Society for the Study of Economic Inequality

Working Paper Series

Combining absolute and relative poverty: Income poverty measurement with two poverty lines

Benoit Decerf

ECINEQ 2020 548

ECINEQ 2020 548
2020 July
www.ecineq.org

Combining absolute and relative poverty: Income poverty measurement with two poverty lines

Benoit Decerf
University of Namur

Abstract

I study income poverty indices in a framework considering two poverty lines: one absolute line capturing subsistence and one relative line capturing social exclusion. In this framework, a set of basic axioms à la Foster and Shorrocks (1991) characterizes the class of hierarchical indices. This is a class of additive indices for which the poverty contribution of any individual depends on both her income and the income standard in her society. The key feature of hierarchical indices is to grant some form of precedence to absolutely poor individuals. These indices always consider that an absolutely poor individual is poorer than an individual who is only relatively poor, regardless of the income standard in their respective societies. Classical indices are not hierarchical, except in trivial cases. As a result, they yield debatable poverty comparisons of societies having different income standards.

Keyword: Income Poverty Measurement, Poverty Line, Relative Poverty, Absolute poverty.

JEL Cassification: D63, I32.

Combining absolute and relative poverty: Income poverty measurement with two poverty lines.*

Benoit Decerf †

June 12, 2020

Abstract

I study income poverty indices in a framework considering two poverty lines: one absolute line capturing subsistence and one relative line capturing social exclusion. In this framework, a set of basic axioms à la Foster and Shorrocks (1991) characterizes the class of *hierarchical* indices. This is a class of additive indices for which the poverty contribution of any individual depends on both her income and the income standard in her society. The key feature of hierarchical indices is to grant some form of precedence to absolutely poor individuals. These indices always consider that an absolutely poor individual is poorer than an individual who is only relatively poor, regardless of the income standard in their respective societies. Classical indices are not hierarchical, except in trivial cases. As a result, they yield debatable poverty comparisons of societies having different income standards.

JEL: D63, I32.
Keywords: Income Poverty Measurement, Poverty Line, Relative Poverty, Absolute poverty.

*Acknowledgments : I express all my gratitude to Francois Maniquet for his many comments and suggestions. I am greatly indebted to John Weymark who helped me improve the exposition of the proof of the main Theorem (even if this proof is still far from the clarity that John Weymark would have achieved, if he had written it entirely by himself, it is at least much better than my own previous version). I am grateful to Mery Ferrando and Martin Van der Linden who extensively commented on earlier versions of this document. I thank all the participants to the 5th ECINEQ conference (2013), the 12th meeting of the Society for Social Choice and Welfare (2014) and later seminars at UCLouvain, KULeuven, U. Laval, U. Bielefeld and the World Bank and in particular C. d'Aspremont, K. Decancq, F. Ferreira, J. Foster, I. Gilboa, S. Jenkins, M. Ravallion, F. Riedel and E. Schokkaert. I am grateful to CORE, my host institution for most of the time during which I worked on this research. All remaining mistakes are of course mine. Fundings from the European Research Council under the EU's Seventh Framework Program (FP/2007-2013) / ERC Grant Agreement n.269831 and from the Fond National de la Recherche Scientifique (Belgium, mandat d'aspirant FC 95720) are gratefully acknowledged.

†University of Namur. benoit.decerf@unamur.be

1 Introduction

There are two different approaches to income poverty measurement: the absolute and the relative approach. An individual is deemed absolutely poor if her income is insufficient to cover her subsistence needs, e.g. being sufficiently nourished or wearing clothes. In a first approximation, the real cost of subsistence does not depend on standards of living. Therefore, the absolute poverty threshold does not dependent on the income standard of the considered society. This is for instance the approach underlying the extreme poverty line of the World Bank, whose threshold is $1.9 per person per day (Ferreira et al., 2016). In turn, an individual is deemed relatively poor if her income is so much smaller than the income standard in her society that she is at risk of social exclusion.[1] The real cost of social participation evolves with standards of living. Therefore, the relative poverty threshold depends on the income standard of the considered society. For instance, most OECD countries use a relative threshold that corresponds to a given fraction of mean or median income.

Unsurprisingly, the main critic raised against any of these two approaches is to ignore either subsistence or social exclusion. On the one hand, any absolute poverty threshold becomes less and less relevant for the identification of the socially excluded as the income standard grows. On the other hand, relative poverty measures often ignore the increase in individual resources that results from growth. Importantly, if a country's growth is such that the income of its poorest citizens becomes sufficient to cover their subsistence needs, its poverty has arguably been reduced, even if these individuals are still socially excluded. Relative measures do not acknowledge such poverty reduction.

There is a need for income poverty measures combining both absolute and relative poverty. Many policymakers, such as the World Bank and the European Commission, aim at reducing both absolute and relative poverty (World Bank, 2015; European Commission, 2015). For such policymakers, using two separate poverty measures, one absolute and one relative, is not a solution. The reason is that two measures would often yield opposing poverty evaluations, in which case no conclusion can be drawn. Such opposing evaluations happen for instance when the income of poor individuals grow, but not as fast as their society's income standard.

The research efforts aimed at combining absolute and relative poverty mostly focus on the design of new poverty lines (Foster, 1998; Ravallion and Chen, 2011; Jolliffe and Prydz, 2017). This strand of research has proposed poverty lines whose threshold depends on the income standard, but is less sensitive to the income standard than the threshold of a relative line. These new lines have the potential to better identify the poor, but they cannot resolve on their own all the limitations associated with poverty measures of either approaches. Importantly, new lines cannot resolve the serious limitation that relative poverty measures should decrease when poor individuals become able to cover their subsistence needs. This

[1] See Ravallion (2008) for a review of the normative foundations of the relative approach to poverty.

limitation is not so much a problem of whom should be identified as poor, but rather a problem of how poor individuals are compared across societies with different income standards. These inter-personal comparisons primarily rely on the index with which a poverty measure is constructed. Indeed, a poverty measure is defined with two components: a poverty line and a poverty index (Sen, 1976). Poverty indices, like the head-count ratio or the poverty-gap ratio, aggregate the contributions to poverty of all individuals in a distribution. The properties of standard poverty indices have been extensively studied under the assumption that the poverty line is *absolute* (Zheng, 1997). Surprisingly, the properties of poverty indices combined with non-absolute poverty lines have never been rigorously studied. Unfortunately, when combined with a poverty line whose threshold depends on the income standard, standard indices provide highly counterintuitive poverty comparisons (Decerf, 2017).

The following example illustrates the counterintuitive comparisons associated with standard indices when comparing Bangladesh and Colombia in 2015. The poverty measure considered is the head-count ratio below the maximum threshold for two poverty lines: the extreme line and the relative line used by the World Bank.[2] Bangladesh has a much smaller income standard than Colombia and their relative thresholds are respectively \$2.4 a day and \$5.5 a day. As these two countries have the same fraction of relatively poor individuals (29%), this measure considers that they have equal poverty.[3] This comparison is highly debatable given that 15% of the population is extremely poor in Bangladesh, against only 5% in Colombia. This debatable comparison is due to the head-count ratio. Under this index, an extremely poor individual earning less than \$1.9 a day in Bangladesh has the same poverty contribution as a Colombian whose income is just below the relative threshold in Colombia, i.e. \$5.5 a day. Importantly, this problem is not solved when using depth-sensitive indices such as the poverty gap ratio. As shown by Decerf (2017), standard FGT indices (Foster et al., 1984) can make the problem even worse: they implicitly consider that some extremely poor in Bangladesh are *less poor* than some only relatively poor in Colombia (see illustration in Section 4).

In this paper, I study poverty indices based on two poverty lines: one absolute line and one relative line. The main contribution is a result characterizing the class of indices satisfying a set of basic axioms à la Foster and Shorrocks (1991). The indices characterized are additive, i.e. they sum the contributions to poverty of all individuals in a distribution. In this class, individual contributions depend on both own income and the income standard. The contribution function determines the inter-personal comparisons that the index makes across different societies. Crucially, the inter-personal comparisons are constrained by the axioms: an absolutely

[2] The relative line used by the World Bank is the Societal poverty line proposed by Jolliffe and Prydz (2017). The Societal line is defined as $z_r(y) = 1 + 0.5\overline{y}$, where $\overline{y}$ denotes median income in \$ a day.

[3] The source of data is Povcalnet: the online tool for poverty measurement developed by the Development Research Group of the World Bank. This tool can be found here: www.iresearch.worldbank.org/PovcalNet.

poor individual must contribute more to poverty than an individual who is only relatively poor, regardless of the income standard in their respective societies. I call these indices *hierarchical* because they grant this form of precedence to the absolutely poor. Hierarchical indices avoid the counter-intuitive comparisons illustrated in the above example. Standard indices are not hierarchical when the absolute threshold is strictly positive.

The main result emphasizes the double role played by poverty indices. The choice of a poverty index determines two different aspects of the poverty measure: its "prioritization" and its inter-personal comparisons. The first aspect concerns the trade-offs made by the measure between the incomes of different poor individuals living in the same society. These trade-offs determine to whom an extra unit of income should be given if the objective is to yield the largest poverty reduction. I call this aspect the "prioritization" inherent to the measure. This well-known aspect has been extensively studied in the literature surveyed by Zheng (1997). The second aspect, which has been much less studied, concerns the implicit inter-personal comparisons that the measure makes across societies with different income standards. As illustrated in the above example, standard indices perform debatable comparisons when combined with a relative line. My main result shows that their debatable inter-personal comparisons are ruled out by a set of basic properties that indices should satisfy when combined with two poverty lines.

The second contribution is an impossibility result showing that hierarchical indices violate basic "prioritization" properties when the two poverty lines cross each other, i.e. when the two lines have the same threshold for some value of the income standard. In particular, hierarchical indices violate the requirement that a progressive balanced transfer among two poor individuals should not increase poverty. This result reveals that one cannot demand too much from an index when both the absolute and relative aspects of income poverty have to be taken into account. It is not possible for an index to simultaneously perform meaningful inter-personal comparisons and provide a defensible prioritization. I shortly discuss three ways out of this impossibility.

The paper is organized as follows. A succinct literature review is provided in Section 2. I present the framework in Section 3. I characterize the family of hierarchical indices in Section 4. I expose an impossibility result for these indices in Section 5. I make some concluding remarks in Section 6.

2 Literature review

The literature on income poverty measurement studies indicators – called poverty measures – that rank income distributions as a function of poverty. Any poverty measure is composed of two elements: a poverty line and an index. In his groundbreaking paper, Sen (1976) proposes a framework allowing to study the properties inherent to these indices. Following Sen, many authors have proposed particular families of indices and characterized their properties. Among other proposals are the indices studied by Foster et al. (1984), Foster and Shorrocks (1991), Kakwani (1980), Chakravarty (1983) or Duclos and Gregoire (2002). The major results derived in this literature are reviewed in Zheng (1997). This paper extends this literature by departing from the assumption that poverty indices are combined with an absolute line.

A small literature launched by Atkinson and Bourguignon (2001) investigates indices that combine the absolute and relative aspects of income poverty. Atkinson and Bourguignon (2001) suggest to use two poverty lines, an absolute line capturing subsistence and a relative line capturing social exclusion. They propose a family of additive indices – which are not hierarchical – but do not study their properties. The same holds for Anderson and Esposito (2013). Finally, Decerf (2017) considers a different framework with one absolute threshold and one "hybrid" relative line whose threshold is everywhere above the absolute threshold. He starts from the assumption that being absolutely poor is worse than being relatively poor and proposes a particular hierarchical index. There are two key differences between Decerf (2017) and this paper. First, the higher severity associated to absolute poverty status is not assumed but derived from fundamental properties. Second, the two lines studied in this paper are different because they cross each other, i.e. they have the same threshold for a given income standard. This definition of the two lines is more in line with the literature and its implementation requires making fewer normative assumptions. Decerf (2017) shows that, when the relative line crosses the absolute line, his index entirely disregards the relative aspect of income poverty. In a recent follow-up paper, Decerf and Ferrando (2020a) show that income poverty has been halved in the developing world over the period 1990-2015, even when accounting for relative poverty.

The design of appropriate absolute or relative poverty lines is still an active area of research (Foster, 1998; Ravallion and Chen, 2011, 2019; Allen, 2017). This paper does not contribute to such design. Rather, my starting point is to consider two poverty lines that cross each other, a premise in agreement with this strand of literature.

3 The framework

Let an income distribution $y = (y_1, \dots, y_{n(y)})$ be a list of non-negative incomes. The number of individuals in distribution y is denoted by $n(y)$.

Two poverty lines are considered. Each of these two lines defines a different poverty status. First, there is an absolute line whose threshold defines the minimal income necessary to cover an individual's subsistence needs. Its absolute threshold is denoted by $z_a \geq 0$. The set of individuals who qualify as absolutely poor in distribution y is $Q_a(y) = \{i \leq n(y) \mid y_i < z_a \text{ or } y_i = 0\}$.[4]

Second, there is a relative line whose threshold defines the minimal

[4] For the particular case $z_a = 0$, this definition implies that individual i is absolutely poor if $y_i = 0$. This convention allows Theorem 1 to cover the particular case $z_a = 0$.

income necessary to be able to participate in social life. The relative poverty threshold is a function of the income standard $\overline{y} = f(y)$. In line with applications, the income standard is either mean income or median income. In the former case $f(y) = \frac{1}{n(y)} \sum y_i$ and in the latter case $f(y) = y_m$, where m denotes the index attached to the median earner in distribution y.[5] The relative line is defined by its threshold function

$$z_r(y) = b + s\overline{y},$$

where $s \in (0, 1)$ defines the slope of the relative line and $b \in [0, z_a(1 - s)]$ defines the lower bound to social participation costs. The line z_r is strongly relative when $b = 0$ and weakly relative when $b > 0$ (Ravallion and Chen, 2011). As I impose that $b \leq z_a(1 - s)$, the relative threshold is weakly smaller than the absolute threshold when the income standard takes value z_a. Importantly, this restriction implies that the two poverty lines *cross* (i.e. have equal thresholds) at a level of income standard $\overline{y}^c$ defined as

$$\overline{y}^c = \frac{1}{s}(z_a - b),$$

which is such that $\overline{y}^c \geq z_a$. This crossing property is necessary for the impossibility result established in Theorem 2. Let $\mathcal{Z}$ denote the set of acceptable threshold functions. The set of individuals who qualify as relatively poor in y is $Q_r(y) = \{i \leq n(y) \mid y_i < z_r(y)\}$. Figure 1 illustrates several pairs of poverty lines.

