Multidimensional poverty measurement
with individual preferences*

January 2014

Koen Decancq

University of Antwerp

Marc Fleurbaey

Princeton University

Francois Maniquet

Université catholique de Louvain

Multidimensional poverty measurement with individual preferences[*]

Koen Decancq[†]   Marc Fleurbaey[‡]   François Maniquet[§]

January 2014

Abstract

We propose a new class of multidimensional poverty indices. Aggregation of the different dimensions relies on individual preferences. The Pareto principle is, therefore, satisfied among the poor. The indices add up individual measures of poverty that are computed as a convex transform of the fraction of the poverty line vector to which the agent is indifferent. The axiomatic characterization of this class retains the usual axioms of focus and subgroup consistency and it introduces additional principles for interpersonal poverty comparisons and for inequality aversion among the poor. We illustrate our approach with Russian survey data between 1995 and 2005. We find that taking preferences into account leads to considerable differences in the identification of the poor compared to standard poverty measures.

JEL Classification: D63, D71.

Keywords: multidimensional poverty measurement, preferences.

[†]CSB, University of Antwerp, CORE, Université catholique de Louvain, and CES, Katholieke Universiteit Leuven. Email: koen.decancq@ua.ac.be.

[‡]Princeton University. Email: mfleurba@princeton.edu.

[§]CORE, Université catholique de Louvain, and University of Warwick. Email: francois.maniquet@uclouvain.be.

1 Introduction

A growing consensus has emerged that well-being is multidimensional, and that income, even suitably deflated for differences in prices, does not qualify as a good proxy for it.[1] There are at least two main reasons. First, private good markets, as well as labor markets, may fail to be competitive, so that individuals suffer from rationing. Second, some relevant goods may not be private and marketable (think of education, security or health).

As a consequence of this growing consensus, poverty is increasingly measured as a multidimensional phenomenon. The common practice consists of defining a threshold in each dimension of well-being, and claiming that an agent is deprived in a dimension if she experiences a lower level than the threshold. Measuring multidimensional poverty then requires a method to aggregate these deprivations across the different dimensions for all poor individuals.

Hence, measuring multidimensional poverty is only possible after two ethical choices have been made. First, the so-called identification issue concerns the demarcation of the set of poor agents. Some researchers adopt the union definition of poverty in which deprivation in at least one dimension is sufficient to qualify as poor[2], others follow the intersection definition and require agents to be below the threshold in all dimensions. Intermediate positions can be taken in which deprivation in a limited number of dimensions is sufficient to qualify as poor (see, for instance, Alkire and Foster [1, 2]). Second, ethical choices about the relative importance of the dimensions (i.e. the weight assigned to the deprivation in each dimension) and whether the dimensions are seen as complements or substitutes have to be made before multidimensional poverty can be measured.

Typically, the researcher measuring poverty makes both ethical choices. For obvious reasons, this practice can be criticized for being arbitrary. Ravallion [29] writes: "... those with a stake in the outcomes will almost certainly be in a better position to determine what weights to apply than the analyst calibrating a measure of poverty."

Turning to the opinions of the poor themselves, a large-scale participatory consultation by the World Bank at the end of the 1990s has indeed endorsed

---

[1] See Kolm [24], Atkinson and Bourguignon [5], Sen [32], Maasoumi [25], and Ravallion [28], among many others.

[2] See, for instance, Tsui [34], Atkinson [4], Bourguignon and Chakravarty [8], Bossert, Chakravarty and D'Ambrosio [7], and Bosmans, Ooghe, and Lauwers [6].

the view that poverty is a multidimensional phenomenon, while at the same time documenting the diversity of views held by the individuals involved (see Narayan et al. [26]). The question now arises whether poverty can effectively be measured as a multidimensional phenomenon, without becoming overly arbitrary on the embedded ethical choices.

In this paper, we therefore propose to use the individuals' own preferences to identify the poor and to aggregate across dimensions. That is, we enrich the model by considering that individuals have possibly different preferences over the different poverty dimensions, and we propose axioms capturing the idea that these preferences should be respected when measuring poverty. Taking preferences into account has the immediate consequence of transforming multidimensional poverty measurement into an aggregation problem of one-dimensional individual well-being levels. That is, the ethical choice of how to weight the goods or assessing their complementarity or substitutability is left to the agents themselves. Respecting preferences also changes the outlook of the identification issue. It does not make sense anymore to think of poverty as deprivation in a number of distinctive dimensions, each with a threshold. Now, the relevant threshold becomes a well-being threshold. An agent is identified as poor if she consumes a bundle of goods that lies in the lower contour set of a preference-specific poverty line vector. Furthermore, the idea of respecting preferences is captured by the requirement that the poverty measure should satisfy a Pareto property among the poor: an increase in the preference satisfaction of a poor agent decreases overall poverty.

A difficulty with the preference-based approach, on the other hand, is that comparison between agents is lost. For instance, we can no longer assume that two agents consuming the same bundle of goods are equally poor. Likewise, we can no longer assume that an identical increase in consumption has the same impact on two agents if they have different preferences. In this paper, we propose solutions to both issues. More specifically, we propose two sets of properties. The first set of properties create some comparability among poor agents. The second set of properties capture the idea that increasing the bundle of goods of an agent decreases poverty more the poorer the agent is.

Our main result is a characterization of all poverty measures that satisfy these properties. Maybe the most surprising result of our inquiry is that the properties force us to define poverty in a very specific way. We obtain that there should exist a single vector of poverty lines, one for each dimension,

3

and an agent qualifies as poor if and only if she prefers the vector of poverty lines over the bundle she is consuming. That is, we endogenize the common practice of defining a poverty line vector, but, at the same time, we introduce a novel identification procedure based on individual preferences. The specific index that evaluates an individual situation is one minus the fraction of the poverty line vector to which the agent is indifferent. This particular way of measuring individual poverty is similar to the ray index proposed by Samuelson [30], the distance function studied by Deaton [10], and the notion of egalitarian-equivalence due to Pazner and Schmeidler [27]. This index has been axiomatically analyzed in the literature on fair social orderings (for a synthetic presentation, see Fleurbaey and Maniquet [14]). One contribution of this paper is to provide a new axiomatic justification for it, in the specific context of poverty measurement.

Taking preferences into account in the measurement of poverty raises new empirical problems. In this paper we illustrate how one can tackle these problems. Using an existing Russian survey data set (RLMS-HSE) between 1995 and 2005, we estimate indifference maps based on a happiness regression. This allows us to compute the proposed poverty indices, to assess the evolution of poverty and to compare our results with other (multidimensional) methods. We find that taking preferences into account leads to important differences in the identification of the poor.

The remainder of the paper is organized as follows. In Section 2, we define the model and derive a representation theorem that extends the classical one-dimensional result of Foster and Shorrocks [17]. In Sections 3 and 4, we introduce our main properties and characterize our poverty measure. Section 5 discusses a generalization towards discrete variables and Section 6 provides an empirical illustration. Section 7 concludes. The proofs are provided in the appendix.

2   The model and a basic result

An economic situation is composed of a set of agents having self-centered preferences over bundles of private goods. We assume that these bundles have ℓ dimensions (which can be interpreted as goods, functionings, ...). We let the set of positive natural numbers, $\mathbb{N}_{++}$, denote the set of all potential agents. An economic situation will always refer to some non-empty and finite set $N \subset \mathbb{N}_{++}$ of agents. Let $\mathcal{N}$ denote the set of non-empty finite

subsets of $\mathbb{N}_{++}$. Each agent $i \in N$ is characterized by her preference relation R_i, an ordering[3] over her consumption set $X \subseteq \mathbb{R}_+^\ell$. We assume that X is convex, compact and contains the ℓ-dimensional 0, the worst possible bundle of goods. The assumption of compactness is not standard.[4] It will give us that all continuous preferences have global maxima over all closed sets.

For two bundles $x_i, x_i' \in X$, we write $x_i R_i x_i'$ to denote that agent i is at least as well off at x_i as at x_i'. The corresponding strict preference and indifference relations are denoted P_i and I_i, respectively. Let $\mathcal{R}$ denote the set of preferences which are continuous, monotonic[5] (that is, for two bundles $x_i, x_i' \in \mathbb{R}_+^\ell$, if $x_i \leq x_i'$, then $x_i' R_i x_i$, and if $x_i \ll x_i'$, then $x_i' P_i x_i$), and convex. We impose an unrestricted domain assumption on $\mathcal{R}$. It contains *all* continuous, monotonic and convex preferences over X. This will play a crucial role in Theorems 2 and 3 below.

Given a set N of agents, an allocation is a list $x_N = (x_i)_{i \in N} \in X^N$ of bundles, where x_i denotes agent i's bundle, for all $i \in N$. For a specified $i \in N$, we sometimes refer to x_N as (x_i, x_{-i}), where x_{-i} stands for the list of bundles of all members of N except i. A *poverty index* is a function $P : \mathcal{S} = \bigcup_{N \subset \mathcal{N}} X^N \times \mathcal{R}^N \to \mathbb{R}_+$, such that $P(x_N, R_N)$ tells us how much poverty we have in allocation x_N when preferences are R_N. We refer to (x_N, R_N) as an economic situation.

We start by addressing the question "who is poor?". In our framework, it amounts to answering the question: with preferences R_i, which bundles make agent i poor? We allow the answer to depend on individual preferences. The only requirement we impose is that poverty is defined in a way that is consistent with agents' preferences. That is, for all $i \in \mathbb{N}_{++}$, all $x_i \in X$, if agent i qualifies as poor (resp., non poor) with bundle x_i, then she also qualifies as poor (resp. non poor) with every bundle she deems equivalent to x_i.

