) \), \( S \subseteq M \). Call the pair \( <M,F> \) a model, and given the model \( <M,F> \), a decomposition of the model is a set of real numbers \( \{C_j(M,F)\}_{j \in M} \), such that, for every \( j \in M \), \( C_j \) represents the contribution of subgroup \( j \)’s poverty to total poverty. A decomposition rule \( C \) is a function which, for every model \( <M,F> \), assigns a set of subgroup poverty contributions \( \{C_j(M,F)\}_{j \in M} \). It would be desirable for the decomposition rule to be (a) exact (which is the requirement that the subgroup poverty contributions add up to total poverty, namely \( \sum_{j \in M} C_j(M,F) = F(M) (\equiv P) \)); and (b) symmetric (which is the requirement that subgroup
contributions are invariant with respect to all possible permutations of subgroup labels.

A decomposition rule which satisfies these properties is the Shorrock-Shapley rule, the content of which can be explained as follows.

The basic idea here is to consider the marginal impact of each of the subgroups’ poverty levels when the subgroups are eliminated in sequence. To do this, we first denote by \( \Lambda \) the set of all one-to-one mappings from \( M \) to itself. Every \( \lambda \in \Lambda \) can then be interpreted as a particular sequence in which the subgroups in \( M \) are eliminated. For every \( j \in M \) and \( \lambda \in \Lambda \), let \( S(j, \lambda) \) [respectively, \( S'(j, \lambda) \)] be the set of subgroups left after eliminating all the subgroups that come before subgroup \( j \) [respectively, the set of subgroups left after eliminating all the subgroups that do not come after subgroup \( j \)] according to the sequence represented by \( \lambda \). That is, \( S(j, \lambda) \equiv \{ k \in \lambda(M) \mid k \geq j \} \) and \( S'(j, \lambda) \equiv \{ k \in \lambda(M) \mid k > j \} \). If we take away \( F(S'(j, \lambda)) \) from \( F(S(j, \lambda)) \), then it makes intuitive sense to interpret the remainder as the marginal impact of subgroup \( j \)’s poverty when \( \lambda \) represents the sequence of subgroup elimination; this marginal impact is denoted by \( C_j^\lambda \).

\[ C_j^\lambda = F(S(j, \lambda)) - F(S'(j, \lambda)) \]

It can be checked that \( \sum_{j \in M} C_j^\lambda = F(M) (\equiv P) \), so (7) is an exact decomposition. To ensure symmetry, one can assign equal weight to the marginal impacts derived from different sequences of subgroup elimination, by taking the simple average of the marginal impacts corresponding to the various sequences of elimination. Specifically, with \( m \) subgroups, the cardinality of the set \( \Lambda \) will be \( m! \); and a symmetric (in the sense of ‘[elimination] path independent’) decomposition rule, \( C^S \), is given by

\[ C_j^S(M,F) = \frac{1}{m!} \sum_{\lambda \in \Lambda} C_j^\lambda \]

where \( C_j^\lambda \) as defined in (7). From (8), it can be verified that \( C^S \) is an exact decomposition rule. Finally, since the decomposition rule is exact, it is meaningful to speak of the ‘proportionate contribution’ of each subgroup’s poverty to total poverty: for every \( j \in M \), this proportionate contribution will be given by \( c_j^S \equiv C_j^S/P \). This is the substance of the SS-decomposition procedure, as applied to the present context.

Shorrocks (1999) has shown that the SS-decomposition of any poverty index belonging to the family \( P_\alpha \), will yield a contribution for the \( j \)th subgroup which is given by

\[ C_j^S(M,F) = t_j P_{\alpha j} \]

This, precisely, is also the customary interpretation. What is the contribution of subgroup \( j \)'s poverty to poverty as measured by an index belonging to the \( \Pi_\alpha \) family when the latter is subjected to SS-decomposition? Given (5), and noting that the FGT indices satisfy Axiom D, it is immediate that

\[ \Pi_\alpha(x;z) = (P_\alpha(x;z))^{1/\alpha} = \left( \sum_{j \in M} t_j P_{\alpha j} \right)^{1/\alpha} = \left( \sum_{j \in M} t_j \Pi_{\alpha j} \right)^{1/\alpha}, \alpha \geq 1 \]