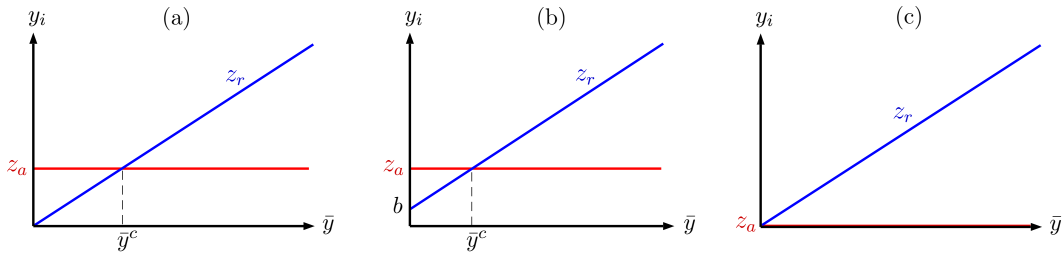

Figure 1: Several pairs of absolute and relative poverty lines.
(a) Positive absolute threshold and strongly relative line. (b) Positive absolute threshold and weakly relative line. (c) Null absolute threshold and strongly relative line.

Together, an individual is poor if her income is below the upper contour of the two lines, which is

$$z(y) = \max\{z_a, z_r(y)\}.$$

The set of individuals who qualify as poor in y is $Q(y) = Q_a(y) \cup Q_r(y)$. The number of poor individuals and the number of absolutely poor individuals are respectively denoted by $q(y)$ and $q_a(y)$. The sets of absolutely poor

[5] If y is sorted in non-decreasing order $y_1 \leq \cdots \leq y_n$, then $m = \frac{1}{2}(n + 1)$ if n is odd and $m = \frac{1}{2}n$ if n is even. This definition of the median when n is even is without loss of generality as the proofs can be easily adapted if m is instead defined as $m = \frac{1}{2}n + 1$.

and relatively poor individuals need not be disjoint. The set of individuals who qualify as only relatively poor in y is $Q(y) \backslash Q_a(y)$. The number of only relatively poor in y is $q(y) - q_a(y)$.

Letting $N = \{n \in \mathbb{N} | n \geq 4\}$, the set of income distributions considered is[6]

$$Y = \left\{ y \in \cup_{n \in N} \mathbb{R}^n_+ \mid \overline{y} \geq z_a \text{ and } \overline{y} > 0 \right\}.$$

This set excludes distributions whose income standard is smaller than the absolute threshold. This restriction is necessary for Theorem 1 to hold when the income standard is median income. Also, this restriction implies that the income of poor individuals is smaller than the income standard, i.e. $y_i < \overline{y}$ for all $i \in Q(y)$.

A poverty index is a real-valued function $P : \mathcal{P} \to \mathbb{R}_+$ that ranks income distributions using the two poverty lines as parameters. In general, a poverty index has a domain of definition $\mathcal{P} = Y \times \mathbb{R}_+ \times \mathcal{Z}$. However, the results can be derived when assuming that the two poverty lines are given. As it makes the results more general,[7] I adopt the following much narrower domain of definition

$$\mathcal{P} = Y \times \{(z_a, s, b, f)\}.$$

The poverty in distribution y is simply denoted by $P(y)$. For any two distributions y and y', there is strictly more poverty in y than in y' if $P(y) > P(y')$, and weakly more if $P(y) \geq P(y')$.

4 Hierarchical poverty indices

I study which indices should be used when comparing poverty in different income distributions. The particularity of the framework is the presence of a second poverty line whose threshold depends on the income standard. The properties I impose on indices acknowledge this presence in two ways. First, the relevance of each of the two poverty lines is established in a separate focus axiom. Second, several properties are restricted to the comparison of income distributions that have equal income standards.

The particularity of poverty indices is that only the situation of poor individuals matters to poverty comparisons. This property that distinguishes poverty indices from other kinds of normative indices, e.g. inequality or mobility indices, is traditionally encapsulated in a focus axiom. This axiom requires that the exact income of individuals earning more than the poverty threshold is irrelevant to some extent. The extent to which their income is irrelevant depends on the poverty status considered. As I consider two kinds of poverty, I impose two separate focus axioms. Each focus axiom is specific to the particular need captured by its associated line.

An individual is absolutely poor if her income is insufficient to meet her subsistence needs. Traditionally, the minimal income necessary to cover

[6] The requirement $\overline{y} \geq z_a$ does not exclude $\overline{y} = 0$ when $z_a = 0$.

[7] The results obtained on the narrower domain also constrain indices defined on the larger domain.

subsistence needs is assumed to be independent of the income of non-poor individuals. *Absolute Focus* requires that, when all poor individuals are absolutely poor, the exact income earned by non-poor individuals is irrelevant. This axiom is a weakening of the classical focus axiom.

Absolute Focus. *For all* $y, y' \in Y$ *with* $n(y) = n(y')$ *and* $Q_a(y) = Q(y) = Q_a(y') = Q(y')$*, if* $y_i = y'_i$ *for all* $i \in Q_a(y)$*, then* $P(y) = P(y')$*.*

In contrast, the focus axiom associated to the relative line does not completely disregard the income of non-poor individuals. The income necessary for an individual to meet her social participation needs depends on the income standard. In turn, the income standard depends on the income of non-poor individuals. *Relative Focus* requires that, when all poor individuals are relatively poor, the exact income earned by non-poor individuals is irrelevant only as long as the income standard is unchanged.

Relative Focus. *For all* $y, y' \in Y$ *with* $n(y) = n(y')$*,* $Q_r(y) = Q(y) = Q_r(y') = Q(y')$ *and* $\overline{y} = \overline{y}'$*, if* $y_i = y'_i$ *for all* $i \in Q_r(y)$*, then* $P(y) = P(y')$*.*

The classical monotonicity property requires that poverty is reduced when a poor individual earns an additional amount of income. *Weak Monotonicity* adds the precondition that the other poor individuals are not affected. This precondition is guaranteed by restricting monotonicity comparisons to distributions that have the same income standard.

Weak Monotonicity. *For all* $y, y' \in Y$ *with* $n(y) = n(y')$*,* $Q(y') \subseteq Q(y)$ *and* $\overline{y} = \overline{y}'$*, if* $y_j < y'_j$ *for some* $j \in Q(y)$ *and* $y_i = y'_i$ *for all* $i \in Q(y') \setminus \{j\}$*, then* $P(y) > P(y')$*.*

Subgroup Consistency is a standard axiom requiring that, if poverty decreases in a subgroup while it remains constant in the rest of the distribution, overall poverty must decline. Sen (1992) questioned the desirability of this axiom by arguing that the incomes in one subgroup may affect poverty in another subgroup. Foster and Sen (1997) recommend not to use this axiom when the index aims at capturing relative aspects of poverty. I subscribe to this point of view. The issue becomes transparent once the channel through which one subgroup affects the other is modeled. In this framework, the incomes in a subgroup impact the income standard, which in turn affects poor individuals in another subgroup. When a poverty line is relative, it is thus not always meaningful to extrapolate the judgments made on a subgroup to the whole population. *Weak Subgroup Consistency* restricts such extrapolations to distributions for which the subgroups have the same income standard. In such cases, the income standard of a subgroup is equal to the income standard of the entire distribution and poverty judgments made on the subgroup are meaningful for the entire distribution.

Weak Subgroup Consistency. *For all* $y^1, y^2, y^3, y^4 \in Y$ *with* $n(y^1) = n(y^3)$*,* $n(y^2) = n(y^4)$ *and* $\overline{y}^1 = \overline{y}^2 = \overline{y}^3 = \overline{y}^4$*, if* $P(y^1) > P(y^3)$ *and* $P(y^2) = P(y^4)$*, then* $P((y^1, y^2)) > P((y^3, y^4))$*.*

The remaining three auxiliary axioms are standard. *Symmetry* requires that individuals' identities do not matter. Working with sorted distributions is therefore without loss of generality.

Symmetry. *For all* $y, y' \in Y$ *with* $n(y) = n(y')$*, if* $y' = y \cdot \pi_{n(y)\times n(y)}$ *for some permutation matrix* $\pi_{n(y)\times n(y)}$*, then* $P(y) = P(y')$*.*

Symmetry implies that individual preferences are irrelevant to the poverty index. This property generates little debate when only the level of own income appears in preferences. If preferences are monotonic, then monotonic indices do not override individual preferences. When both the level of own income and the relative situation matter, the monotonicity of preferences does not entirely define individual preferences and *Symmetry* explicitly requires to completely disregard these preferences. This form of paternalism can be defended on the ground that it prevents poverty indices from giving priority to individuals that are more other-regarding.[8]

Next axiom requires indices to be continuous in incomes. Such continuity requirement is important in empirical applications in order to avoid that measurement errors have an excessive impact on poverty judgments. *Weak Continuity* requires indices to be continuous in all incomes, but only for distributions whose income standard is larger than the income standard at which the two lines cross.

Weak Continuity. *For all* $y \in Y$ *with* $\overline{y} > \overline{y}^c$*,* P *is continuous in* y*.*

In the absence of any restriction, the continuity property would be incompatible with the above axioms. As shown in Lemma 2 in Appendix 7.1, indices that are continuous on the whole domain Y cannot simultaneously satisfy *Absolute Focus* and *Weak Monotonicity* when $z_a > 0$.

Finally, *Replication Invariance* specifies how to compare poverty across distributions of different population sizes. If a distribution is obtained by replicating another distribution several times, then the two distributions have equal poverty. Formally, for any $k \in \mathbb{N}$, the $k-$replication of a distribution y is the distribution $y^{\times k} = (y, \ldots, y)$ for which $n(y^{\times k}) = kn(y)$.

Replication Invariance. *For all* $y \in Y$ *and* $k \in \mathbb{N}$*, we have* $P(y) = P(y^{\times k})$*.*

Lemma 1 shows that not all of these axioms are independent.

Lemma 1. *If the income standard is either mean income or median income, and if* P *satisfies Absolute Focus, Weak Subgroup Consistency and Symmetry, then* P *satisfies Relative Focus.*

Proof. See Appendix 7.2. ■

As *Relative Focus* is implied by the other axioms, it can be dropped from the statement of Theorem 1 (see below). This theorem states that these axioms jointly characterize the family of hierarchical poverty indices.

[8] For an illustration of the issue, consider two poor individuals living in the same society. Assume that individual 1 has a smaller income than individual 2 but individual 2 has preferences that are more affected by relative income than the preferences of individual 1. If individual preferences matter for the poverty index, it could be that the contribution to poverty of individual 2 is larger than that of individual 1. Hence, individual 2 is considered more poor than individual 1. Such conclusion is debatable given that both individuals would agree that individual 2 is better-off than individual 1.

Definition 1 (Hierarchical poverty indices).
P is a hierarchical poverty index if

$$P(y) = \frac{1}{n(y)} \sum_{i=1}^{n(y)} p(y_i, \overline{y}), \tag{1}$$

where the poverty contribution function $p : \mathbb{R}_+ \times [z^a, \infty) \to [0, 1]$ *satisfies*

(i) $p(0, \overline{y}) = 1$ *and* $p(y_i, \overline{y}) = 0$ *if* $i \notin Q(y)$,

(ii) p *is strictly decreasing in its first argument if* $i \in Q(y)$,

(iii) p *is continuous in both its arguments if* $\overline{y} > \overline{y}^c$,

(iv) p *is constant in its second argument if* $i \in Q_a(y)$.

Hierarchical indices are additive indices. Therefore, they sum the poverty contributions of all individuals in a distribution. As usual, non-poor individuals contribute zero (i), the contribution of any poor individual is decreasing in her income (ii) and the contribution function is continuous on some domain (iii). More importantly, the contribution of any poor individual depends on her income, as for standard additive indices (Foster and Shorrocks, 1991), but also on the income standard. These two variables summarize the relevant aspects of a poor individual's situation. In this sense, the pair $(y_i, \overline{y})$ defines the "bundle" consumed by the poor individual i. The contribution function ranks all the bundles that poor individuals may consume.[9] Therefore, the contribution function implicitly compares poor individuals across societies with different levels of income standards.

The key feature of hierarchical indices is that they compare poor individuals living in different societies in a specific way. The constraints imposed on these comparisons are revealed by a graphical representation of the *iso-poverty map* defined by the index. An iso-poverty map is a collection of iso-poverty curves. An iso-poverty curve is the set of bundles associated to the same poverty contribution. Implicitly, two individuals whose bundles are on the same iso-poverty curve are deemed equally poor by the contribution function. Figure 2 shows iso-poverty maps satisfying the constraints imposed by restriction (iv). This restriction requires that the contribution of *absolutely* poor individuals only depends on their own income. As a result, their iso-poverty curves are flat and do not cross the absolute threshold. This implies that the bundle of an absolutely poor individual is always on a lower iso-poverty curve than the bundle of an individual who is only relatively poor. In other words, an individual who is absolutely poor must be deemed *poorer* than an individual who is only relatively poor, regardless of the income standards in their respective societies. An extremely poor individual in Bangladesh must be deemed poorer than an individual who is only relatively poor in Colombia, even if the income standard in Colombia is larger. I call these indices "*hierarchical*"

[9] Formally, the contribution function defines a complete *ordering* on the space of bundles.

because they grant a particular form of precedence to the absolute poverty status. Below the absolute threshold, the relative aspect of income poverty is irrelevant. It is only when an individual is not absolutely poor that the relative aspect of her poverty becomes relevant.

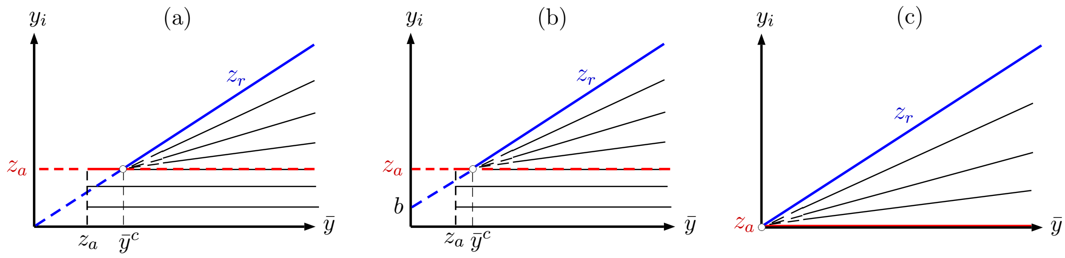

Figure 2: Iso-poverty maps of hierarchical indices.
Note: The plain lines are iso-poverty curves. (a) Positive absolute threshold and strongly relative line. (b) Positive absolute threshold and weakly relative line. (c) Null absolute threshold and strongly relative line.