Our first axiom captures this idea. We call it *Focus*, by reference to Sen's Focus axiom (Sen [31]). The axiom requires that the poverty index, at the individual level, be independent of any change in the situation of a non-poor agent.

Axiom 1 FOCUS
There is a function $z : \mathcal{R} \to X$ such that for all $(x_N, R_N) \in \mathcal{S}$, $i \in N$, $x_i' \in X$,

---

[3] An ordering is a complete transitive binary relation.

[4] A similar assumption is imposed in, for instance, Bourguignon and Chakravarty [8].

[5] The three vector inequalities are denoted $\leq$, $<$ and $\ll$.

$R'_i \in \mathcal{R}$, *if*

$$x_i \, P_i \, z(R_i), \; x'_i \, P'_i \, z(R'_i)$$

then

$$P((x'_i, x_{-i}), (R'_i, R_{-i})) = P(x_N, R_N).$$

We now adapt the classical requirement that an improvement in the situation of one agent cannot increase poverty. The core of our contribution in this paper is to use individual preferences to evaluate improvements in terms in individual well-being. We propose to apply a Pareto axiom, restricted to the poor: If the preference satisfaction of all poor agents weakly increases, then poverty weakly decreases. If, in addition, the preference satisfaction of at least one poor agent strictly increases, then poverty strictly decreases.

Axiom 2 PARETO AMONG THE POOR
For all $(x_N, R_N), (x'_N, R_N) \in \mathcal{S}$, *if for all* $i \in N$ *such that* $z(R_i) \, P_i \, x_i$, $x'_i \, R_i \, x_i$, *then*

$$P(x'_N, R_N) \leq P(x_N, R_N).$$

If, in addition, there is $j \in N$ *such that* $z(R_j) \, P_j \, x_j$ *and* $x'_j \, P_j \, x_j$, *then*

$$P(x'_N, R_N) < P(x_N, R_N).$$

The next two axioms are standard axioms of the classical poverty measurement theory that do not need much adjustment to our framework. The next axiom is *Subgroup Consistency*. It requires that overall poverty decreases if it decreases in a subgroup of the population and does not decrease among the other agents.

Axiom 3 SUBGROUP CONSISTENCY
For all $(x_N, R_N), (y_M, R_M), (y'_M, R'_M) \in \mathcal{S}$, $P(y_M, R_M) \geq P(y'_M, R'_M)$ *if and only if*

$$P((x_N, y_M), (R_N, R_M)) \geq P((x_N, y'_M), (R_N, R'_M)).$$

This *Subgroup Consistency* axiom is a powerful decomposability requirement. Observe that the decomposition it proposes does not allow the separation of the bundle from the preferences of an agent. That is the only difference with the classical axiom.

We also impose the axiom of *Continuity*. It requires continuity of the poverty index with respect to agents' bundles.

Axiom 4 CONTINUITY
For all $N \in \mathcal{N}$, $i \in N$, $P(x_N, R_N)$ is continuous in x_i.

The above four axioms enable us to derive the representation theorem that we will use in the remaining of the paper. This result can even be simplified if the axiom of *Replication Invariance* is added. It requires that the poverty measure remains the same if the population is replicated and each replica of the current population exhibits the same characteristics as the current one. We need the following terminology. Let $r \in \mathbb{N}_{++}$ be a positive integer. The economic situation (x_N^r, R_N^r) is a replica of (x_N, R_N) if the set of agents is r times larger than N and is partitioned in r subgroups, one of which is N, and each subgroup has the same distribution of goods and preferences as N.

Axiom 5 REPLICATION INVARIANCE
For all $(x_N, R_N) \in \mathcal{S}$, all $r \in \mathbb{N}_{++}$,

$$P(x_N, R_N) = P(x_N^r, R_N^r).$$

We are now equipped to state and prove the following representation result. It is a generalization of a result obtained by Foster and Shorrocks (1991) in the standard one-dimensional framework. If we gather the above axioms, the resulting poverty index needs to be additively separable in individual characteristics (x_i, R_i).

Theorem 1 *A poverty index P satisfies* Focus, Pareto among the Poor, Subgroup Consistency, *and* Continuity *if and only if there exist*

- *a continuous function $G : \mathbb{R} \times \mathcal{N} \to \mathbb{R}$, strictly increasing in its first argument,*

- *for all $N \in \mathcal{N}$, for all $i \in N$, a function $\phi_i^N : X \times \mathcal{R} \to \mathbb{R}$ such that ϕ_i^N is continuous in its first argument, $\phi_i^N(x_i, R_i) > \phi_i^N(x_i', R_i)$ whenever $z(R_i) \, R_i \, x_i' \, P_i \, x_i$, and $\phi_i^N(x_i, R_i) = 0$ whenever $x_i \, R_i \, z(R_i)$,*

so that for all $(x_N, R_N) \in \mathcal{S}$,

$$P(x_N, R_N) = G\left[\sum_{i \in N} \phi_i^N(x_i, R_i), N\right].$$

Moreover, if Replication Invariance *is added to the axioms, the poverty index can be simplified into*

$$P(x_N, R_N) = G\left[\frac{1}{|N|}\sum_{i \in N}\phi_i(x_i, R_i)\right],$$

for a continuous and strictly increasing function $G : \mathbb{R} \to \mathbb{R}$.

The proofs of all theorems are given in Section 5. Theorem 1 illustrates how taking preferences into account affects the definition of a multidimensional poverty index. The first consequence is that we return to a one-dimensional individual measure of poverty, i.e., the $\phi_i(x_i, R_i)$ measure. Preferences provide a powerful way of aggregating several dimensions into one complete order, and, therefore, into a one-dimensional individual measure. The second consequence is that we lose the comparability that is offered by restricting attention to quantities of goods alone. When preferences are ignored, two agents consuming the same bundle of goods can be assumed to experience the same poverty. This is no longer true once we aggregate dimensions with possibly heterogenous preferences. We cannot resort to classical one-dimensional poverty measurement either, as levels of preference satisfaction are not interpersonally comparable. What we do in the next sections, therefore, is to study a way of comparing poverty levels across individuals —more precisely, across preferences.

3   Inter-preference poverty comparisons

First, we introduce an axiom that puts a natural constraint on inter-preference comparisons of poverty. It requires that an agent moving to an indifference curve below her initial curve (by a change of bundle and, possibly, preferences) should be considered poorer. This requirement can be justified in two ways. First, it involves a respect for preferences that extends the Pareto principle to inter-preference comparisons. When an individual has two (initial and final) indifference curves that do not cross, not only does she have a stable judgment about where the best bundle is, but would sustain this judgment for any other pair of bundles giving her the same levels of preference satisfaction as the contemplated ones. This axiom forces the poverty index to agree with this robust ranking of situations. Second, it is also an informational simplicity requirement. By introducing preferences into the picture,

8

we allow a poverty index to depend on much more information than in the classical theory of poverty measurement. This axiom limits the information that can be taken from preferences to evaluate an economic situation to the indifference curve containing the actual bundle.

To state the axiom, we need to introduce some additional terminology. For all $i \in \mathbb{N}_{++}$, $R_i \in \mathcal{R}$ and $x_i \in X$, we let $L(x_i, R_i) \subseteq X$ denote the (closed) lower contour set of R_i at x_i, that is, the set of bundles that agent i deems not better than x_i when she has preferences R_i. In a similar way, we define the upper contour set, $U(x_i, R_i)$, and the indifference surface $I(x_i, R_i) = L(x_i, R_i) \cap U(x_i, R_i)$.

Axiom 6 Nested Contours
For all $N \subset \mathbb{N}_{++}$, $(x_N, R_N) \in \mathcal{S}$, $i \in N$, $x_i' \in X$, $R_i' \in \mathcal{R}$, if

$$U(x_i, R_i) \cap L(x_i', R_i') = \varnothing,$$

then

$$P((x_i', x_{-i}), (R_i', R_{-i})) \geq P(x_N, R_N).$$

In combination with *Continuity*, *Nested Contours* implies that when

$$I(x_i, R_i) = I(x_i', R_i'),$$

then

$$P((x_i', x_{-i}), (R_i', R_{-i})) = P(x_N, R_N).$$

Therefore, adding *Nested Contours* to the axioms already present in Theorem 1 implies that the second argument of the ϕ_i functions can be simplified from R_i to $I(x_i, R_i)$. The results of this section build on this fact.

There are two ways of introducing inter-preference poverty comparisons. The first way consists of extending nested contours from one-to-one comparisons to one-to-two comparisons. The second way consists of comparing preferences and stating that having some preferences, everything else equal, making poverty more severe than other preferences. In this section, we present one axiom of each type. Our main (and somehow surprising) result is that, starting from the representation result of Theorem 1, both axioms are equivalent.

We begin with the idea of extending the underlying principle of *Nested Contours*. This is illustrated in Figure 1. *Nested Contours* implies that a

9

situation (x_i, R_i) can have lower poverty than another situation (x_i', R_i') only if the lower contour set of the former is not included in the lower contour set of the latter. Now introduce a third situation (x_i'', R_i''), such that $L(x_i, R_i)$ is included in the union of the two other lower contour sets.