For specificity, and by way of illustration, we consider the simple case of \( m = 2 \). There are two possible sequences of subgroup elimination available now: In the first sequence, subgroup 1 is first eliminated, followed by subgroup 2; in the second
sequence, subgroup 2 is first eliminated, followed by subgroup 1. For the first elimination sequence, the contribution of subgroup 1 is given by \( F(M) - F(\{2\}) \), and for the second elimination sequence, the contribution of subgroup 1 is given by \( F(\{1\}) \); averaging over the two sequences gives the SS-contribution of group 1: 
\[
C_1^S = \frac{1}{2}[F(M) - F(\{2\}) + F(\{1\})].
\]
Noting that \( F(M) \equiv \Pi_\alpha \), \( F(\{2\}) \equiv (t_2)^{1/\alpha} \Pi_\alpha \), and \( F(\{1\}) \equiv (t_1)^{1/\alpha} \Pi_\alpha \), we have

\[
(11) \quad C_1^S = \frac{1}{2} [\Pi_\alpha - (t_2)^{1/\alpha} \Pi_\alpha + (t_1)^{1/\alpha} \Pi_\alpha]
\]

By virtue of the exactness property of the SS-decomposition rule, the contribution of subgroup 2 will just be total poverty less subgroup 1’s contribution; or, in view of (11)

\[
(12) \quad C_2^S = \frac{1}{2} [\Pi_\alpha - (t_1)^{1/\alpha} \Pi_\alpha + (t_2)^{1/\alpha} \Pi_\alpha]
\]

Given (11) and (12), we are assured that the use of an index belonging to the \( \Pi_\alpha \) family is in no way compromised by the possibility that we may not be able to infer subgroup contributions to total poverty: the Shorrocks-Shapley approach to decomposition has laid that problem to rest. Indeed, as Shorrocks (1999: 1) says:

[A] problem with conventional procedures is that they often place constraints on the kinds of poverty and inequality indices which can be used. Only certain forms of indices yield a set of contributions that sum up to the amount of poverty or inequality that requires explanation.

Finally, if one is persuaded that there is a case for favouring \( \Pi_\alpha \) over \( P_\alpha \), then this could also have non-trivial implications for between-group allocations of anti-poverty budgets when such allocations are made, in line with a not unreasonable rule-of-thumb, according to the proportionate contributions of subgroups to aggregate poverty. By way of illustration, consider again the case in which \( m = 2 \) and, for specificity, \( \alpha = 2 \). Then, if \( P_{21} = 0.05 \), \( P_{22} = 0.025 \), and \( t_1 = 0.1 \), using (9) and (11) will enable confirmation of the fact that the proportionate contribution of subgroup 1 to total poverty is (a) 18.18 per cent when poverty is measured by \( P_2 \), and (b) 26.08 per cent when poverty is measured by \( \Pi_2 \). The difference between getting an 18 per cent and a 26 per cent share of the budgetary outlay is presumably non-negligible.

5 Summation

To summarize and conclude: a family of poverty indices \( \Pi_\alpha \) which turn out to be just the \( \alpha \)th roots of the corresponding members of the Foster-Greer-Thorbecke family \( P_\alpha \), has been derived as a set of plausible, normalized (Minkowski) distance functions. The re-scaled version \( \Pi_\alpha \) of \( P_\alpha \), unlike the latter, is not ‘decomposable’ in the usually accepted sense of that term, that is, it cannot be written as a population share weighted sum of subgroup poverty levels. This, however, need not be a problem for \( \Pi_\alpha \): Shorrock’s (1999) work on what one may call Shorrooms-Shapley decomposition furnishes a completely general solution to the decomposition problem, in terms of which subgroup contributions to total poverty can be computed even for poverty indices (like \( \Pi_\alpha \)) which cannot be expressed as population share weighted sums of subgroup poverty levels. This frees up poverty analysts to avail themselves of such advantages—
in terms of directness of interpretation and the convenience afforded by a re-calibration of the scale on which poverty is measured—as may be had from a ‘concavification’ of the $P_\alpha$ indices in terms of the $\Pi_\alpha$ indices. Finally, the switch from $P_\alpha$ to $\Pi_\alpha$ could also make for a non-trivial difference at the margin if subgroup contributions to poverty are invested with more than ordinal significance.

References