Importantly, standard poverty indices are not hierarchical when the absolute threshold is strictly positive. Consider for instance the FGT indices (Foster et al., 1984), which are pervasive in empirical applications. In the presence of two lines, FGT indices are usually defined as

$$\hat{p}_\alpha(y_i, \overline{y}) = \left(1 - \frac{y_i}{\max(z_a, z_r(y))}\right)^\alpha \tag{2}$$

where $\alpha \geq 0$ is the poverty aversion parameter (Atkinson and Bourguignon, 2001). Equation (2) reveals that the inter-personal comparisons performed by these indices only depend on the normalized income, i.e. on own income divided by the relevant poverty threshold. Two individuals with the same normalized incomes are considered equally poor. Coming back to the example developed in the Introduction, an individual earning half the poverty threshold in Bangladesh, where $z\left(y^B\right) = \$2.4$ a day, is considered as poor as an individual earning half the poverty threshold in Colombia, where $z\left(y^C\right) = \$5.5$ a day. This comparison is made even if the former is extremely poor while the latter is only relatively poor. To fix ideas, given the respective relative thresholds, standard FGT indices implicitly consider that an individual earning \$1.5 a day in Bangladesh is *less poor* than an individual earning \$3 a day in Colombia. This shows that standard FGT indices are not hierarchical. I emphasize that it is not a mere theoretical issue. When they are combined with relative poverty lines, non-hierarchical indices regularly lead to highly counterintuitive poverty comparisons (Decerf, 2017).

It is however possible to define a hierarchical version of FGT indices. Consider the following family of indices for which the contribution of any poor individual i in distribution y is given by[10][11]

$$p_{\alpha\lambda}(y_i, \overline{y}) = \begin{cases} \left(1 - \lambda \frac{y_i}{z_a}\right)^{\alpha} & \text{if } \; y_i < z_a, \\ \\ \left(1 - \lambda - (1-\lambda)\frac{y_i - z_a}{z_r(y) - z_a}\right)^{\alpha} & \text{if } \; z_a \leq y_i < z_r(y), \end{cases} \tag{3}$$

where α is the poverty aversion parameter and λ is a parameter tuning the poverty contribution of an individual earning exactly z_a. Any index in the family proposed in Equation (3), which is denoted by $P_{\alpha\lambda}$, is defined by a pair of values for the two parameters α and λ. As can be easily verified, when $z_a > 0$, $P_{\alpha\lambda}$ is a hierarchical index if $\alpha > 0$ and $\lambda \in (0, 1)$.

Theorem 1. *If the income standard is either mean income or median income, then the following two statements are equivalent.*

1. *P satisfies Absolute Focus, Weak Monotonicity, Weak Subgroup Consistency, Symmetry, Weak Continuity and Replication Invariance.*

2. *P is ordinally equivalent to a hierarchical poverty index.*

Proof. See Appendix 7.3. ■

Theorem 1 has three important implications. First, indices based on two lines should be additive, which means that they sum individual contributions to poverty. Importantly, this additive separability result is obtained with a *weak* version of the Subgroup Consistency axiom of Foster and Shorrocks (1991). This weak version restricts its application to subgroups that have the same income standard. As a result, *Weak Subgroup Consistency* escapes the critique of Foster and Sen (1997). Surprisingly, this weak property is sufficient to obtain additive separability and Theorem 1 constitutes a firm foundation for additive indices, provided individual contributions depend on both own income and the income standard.

Second, because the contribution function depends on both own income and the income standard, this function implicitly compares poor individuals across societies with different income standards. In other words, the selection of a poverty index determines the inter-personal comparisons performed by the poverty measure. It is well-known that the selection of a poverty index determines the trade-offs that the poverty measure makes between the incomes of different poor individuals living in the same society. This aspect, which I call the "prioritization" inherent to the index, has been widely studied (Zheng, 1997). However, the inter-personal comparison aspect is much less studied. Decerf (2017) emphasizes the importance of these inter-personal comparisons.

The last key implication of Theorem 1 is that poor individuals living in different societies must be compared in a specific way. An absolutely

[10] The expression given in Equation (3) is valid for all poor individuals. However, when $\overline{y} \leq \overline{y}^c$ and $y_i = z_a$, individual i is non-poor and her contribution is then equal to zero.

[11] When $z_a = 0$, the first part of Equation (3) is not well-defined. However, this first part is irrelevant when $z_a = 0$ because no individual has an income $y_i < 0$.

poor individual must be considered more poor than an individual who is only relatively poor, regardless of the income standards in their respective societies. The normative view according to which having absolute poverty status is more severe than having only relative poverty status is largely shared. It has been expressed in the literature (Atkinson and Bourguignon, 2001; Decerf, 2017) and is largely shared in the population, as appeared from questionnaire studies run all over the world by Corazzini et al. (2011). Theorem 1 provides a normative foundation for this view. As revealed by the proof, the two axioms that jointly imply the hierarchical structure of these indices are *Absolute Focus* and *Weak Subgroup Consistency*.

Two remarks are in order. First, Theorem 1 still holds in the special case $z_a = 0$, i.e. when the index only considers one (relative) poverty line. Despites the critics expressed against such widely used additive measures (Sen, 1992; Foster and Sen, 1997), indices based on a relative line have never been rigorously studied. This theorem thus provides a normative foundation for additive indices used in combination with a relative line. Second, Theorem 1 can be extended to relative lines whose threshold function is not linear. This theorem applies as long as the relative line is continuous and monotonically increasing with the income standard and its threshold is always smaller than the income standard.

5 An impossibility result

Theorem 1 places no restriction on the shape that the contribution function takes for a fixed level of income standard, i.e. no restriction on the prioritization inherent to the index. The only requirement is that the contribution function decreases continuously in own income. Restrictions on its shape emerge from axioms constraining how the index must trade-off the incomes of different poor individuals. I consider two such axioms and show that, unfortunately, hierarchical indices violate at least one of them when the absolute threshold is strictly positive.

The first property, *Transfer*, is a classical axiom requiring that a Pigou-Dalton transfer taking place between two poor individuals never unambiguously increases poverty.

Transfer. *For all $y, y' \in Y$ with $n(y) = n(y')$, $Q(y) = Q(y')$ and $\delta > 0$, if $y_j - \delta = y'_j > y'_k = y_k + \delta$ for some $j, k \in Q(y)$, $y'_i = y_i$ for all $i \neq j, k$ and $\overline{y}' = \overline{y}$, then $P(y) \geq P(y')$.*

As is well-known, poverty indices satisfying *Transfer* are based on convex contribution functions.

The second property, *Strong Monotonicity*, considers an increase in the income of some poor individual. When the income standard is mean income, an increase in the income of a poor individual has opposing effects. On the one hand, her poverty contribution decreases. This direct effect reduces poverty. On the other hand, mean income increases.[12] If the poverty

[12] Observe that the larger the number of individuals, the lower is the impact of the increase in income on mean income and, hence, on the poverty contributions of others.

threshold increases, then the poverty contributions of poor individuals may increase. Moreover, some individuals who were non-poor might become relatively poor. *Strong Monotonicity* requires that these indirect adverse effects are dominated by the direct effect. Indices satisfying this property never deem that a policy whose unique impact is to decrease the income of some *poor* individuals reduces poverty.

Strong Monotonicity. *For all* $y, y' \in Y$ *with* $n(y) = n(y')$*, if* $y_j < y'_j$ *for some* $j \in Q(y) \cup Q(y')$ *and* $y_i = y'_i$ *for all* $i \neq j$*, then* $P(y) > P(y')$*.*

Observe that, when the income standard is median income, last axiom is equivalent to *Weak Monotonicity*. The reason is that the income of poor individuals does not impact median income in the domain of distributions considered.

Theorem 2 (see below) shows that all hierarchical indices violate *Transfer*. Moreover, if the income standard is mean income, then hierarchical indices also violate *Weak Monotonicity*.

I illustrate the violation of these axioms for a particular hierarchical index, but the intuition provided is general. Consider $P_{1\lambda}$, which is defined by Equation (3) when assuming $\alpha = 1$, i.e.

$$p_{1\lambda}(y_i, \overline{y}) = \begin{cases} 1 - \lambda \frac{y_i}{z_a} & \text{if } \; y_i < z_a, \\ \\ 1 - \lambda - (1-\lambda)\frac{y_i - z_a}{z_r(y) - z_a} & \text{if } \; z_a \leq y_i < z_r(y), \end{cases}$$

First, $P_{1\lambda}$ violates *Transfer* when the income standard is sufficiently close to $\overline{y}^c$, the value for which the two poverty lines have the same threshold. Let individual i be absolutely poor and individual j be only relatively poor. *Transfer* implies that a balanced transfer of an amount ϵ from j to i cannot increase poverty. Given the definition of $p_{1\lambda}$, we have

$$\frac{\partial p_{1\lambda}(y_i, \overline{y})}{\partial y_i} = -\frac{\lambda}{z_a} \qquad \text{and} \qquad \frac{\partial p_{1\lambda}(y_i, \overline{y})}{\partial y_j} = -\frac{1-\lambda}{z_r(y) - z_a}.$$

Thus, when i receives an amount ϵ of income, her contribution to poverty decreases by $\epsilon\frac{\lambda}{z_a}$, regardless of her exact income. In contrast, when j reduces her income by an amount ϵ, her contribution increases by $\epsilon\frac{1-\lambda}{z_r(y)-z_a}$, regardless of her exact income. *Transfer* is satisfied when the variation in j's contribution is smaller than the variation in i's contribution. The reason why this axiom is violated is that the variation in j's contribution depends on the difference between the two thresholds, whereas it is not the case for i's contribution. Provided the income standard is sufficiently close to $\overline{y}^c$, the difference between the two thresholds tends to zero, and the variation in j's contribution becomes larger than the variation in i's contribution.

Figure 3 graphically illustrates the issue. This figure shows the contribution function at two different levels of income standards. The contribution of relatively poor individuals depends on the income standard, while the contribution of absolutely poor individuals is independent of the income standard. When the income standard is sufficiently close to $\overline{y}^c$, as

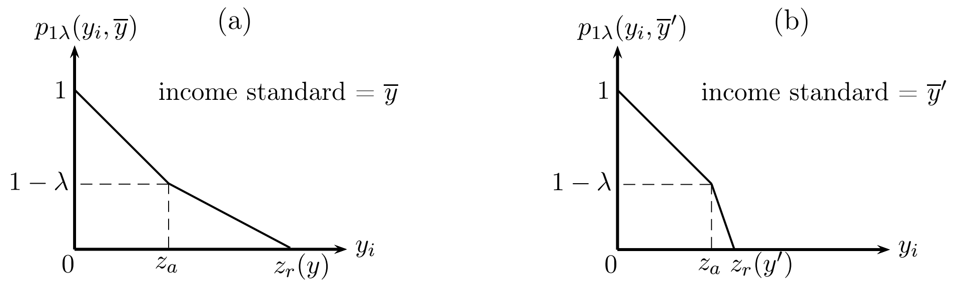

Figure 3: Hierarchical index $P_{1\lambda}$ violates *Transfer*.
Note: (a) At the income standard $\overline{y}$, the contribution function is convex. (b) At the income standard $\overline{y}'$, the contribution function is concave.

is the case in panel (b), the contribution function becomes concave in own income, implying a violation of *Transfer*.

The violation of *Strong Monotonicity* is based on a similar intuition. Consider a distribution with at least one absolutely poor individual and a second individual who is non-poor but whose income is exactly equal to the relative threshold. When the income of the absolutely poor increases, mean income increases. If the income standard is mean income, then the relative threshold increases and the second individual becomes relatively poor. Again, the increase in the contribution of the second individual can be made larger than the decrease in the contribution of the absolutely poor, provided the relative threshold is sufficiently close to the absolute threshold.

Theorem 2 shows that all hierarchical indices fail some basic property.

Theorem 2. *Let $z_a > 0$ and let P be a hierarchical index.*

1. *If the income standard is either mean income or median income, then P violates Transfer.*

2. *If the income standard is mean income, then P violates Strong Monotonicity.*

Proof. See Appendix 7.4. ■

Theorems 1 and 2 jointly constitute a negative result. One cannot demand too much from a poverty index when both the absolute and relative aspects of income poverty have to be taken into account. Any index based on two poverty lines either ignores the higher severity associated to having absolute poverty status or fails some basic property. In this sense, any index must make a trade-off between plausible inter-personal comparisons and defensible "prioritization".

There are at least three possible ways to deal with this normative trade-off. The first way is to escape the impossibility stated in Theorem 2 by changing the definition of the two poverty lines or the domain of distributions Y. The impossibility holds because the relative line crosses the absolute line on the domain Y. Decerf (2017) considers a different framework, with two poverty lines that do not cross. The two lines do not cross because the relative threshold is assumed to be always larger than the absolute threshold. In this alternative framework, the hierarchical index proposed by Decerf (2017) satisfies both *Transfer* and *Strong Monotonicity*, even when the income standard is mean income. The problem with this framework is that it is at odds with the literature considering two poverty lines. For instance, the Societal poverty line recently adopted by the World Bank does cross its extreme poverty threshold (World Bank, 2018). Furthermore, even when taking a very low absolute line, as the extreme line of the World Bank, and a rather high relative line, as the Societal line (Jolliffe and Prydz, 2017), many developing countries have their income standards smaller than $\overline{y}^c$.[13] Therefore, this first way out is not always a viable option, and one must then compromise on either plausible inter-personal comparisons or defensible prioritization.

The second way out is to compromise on defensible prioritization. The solution here is to use hierarchical indices that satisfy weakenings of *Transfer* and *Strong Monotonicity*. This is for instance the route followed by Decerf and Ferrando (2020b).

The third way out is to compromise on plausible inter-personal comparisons. There exist non-hierarchical indices that satisfy both *Transfer* and *Strong Monotonicity*, even when the income standard is mean income. This is for instance the case of the poverty gap ratio, as defined by Equation (2) when assuming $\alpha = 1$.

I believe that the minimal form of precedence granted to absolute poverty by hierarchical indices is conceptually more fundamental than aggregation axioms such as *Transfer* or *Strong Monotonicity*. The reason is that inter-personal comparisons may relate to the welfare-consistency of the poverty measure.[14] As a result, non-hierarchical indices regularly lead to debatable poverty comparisons (Decerf, 2017). In contrast, aggregation axioms relate to the comparisons of gains and losses made by different poor individuals. Therefore, these axioms "merely" relate to the fairness of the index. For these reasons, I believe that the second way out is in general preferable to the third one.