It is not obvious that (x_i, R_i) should be considered a worse situation than either of the other situations, because, as in the figure, the lower contour set at (x_i, R_i) may contain bundles which are not in one of the other lower contour sets. From *Nested Contours* we know that if $L(x_i, R_i)$ was included in the *intersection* of the other two lower contour sets, then (x_i, R_i) would be at least as bad as the *worst* of the other two. What the new axiom requires is that if $L(x_i, R_i)$ is included in the *union* of the two other lower contour sets, then (x_i, R_i) is not better than the *best* of the other two.

What is interesting about this axiom is that it leaves unspecified the type of bundles or preferences that induce more poverty, that is, no restriction is imposed on the *value* of any $P(x_i, R_i)$. It only extends the possibilities of comparisons of lower contour sets.
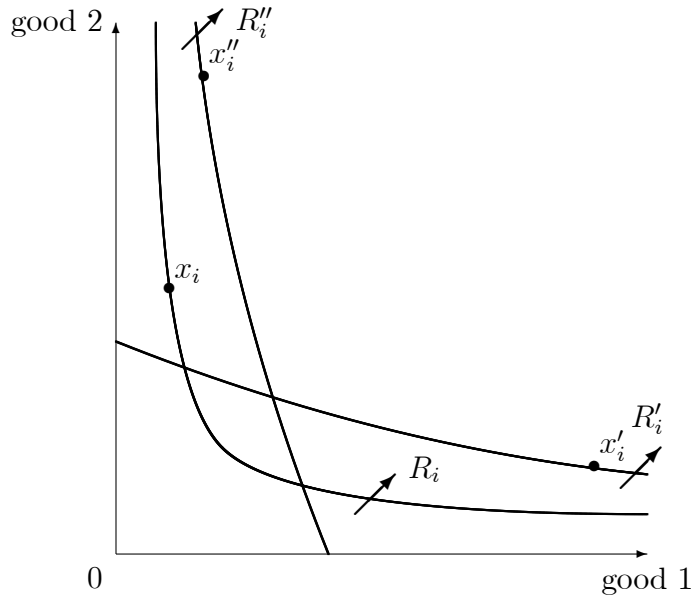

Figure 1: An illustration of *Nested Unions*: $P(x_i, R_i) \geq \min\{P(x_i', R_i'), P(x_i'', R_i'')\}$

Axiom 7 NESTED UNIONS

For all $N \subset \mathbb{N}_{++}$, $i \in N$, $(x_N, R_N) \in \mathcal{S}$, if for $x_i', x_i'' \in X$ and $R_i', R_i'' \in \mathcal{R}$,

$$L(x_i, R_i) \subseteq L(x_i', R_i') \cup L(x_i'', R_i''),$$

then

$$P\left(x_N, R_N\right) \geq \min\left\{P\left((x_i', x_{-i}), (R_i', R_{-i})\right), P\left((x_i'', x_{-i}), (R_i'', R_{-i})\right)\right\}.$$

Recall that *Nested Contours* implies that two situations with the same indifference set must be equally poor. We obtain a similar implication here, if *Nested Unions* is combined with *Nested Contours*. Observe from the figure that the union of two lower contour sets is the lower contour set of a convex preference, and that by *Nested Contours*, this corresponds to a situation that is at least as good as the best of the two. The new axiom therefore implies that it must be just as good as the best of the two.

The second way of constructing interpersonal poverty comparison consists of directly comparing preferences at a given bundle. We already know that the axioms of the previous section force us to conclude that two agents with different preferences may have different poverty levels at the same bundle. This means that poverty at some bundle of goods may be more or less severe depending on agents' preferences.

We consider that a preference relation qualifies as "the" worst preferences at a bundle if an agent with such preferences is at least as poor as with any other preferences. Worst preferences necessary exist at all bundles. We require that the ethical judgments that govern the identification of worst preferences at each bundle be consistent across bundles in the following way: there should exist some preference relation that is the worst preference at all bundles. This requirement is weaker than it may look at first sight. We may think that the decision that a particular preference relation is the worst one at some bundle should only depend on the local information around that bundle. The axiom is consistent with this idea. Simply, it requires that all those local requirements be consistent with each other in the sense that they should all be integrable into one preference relation.

Formally, *Worst Preferences* requires that there exist some preferences R^w such that whenever an agent's preferences change to R^w, everything else remaining constant, poverty weakly increases.[6]

_____

[6]One could also introduce a mirror axiom of *Best Preferences*, but the focus on the

Axiom 8 WORST PREFERENCES
There is $R^w \in \mathcal{R}$ such that for all $(x_N, R_N) \in \mathcal{S}$, all $i \in N$,

$$P((x_i, x_{-i}), (R^w, R_{-i})) \geq P(x_N, R_N).$$

We proceed by illustrating the axiom with an example. An ethical observer may consider that the worst preferences at a given bundle are the preferences such that preference satisfaction only increases after an increase of the good of which this bundle is most deprived. This is illustrated in Figure 2. Two agents are consuming the same bundle $x_j = x_k$. At this bundle, suppose that the ethical observer considers that the largest deprivation is in good 1. Agent j is able to substitute good 2 for good 1, whereas agent k is not. As a consequence, this ethical observer will consider that agent k has the worst preferences. This reasoning leads to adopting some Leontief preferences as the worst preferences. The locus of all cusps in the Leontief preferences partitions the consumption set into regions in which one (and only one) good is incurring the strongest deprivation.

The main message of this section is that the two approaches to constructing interpersonal poverty comparison are equivalent. Moreover, we can characterize the family of poverty indices that satisfy the additional axioms. Before we state and prove the theorem, let us illustrate this family. Figure 3 represents the indifference curve that partitions the consumption set of an agent with worst preferences into the bundles with which she qualifies as poor (below the curve through $z(R^w)$) and those with which she does not qualify as poor (above that curve). Let us look at agent j. Bundle z lies at the tangency between the indifference curve through $z(R^w)$ and one indifference curve for j. In the graph, this bundle is unique. If it were not, then agent j would be indifferent among all bundles at the tangency between one of his indifference surface and this particular indifference curve for the worst preferences. We look at three bundles for agent j, x_j, x_j', x_j'', such that $x_j'' \, P_j \, x_j' \, I_j \, z \, P_j \, x_j$. We show that $z(R_j) \, I_j \, z$. We cannot have $z(R_j) \, I_j \, x_j$. Indeed, by the unrestricted domain assumption, there exists some $R_j' \in \mathcal{R}$ such that $I(x_j, R_j)$ and $I(z(R^w), R^w)$ are indifference curves for R_j' as well. By *Nested Contours* and *Continuity*, we should have $\phi_j(x_j, R_j') = \phi_j(x_j, R_j)$ and $\phi_j(z(R^w), R_j') = \phi_j(z(R^w), R^w)$. By *Focus* and *Continuity*, we should

---

worst possible case appears more relevant in the context of poverty. Moreover, as it turns out, an axiom of *Best Preferences* is incompatible with the axioms considered in the next section.
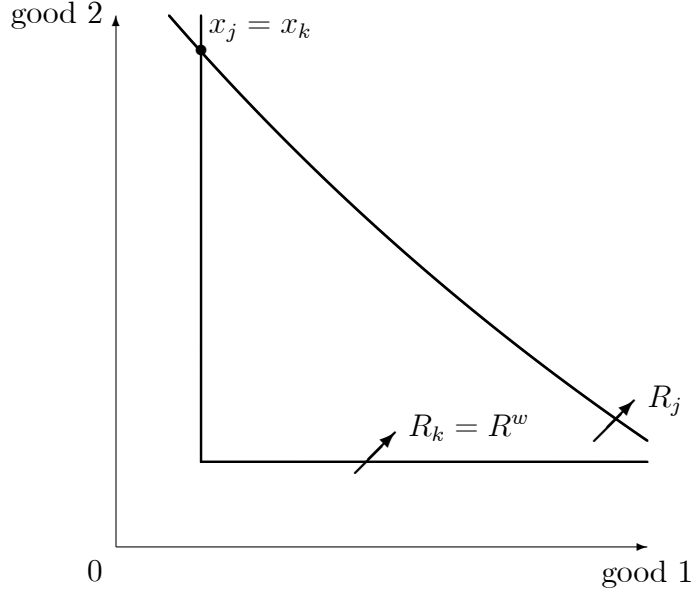
Figure 2: An illustration of *Worst preferences*: $P(x_k, R_k) \geq P(x_j, R_j)$

have $\phi_j(x_j, R_j) = \phi_j(z(R^w), R^w) = 0$. By transitivity, we should have $\phi_j(x_j, R'_j) = \phi_j(z(R^w), R'_j)$, violating *Pareto among the Poor*. We cannot have $z(R_j) I_j x''_j$ either. Indeed, there necessarily exists a bundle z' such that $z' \gg z$ and $x''_j P_j z'$. By *Pareto among the Poor*, $\phi_j(z', R_j) > \phi_j(x''_j, R_j)$. By *Focus*, on the other hand, $\phi_j(z'^w) = \phi_j(z, R^w) = 0$, violating *Worst Preferences*. We are therefore left with $z(R_j) I_j z$. This reasoning does not depend on our choice of R_j, but only on the shape of R^w.

What this shows is that $z(R_j)$ must lie on the indifference curve of R_j that is tangent from below to the indifference curve of R^w through $z(R^w)$ (such a tangent indifference curve always exists because X is compact). Now, this reasoning can be extended to any other indifference curve of R^w below the one containing $z(R^w)$. Assume x^w is such that $\phi_j(x^w, R^w) = \phi > 0$. By a similar argument, we can deduce that $\phi(x_j, R_j) = \phi$ if $I(x_j, R_j)$ is the indifference curve of R_j that is tangent from below to $I(x^w, R^w)$. This characterizes the family of poverty indices that we obtain by adding either *Worst Preferences* or *Nested Unions* to the list of axioms we already imposed.