Of course, the particular application one has in mind can influence the choice between alternative escape routes. The difference between the inter-personal comparisons performed by standard and hierarchical indices is large when comparing societies with significantly different income standards. However, this difference becomes smaller as the distance between

[13] Recall that the Societal line is defined as $z_r(y) = 1+0.5\overline{y}$, where the income standard is median income. This weakly relative line has a high intercept ($1 a day), implying a high relative threshold. In 2015, the latest reference year available in Povcalnet data, at least nine sub-Saharan countries have their Societal threshold below the extreme poverty threshold.

[14] In order to get an intuition, assume that the precedence granted to the absolutely poor corresponds to individual preferences. Under this assumption, individuals prefer to have the possibility to satisfy their subsistence needs, even if it comes at the cost of a worse relative situation. Then, the problem with non-hierarchical indices is that they sometimes conclude that poverty is reduced even when all poor individuals deem that their situation has worsened. For instance, this could happen after an inequality-reducing recession. Such recession may decrease the income of poor individuals below the absolute threshold while improving their relative situation.

their income standards is reduced. So, if the application only requires comparing societies with similar income standards, then the third route might be preferable to the second one. This is the for instance the case if one is interested in the poverty trend of a country that does not experience growth, or when comparing the poverty in countries sharing similar standards of living. Yet, for many important applications such as measuring global income poverty or assessing the poverty trend of a developing country experiencing significant growth, the second way out seems to offer a more compelling solution.

6 Concluding remarks

Income inequality has recently attracted increasing attention. Abstracting from the impacts that inequality may have on behavior, there exists two main normative reasons why one may care about inequality. The first is fairness. An ethical observer may prefer more equal distributions of resources. The second is that inequality may have intrinsic value for the concerned individual. For instance, their preferences may depend on both their absolute income and their relative income. Alternatively, the social functionings provided by a given amount of resources may depend on the society's income standard (Sen, 1992). The second reason is the mainstream foundation used to defend relative poverty lines. Any poverty measure endorsing such foundation must *first* aggregate the absolute and relative aspects of income at the individual level and *second* aggregate individual contributions over the whole population, as proposed by Atkinson and Bourguignon (2001). As these authors suggest, taking onboard the relative aspect of poverty is not only a matter of picking the right poverty line(s) but also a matter of selecting an appropriate index. Following their approach, Decerf (2017) stresses the importance of the iso-poverty maps inherent to poverty measures. In this paper, I show that indices based on two poverty lines should be based on an iso-poverty map that grants some precedence to the absolute aspect of income poverty. This result further emphasizes the key role played by iso-poverty maps, which has been little studied in the literature.

7 Appendix

7.1 Incompatibility between *Absolute Focus*, *Weak Monotonicity* and *Continuity*

Continuity. *For all $y \in Y$, P is continuous in y.*

Lemma 2. *If $z_a > 0$ and the income standard is either mean income or median income, then no index P satisfies Absolute Focus, Weak Monotonicity and Continuity.*

Proof. Take any $\overline{y}$ and $\overline{y}'$ such that $z_a < \overline{y} \leq \overline{y}^c < \overline{y}'$. For any $\epsilon \in [0, z_a]$, I define two distributions $y^\epsilon, y^{\epsilon'} \in Y$ in the following way: $n(y^\epsilon) = n(y^{\epsilon'}) = 4$,

$\overline{y}^\epsilon = \overline{y}$, $\overline{y}^{\epsilon'} = \overline{y}'$, $y_1^\epsilon = y_1^{\epsilon'} = z_a - \epsilon$ and $y_2^\epsilon = y_3^\epsilon = y_4^\epsilon$ and $y_2^{\epsilon'} = y_3^{\epsilon'} = y_4^{\epsilon'}$. If the income standard is mean income, this implies that for all $i \in \{2, 3, 4\}$ we have $y_i^\epsilon = \frac{1}{3}(4\overline{y} - y_1^\epsilon)$ and $y_i^{\epsilon'} = \frac{1}{3}(4\overline{y}' - y_1^{\epsilon'})$. If the income standard is median income, this implies that for all $i \in \{2, 3, 4\}$ we have $y_i^\epsilon = \overline{y}$ and $y_i^{\epsilon'} = \overline{y}'$. Either way, we have for all $i \in \{2, 3, 4\}$ that $y_i^\epsilon \geq \overline{y}$ and $y_i^{\epsilon'} \geq \overline{y}'$, which implies that $i \notin Q(y^\epsilon)$ and $i \notin Q(y^{\epsilon'})$ because $z_a < \overline{y} < \overline{y}'$ and $b \leq z_a(1 - s)$. As $z_a > 0$, we have for all $\epsilon > 0$ that $Q(y^\epsilon) = Q_a(y^\epsilon) = Q(y^{\epsilon'}) = Q_a(y^{\epsilon'}) = \{1\}$ and therefore $P(y^\epsilon) = P(y^{\epsilon'})$ by *Absolute Focus* because $y_1^\epsilon = y_1^{\epsilon'}$.

Consider the distribution $y'' = (\overline{y}', \overline{y}', \overline{y}', \overline{y}')$, which is such that $\overline{y}'' = \overline{y}'$ and $Q(y'') = \varnothing$. For $\epsilon = 0$, we have $Q(y^0) = \varnothing$ because $y_1^0 = z_a \geq z_r(y^0)$, and therefore $P(y^0) = P(y'')$ by *Absolute Focus*. However, for $\epsilon = 0$, we have $Q(y^{0'}) = \{1\}$ because $y_1^{0'} = z_a < z_r(y^{0'})$, and therefore $P(y^{0'}) > P(y'')$ by *Weak Monotonicity* because $\overline{y}^{0'} = \overline{y}''$ and $y_1^{0'} < z_r(y^{0'}) < y_i''$ for all i.

As $P(y^\epsilon) = P(y^{\epsilon'})$ for all $\epsilon > 0$, we must have $P(y^0) = P(y^{0'})$ by *Continuity*. This yields the desired contradiction because $P(y^0) = P(y'') < P(y^{0'})$. ■

7.2 Proof for Lemma 1

Consider any two $y, y' \in Y$ with $n(y) = n(y')$, $Q_r(y) = Q(y) = Q_r(y') = Q(y')$ and $\overline{y} = \overline{y}'$, $y_i = y_i'$ for all $i \in Q_r(y)$. I must show that $P(y) = P(y')$ if P satisfies *Absolute Focus*, *Weak Subgroup Consistency* and *Symmetry*.

If $\overline{y} \leq \overline{y}^c$, then $z_r(y) \leq z_a$ and we have by assumption that $Q_a(y) = Q(y) = Q_a(y') = Q(y')$. By *Absolute Focus* we have $P(y) = P(y')$, the desired result.

I consider the remaining case $\overline{y} > \overline{y}^c$. Assume that both distributions are sorted in non-decreasing order, i.e. $y_1 \leq y_2 \leq \cdots \leq y_{n(y)}$ and the same for y'. By *Symmetry*, this assumption is without loss of generality.

If $Q_r(y) = \varnothing$, then we have $P(y) = P(y')$ by *Absolute Focus*. I consider the remaining case $Q_r(y) \neq \varnothing$. Let $y^p = (y_1, \ldots, y_{q(y)})$ and $y^{np} = (y_{q(y)+1}, \ldots, y_{n(y)})$ be distributions such that $y = (y^p, y^{np})$ and all individuals in y^p have their income below $z_r(y)$ while all individuals in y^{np} have their income above $z_r(y)$. Let $y^{p'} = (y_1', \ldots, y_{q(y)}')$ and $y^{np'} = (y_{q(y)+1}', \ldots, y_{n(y)}')$ be distributions such that $y' = (y^{p'}, y^{np'})$ and all individuals in $y^{p'}$ have their income below $z_r(y)$ while all individuals in $y^{np'}$ have their income above $z_r(y)$. By assumption on y and y', we have $y^p = y^{p'}$.

Consider a distribution x^c such that for distributions $x = (y^{np}, x^c) \in Y$ and $x' = (y^{np'}, x^c) \in Y$ we have $\overline{x} = \overline{y}$, $\overline{x}' = \overline{y}$ and $Q_r(x) = Q_r(x') = \varnothing$. In words, the distribution x^c is used to construct two distributions x and x' that have the appropriate income standard $\overline{y}$ and all individuals in x and x' are non-poor.

Assume that such distribution x^c exists (I prove that it exists below). By *Symmetry* we have

$$P(\underbrace{y^p, y^{np'}}_{y'}, \underbrace{y^{np}, x^c}_{x}) = P(\underbrace{y^p, y^{np}}_{y}, \underbrace{y^{np'}, x^c}_{x'}).$$

I show that $P(y) \neq P(y')$ would lead to a contradiction. Consider the contradiction assumption $P(y) > P(y')$. As $n(y^{np}, x^c) = n(y^{np'}, x^c)$ and these two distributions have no poor individuals, we have $P(y^{np}, x^c) = P(y^{np'}, x^c)$ by *Absolute Focus*. Then, the preconditions for *Weak Subgroup Consistency* are all met, i.e. $n(y) = n(y')$, $n(x') = n(x)$ and $\overline{y} = \overline{y}' = \overline{x} = \overline{x}'$, $P(y) > P(y')$ and $P(x) = P(x')$, and this axioms implies

$$P(\underbrace{y^p, y^{np'}}_{y'}, \underbrace{y^{np}, x^c}_{x}) < P(\underbrace{y^p, y^{np}}_{y}, \underbrace{y^{np'}, x^c}_{x'}),$$

in contradiction with *Symmetry*. The alternative contradiction assumption $P(y) < P(y')$ leads to a similar impossibility.

Finally, I show that distribution x^c exists. Given that $\overline{y} > \overline{y}^c$, we have that $\overline{y} > z_r(y)$ by the assumptions on the two lines: $z_a \leq \overline{y}^c$ and $s < 1$. If the income standard is median income, then x^c can be constructed such that $n(x^c) = n(y)$ and all individuals in distribution x^c earn the same income equal to $\overline{y}$. Consider now the case for which the income standard is mean income. In that case, we have $\overline{y}^{np} = \overline{y}^{np'}$ because $\overline{y} = \overline{y}'$ and $y^p = y^{p'}$. However, we have $\overline{y} < \overline{y}^{np} = \overline{y}^{np'}$. As $\overline{y} > z_r(y)$ and there is no upper limit on the number of individuals in x^c, it is clearly possible to find an appropriate x^c for which all individuals in distribution x^c earn the same income a with $z_r(y) < a < \overline{y}$ and x and x' have mean income equal to $\overline{y}$.

7.3 Proof of Theorem 1

The proof that statement 2 implies statement 1 is straightforward and therefore omitted. A complete proof is provided for the converse implication for the case in which the income standard is the mean income. Following that proof, I explain how to adapt this proof so as to apply to the case in which the income standard is the median income.

A. The income standard is the mean income.

Consider any poverty index P satisfying the axioms listed in statement 1. By Lemma 1, P satisfies *Relative Focus* on top of the axioms listed in statement 1. The proof that statement 1 implies statement 2 is based on a version of Gorman's Theorem on additively separable functions.[15] To state it, I need a preliminary definition.

Separability. *Let $L = \{1, \ldots, l\}$, where $l \geq 3$. The function $P^* : \times_{i=1}^{l}[0,1] \rightarrow \mathbb{R}$ is* separable *if for all $\hat{L} \subseteq L$ with $\varnothing \neq \hat{L} \neq L$ and all $u, v, u', v' \in \times_{i=1}^{l}[0,1]$, if $u_i = v_i$ and $u'_i = v'_i$ for all $i \in \hat{L}$ and $u_j = u'_j$ and $v_j = v'_j$ for all $j \in L \setminus \hat{L}$, then*

$$P^*(u) \geq P^*(v) \Leftrightarrow P^*(u') \geq P^*(v'). \tag{4}$$

[15] Theorem 2 in Gorman (1968) is more general than the version presented in Theorem 1. For instance, Gorman (1968)'s theorem does not require that each sector be a real interval. It only requires that each sector has a countably dense subset and is arc-connected.

Theorem 1. (Gorman, 1968, Theorem 2) *Let $L = \{1, \dots, l\}$, where $l \geq 3$. If the function $P^* : \times_{i=1}^{l}[0,1] \to \mathbb{R}$ is continuous, strictly increasing in its arguments and separable, then for any $\nu \in \times_{i=1}^{l}[0,1]$,*

$$P^*(\nu) = \tilde{F}\left(\sum_{i=1}^{l} \tilde{\phi}_i(\nu_i)\right), \tag{5}$$

where $\tilde{F}$ and $\tilde{\phi}_i$ are continuous and strictly increasing functions.

Gorman's Theorem cannot be directly applied here for a number of reasons. First, the function P is defined on a space Y that is not a Cartesian product of intervals. Second, P need not be continuous everywhere because *Weak Continuity* only requires P to be continuous when the income standard is larger than $\overline{y}^c$. Third, P need not be strictly increasing in all of its arguments because all distributions in Y have at least one non-poor individual and the exact incomes of non-poor individuals are often irrelevant by *Absolute Focus* and *Relative Focus*.[16] Fourth, the arguments of P need not be separable because changing the value of any income can affect the value of the income standard, and its value is used in the preconditions of several axioms. Nevertheless, I show that Gorman's Theorem plays a fundamental role in the proof of Theorem 1 because it applies on a subset of the domain Y.

The proof is organized in five main claims.

In Claim 1, attention is restricted to a subdomain of Y with a fixed population of size $n \geq 4$, a fixed income standard $\overline{y}^* > \overline{y}^c$ and for which no individual has more income than individual n.[17] As individual n is necessarily non-poor, her income has no impact on P by *Relative Focus*. This allows me to associate each distribution y in this subdomain with a distribution in $\times_{i=1}^{n-1}[0,1]$ (with dimension $n-1$) and define a function P^* on $\times_{i=1}^{n-1}[0,1]$ whose value at any $(n-1)$-dimensional distribution in its domain is equal to the value of the n-dimensional distribution that generates it. Finally, I show that P^* is additively separable by proving that P^* satisfies the three assumptions of Gorman's Theorem.

In Claim 2, I show that P^* has the same additively separable functional form as that derived in Claim 1 for *any* fixed population size larger than 4, but the functions used in it are specific to the particular population size considered.

Claim 2 applies to each population size separately. In Claim 3, I show how the functional forms of P for different population sizes are related. Adapting the reasoning of Foster and Shorrocks (1991) and using Claim 2, I show that *Replication Invariance* implies that P is an increasing function

[16] For example, this is the case for income distributions whose income standard is smaller than $\overline{y}^c$ or when a progressive balanced transfer takes place between two non-poor individuals.