Theorem 2 *Let P be a poverty index in the family characterized in Theorem*
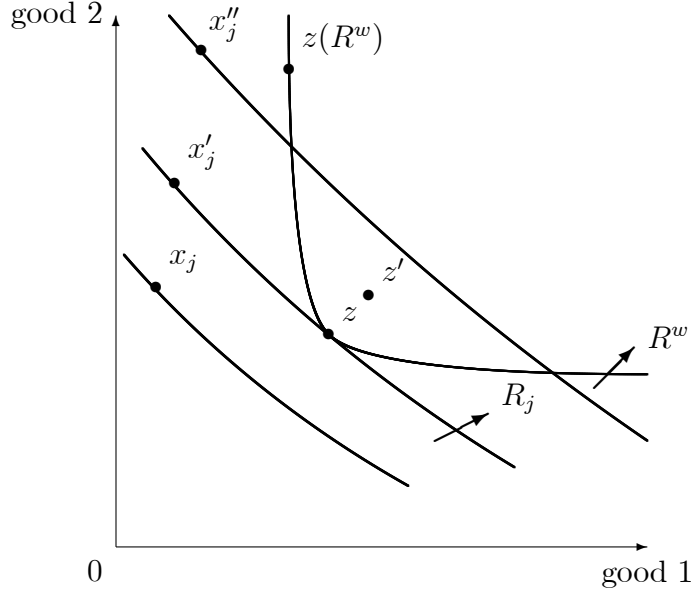
13

Figure 3: The poverty threshold $z(R_j)$ needs to be equivalent to z

1, satisfying Nested Contours. *The following three claims are equivalent.*

1. *P satisfies* Nested Unions.

2. *P satisfies* Worst Preferences.

3. *There exist*

- $R^w \in \mathcal{R}$,
- *a continuous and strictly increasing function $G : \mathbb{R} \to \mathbb{R}$,*
- *a function $\phi : X \times \mathcal{R} \to \mathbb{R}$ such that*
  - *ϕ is continuous in its first argument,*
  - *$\phi(0, R^w) = 1$, $\phi(x_i, R^w) = 0$ for all $x_i \in X$ such that $x_i \, R^w \, z(R^w)$, and $\phi(x_i, R^w) > \phi(x_i', R^w)$ whenever $z(R^w) \, R^w \, x_i' \, P^w \, x_i$, and*
  - *$\phi(x_i, R_i) = \phi(x_i', R^w)$ if and only if $L(x_i, R_i) \cap U(x_i', R^w) \neq \emptyset$ and $intL(x_i, R_i) \cap intU(x_i', R^w) = \emptyset$ (that is, $I(x_i, R_i)$ is tangent to $I(x_i', R^w)$ from below),*

14

so that for all $(x_N, R_N) \in \mathcal{S}$,

$$P(x_N, R_N) = G \left[\frac{1}{|N|} \sum_{i=1}^{|N|} \phi(x_i, R_i) \right].$$

4   Poverty Sensitivity

Theorem 2 above tells us how to construct functions $\phi(\cdot, R_i)$ for all $R_i \in \mathcal{R}$, once $\phi(\cdot, R^w)$ has been defined. How to define these ϕ functions more completely is the topic of this section.

We study two axioms in this section. They are consistent with the idea that providing more resources to a poor agent decreases poverty more the poorer the agent is. We begin with a natural but strong axiom capturing this idea. It requires that if two agents consume bundles having the property that one bundle dominates the other one in each dimension, then poverty decreases more whenever the agent with the smallest bundle receives an additional amount of resources than if this increment of resources is assigned to the agent with the largest bundle.

Axiom 9 POVERTY SENSITIVITY
For all $(x_N, R_N), (x'_N, R_N), (x''_N, R_N) \in \mathcal{S}$, *if there exist* $j, k \in N$ *such that* $x_j = x''_j \gg x_k = x'_k$, $x'_j = x_j + \mu(x_j - x_k)$, $x''_k = x_k + \mu(x_j - x_k)$, *for some* $\mu \in (0, 1)$, *and for all* $i \neq j, k : x_i = x'_i = x''_i$, *then*

$$P(x_N, R_N) - P(x''_N, R_N) > P(x_N, R_N) - P(x'_N, R_N).$$

Note that the requirement of the axiom is equivalent to $P(x''_N, R_N) < P(x'_N, R_N)$. As a result, we could equivalently write the axiom as a transfer principle, that is, with two economic situations $(x_N, R_N), (x'_N, R_N)$ such that $x_j \gg x_k$, $x'_j = (1 - \mu)x_j + \mu x_k$ and $x'_k = (1 - \mu)x_k + \mu x_j$ for some $\mu \in (0, \frac{1}{2})$ (that is, bundles x'_j and x'_k are convex combinations of x_j and x_k), and then require $P(x_N, R_N) > P(x'_N, R_N)$. This transfer principle is weaker than the standard multidimensional transfer principle considered in the literature on multidimensional inequality (See, Weymark [36], among others). In that literature, typically $\mu \in (0, 1)$ allowing for rank reversals and, more importantly, there is no requirement of component-wise dominance imposed between bundles of donor and recipient so that transfers may go in opposite directions for different dimensions.

Poverty Sensitivity defines who, between agents j and k, is poorer than the other only as a function of the bundles they consume. It turns out that this is incompatible with our general project of taking preferences into account. Indeed, this axiom is incompatible with *Pareto among the Poor*. This finding is actually reminiscent of a result proven in Fleurbaey and Trannoy [15].

Lemma 1 *No poverty index P satisfies* Pareto among the Poor *and* Poverty Sensitivity.

The following graphical sketch of the proof will help us understand where the difficulty comes from and how to circumvent it. In Figure 4, two pairs of economic situations are represented. When agents j and k consume x_j and x_k, agent k is poorer, so that, by *Poverty Sensitivity*, there is less poverty in (x'_j, x'_k) than in (x_j, x_k). When agents j and k consume x''_j and x''_k, agent j is poorer, so that there is less poverty in (x'''_j, x'''_k) than in (x''_j, x''_k). The problem is that (x_j, x_k) and (x'''_j, x'''_k) are Pareto equivalent, and (x'_j, x'_k) and (x''_j, x''_k) are Pareto equivalent, too. By a successive application of *Pareto among the Poor* and *Poverty Sensitivity*, we should have $P(x_N, R_N) > P(x'_N, R_N) = P(x''_N, R_N) > P(x'''_N, R_N) = P(x_N, R_N)$, which entails a contradiction.

The difficulty is related to the fact that the relationship "is poorer than" defined in terms of component-wise dominance between bundles (condition $x_j \gg x_k$ in the axiom) is too loose. With this definition, agent k qualifies as poorer than agent j at (x_j, x_k) but richer at (x'''_j, x'''_k) whereas their respective well-being remains the same.

To avoid the difficulty revealed by Lemma 1, we propose to weaken *Poverty Sensitivity*. We weaken it in two directions. First, we allow the set of bundles among which the transfer among poor agents is required to decrease poverty to be a subset of X. This subset is required to be convex, so that the bundles obtained after the equalizing transfer are also in the set, guaranteeing that a further transfer reducing inequality is still considered desirable.

Second, we further require that the agent who is consuming more before the transfer is also considered richer than the other agent (in the precise sense that if she were the only agent in society, poverty would be lower than if the receiving agent were the only agent in society). This guarantees that an agent consuming strictly more of all goods than another agent is not necessarily qualified as richer, which is consistent with our general objective of making
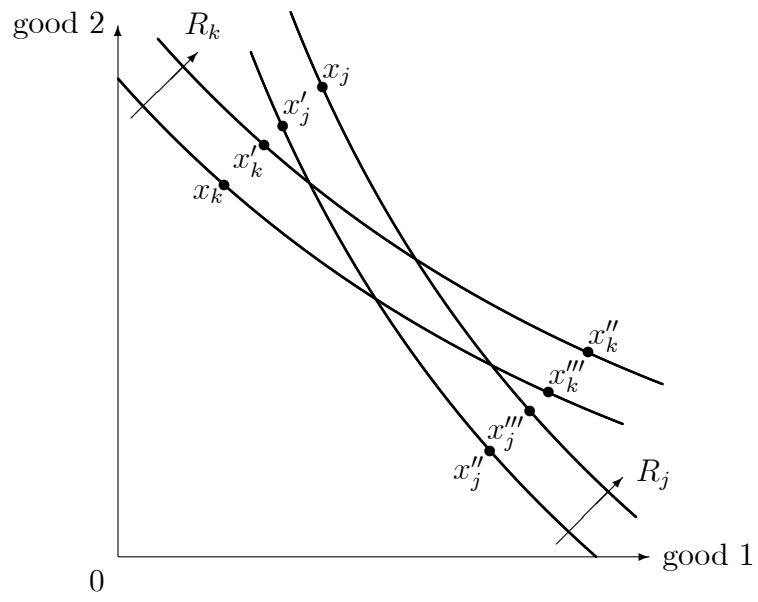
Figure 4: *Pareto Efficiency among the Poor* and *Poverty Sensitivity* are incompatible

our poverty judgments depend on preferences.

Either restriction would allow us to escape the difficulty revealed by Lemma 1. We choose to impose both restrictions, in order to obtain a reasonably weak axiom. As our last result will prove, combining this axiom with the axioms we already defined will allow us to single out a unique family of indices.

We need the following terminology. Let $\mathcal{T}$ be defined by: for all $T \in \mathcal{T}$, (i) $T \subset X$, (ii) T is convex, and (iii) for all $R_i \in \mathcal{R}$, all $x_i \in X$, $I(x_i, R_i) \cap T \neq \emptyset$.

Axiom 10 $\mathcal{T}$-POVERTY SENSITIVITY
There exists $T \in \mathcal{T}$ such that for all $(x_N, R_N), (x'_N, R_N) \in \mathcal{S}$, if there exists $j, k \in N$ such that

- $x_j, x'_j, x_k, x'_k \in T$,

- $x_j \gg x_k$, $x'_j = (1 - \mu)x_j + \mu x_k$ and $x'_k = (1 - \mu)x_k + \mu x_j$ for some $\mu \in (0, \frac{1}{2})$,

- $P(x_j, R_j) < P(x_k, R_k)$, and

- *for all $i \neq j, k : x_i = x'_i$,*

then
$$P(x_N, R_N) > P(x'_N, R_N).$$

The following theorem proves that it is possible to add $\mathcal{T}$-*Poverty Sensitivity* to the list of axioms of Theorem 2. The theorem also characterizes the consequence of $\mathcal{T}$-*Poverty Sensitivity* on the choice of R^w. Only one preference relation can be chosen as the worst, i.e., the Leontief preferences with the cusps along a particular ray. This finding is equivalent to identifying a single poverty line vector z such that an agent is poor if and only if she consumes a bundle to which she prefers z. Individual poverty, as measured by the ϕ function, can be any decreasing and convex (that is, inequality averse) function computed from the fraction λ of z to which an agent is indifferent. The theorem is silent, though, on the choice of z, the poverty line vector, and the shape of the decreasing and convex transformation function ϕ, which remain the degrees of freedom of the ethical observer.

Theorem 3 *Let P be a poverty index in the family characterized in Theorem 2. It satisfies $\mathcal{T}$-Poverty Sensitivity if and only if there exist*

- $z \in \mathbb{R}_{++}^{\ell}$ such that $x_i R^w x'_i \Leftrightarrow \min_{l \in \{1, \ldots, \ell\}} \frac{x_{il}}{z_l} \geq \min_{l \in \{1, \ldots, \ell\}} \frac{x'_{il}}{z_l}$,

- a continuous and strictly increasing function $G : [0, 1] \to \mathbb{R}$,

- a continuous, decreasing and convex function $\phi : [0, 1] \to [0, 1]$

so that for all $(x_N, R_N) \in \mathcal{S}$,

$$P(x_N, R_N) = G\left[\frac{1}{|N|} \sum_{i=1}^{|N|} \phi(1 - \min\{1, \lambda(x_i, R_i)\})\right],$$

where $\lambda(x_i, R_i) = \lambda$ if and only if $x_i I_i \lambda z$.

Here is the intuition of why only Leontief preferences can be used as worst preferences. Let us consider Figure 5. Let us assume that R^w is not Leontief and $\mathcal{T}$-Poverty Sensitivity is satisfied with respect to some ray from the origin. By Theorem 2, we already know that $\phi(\tilde{x}_j, R^w) = \phi(x_j, R_j)$ and $\phi(x'_j, R^w) = \phi(x''_j, R_j)$. $\mathcal{T}$-Poverty Sensitivity implies that

$$\phi(x'_j, R_j) - \phi(x_j, R_j) \geq \phi(x'''_j, R^w) - \phi(x''_j, R^w),$$

which is equivalent to

$$\phi(x''_j, R^w) - \phi(\tilde{x}_j, R^w) \geq \phi(x'''_j, R^w) - \phi(x''_j, R^w). \tag{1}$$

The problem comes from the fact that the distance between the indifference curves of R^w though $\tilde{x}_j$ and x''_j can be made arbitrarily small, for a fixed $x'''_j - x''_j$. In other words, we can choose R_j so as to have an arbitrarily small $\phi(x''_j, R^w) - \phi(\tilde{x}_j, R^w)$ compared to a fixed $\phi(x'''_j, R^w) - \phi(x''_j, R^w)$. As a consequence, no continuous function ϕ can satisfy Eq. 1. This proves that $\mathcal{T}$-Poverty Sensitivity cannot be satisfied for T as represented in the figure. It turns out that whatever T, a construction like the one in the figure can be made, unless R^w is of the Leontief type and T is precisely the ray at which R^w has cusps.

We end this section with a discussion of the poverty line vector z. In the multidimensional literature, that vector consists in the dimension-wise thresholds below which an individual is viewed as deprived in that dimension. In our approach, however, the choice of the poverty line vector is concomitant to the choice of the Leontief preferences that are considered the
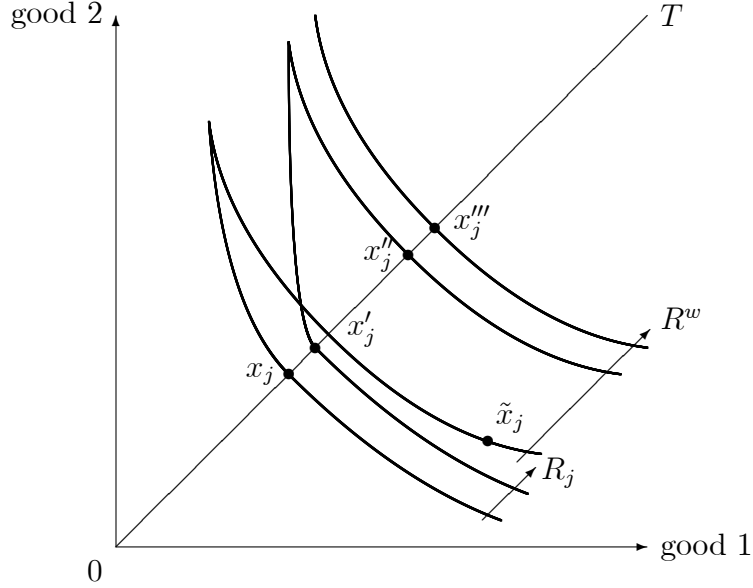
19

Figure 5: Why worst preferences need to be Leontief preferences

worst preferences. Choosing a poverty line vector, therefore, amounts to dividing the consumption set in as many subspaces as there are goods, where one good can be considered as least abundant or most deprived. The poverty line vector and corresponding division of the consumption space should then be chosen so that an agent who is unable to trade-off between the most deprived good and any other good (i.e. who has Leontief preferences that only depend on the consumption of the most deprived good in the respective subspace) is ceteris paribus believed to be the worst-off.

5   Generalization to discrete variables

In practice, some variables that one may want to include in a multidimensional analysis of poverty are discrete in nature (see, amongst others, Alkire and Foster [1], Bossert et al. [7] and Bosmans et al. [6]). In this section, we briefly and informally describe therefore how the results obtained in the preceding sections generalize to the case in which agents' consumption bundles are composed of goods that come in discrete quantities, in addition to the

20

goods represented by the continuous variables of our previous model.

To simplify the exposition, let us only add one binary variable, for instance an indicator whether the individual is being unemployed or not). That is, the consumption set is now $X \subseteq \mathbb{R}_+^\ell \times \{0, 1\}$, and we describe the consumption of an agent by a list (x_i, d_i) such that $x_i \in \mathbb{R}_+^\ell$ and $d_i \in \{0, 1\}$. The discussion of this section generalizes immediately to the more general case with several additional variables, and with discrete rather than binary variables, so that the consumption set will be a finite product of Euclidean spaces.

To generalize our previous results we require that individual preferences are able to compare bundles in which $d_i = 0$ with bundles in which $d_i = 1$. A simple way of doing so is by assuming that the divisible goods (which include food, for instance) are necessary to survive, with the consequence that whether the binary good is consumed or not does not matter for someone whose consumption of the necessary goods is zero. Formally, this is the assumption that for all admissible preferences $R_i \in \mathcal{R}$, if $x_i = (0, \ldots, 0)$, then $(x_i, 0) \, I_i \, (x_i, 1)$. This simplifying assumption is not necessary for the generalization to hold true, but we find it sufficiently realistic to use it in this brief section.

None of axioms 1 to 8 requires any rewriting, and Theorems 1 and 2 are still valid. This can be checked by going throughout the proofs that are provided in the appendix and observing that all steps are still valid if the model is changed to accommodate an additional binary variable. Without entering into the details, let us simply point out that Theorem 1 continues to hold because our assumption that divisible goods are necessary guarantees that individual poverty measures range in an interval, and all the other steps follow. Theorem 2 is not affected by the change in the definition of the consumption set, provided the arguments involving convex upper contour sets are now rewritten to bear on upper contour sets that are the union of sets that are convex in the two Euclidean spaces (corresponding to the spaces in which $d_i = 0$ and that in which $d_1 = 1$).

This means that the consequence of imposing axioms 1 to 8 when agents have this new consumption set still amounts to choosing some worst preferences and evaluating poverty of any agent by looking at the lowest possible poverty level according to the worst preferences associated to a bundle in the lower contour sets of this agent's consumption bundle, as required by Theorem 2.