[17] By *Weak Continuity*, the requirement that $\overline{y}^* > \overline{y}^c$ ensures that P is continuous. By *Symmetry*, it is without loss of generality to assume that no income exceeds that of individual n.

of the *average* amount that incomes fall below the relative income threshold, where these shortfalls are measured as a fraction of the threshold.

In Claim 4, I show that P has the same functional form derived in Claim 3 for *any* fixed income standard $\overline{y} > \overline{y}^c$ but with the functions entering its expression specific to the particular income standard considered.

In Claim 5, a particular poverty index P° is defined that has the functional form of a hierarchical index. I first show that for any distribution y whose income standard is weakly smaller than $\overline{y}^c$, P° is ordinally equivalent to the expression for P derived in Claim 4. In this step of the argument, *Absolute Focus* plays a fundamental role in extending the expression for P obtained in Claim 4 to distributions with income standards not in the domain for which Claim 4 applies. I then show that P° is ordinally equivalent to the expression for P derived in Claim 4 for any distribution whose income standard is larger than $\overline{y}^c$. Finally, I show that P° satisfies the four properties of a hierarchical index. *Absolute Focus* is once again used in a fundamental way to show that P° satisfies property (iv) in the definition of a hierarchical index.

Proceeding more formally, consider any $n \geq 4$. As will become clear below, the minimal size to be able to use Gorman's Theorem is 4. Let $\overline{y}^*$ be a value of mean income with $\overline{y}^* > \overline{y}^c$. Let Y^* denote the subset of distributions of size n whose mean income is equal to $\overline{y}^*$ and for which no individual has more income than individual n. That is,

$$Y^* = \{y \in Y | n(y) = n \text{ and } \overline{y} = \overline{y}^* \text{ and } y_i \leq y_n \text{ for all } i \leq n\}.$$

For notational convenience, let $z^* = z_r(y)$ denote the relative threshold associated with a distribution y with mean income $\overline{y}^*$. Because $\overline{y}^* > \overline{y}^c$, we have by assumption that $\overline{y}^* > z^a$ and, therefore, $\overline{y}^* > z^*$. This in turn implies that the richest individual is non-poor; that is, $y_n > z^*$ for all $y \in Y^*$.

Consider the function $D : \mathbb{R}_+ \to [0, 1]$ defined by

$$D(w) = \begin{cases} 1 - \frac{w}{z^*} & \text{if } \ w \in [0, z^*], \\ \\ 0 & \text{if } \ w > z^*. \end{cases}$$

Next, consider its "inverse" function $D^- : [0, 1] \to [0, z^*]$ defined by[18]

$$D^-(w) = z^*(1 - w).$$

The functions D and D^- enter the construction of two mappings, M and M^-, which are then used to define a function P^* that satisfies the assumptions of Gorman's Theorem. First, $M : Y^* \to \times_{i=1}^{n-1}[0, 1]$ is defined by

$$M(y) = (D(y_1), \ldots, D(y_{n-1})).$$

Then, $M^- : \times_{i=1}^{n-1}[0, 1] \to Y^*$ is defined by

$$M^-(\nu) = \left(D^-(\nu_1), \ldots, D^-(\nu_{n-1}), n\overline{y}^* - \sum_{k=1}^{n-1} D^-(\nu_k) \right).$$

[18]Strictly speaking, the function D^- is not the inverse of D because the range of D^- is a subset of the domain of definition of D.

The definition of M^- is such that $M^-(\nu) \in Y^*$ for all ν and, thus, the distribution $M^-(\nu)$ has n components. By the definition of its n^{th} component, $M^-(\nu)$ has an income standard equal to $\overline{y}^*$. Furthermore, its n^{th} component is the largest because all of its other components are by definition weakly smaller than z^* and $z^* < \overline{y}^*$. Consequently, we have $M^-(\nu) \in Y^*$.

The mapping M^- and the index P are used to define a function $P^* : \times_{i=1}^{n-1}[0,1] \to [0,1]$. For all $\nu \in \times_{i=1}^{n-1}[0,1]$, the function P^* is defined by

$$P^*(\nu) = P(M^-(\nu)). \tag{6}$$

The function P^* is well-defined because $M^-(\nu) \in Y^*$ for all $\nu \in \times_{i=1}^{n-1}[0,1]$ and P is defined on all of Y^*.

For any two $y, y' \in Y^*$, I show that

$$P(y) \geq P(y') \Leftrightarrow P^*(M(y)) \geq P^*(M(y')). \tag{7}$$

To do so, I show that for all $y \in Y^*$, we have

$$P(y) = P^*(M(y)). \tag{8}$$

By the definition of P^*, this is equivalent to $P(y) = P(M^-(M(y)))$ for all $y \in Y^*$. Even if some $y' \in Y^*$ are such that $y' \neq M^-(M(y'))$, we still have $P(y') = P(M^-(M(y')))$ for such distributions by properties 2 and 3 established for M and M' in Lemma 3.

Lemma 3. *The mappings M and M' satisfy the following three properties:*

1. *M and M' are continuous,*
2. *$P(y) = P(y')$ for any two $y, y' \in Y^*$ such that $M(y) = M(y')$,*
3. *$M(M^-(\nu)) = \nu$ for all $\nu \in \times_{i=1}^{n-1}[0,1]$.*

Proof. (1) The mapping M is continuous on Y^* because function D is continuous on its domain. The mapping M' is continuous on the product space because function D^- is continuous on its domain and so is the n^{th} component of the distribution $M^-(\nu)$.

(2) As individual n is non-poor, we have by the construction of M that $M(y) = M(y')$ only if $Q(y) = Q(y')$ and $y_i = y_i'$ for all $i \in Q(y)$. As all distributions in Y^* have the same income standard $\overline{y}^*$, whose associated relative threshold z^* is such that $z_a < z^*$, *Relative Focus* implies that $P(y) = P(y')$.

(3) By construction we have $D(D^-(w)) = w$ for all $w \in [0,1]$. Therefore, the definitions of M and M' imply this property. ∎

Claim 1. *For fixed $n \geq 4$ and fixed $\overline{y}^*$ such that $\overline{y}^* > \overline{y}^c$,*

$$P^*(\nu) = \tilde{F}\left(\sum_{i=1}^{n-1} \tilde{\phi}(\nu_i)\right) \tag{9}$$

for all $\nu \in \times_{i=1}^{n-1}[0,1]$, where $\tilde{F}$ and $\tilde{\phi}$ are continuous and strictly increasing functions.

To prove Claim 1, I first show that P^* is continuous, strictly increasing in its arguments and separable, and then appeal to Gorman's Theorem.

Lemma 4. *P^* is continuous.*

Proof. Because $\overline{y}^* > \overline{y}^c$, by *Weak Continuity*, P is continuous on $Y^* \subseteq Y$. Because $P^*(\nu) = P(M^-(\nu))$ for all $\nu \in \times_{i=1}^{n-1}[0,1]$ (by definition) and M^- is continuous on $\times_{i=1}^{n-1}[0,1]$ (property 1 in Lemma 3), P^* is continuous. ■

Lemma 5. *P^* is strictly increasing in its arguments.*

Proof. Consider any $i \in \{1, \ldots, n-1\}$, any $(\nu_1, \ldots, \nu_{i-1}, \nu_{i+1}, \ldots, \nu_{n-1}) \in \times_{j=1}^{i-1}[0,1] \times_{j=i+1}^{n-1}[0,1]$ and any $\nu_i, \nu_i' \in [0,1]$ with $\nu_i > \nu_i'$. I show that

$$P^*(\underbrace{\nu_1, \ldots, \nu_{i-1}, \nu_i, \nu_{i+1}, \ldots, \nu_{n-1}}_{=\,\nu}) > P^*(\underbrace{\nu_1, \ldots, \nu_{i-1}, \nu_i', \nu_{i+1}, \ldots, \nu_{n-1}}_{=\,\nu'}).$$

Let $y = M^-(\nu)$ and $y' = M^-(\nu')$. By the construction of the mapping M^-, we have $\overline{y} = \overline{y}'$ and $y_i < y_i'$ with $i \in Q(y)$ and $y_j = y_j'$ for all $j \in Q(y') \backslash \{i\}$. By *Weak Monotonicity*, we have $P(y) > P(y')$, which by Equation (7) implies that $P^*(\nu) > P^*(\nu')$.

Because individual n is never poor on Y^*, by *Relative Focus*, the value of P is independent of the value of y_n. Because M does not associate any argument to individual n, P^* is therefore strictly increasing in all of its arguments. ■

Lemma 6. *P^* is separable.*

Proof. Consider any $\hat{L} \subseteq L$ with $\varnothing \neq \hat{L} \neq L$. Without loss of generality, suppose that $\hat{L} = \{1, \ldots, j\}$. Let $j^\circ = (n-1) - j$. Consider any $u, v \in \times_{i=1}^{j}[0,1]$ and any $w, t \in \times_{i=1}^{j^\circ}[0,1]$. I need to show that

$$P^*(u,w) \geq P^*(v,w) \Leftrightarrow P^*(u,t) \geq P^*(v,t).$$

In order to establish this inequality, I construct four specific distributions $y^{1'''}, y^{2'''}, y^{3'''}, y^{4'''} \in Y \cap \mathbb{R}^{3n}$ that satisfy two properties. First, each of them can be partitioned into three subsets with $3j$, $3j^\circ$ and 1 components whose means incomes are equal to $\overline{y}^*$. Second,

$$\begin{aligned} P(y^{1'''}) = P^*(u,w), \quad P(y^{2'''}) = P^*(v,w), \\ P(y^{3'''}) = P^*(u,t), \quad P(y^{4'''}) = P^*(v,t) \end{aligned} \tag{10}$$

and

$$P(y^{1'''}) \geq P(y^{2'''}) \Leftrightarrow P(y^{3'''}) \geq P(y^{4'''}). \tag{11}$$

This is done in two steps.

Step 1. In this step, I construct $y^{1'''}, y^{2'''}, y^{3'''}, y^{4'''} \in Y \cap \mathbb{R}^{3n}$ that satisfy (10) and which can be partitioned into three subsets of components as described above.

Substep 1.1. Let $y^1, y^2, y^3, y^4 \in Y^*$ be defined by setting $y^1 = M^-(u,w)$, $y^2 = M^-(v,w)$, $y^3 = M^-(u,t)$, and $y^4 = M^-(v,t)$.

The next four substeps aim at constructing from y^1 a particular income distribution $y^{1'''}$ with $P(y^{1'''}) = P(y^1)$ for which the elements of a particular partition of $y^{1'''}$ all have their mean income equal to $\overline{y}^*$, which is a precondition for applying *Weak Subgroup Consistency*.

Substep 1.2. I now partition y^1 into three distributions y^u, y^w and y^1_n such that

$$y^1 = (\underbrace{y^1_1, \ldots, y^1_j}_{= \, y^u}, \underbrace{y^1_{j+1}, \ldots, y^1_{n-1}}_{= \, y^w}, y^1_n),$$

where $u = (D(y^1_1), \ldots, D(y^1_j))$ and $w = (D(y^1_{j+1}), \ldots, D(y^1_{n-1}))$. By the definition of M^-, all $i \notin Q(y^1) \cup \{n\}$ have income $y^1_i = z^*$ and individual n has income

$$y^1_n = n\overline{y}^* - \sum_{k=1}^{n-1} y^1_k. \tag{12}$$

For the distributions y^u, y^w and y^1_n, it need not be the case that $\overline{y}^u = \overline{y}^w = y^1_n = \overline{y}^*$. The next substeps modify y^u, y^w and y^1_n so that the resulting distributions do satisfy these equalities.

Substep 1.3. I now construct a distribution $y^{1'}$ that is a 3-fold replication of y^1. Specifically, $y^{1'}$ is given by

$$y^{1'} = (y^1, y^1, y^1) = (y^u, y^w, y^1_n, y^u, y^w, y^1_n, y^u, y^w, y^1_n).$$

We have $\overline{y}^{1'} = \overline{y}^1$ as replication does not affect the mean. By *Replication Invariance*, $P(y^{1'}) = P(y^1)$.

Substep 1.4. Next, I construct a distribution $y^{1''}$ from $y^{1'}$ for which $P(y^{1''}) = P(y^{1'})$ by implementing particular transfers among the three individuals whose incomes are y^1_n.

Letting $n(u) = j$ and $n(w) = n - (j + 1)$ denote the respective sizes of u and w. Three incomes a_u, a_w and a_x are defined by setting

$$a_u = (3n(u) + 1)\overline{y}^* - 3\sum_{k=1}^{j} y^1_k, \tag{13}$$

$$a_w = (3n(w) + 1)\overline{y}^* - 3\sum_{k=j+1}^{n-1} y^1_k, \tag{14}$$

$$a_x = 3y^1_n - a_u - a_w. \tag{15}$$

It is now shown that $a_u > z^*$, $a_w > z^*$ and $a_x = \overline{y}^*$. Recall that all non-poor individuals in y^u have incomes equal to z^*. Therefore, we have that $3\sum_{k=1}^{j} y^1_k \leq 3n(u)z^*$. Given that $\overline{y}^* > z^*$, it then follows that $a_u > z^*$. The same reasoning shows that $a_w > z^*$. Finally, to show that $a_x = \overline{y}^*$, I substitute the expressions for y^1_n, a_u and a_w in (12), (13) and (14) into (15) and use the identity $n(u) + n(w) + 1 = n$.

The distribution

$$y^{1''} = (y^u, y^w, a_u, y^u, y^w, a_w, y^u, y^w, a_x)$$

is obtained from $y^{1'}$ by implementing balanced transfers among the three non-poor individuals whose income is y^1_n. As balanced transfers do not affect the mean, $\overline{y}^{1''} = \overline{y}^{1'} = \overline{y}^*$. As $Q(y^{1''}) = Q(y^{1'})$ and $y_i^{1''} = y_i^{1'}$ for all $i \in Q(y^{1'})$, we have by *Relative Focus* that $P(y^{1''}) = P(y^{1'})$.

Substep 1.5. I now permute the components of $y^{1''}$ so that the components indexed by u precede the components indexed by w which, in turn, precede the component indexed by x. Let

$$y^{1'''} = (y^u, y^u, y^u, a_u, y^w, y^w, y^w, a_w, a_x).$$

By *Symmetry*, we have $P(y^{1'''}) = P(y^{1''})$. The distribution $y^{1'''}$ is partitioned into three distributions y^{3u}, y^{3w} and a_x as follows:

$$y^{1'''} = (\underbrace{y^u, y^u, y^u, a_u}_{=\, y^{3u}}, \underbrace{y^w, y^w, y^w, a_w}_{=\, y^{3w}}, a_x).$$

I have already shown that $a_x = \overline{y}^*$. It follows from (13) and (14) that $\overline{y}^{3u} = \overline{y}^{3w} = \overline{y}^*$ as well.