The generalization is slightly more complicated when one adds $\mathcal{T}$-*Poverty*

Sensitivity. How should we define the set of bundles $\mathcal{T}$ in which transfers will be said to decrease poverty? The easiest answer consists probably in requiring again this set to be convex. As a consequence, set $\mathcal{T}$ is either in the space defined by $d_i = 0$ or in the space $d_i = 1$. Unsurprisingly, if we choose the space $d_i = 1$, then the worst preferences must be defined by the following two requirements: 1) they must be Leontief in that space along a ray from 0 to a bundle z, and 2) they must have the strongest possible preference for bundles in the space $d_i = 1$ over bundles in the space $d_i = 0$, which means that all other admissible preferences must require a lower increase in x_i to leave the agent indifferent between a bundle with $d_i = 1$ and one with $d_i = 0$.

Intuitively, this is reminiscent of what we found in the previous section. Indeed, the worst preferences need to be the ones associated to the lowest ability to trade-off between the dimension in which one agent is deprived and the other dimensions. Here, if transfers are evaluated in the space $d_i = 1$, it must be the case that the worst preferences are the ones that suffer most from having $d_i = 0$ compared to $d_i = 1$.

Theorem 3, above, tells us that individual situations have to be evaluated with the fraction of the poverty line vector to which an agent is indifferent. The theorem, on the other hand, does not tell us anything about how to choose the poverty line. Once we generalize Theorem 3 by adding a binary consumption good (and, more generally, by adding discrete goods), we obtain that individual situations have to be evaluated with the fraction of the poverty line vector, combined with a reference consumption of the binary good, to which an agent is indifferent. As a consequence, we still have to choose the poverty line, but, in addition to it, we now have to choose one reference value for the binary variable (or, more generally, one reference value of the discrete variables) so as to choose the space in which individual situations are evaluated.

6 Multidimensional poverty in Russia

In this section we illustrate the implementation of the poverty index using data from the Russian Longitudinal Monitoring Survey (RLMS-HSE) between 1995 and 2005. This period was particularly turbulent due to the fast Russian transition towards a market economy and the severe financial crisis of August 1999. These changes had far-reaching effects on the monetary and non-monetary outcomes of the Russian citizens. To capture some of these

effects, we include in our poverty analysis four "goods": a measure of equiv-alized household expenditures, health, housing quality and unemployment.[7] The first three dimensions are (approximately) continuous in nature, whereas the unemployment status is binary. This choice enables us to illustrate the theoretical framework of the previous sections. The poverty bundle z is set at 60% of the bundle that consists of the pooled median value in expenditures, health and housing and the reference value for the binary unemployment variable is not being unemployed. We do not think that these are the only possible choices, but we believe that this selection of the list of dimensions, the poverty bundle and the reference value are reasonable for this illustration.

To compute preference-sensitive multidimensional poverty, additional in-formation is needed on the ordinal preferences of the respondents over the four dimensions. Various approaches can be followed to estimate them. First, one can simply ask the respondent's opinions on the most appropriate trade-offs between the dimensions of poverty. Such a stated-preference procedure may be cognitively demanding for the respondents, however. Second, pref-erences can be estimated from observed behavior. Yet, revealed preference methods are only applicable to dimensions over which individuals make ac-tual choices and furthermore they may incorporate all sorts of decision errors. Third, preferences of different socio-demographic groups can be estimated based on self-reported life satisfaction information (see Fleurbaey et al. [16], for a similar approach). Given the data availability in the RLMS-HSE, the latter approach seems most attractive, although the assumption of preference homogeneity within socio-demographic groups is obviously a strong one.

Life satisfaction is measured in the RLMS-HSE by the following question: *"To what extent are you satisfied with your life in general at the present*

---

[7]We use the square root of household size as equivalence scale for the real household expenditures (the reference year is 1992). The measure of individual health is a composite index that consists of various objective health indicators such as indicators of diabetes, heart attack, aneamia, hospitalization and recent operations. The weights of the indicators in the health index are derived from the coefficients of an ordered logit regression using the self-assessed health as explained variable (for a similar approach, see van Doorslaer and Jones [35]). The measure of housing quality is the predicted value of a logarithmic hedonic housing regression based on self-reported housing values and a series of housing characteristics after controlling for regional price differences, time trends and the size of the household (by using again the square root equivalence scale). Estimation results are available upon request. The measure of unemployment is a binary indicator that takes 1 if the respondent is unemployed. The unemployment rate in the considered period fluctuates around 8%.

time?", with answers on a five point-scale ranging from "not at all satisfied" to "fully satisfied". The self-reported life satisfaction of individual i in period t is denoted (S_{it}). We start from a standard happiness regression with life satisfaction as the explained variable and a series of usual explanatory variables, including the vector of (transformed) individual outcomes for the four dimensions of poverty (X_{it}), a time trend (γ_t) and some observable socio-demographic characteristics (Z_{it}) such as education, social status, marital status, average expenditures and employment level in a small geographical reference group, and the presence of wage arrears, which used to be a common phenomenon during the late nineties in Russia. As unobservable personality traits are likely to influence self-reported life-satisfaction, we control for these time-invariant factors by including individual fixed effects (α_i) in the regression. Such a model would lead to identical preferences for all respondents. We allow for preference heterogeneity by including interaction effects between the outcome vector and a vector of five dummies (D_{it}), which capture the socio-demographic background of the individuals, i.e., whether they are young (below the age of 33), male, in a rural area, finished higher education and have a minority status. This leads to the following model:

$$S_{it}^* = \alpha_i + \gamma_t + (\beta + \Lambda D_{it})' X_{it} + \delta' Z_{it} + v_{it},$$

where S_{it}^* is a latent satisfaction variable, β and δ are vectors of direct effects and Λ a matrix with interaction effects to be estimated. The idiosyncratic error term v_{it} is assumed to follow a logistic distribution function. We observe the reported life satisfaction $S_{it} = k$ for k in $\{1, 2, \ldots, 5\}$ if the latent life satisfaction (S_{it}^*) lies within an interval between η_{k-1} and η_k :

$$S_{it} = k \text{ if } \eta_{k-1} < S_{it}^* \leq \eta_k.$$

The thresholds η_k are allowed to depend on the individual fixed effects, the observable socio-demographic characteristics and the time trend (for more details, see Jones and Schurer [22]). Finally, to allow for non-perfect substitutability between the three continuous dimensions of poverty, the outcomes of the individuals in these dimensions are transformed by a so-called Box-Cox transformation. For each dimension j in {expenditures, health, housing}, we have that:

$$X_{it}^j = \begin{cases} \left[\left(Y_{it}^j\right)^{\varepsilon_j} - 1\right]/\varepsilon_j & \text{when } \varepsilon_j \neq 0 \\ \log\left(Y_{it}^j\right) & \text{when } \varepsilon_j = 0, \end{cases}$$

24

where Y_{it}^j is the observed outcome of individual i in period t in dimension j. For the binary unemployment variable, $Y_{it} = X_{it}$. The three dimension-specific transformation parameters ε_j are chosen over a fine grid to maximize the overall fit of the model (Box and Cox [9]).

To compute the poverty index proposed in this paper, ordinal information on the indifference maps for the different socio-demographic groups is needed. This information is contained in the coefficients ε^j, β and Λ. The relevant estimation results are given in Table 1.[8] To reach a parsimonious and tractable model, the least significant interaction effects in Λ have been dropped consequently until all the remaining interaction effects are significant at the 10% level. All reported standard errors are corrected for clustering at the household level. The likelihood maximizing Box-Cox transformation parameters are respectively $-0,055; 0,485$ and $-0,356$ for expenditures, health, and housing quality. The approximately logarithmic transformation of expenditures is a common finding in the happiness literature, suggesting important decreasing marginal returns of expenditures (see, amongst others, Layard et al. [23]). The marginal returns for health and housing are found to be respectively less and more decreasing.

The (McFadden) pseudo R-squared of our estimation is around 0,073, which is comparable to other studies using panel data (see, for instance Graham et al. [20]). Yet, its magnitude highlights that only a small part of the variation in life satisfaction can actually be explained.[9] One may wonder at this point how problematic this finding is for our approach. Our starting point has been that people have different opinions on how to trade-off dimensions of well-being. If the estimated preferences were only meant to capture such heterogeneity in opinions or in behavior, then the relatively small share of the explained variation in life satisfaction should be a source of dissatisfaction, as we should aim at approaching the actual preferences of people as close as possible. One may argue, however, that actual preferences are too

---

[8] We follow the estimation method suggested by Jones and Schurer [22] that approximates the approach proposed by Ferrer-i-Carbonell and Frijters [13] and applied by Frijters et al. [19] to the RLMS-HSE.

[9] However, when considering for each respondent all pairwise comparisons between the ranking of available waves for the reported life satisfaction and the computed λ-values (as defined in Theorem 3), we find inconsistencies in only 4,8% of the comparisons, i.e., when $S_{it} > S_{it'}$, but $\lambda_{it} < \lambda_{it'}$. This suggests that, in spite of assuming identical preferences within socio-demographic subgroups, the λ-values based on these estimations are broadly consistent with the ordinal information in the reported life satisfaction at the individual level.

Table 1: Happiness equation

	life satisfaction	
expenditures transformed ($\varepsilon = -0,055$)	0,440***	(0,0351)
health transformed ($\varepsilon = 0,485$)	0,557***	(0,124)
house transformed ($\varepsilon = -0,356$)	0,259*	(0,128)
unemployed	-0,172	(0,123)
young × health	-0,315*	(0,159)
young × unemployed	0,190*	(0,0881)
male × health	0,401*	(0,179)
male × unemployed	-0,361***	(0.0882)
rural × house	0,397*	(0,169)
higher educated × expenditures	0,0441**	(0,0146)
higher educated × unemployed	-0,201+	(0,107)
higher educated × house	-0,240*	(0,109)
minority × health	0,655*	(0,263)
minority × expenditures	-0,382***	(0,0849)
N	53873	
pseudo R^2	0,073	

Clustered standard errors in parentheses. Coefficients are obtained after controling for education level, social status, marital status, reference group expenditures, reference group employment level, the presence of wage arrears and year dummies.

+ $p < 0,10$, * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$

idiosyncratic to be normatively compelling as people may make mistakes, for instance. Consequently, the actual preferences should be laundered before they are used in a normative judgement. This is precisely what the estimation carries out. We replace the actual individual preferences with some average preferences of the group to which the individual belongs, so that we end up only taking account of facts like the relative concern of elderly people for their health condition, the relative low worry of higher educated people on housing conditions, and so on. If the sample size would have allowed us to increase the number of these groups and to take account of more relevant characteristics, that would certainly have been desirable, but arguably not to the point that we should take the actual preferences of each agent individually.

Figure 6 illustrates the indifference maps in the expenditures-health space for two groups: the young, higher educated women (their outcomes are depicted with a diamond) and the old, lower educated men (the triangles in the figure). In line with expectations, we see that the elder individuals are on average in worse health. The illustrated indifference maps show that preferences of the latter group (depicted by the solid curves) are generally steeper, meaning that their willingness to pay for an increase in health is also higher.

The approach discussed in the previous sections identifies individuals as poor whenever they consider themselves worse off than with the poverty bundle z (depicted by the black dot on Figure 6). Consider, for instance, the young, higher educated women (the diamonds) situated in the south-east of the poverty bundle between both indifference curves through z. These individuals consider themselves to be poor, i.e., worse off than the poverty bundle. However, if these individuals had the steeper (solid) indifference map, then they would not have considered themselves to be poor. This example illustrates how taking preferences into account also matters empirically for the identification of the poor.

A popular way to measure the incidence of poverty is by looking at the share of the total population who is identified as poor, the so-called headcount rate. The first column of Table 2 presents the evolution of the headcount rates according to the preference-sensitive method proposed in this paper. Poverty almost doubles in the first three waves until it reaches its peak in 1998, and then it decreases steadily to its initial level. In the second column, these findings are compared to another multidimensional poverty measure using the counting approach (see Atkinson [4] and Alkire and Foster [1, 2]). In the counting approach, the identification of the poor is based on the number

Green diamonds and dashed indifference curves : young, higher educated women
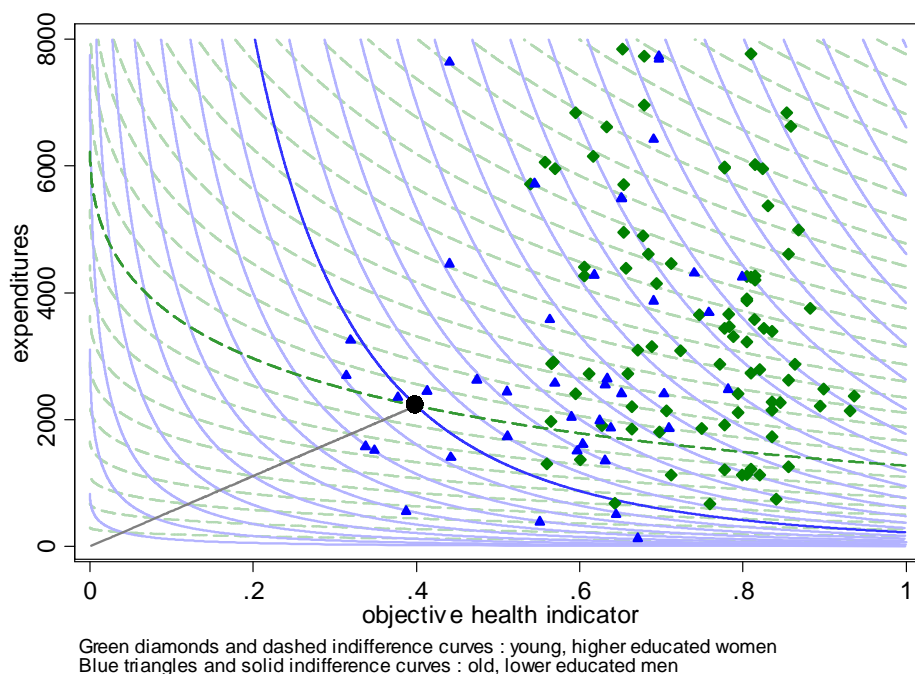Blue triangles and solid indifference curves : old, lower educated men

Figure 6: Indifference map of the poor in 2000

of dimensions for which the individual falls below the threshold. In this illustration we keep the dimensionwise thresholds at 60% of the median value and consider individuals as poor when they are below the threshold for at least two out of the four dimensions. Also headcount rates obtained with the counting method have peaked around the financial crisis of the late 90s, and have thereafter returned to their pre-crisis level. Finally, we compare the results to the traditional one-dimensional headcount poverty measures that use equivalized household expenditures as indicator and a threshold set at 60% of the (pooled) median expenditures. Again a similar pattern can be discerned over time, with the expenditures returning much faster to their pre-crisis level than the multidimensional measures. The slower recovery of the multidimensional measures may be due to transition-induced disruptions of the health care system, for instance.

As suggested by Figure 6 already, the relatively similar trend of the three different headcount measures may mask important differences in the composition of the poor subpopulation. Table 3 illustrates these differences. Each

Table 2: Multidimensional headcount rates

year	preference sensitive	counting approach	expenditure poverty
1995	12,6%	11,6%	23,4%
1996	15,7%	14,1%	28,4%
1998	21,8%	17,8%	39,7%
2000	18,6%	16,1%	32,4%
2001	15,4%	14,1%	26,3%
2002	15,0%	13,2%	23,2%
2003	13,6%	12,2%	20,4%
2004	13,8%	12,3%	19,9%
2005	12,0%	11,5%	17,4%

column gives the average characteristics of the poor in 2000 according to one of the identification methods. The poor according to the preference-sensitive method proposed in this paper are relatively old, predominantly female and in bad health. These are the individuals who give a large relative weight to the dimension on which they score badly. The counting method, on the other hand, identifies more people as poor who are unemployed, living in a relatively low quality house and who are lower educated. The poor according to the standard expenditure measure have low expenditures, are more often male, and are younger in general. As different people are indeed identified as poor according to the three methods, it is clear that the question whether and how to take their preferences into account may have strong implications for the design of poverty alleviation programs, for instance.

In Figure 7, we consider the overlap between the 16,1% worst off according to each of the three methods in 2000. For this analysis, we focus on equally sized groups of the poorest individuals determined by a cut-off set at the minimal poverty rate in 2000 across the three methods, rather than on the poor as identified by the three methods as the latter groups have rather different sizes (as can be seen in the first row of Table 2). Slightly more than one third of the individuals (6,6%) are identified as belonging to the bottom 16,1% according to all methods. All other individuals are considered worst off by at least one method, but not by all methods. One finds remarkably little overlap between both multidimensional methods given that the same poverty

Table 3: Portrait of the poor in 2000

	preference sensitive	counting approach	expenditure poverty
number of poor (in %)	18,6	16,1	32,4
expenditure (in rubbles)	1655	1541	1403
health (on 0-1 scale)	0,52	0,56	0,62
house (in 100.000 rubbles)	2,40	2,04	2,52
unemployed (in %)	24,5	26,8	11,2
life satisfaction	1,97	1,96	2,09
male (in %)	32,1	38,6	40,3
young (in %)	25,7	29,2	33,1
higher educated (in %)	57,3	50,8	59,6
rural (in %)	35,2	46,1	34,0
minority (in %)	9,6	17,1	13,6

bundle z is used. This finding stresses once more the empirical implications of taking the preferences of the poor into account in the measurement of poverty.

Finally, we illustrate how one easily tests the partial poverty ordering that one obtains if one seeks unanimity among all the indexes belonging to the class proposed by Theorem 3. This ordering requires a unanimous agreement among all continuous, decreasing and convex functions ϕ. We use the so-called "Three I's of Poverty" (TIP) curve as a tool to check whether all such ϕ functions will order the multidimensional distributions in the same way. A TIP-curve plots the cumulative poverty gaps for the population ranked from poor to rich individuals. Jenkins and Lambert [21] have shown that whenever the TIP-curve of one distribution is everywhere above the TIP-curve of another distribution, then there is unanimous agreement in the class of all considered poverty measures that poverty is higher in the first distribution (see also Zheng [37]). Their result immediately extends to our class of indexes. In Figure 8 we have depicted for each wave the TIP-curve of the computed λ-values. As these curves become flat above the poverty line, we only depict their leftmost part. Clearly, the TIP-curve of 1998 is everywhere above the other curves. In other words, 1998 is unambiguously the year with most multidimensional poverty according to all poverty measures that
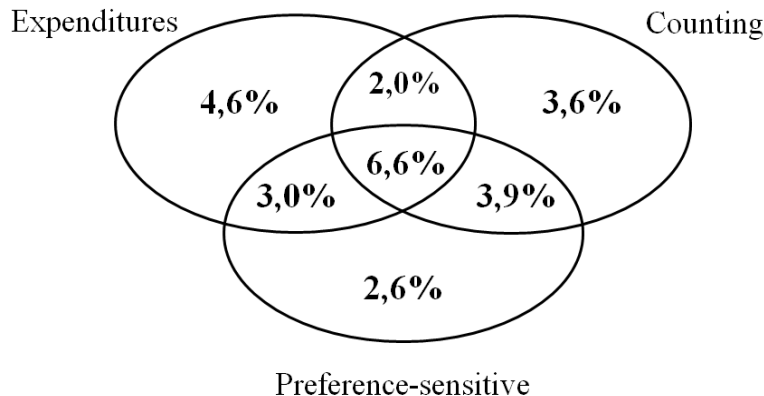
Figure 7: Overlap between the bottom 16,1% according to different metrics

satisfy the axioms listed in Theorem 3. On the other hand, we see that the curve of 2005 is everywhere below the curve of 1995, indicating that poverty decreased over the considered period —though we do not test the statistical robustness of this statement.