Substep 1.5. Finally, Substeps 1.1–1.5 are repeated for y^2, y^3 and y^4 to obtain the distributions $y^{2'''} = (y^{3v}, y^{3w}, a_x)$, $y^{3'''} = (y^{3u}, y^{3t}, a_x)$ and $y^{4'''} = (y^{3v}, y^{3t}, a_x)$. In summary, I have shown that

$$\begin{aligned}
y^{1'''} &= (y^{3u}, y^{3w}, a_x) \qquad && \text{with } P(y^{1'''}) = P^*(u, w),\\
y^{2'''} &= (y^{3v}, y^{3w}, a_x) && \text{with } P(y^{2'''}) = P^*(v, w),\\
y^{3'''} &= (y^{3u}, y^{3t}, a_x) && \text{with } P(y^{3'''}) = P^*(u, t),\\
y^{4'''} &= (y^{3v}, y^{3t}, a_x) && \text{with } P(y^{4'''}) = P^*(v, t),
\end{aligned}$$

where by construction $\overline{y}^{3u} = \overline{y}^{3w} = \overline{y}^{3v} = \overline{y}^{3t} = a_x = \overline{y}^*$. In other words, these distributions satisfy (10) and all of them can be partitioned into the requisite sized sets of components whose mean incomes equal $\overline{y}^*$.

Step 2. In this step, I show that that the distributions $y^{1'''}$, $y^{2'''}$, $y^{3'''}$ and $y^{4'''}$ satisfy (11).

The inequality in (11) can be rewritten as

$$P\left((y^{3u}, y^{3w}, a_x)\right) \geq P\left((y^{3v}, y^{3w}, a_x)\right) \Leftrightarrow P\left((y^{3u}, y^{3t}, a_x)\right) \geq P\left((y^{3v}, y^{3t}, a_x)\right).$$

It is sufficient to prove that

$$P\left((y^{3u}, y^{3w}, a_x)\right) \geq P\left((y^{3v}, y^{3w}, a_x)\right) \Rightarrow P\left((y^{3u}, y^{3t}, a_x)\right) \geq P\left((y^{3v}, y^{3t}, a_x)\right)$$

as the converse implication is obtained by the same argument.

Recall that the distributions I am concerned with all have mean income equal to $\overline{y}^*$. By assumption, $P\left((y^{3u}, y^{3w}, a_x)\right) \geq P\left((y^{3v}, y^{3w}, a_x)\right)$. If $P(y^{3v}) > P(y^{3u})$, because $P\left((y^w, a_x)\right)$ trivially equals $P\left((y^w, a_x)\right)$, *Weak Subgroup Consistency* would be violated.[19] Hence,

$$P(y^{3u}) \geq P(y^{3v}).$$

[19]By construction, we have $n(y^{3u}) = n(y^{3v}) \geq 4$.

Two cases can arise.

Case 1: $P(y^{3u}) > P(y^{3v})$. In this case, *Weak Subgroup Consistency* implies that

$$P\left((y^{3u}, y^{3t}, a_x)\right) > P\left((y^{3v}, y^{3t}, a_x)\right).$$

Case 2: $P(y^{3u}) = P(y^{3v})$. I show by contradiction that this case is such that $P((y^{3u}, y^{3t}, a_x)) \geq P((y^{3v}, y^{3t}, a_x))$. Assume to the contrary that we have

$$P\left((y^{3u}, y^{3t}, a_x)\right) < P\left((y^{3v}, y^{3t}, a_x)\right)$$

Because $P(y^{3u}) = P(y^{3v})$, *Weak Subgroup Consistency* implies that

$$P\left((y^{3u}, y^{3t}, a_x, y^{3v})\right) < P\left((y^{3v}, y^{3t}, a_x, y^{3u})\right).$$

This is a contradiction because the two distributions have equal poverty by *Symmetry*.

The two cases establish that $P((y^{3u}, y^{3t}, a_x)) \geq P((y^{3v}, y^{3t}, a_x))$. As noted earlier, this conclusion is sufficient to establish (11). Using the equivalences in (10), it now follows that P^* is separable. ■

I have shown that P^* satisfies the three properties necessary for Theorem 1, namely that it is continuous, strictly increasing in its arguments, and separable. Because $n \geq 4$, we have that $n - 1 \geq 3$ and so function P^* has enough arguments to apply Theorem 1. Hence, for all $\nu \in \times_{i=1}^{n-1}[0, 1]$,

$$P^*(\nu) = \tilde{F}'\left(\sum_{i=1}^{n-1} \tilde{\phi}_i(\nu_i)\right),$$

where $\tilde{F}'$ and $\tilde{\phi}_i$ are continuous and strictly increasing functions.

The function $\tilde{\phi}_i$ may depend on which value of i is considered. By *Symmetry*, P is invariant to a permutation of its first $n - 1$ arguments on Y^*. Because the function D^- does not depend on i, it then follows from the definition of P^* that P^* is a symmetric function. Therefore, we must have $\tilde{\phi}_i(\nu_i) = \tilde{\phi}(\nu_i) + g(i)$ for some functions $\tilde{\phi}$ and g. Defining the function $\tilde{F}(x) := \tilde{F}'(x + \sum_i g(i))$, a translation of the function $\tilde{F}'$, I conclude that (9) holds for all $\nu \in \times_{i=1}^{n-1}[0, 1]$ for some continuous and strictly increasing functions $\tilde{F}$ and $\tilde{\phi}$.

This concludes the proof of Claim 1.

Claim 1 is now extended to arbitrary $n \geq 4$.

Claim 2. *For any $n \geq 4$ and fixed $\overline{y}^*$ such that $\overline{y}^* > \overline{y}^c$,*

$$P^*(\nu) = \tilde{F}_n\left(\sum_{i=1}^{n-1} \tilde{\phi}_n(\nu_i)\right) \tag{16}$$

for all $\nu \in \times_{i=1}^{n(\nu)-1}[0, 1]$, where $\tilde{F}_n$ and $\tilde{\phi}_n$ are continuous and strictly increasing functions.[20]

[20]In (16), $n = n(\nu) + 1$. For notational simplicity, the dependence of n on ν is suppressed.

The expression for P^* in (9) is valid for all $y \in Y$ for which $n(y)$ is equal to the specified value of n and $\overline{y} = \overline{y}^*$ for the specified value of $\overline{y}^*$ provided that $n \geq 4$ and $\overline{y}^* > \overline{y}^c$. For this value of $\overline{y}^*$, the functional forms of $\tilde{F}$ and $\tilde{\phi}$ are therefore dependent on the choice of n. Claim 2 follows from Claim 1 by taking account of this dependence.

Claim 3. *For fixed $\overline{y}^*$ such that $\overline{y}^* > \overline{y}^c$, for all $y \in Y$ for which $n(y) \geq 4$ and $\overline{y} = \overline{y}^*$,*

$$P(y) = G\left(\frac{1}{n(y)}\sum_{i=1}^{n(y)} d(y_i)\right), \tag{17}$$

where G is continuous and strictly increasing and $d : \mathbb{R}_+ \to [0,1]$ is continuous and strictly decreasing on $[0, z^*]$ with $d(w) = 0$ if $w \geq z^*$ and $d(0) = 1$.

The proof of Claim 3 proceeds in three steps. In the first step, the functions $\tilde{F}_n$ and $\tilde{\phi}_n$ in (16) are transformed so that any term in the summation has a value of 0 if the corresponding income is 0. This is done in such a way that the value of $P^*(\nu)$ is unaffected. This is a harmless normalization. In the second step, I modify the reasoning of Foster and Shorrocks (1991) in order to show that the (transformed versions of) $\tilde{F}_n$ and $\tilde{\phi}_n$ are independent of n. In the third step, using the definition of P^* in terms of P given in (6), I show that P has the functional form given in (17) with the restrictions stated in Claim 3.

Step 1. As a normalization, the following transforms F_n and ϕ_n of $\tilde{F}_n$ and $\tilde{\phi}_n$ are adopted:

$$F_n(w) = \tilde{F}_n\left[w + (n-1)\tilde{\phi}_n(0)\right],$$
$$\phi_n(\nu_i) = n\left[\tilde{\phi}_n(\nu_i) - \tilde{\phi}_n(0)\right].$$

Using these transforms, the expression for P^* in (17) may be rewritten as

$$P^*(\nu) = F_n\left(\frac{1}{n}\sum_{i=1}^{n-1}\phi_n(\nu_i)\right),$$

where $\phi_n(0) = 0$.

Let $I^+(\nu)$ denote the set of i for which $\nu_i > 0$. By the definition of the mapping M^-, we have $I^+(\nu) = Q(M^-(\nu))$. For any $i \leq n-1$ such that $i \notin I^+(\nu)$, $\nu_i = 0$ and, therefore, $\phi_n(\nu_i) = 0$. So, the preceding equation may be rewritten as

$$P^*(\nu) = F_n\left(\frac{1}{n}\sum_{i \in I^+(\nu)}\phi_n(\nu_i)\right), \tag{18}$$

where F_n and ϕ_n are continuous and strictly increasing with $\phi_n(0) = 0$.

Step 2. In this step, I use *Replication Invariance* to prove that the functions F_n and ϕ_n do not depend on n.

For a given $\overline{y}^* > \overline{y}^c$, the set Y and the functions P^*, M, and M^- have been defined for a fixed value of n. I now need to extend these definitions so that they apply to all of the relevant values of n. To economize on notation, I suppress their dependence on n.[21]

From the previous step, we have $\phi_n : [0,1] \to [0, a_n]$ with $\phi_n(0) = 0$ for all $n \geq 4$, where a_n is the largest value that the function ϕ can attain when the population size is n. Consider any $y \in Y^*$ with $n(y) = 4$ such that all individuals except perhaps individual 1 are non-poor in y. Hence, $q(y) \leq 1$. Let $\nu = M(y) = (t, 0, 0)$. By an appropriate choice of y, t can take on any value in $[0, 1]$. Let $y^{\times k} = (y, \ldots, y)$, which is a k-replication of the distribution y. Also, let $\nu' = M(y^{\times k}) = (t, 0, 0, 0, t, 0, 0, 0, \ldots, t, 0, 0)$. The distribution ν' has $3k-1$ zeros and k t's. The size of ν is $n(\nu) = 3$, the size of ν' is $n(\nu') = 4k-1$ and the size of $y^{\times k}$ is s, where $s = n(\nu')+1 = 4k$.

By *Replication Invariance*, we have $P(y) = P\left(y^{\times k}\right)$, which by (the extended version of) (8) is equivalent to $P^*(M(y)) = P^*(M(y^{\times k}))$, which in turn is equivalent to $P^*(\nu) = P^*(\nu')$ by the definitions of ν and ν'. Letting $F = F_4$ and $\phi = \phi_4$ and using the fact that $\phi_n(0) = 0$, it follows that

$$P^*(\nu) = F\left[\frac{1}{4}\phi(t)\right] = F_s\left[\frac{k}{4k}\phi_s(t)\right] = P^*(\nu'),$$

for all $t \in [0, 1]$. Rearranging the second equality, I obtain

$$\phi_s(t) = 4F_s^{-1}\left[F\left(\frac{1}{4}\phi(t)\right)\right].$$

Let $H_s(w) = F_s^{-1}(F(w))$. The function H_s is continuous and strictly increasing on $[0, a_s]$. Using H_s, the preceding equation may be rewritten as

$$\phi_s(t) = 4H_s\left(\frac{1}{4}\phi(t)\right). \tag{19}$$

If $t = 0$, (19) implies that $H_s(0) = 0$ because $\phi(0) = \phi_4(0) = 0$. Note that for $s = 4$, we have $H_4(w) = F^{-1}(F(w)) = w$.

Now consider any $y' \in Y^*$ with $n(y') = 4$ such that all individuals except perhaps individuals 1 and 2 are non-poor in y. Hence, $q(y') \leq 2$. Let $\nu'' = M(y') = (t, u, 0)$. By an appropriate choice of y, t and u can take on any values in $[0, 1]$. Let $y^{\times k'} = (y', \ldots, y')$, which is a k-replication of y'. Let $\nu''' = M(y^{\times k'}) = (t, u, 0, 0, t, u, 0, 0, \ldots, t, u, 0)$. The distribution ν''' k t's, k u's and $2k-1$ zeros.

Applying *Replication Invariance* once again, we have that $P^*(\nu'') = P^*(\nu''')$ and, hence, that $F_s^{-1}[P^*(\nu'')] = F_s^{-1}[P^*(\nu''')]$. Using (18) and (19) together with the fact that $\phi(0) = \phi(0) = 0$, simple algebra establishes that

$$H_s\left(\frac{1}{4}\phi(t) + \frac{1}{4}\phi(u)\right) = H_s\left(\frac{1}{4}\phi(t)\right) + H_s\left(\frac{1}{4}\phi(u)\right).$$

With a change of variables, I obtain the following Jensen equation:

$$H_s(w + w') = H_s(w) + H_s(w').$$

[21]Thus, for example, $M(y) \in \times_{i=1}^{n(y)-1}[0, 1]$ if y has size $n(y)$ and the domains of P^* and M^- are $\cup_{l \in N'} \times_{i=1}^{l} [0, 1]$, where $N' = \{n \in \mathbb{N} | n \geq 3\}$.

Because H_s is defined on an interval of $\mathbb{R}$ and it is continuous and strictly increasing, all solutions to this Jensen equation have the form

$$H_s(w) = \alpha_s w + \beta_s,$$

where $\alpha_s > 0$. Because $H_s(0) = 0$, we have $\beta_s = 0$.

For any $\nu \in \times_{i=1}^{4k-1}[0,1]$, by substituting (19) into (18) and using the definition of H_s, I obtain

$$\begin{aligned} F^{-1}[P^*(\nu)] &= F^{-1}\left[F_s\left(\frac{1}{4k}\sum_{i\in I^+(\nu)} 4H_s\left(\frac{1}{4}\phi(\nu_i)\right)\right)\right] \\ &= H_s^{-1}\left(\frac{1}{4k}\sum_{i\in I^+(\nu)} 4H_s\left(\frac{1}{4}\phi(\nu_i)\right)\right). \end{aligned}$$

Using the fact that $H_s(w) = \alpha_s w$, it then follows that

$$F^{-1}[P^*(\nu)] = \frac{1}{4k}\sum_{i\in I^+(\nu)} \phi(\nu_i).$$

Therefore, for any $y \in Y^*$ with $n(y) = s$ and $\overline{y} = \overline{y}^*$ and its image $\nu = M(y)$, we have

$$P^*(\nu) = F\left(\frac{1}{s}\sum_{i\in I^+(\nu)} \phi(\nu_i)\right). \tag{20}$$

Equation (20) is also valid for all $n(y) \geq 4$ for which $\overline{y} = \overline{y}^*$, not just for $n(y) = s = 4k$ for some positive integer k, as the same reasoning as above can be applied using $n(y)$ and the least common multiple of $n(y)$ and 4.

I now subject ϕ and F to the transformations φ and G, respectively, defined by setting $\varphi(\nu_i) = \frac{\phi(\nu_i)}{\phi(1)}$ and $G(w) = F\big(w\phi(1)\big)$. Using these transformations in (20), we have for all $y \in Y^*$ for which $\overline{y} = \overline{y}^*$ and $n(y) \geq 4$,

$$P^*(\nu) = G\left(\frac{1}{n(y)}\sum_{i\in I^+(\nu)} \varphi(\nu_i)\right),$$

where $\nu = M(y)$. Furthermore, G and φ are continuous and strictly increasing functions and $\varphi : [0,1] \to [0,1]$ is a bijection with $\varphi(0) = 0$ and $\varphi(1) = 1$. Noting that $D(y_i) = 0$ for any $i \notin Q(y)$, the preceding equation can be written as

$$P^*(M(y)) = G\left(\frac{1}{n(y)}\sum_{i=1}^{n(y)} d(y_i)\right), \tag{21}$$

where G is continuous and strictly increasing and $d : \mathbb{R}_+ \to [0,1]$ is continuous and strictly decreasing on $[0, z^*]$ with $d(w) = 0$ if $w \geq z^*$ and $d(0) = 1$. By (8), $P(y) = P^*(M(y))$. Thus, it has thus been shown that (17) is satisfied for $y \in Y^*$

Recall that for $y \in Y^*$, all incomes are no larger than the last income in the distribution y. By *Symmetry*, this proviso can be dropped. Thus, the requirement that $y \in Y^*$ can be replaced by $y \in Y$, which completes the proof of Claim 3.

Claim 3 is now extended to arbitrary $\overline{y} > \overline{y}^c$.

Claim 4. *For all $y \in Y$ for which $\overline{y} > \overline{y}^c$,*

$$P(y) = G_{\overline{y}}\left(\frac{1}{n(y)}\sum_{i=1}^{n(y)} d_{\overline{y}}(y_i)\right), \tag{22}$$

where $G_{\overline{y}}$ is continuous and strictly increasing and $d_{\overline{y}} : \mathbb{R}_+ \to [0,1]$ is continuous and strictly decreasing on $[0, z_r(y)]$ with $d_{\overline{y}}(w) = 0$ if $w \geq z_r(y)$ and $d_{\overline{y}}(0) = 1$.

The expression for P in (17) presupposes that $\overline{y} = \overline{y}^*$ for some arbitrary, but fixed, value of $\overline{y}^* > \overline{y}^c$. The functional forms of G and d in this equation are therefore dependent of the value of $\overline{y}$. Claim 4 follows from Claim 3 by taking account of this dependence.

To complete the proof of Theorem 1, Claim 4 is used to show that the index P is ordinally equivalent to a hierarchical poverty index. That is, P is a strictly increasing transform of an index satisfying (1).

Claim 5. *For all $y \in Y$, P is ordinally equivalent to a hierarchical poverty index.*

Consider a fixed value of $\overline{y}^*$ for which $\overline{y}^* > \overline{y}^c$. For notational convenience, let $d_* = d_{\overline{y}^*}$ and $G_* = G_{\overline{y}^*}$, where $d_{\overline{y}^*}$ and $G_{\overline{y}^*}$ are the functions in Claim 4 for this value of mean income. Let $Y^c = \{y \in Y | \overline{y} > \overline{y}^c\}$ be the subset of Y on which all distributions have mean income larger than $\overline{y}^c$.

For all $y \in Y$, let

$$P^\circ(y) = \frac{1}{n(y)}\sum_{i=1}^{n(y)} p^\circ(y_i, \overline{y}), \tag{23}$$

where (i) for all $y \in Y \backslash Y^c$, $p^\circ(y_i, \overline{y}) = d_*(y_i)$ for all $i \in Q(y)$ and $p^\circ(y_i, \overline{y}) = 0$ for all $i \notin Q(y)$ and (ii) for all $y \in Y^c$, $p^\circ(y_i, \overline{y}) = d_{\overline{y}}(y_i)$. Note that $P^\circ(y)$ has the functional form of a hierarchical poverty index as given in (1). To prove Claim 5, it is shown that $P(y) = G_*(P^\circ(y))$ separately for $y \in Y^c$ and $y \in Y \backslash Y^c$, and then it is shown that the poverty contribution function p° used to define P° satisfies the four properties specified in the definition of a hierarchical poverty index.

Lemma 7. *For all $y \in Y \backslash Y^c$, $P(y) = G_*(P^\circ(y))$.*

Proof. By definition, any distribution $y \in Y \backslash Y^c$ is such that $\overline{y} \leq \overline{y}^c$ and, hence, all of its poor individuals (if any) are absolutely poor. Consider any $y \in Y \backslash Y^c$ with $n \notin Q(y)$. Assuming that $n \notin Q(y)$ is without loss of generality as all distributions in Y have at least one non-poor individual. Consider another distribution $y' \in Y^c$ constructed from y such that $n(y') = n(y)$, $y'_i = y_i$ for all $i \leq Q(y)$, $y'_j = z^*$ for all $j \notin Q(y) \cup \{n\}$ and $y'_n = n(y)\overline{y}^* - \sum_{k=1}^{n(y)-1} y'_k$. By construction, we have $\overline{y}' = \overline{y}^*$.

By the definition of $Y \backslash Y^c$, $Q(y) = Q_a(y)$ because $\overline{y} \leq \overline{y}^c$ implies that $z_a \geq z_r(y)$. The construction of y' then implies that $Q(y') = Q_a(y') = Q(y) = Q_a(y)$ and $y'_i = y_i$ for all $i \in Q(y)$. By *Absolute Focus*, we have $P(y') = P(y)$.

Because $P(y) = P(y')$, $n(y) = n(y')$, $Q(y) = Q(y')$, $\overline{y}' = \overline{y}^*$, $y'_i = y_i$ for all $i \in Q(y)$ and $d_*(y'_i) = 0$ for all $i \notin Q(y)$, by (22) for $\overline{y} = \overline{y}^*$, we have

$$\begin{aligned} P(y) = P(y') &= G_* \left(\frac{1}{n(y')} \sum_{i=1}^{n(y')} d_*(y'_i) \right) \\ &= G_* \left(\frac{1}{n(y)} \sum_{i \in Q(y)} d_*(y_i) \right). \end{aligned} \tag{24}$$

Using the definition of the poverty contribution function p° in (23), the expression on the right-hand side of (24) is $G_*(P^\circ(y))$. ■

Lemma 8. *For all $y \in Y^c$, $P(y) = G_*(P^\circ(y))$.*

Proof. By definition, any distribution $y \in Y^c$ is such that $\overline{y} > \overline{y}^c$. Consider any $y \in Y^c$ with $n \notin Q(y)$. Assuming that $n \notin Q(y)$ is again without loss of generality. The value of $P(y)$ is given by (22).

I want to show that $G_{\overline{y}}(w) = G_*(w)$ for all $w \in [0, 1)$. To do this, it is sufficient to show that this equality holds for all rational numbers in $[0, 1)$ as then the conclusion that it holds on all of $[0, 1)$ follows from the continuity of the functions $G_{\overline{y}}$ and G_*

Consider any rational number $\rho \in [0, 1)$. Because ρ is rational, it can be expressed as $\rho = \frac{q}{n}$ with $q, n \in \mathbb{N}$, $q < n$ and $n \geq 4$. I need to show that $G_{\overline{y}}(\rho) = G_*(\rho)$.

I construct two distributions y' and y'' for which $\overline{y}' = \overline{y}$ and $\overline{y}'' = \overline{y}^*$, with both distributions having a fraction $\rho = \frac{q}{n}$ of poor individuals, all of whom have zero income. The distribution y' is chosen so that $n(y') = n$, $q(y') = q$, $y'_i = 0$ for all $i \in Q(y')$, $y'_j = z_r(y')$ for all $j \notin Q(y') \cup \{n\}$ and $y'_n = n(y')\overline{y} - \sum_{k=1}^{n(y')-1} y'_k$. By construction, $\overline{y}' = \overline{y}$ and, thus, by (22), we have $P(y') = G_{\overline{y}}\left(\frac{q}{n}\right)$ because $d_{\overline{y}}(0) = 1$. The distribution y'' is chosen so that $n(y'') = n$, $q(y'') = q$, $y''_i = 0$ for all $i \in Q(y'')$, $y''_j = z^*$ for all $j \notin Q(y'') \cup \{n\}$ and $y''_n = n(y'')\overline{y}^* - \sum_{k=1}^{n(y'')-1} y''_k$. By construction, $\overline{y}'' = \overline{y}^*$ and, thus, by (22), we have $P(y'') = G_*\left(\frac{q}{n}\right)$. Furthermore, we also have $Q_a(y') = Q(y') = Q_a(y'') = Q(y'')$ and $y'_i = y''_i$ for all $i \in Q(y')$. Therefore, by *Absolute Focus*, $P(y') = P(y'')$ and, hence, $G_{\overline{y}}\left(\frac{q}{n}\right) = G_*\left(\frac{q}{n}\right)$, which is the desired result.

Because the functions $G_{\overline{y}}$ and G_* are identical, by Claim 4,

$$P(y) = G_* \left(\frac{1}{n(y)} \sum_{i=1}^{n(y)} d_{\overline{y}}(y_i) \right), \tag{25}$$

where $d_{\overline{y}}$ is continuous and strictly decreasing on $[0, z_r(y)]$, $d_{\overline{y}}(w) = 0$ when $w \geq z_r(y)$ and $d_{\overline{y}}(0) = 1$. Letting $p^\circ(y_i, \overline{y}) = d_{\overline{y}}(y_i)$, the expression on the right-hand side of (25) is $G_*(P^\circ(y))$. ∎

Lemma 9. *The poverty contribution function p° used to define the index P° satisfies the four properties of a poverty contribution function in the definition of a hierarchical poverty index.*

Proof. Property (i). For $y \in Y \backslash Y^c$, $p^\circ(0, \overline{y}) = 1$ because $d_*(0) = 1$ and $p^\circ(y_i, \overline{y}) = 0$ if $i \notin Q(y)$ by definition. For $y \in Y^c$, $p^\circ(0, \overline{y}) = 1$ because $d_{\overline{y}}(0) = 1$. We also have $p^\circ(y_i, \overline{y}) = 0$ if $i \notin Q(y)$ because $d_{\overline{y}}(w) = 0$ when $w \geq z_r(y)$.

Property (ii). For all $y \in Y$, p° is strictly decreasing in its first argument if $i \in Q(y)$ because both d_* and $d_{\overline{y}}$ are strictly decreasing on $[0, z_r(y)]$.

Property (iii). This property only applies if $y \in Y^c$; that is, when $\overline{y} > \overline{y}^c$. For such y, I need to show that p° is continuous in both of its arguments. It is continuous in its first argument because $d_{\overline{y}}$ is continuous in y_i.

Because P is continuous in y by *Weak Continuity* and G_* is continuous, we also have that P° is continuous in y. The function p° does not depend on the identity of the individual whose income is used in its first argument. Thus, I can suppose that $i \neq n$. We can further suppose that $y_j = y_i$ for all $j \neq n$. That is, I can suppose that $y = (y_i, \ldots, y_i, y_n)$. We have

$$P^\circ(y_i, \ldots, y_i, y_n) = \left[\frac{n(y) - 1}{n(y)} \right] p^\circ(y_i, \overline{y}) \tag{26}$$

because individual n is non-poor and, therefore, $p^\circ(y_n, \overline{y}) = 0$. Variations in $\overline{y}$ are achieved by only varying y_n. Because the mean of a distribution is a continuous function of its arguments and P° is a continuous function, it follows from (26) that p° is continuous in its second argument.

Property (iv). For $y \in Y \backslash Y^c$, the function d_* used to define the value of the function p° does not depend on the mean income. Thus, in this case, p° is constant in its second argument if $i \in Q_a(y)$.

For $y \in Y^c$, I need to show that $p^\circ(w, \overline{y}) = d_*(w)$ for all $w \in [0, z_a)$. Consider any $w \in [0, z_a)$ and any two distributions $y', y'' \in Y^c$ such that $n(y') = n(y'')$, $q(y') = q(y'') = 1$, $Q_a(y') = Q(y') = Q_a(y'') = Q(y'')$, $y'_i = y''_i = w$ for the only $i \in Q(y)$ and the remaining incomes of y' and y'' are such that $\overline{y}' = \overline{y}$ and $\overline{y}'' = \overline{y}^*$. By *Absolute Focus*, we have $P(y') = P(y'')$. By Lemma 8 and (23), it then follows that $p^\circ(w, \overline{y}) = d_*(w)$, which is independent of $\overline{y}$. ∎

This concludes the proof of Theorem 1.

B. The income standard is the median income.

The proof of Theorem 1 when the income standard is the median y_m of the distribution y has the same basic proof strategy as that for the case in which the income standard is the mean income. For this reason, I only describe some of the main ways that the proof for the mean income standard case needs to be modified so as to apply to the median income standard case.

Recall that it has been assumed that $n \geq 4$, so all distributions in Y are of at least size 4. To establish the analogues of Claims 1–4, attention is first restricted to distributions in Y with an even number of individuals for which (i) the value of the income standard $\overline{y}$ (i.e, the median income) is a fixed value $\overline{y}^*$ and (ii) all individuals $i \leq m$ have incomes no larger than individual m and all individuals $j \geq m$ have incomes no smaller than individual m.[22] Letting $\mathbb{E}$ denote the set of positive even numbers, this is the set

$$Y^*_{even} = \{y \in Y | n(y) \in \mathbb{E},\ \overline{y} = \overline{y}^*,\ y_i \leq y_m \leq y_j \text{ for all } i \leq m \leq j\}.$$

The mapping $M^m : Y^*_{even} \to \times_{i=1}^{m-1}[0,1]$ is defined by setting

$$M^m(y) = (D(y_1), \ldots, D(y_{m-1})).$$

A distribution y with size $n(y)$ has an image $M^m(y)$ of size $m-1$. By assumption, any distribution $y \in Y$ is such that $\overline{y} \geq z_a$, and so there are at most $m-1$ poor individuals when the income standard is the median income. By *Relative Focus*, the exact incomes of the other individuals matter only in so far as they help determine the median income y_m. The mapping $M^{m-} : \times_{i=1}^{m-1}[0,1] \to Y^*_{even}$ is defined by setting

$$M^{m-}(\nu) = \left(D^-(\nu_1), \ldots, D^-(\nu_{m-1}), y_m, \ldots, y_m\right),$$

which features $m+1$ individuals earning income y_m.

The definition of the mapping M^m can be used to explain why initially only even-sized distributions are considered. If the definition of M^m is extended so as to also apply to odd-sized distributions, the distributions $y \in Y^*_{even}$ and (y, y_m) (whose size is $n(y)+1$) have the same median income and so would have the *same* image $M^m(y) = M^m(y, y_m)$. However, these two distributions do not exhibit the same degree of poverty because (y, y_m) has one additional non-poor individual, the individual whose income is y_m. By limiting the domain of the mapping M^m to even-sized distributions, this issue does not arise.

Once (17) has been established for distributions in Y^*_{even} using the median as the income standard, it can be extended to distributions $y \in Y \backslash Y^*_{even}$ using *Replication Invariance*. This is done for all $y \in Y \backslash Y^*_{even}$ by setting

$$P(y) = P(y^{\times 2}),$$

[22] As for the case of a mean income standard, distributions with a fixed size n (i.e., $n(y) = n$) are considered before those in which $n(y)$ is allowed to vary. Strictly speaking, when $n(y)$ is allowed to vary, the median individual should be indexed by $m(y)$.

where $P(y^{\times 2})$ is given by (17) because $y^{\times 2} \in Y^*_{even}$.

When the income standard is the median income, the procedure for constructing $y^{1'''}$ from y^1 in Step 1 of the proof of Claim 1 needs modifying. In Substep 1.2, y^1 is now partitioned into three distributions y^u, y^w and y^a for which

$$y^1 = (\underbrace{y_1^1, \dots, y_j^1}_{= y^u}, \underbrace{y_{j+1}^1, \dots, y_{m-1}^1}_{= y^w}, \underbrace{y_m^1, \dots, y_n^1}_{= y^a}),$$

where $u = (D(y_1^1), \dots, D(y_j^1))$, $w = (D(y_{j+1}^1), \dots, D(y_{m-1}^1))$ and the $m+1$ incomes in y^a are all equal to y_m^1. I further partition y^a into y^{au} and y^{aw} with $n(y^{au}) = n(y^u) + 1$ and $n(y^{aw}) = n(y^w) + 1$. Recall that, by definition, $m = \frac{n(y)}{2}$ when $n(y)$ is even, so $n(y) = 2n(y^u) + 2n(y^w) + 2$, as required. In Substep 1.3, the distribution $y^{1'}$ is a 2-replication of y^1 (rather than a 3-replication in the case of the mean income standard). Thus, the distributions $y^{2u} = (y^u, y^u, y^{au}, y^{au})$ and $y^{2w} = (y^w, y^w, y^{aw}, y^{aw})$ which are now used in Substep 1.5 both (i) have at least the minimal size of 4 to be a distribution in the domain for P and (ii) have a median income equal to y_m^1.

Finally, note that the construction of any distribution y for which $\overline{y} = \overline{y}^*$ is simpler when the income standard is the median rather than the mean because it is sufficient to let $y_j = \overline{y}^*$ for all j for which $m \leq j \leq n(y)$.

7.4 Proof of Theorem 2

Take any hierarchical index P, i.e take any contribution function p that satisfies the properties stated in Equation (1). As $z_a > 0$, we have that the absolute and relative lines cross at an income standard $\overline{y}^c$ such that $\overline{y}^c \geq z_a > 0$.

Claim 1: P violates *Transfer*.

Consider a value of mean income $\overline{y}^* = \overline{y}^c + \frac{z_a}{2s}$, which is such that $z_r(\overline{y}^*) = \frac{3}{2} z_a$.[23] I consider two cases.

Case 1: $p(z_a, \overline{y}^*) \geq \frac{1}{2}$.

This case is such that $p(0, \overline{y}^*) - p(\frac{z_a}{2}, \overline{y}^*) < \frac{1}{2}$. Indeed, we have that $p(0, \overline{y}^*) = 1$ and p is strictly decreasing in its first argument on $[0, z_r(\overline{y}^*)]$. By the continuity of p in its first argument on $[0, z_r(\overline{y}^*)]$, there exists an income level r^* with $z_a < r^* < z_r(\overline{y}^*)$ such that

$$p(r^*, \overline{y}^*) < p(z_a, \overline{y}^*) - \left(p(0, \overline{y}^*) - p\left(\frac{z_a}{2}, \overline{y}^*\right) \right). \tag{27}$$

Consider now two distributions $y, y' \in Y$ with $n(y) = n(y')$, $Q(y) = Q(y') = \{1, 2\}$ and whose income standard is equal to $\overline{y}^*$. Distribution y is such that $y_1 = 0$, and $y_2 = r^*$, while distribution y' is such that $y_1' = r^* - z_a$, and $y_2' = z_a$. Distribution y' is obtained from y by a progressive transfer of

[23] When writing $z_r(\overline{y}^*)$, I slightly abuse notation by denoting the argument of the threshold function to be the income standard rather than the whole distribution.

an amount $r^* - z_a$ from the relatively poor individual 2 to the absolutely poor individual 1. As $r^* < z_r(\overline{y}^*)$ and $z_r(\overline{y}^*) = \frac{3}{2}z_a$, we have $r^* - z_a < \frac{z_a}{2}$, and thus inequality (27) implies that

$$\underbrace{p(0,\overline{y}^*) - p(r^* - z_a, \overline{y}^*)}_{\Delta^a} < \underbrace{p(z_a,\overline{y}^*) - p(r^*,\overline{y}^*)}_{\Delta^r}.$$

Δ^a captures the decrease in poverty contribution of individual 1 consecutive to the progressive transfer. In turn, Δ^r captures the increase in poverty contribution of individual 2. As the latter is larger, we have by Equation (1) that $P(y') > P(y)$, a contradiction to *Transfer*.

Case 2: $p(z_a, \overline{y}^*) < \frac{1}{2}$.

Let $a^* > 0$ be the level of income for which $p(a^*, \overline{y}^*) = 1 - p(z_a, \overline{y}^*)$. We have $0 < a^* < z_a$ as p is decreasing in its first argument. Let $\overline{y}^{**} = \overline{y}^c + \frac{a^*}{2s}$, which is such that $z_r(\overline{y}^{**}) = z_a + \frac{a^*}{2}$. As p is constant in $\overline{y}$ for all $a < z_a$, we have $p(a^*, \overline{y}^{**}) = p(a^*, \overline{y}^*)$. As p is decreasing in its first argument, we have $p(\frac{a^*}{2}, \overline{y}^{**}) > p(a^*, \overline{y}^{**})$. We also have $p(z_a, \overline{y}^{**}) = p(z_a, \overline{y}^*)$ since $\overline{y}^{**} > \overline{y}^c$. This follows from the fact that (iv) $p(a, \overline{y})$ is constant in $\overline{y}$ for all $a < z_a$ and (iii) p is continuous in its first argument when the income standard is larger than $\overline{y}^c$. Together, we have that $p(z_a, \overline{y}^{**}) > p(0, \overline{y}^{**}) - p(\frac{a^*}{2}, \overline{y}^{**})$. By the continuity of p in its first argument on $[0, z_r(\overline{y}^{**})]$, there exists an income level r^{**} with $z_a < r^{**} < z_r(\overline{y}^{**})$ such that

$$p(r^{**}, \overline{y}^{**}) < p(z_a, \overline{y}^{**}) - \left(p(0, \overline{y}^{**}) - p\left(\frac{a^*}{2}, \overline{y}^{**}\right)\right). \tag{28}$$

As in case 1, I can construct two distributions $y, y' \in Y$ with $n(y) = n(y')$, $Q(y) = Q(y') = \{1, 2\}$, whose income standard is equal to $\overline{y}^{**}$, with $y_1 = 0$, and $y_2 = r^{**}$ and $y'_1 = r^{**} - z_a$, and $y'_2 = z_a$. As $r^{**} < z_r(\overline{y}^{**})$ and $z_r(\overline{y}^{**}) = z_a + \frac{a^*}{2}$, we have $r^{**} - z_a < \frac{a^*}{2}$. Therefore, $p(r^{**} - z_a, \overline{y}^{**}) > p\left(\frac{a^*}{2}, \overline{y}^{**}\right)$ and inequality (28) implies that

$$\underbrace{p(0,\overline{y}^{**}) - p(r^{**} - z_a, \overline{y}^{**})}_{\Delta^a} < \underbrace{p(z_a,\overline{y}^{**}) - p(r^{**},\overline{y}^{**})}_{\Delta^r}.$$

As Δ^r is larger than Δ^a, we have by Equation (1) that $P(y') > P(y)$, a contradiction to *Transfer*.

Claim 2: P violates *Strong Monotonicity* when the relative line is mean-sensitive.

Take any distribution $y \in Y$ with $y_1 = 0$, $y_2 = z_a$ and $y_j = \frac{1}{n(y)-2}\left(n(y)\overline{y}^c - y_1 - y_2\right)$ for all $j \geq 3$. As the income standard is mean income, we have $\overline{y} = \overline{y}^c$. By construction, only individual 1 is poor in y as $z_r(y) = b + s\overline{y}^c = z_a$.

Construct distribution y^ϵ from y by increasing the income of individual 1 to $y_1^\epsilon = \epsilon$ where $0 < \epsilon < z_a$. This increment implies that $\overline{y}^\epsilon > \overline{y}^c$. By the definition of the relative line, we have $z_r(y^\epsilon) > z_r(y) = z_a$. Therefore, individual 2 is relatively poor in distribution y^ϵ since $y_2^\epsilon = z_a$. Importantly, individual 2's contribution $p(z_a, \overline{y}^\epsilon) > 0$ is constant in the size of ϵ (this follows from properties (iii) and (iv) of p, as explained in Case 2 of Claim 1).

I show that for sufficiently small ϵ, we have $P'(y^\epsilon) - P'(y) > 0$, which implies that P violates *Strong Monotonicity*. As individual 2 is non-poor in distribution y, we have from Equation (1) that

$$P'(y^\epsilon) - P'(y) = \frac{1}{n(y)} \left(p(z_a, \overline{y}^\epsilon) + \underbrace{p\left(\epsilon, \overline{y}^\epsilon\right) - p\left(0, \overline{y}\right)}_{\Delta_\epsilon} \right).$$

By the definition of p, we have $p\left(0, \overline{y}^\epsilon\right) = p\left(0, \overline{y}\right) = 1$. As p is continuous in its first argument, there exists an $\epsilon > 0$ such that $-\Delta_\epsilon < p(z_a, \overline{y}^\epsilon)$, which implies that $P'(y^\epsilon) - P'(y) > 0$, the desired result.

References

Allen, R. C. (2017). Absolute poverty: When necessity displaces desire. *American Economic Review*, 107(12):3690–3721.

Anderson, E. and Esposito, L. (2013). On the joint evaluation of absolute and relative deprivation. *The Journal of Economic Inequality*, 12(3):411–428.

Atkinson, A. and Bourguignon, F. (2001). Poverty and Inclusion from a World Perspective. In Stiglitz, J. and Muet, P.-A., editors, *Governance, equity and global markets*. Oxford University Press, New York.

Chakravarty, S. R. (1983). A New Index of Poverty. *Mathematical Social Science*, 6:307–313.

Corazzini, L., Esposito, L., and Majorano, F. (2011). Exploring the Absolutist vs Relativist Perception of Poverty using a Cross-Country Questionnaire Survey. *Journal of Economic Psychology*, 32(2):273–283.

Decerf, B. (2017). Why not consider that being absolutely poor is worse than being only relatively poor? *Journal of Public Economics*, 152:79–92.

Decerf, B. and Ferrando, M. (2020a). Income poverty has been halved in the developing world, even when accounting for relative poverty. *Mimeo.*

Decerf, B. and Ferrando, M. (2020b). The extended head-count ratio: a new index combining absolute and relative income poverty. *Mimeo.*

Duclos, J. Y. and Gregoire, P. (2002). Absolute and relative deprivation and the measurement of poverty. *Review of Income and Wealth*, 48(4):471–492.

European Commission (2015). *Portfolio of Indicators for the Monitoring of the European Strategy for Social Protection and Social Inclusion.* European Commission, Brussels.

Ferreira, F. H., Chen, S., Dabalen, A., Dikhanov, Y., Hamadeh, N., Jolliffe, D., Narayan, A., Prydz, E. B., Revenga, A., Sangraula, P., Serajuddin, U., and Yoshida, N. (2016). A global count of the extreme poor in 2012: data issues, methodology and initial results. *Journal of Economic Inequality*, 14(2):141–172.

Foster, J., Greer, J., and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures. *Econometrica*, 52(3):761–766.

Foster, J. and Shorrocks, A. (1991). Subgroup Consistent Poverty Indices. *Econometrica*, 59(3):687–709.

Foster, J. E. (1998). Absolute versus Relative Poverty. *American Economic Review*, 88(2):335–341.

Foster, J. E. and Sen, A. (1997). On Economic Inequality after a Quarter Century. In Sen, A., editor, *On Economic Inequality*. Clarendon Press, Oxford, rev. ed. edition.

Gorman, W. M. (1968). Symposium on Aggregation: The Structure of Utility Functions. *Review of Economic Studies*, 35(4):367–390.

Jolliffe, D. and Prydz, E. B. (2017). Societal Poverty: A Relative and Relevant Measure.

Kakwani, N. C. (1980). On a Class of Poverty Measures. *Econometrica*, 48(2):437–446.

Ravallion, M. (2008). On the Welfarist Rationale for Relative Poverty Lines. *Policy Research Working Paper 484*, 4486(January).

Ravallion, M. and Chen, S. (2011). Weakly relative poverty. *Review of Economics and Statistics*, 93(4):1251–1261.

Ravallion, M. and Chen, S. (2019). Global poverty measurement when relative income matters. *Journal of Public Economics*, 117.

Sen, A. (1976). Poverty: an Ordinal Approach to Measurement. *Econometrica*, 44(2):219–231.

Sen, A. (1992). *Inequality Reexamined*. Harvard University Press, Cambridge.

World Bank (2015). *A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals*. World Bank, Washington, DC.

World Bank (2018). *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle*. World Bank, Washington, DC.

Zheng, B. (1997). Aggregate Poverty Measures. *Journal of Economic Surveys*, 11(2):123–162.