7   Conclusion

Measuring multidimensional poverty requires aggregating across dimensions and across agents. In this paper, we have studied the consequences of aggregating across dimensions at the level of each agent by taking the agent's preferences as the aggregation device. This approach forced us to find new ways of aggregating across agents, as individual levels of preference satisfaction cannot readily be compared. By introducing ways to build interpersonal poverty comparisons, we have been able to provide and characterize a family of poverty indices. These poverty indices aggregate individual measures of poverty that are a convex transformation of the fraction of the poverty line vector to which the individual herself is indifferent.

   We have illustrated how the approach proposed in this paper can be implemented using existing Russian survey data from RLMS-HSE and we found some remarkable differences with standard (multidimensional) poverty measures. By taking preferences of the poor into account, different people are indeed identified as poor. The data that are needed to apply our approach are clearly more demanding than what is required to apply the other indices
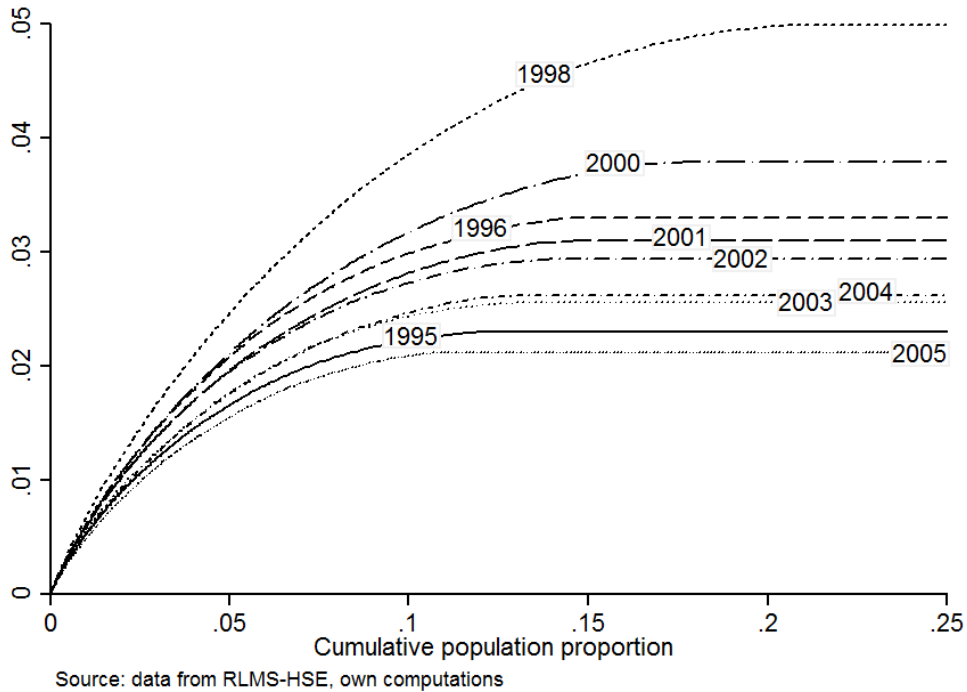
Source: data from RLMS-HSE, own computations

Figure 8: TIP-curves for the different waves

proposed in the literature. For instance, the counting approach is remarkably parsimonious in the required data (see Alkire and Foster [1]), whereas our approach requires a tailored data set that allows identification of the preferences in a wide set of dimensions. We believe that the data requirement of our approach is the price to pay to develop an attractive way to measure multidimensional poverty without relying on arbitrary weights or arbitrary assumptions on the nature of the goods.

Exploring the policy implications of our approach to poverty measurement is beyond the scope of this paper. We will just highlight three points. First, the different population that is identified as poor under the new method should induce redirecting specific policies, e.g., those that are locally targeted, or that are targeted to specific age groups. In our empirical example, a greater concern for the unemployed, the elderly, and women would be warranted. Second, the ray measure of individual poverty proposed here is such that information about preferences is important primarily for the people who are not poor in all dimensions. In contrast, the situations of those who are

32

below the poverty threshold and suffer similar deprivations in all dimensions can be quite accurately assessed without inquiring about their preferences. Although this seems a natural result, it is easy to conceive other measures that rely on preferences and do not have this property. Third, the evaluation of policies which have conflicting effects on different dimensions (e.g., improve health at some cost on income, or conversely) can be assessed in a more appealing way when population preferences are incorporated.

References

[1] Alkire S. and J. Foster 2011, "Counting and multidimensional poverty measurement", *Journal of Public Economics* 95: 476-487.

[2] Alkire S. and J. Foster 2011b, "Understandings and misunderstandings of multidimensional poverty measurement", *Journal of Economic Inequality* 9: 289-314.

[3] Anand S. and A. K. Sen 1997, "Concepts of human development and poverty: a multidimensional perspective", Human Development Papers, UNDP, New York.

[4] Atkinson A.B. 2003, "Multidimensional deprivation: contrasting social welfare and counting approaches", *Journal of Economic Inequality* 1: 51-65.

[5] Atkinson A.B. and F. Bourguignon 1982, "The comparison of multidimensioned distributions of economic status", *Review of Economic Studies* 49: 183-201.

[6] Bosmans K., E. Ooghe and L. Lauwers 2013, "Prioritarian poverty comparisons with cardinal and ordinal attributes", CES discussion papers 13.10 Katholieke Universiteit Leuven.

[7] Bossert W., S. R. Chakravarty and C. D'Ambrosio 2013, "Multidimensional Poverty and Material Deprivation with Discrete Data", *Review of Income and Wealth* 59: 29-43.

[8] Bourguignon F. and S. R. Chakravarty 2003, "The measurement of multi-dimensional poverty", *Journal of Economic Inequality* 1: 25-49.

[9] Box G. E. P. and D. R. Cox 1964, "An analysis of transformations" *Journal of the Royal Statistical Society. Series B* 26: 211-252.

[10] Deaton A. 1979, "The distance function in consumer behaviour with applications to index numbers and optimal taxation," *Review of Economic Studies* 46: 391–405.

[11] Duclos J.-Y., D. E. Sahn and S.D. Younger 2006, "Robust multidimensional poverty comparisons", *Economic Journal* 116: 943-968.

[12] Ferreira F. H. G. and M. A. Lugo 2013, "Multidimensional Poverty Analysis: Looking for a Middle Ground", *The World Bank Research Observer* 28: 220-235.

[13] Ferrer-i-Carbonell A. and P. Frijters 2004, "How important is methodology for the estimates of the determinants of happiness?", *Economic Journal* 114: 641–659.

[14] Fleurbaey M., F. Maniquet 2011, *A Theory of Fairness and Social Welfare*, Cambridge: Cambridge University Press.

[15] Fleurbaey M. and A. Trannoy 2003, "The impossibility of a Paretian egalitarian", *Social Choice and Welfare* 21: 243-263.

[16] Fleurbaey M., E. Schokkaert and K. Decancq 2009, "What good is happiness?", *CORE discussion paper* 2009/17.

[17] Foster J. and A. F. Shorrocks 1991, "Subgroup consistent poverty indices", *Econometrica* 59: 687-709.

[18] Foster J., J. Greer and E. Thorbecke 1984, "A class of decomposable poverty measures", *Econometrica* 52: 761-766.

[19] Frijters P., I. Geishecker, J.P. Haisken-DeNew and M. Shields 2006, "Can the large swings in Russian life satisfaction be explained by ups and downs in real incomes?", *Scandinavian Journal of Economics* 108: 433–458.

[20] Graham, C., A. Eggers and S. Sukhtankar 2004, "Does happiness pay? An exploration based on panel data from Russia" *Journal of Economic Behavior and Organization* 55: 319–342.

[21] Jenkins, S. P. and P. J. Lambert 1997, "Three I's of poverty curves, with an analysis of UK poverty trends" *Oxford Economic Papers* 49: 317-327.

[22] Jones A. M. and S. Schurer 2011, "How does heterogeneity shape the socioeconomic gradient in health satisfaction?" *Journal of Applied Econometrics* 26: 549-579.

[23] Layard, R., G. Mayraz and S. Nickell 2008, "The marginal utility of income" *Journal of Public Economics* 92: 1846 - 1857.

[24] Kolm S. C. 1977, "Multidimensional egalitarianisms", *Quarterly Journal of Economics* 91: 1-13.

[25] Maasoumi E. 1986, "The measurement and decomposition of multidimensional inequality", *Econometrica* 54: 991-997.

[26] Narayan, D., R. Patel, K. Schafft, A. Rademacher and S. Koch-Schulte 2000, "*Voices of the Poor: Can Anyone Hear Us?*" New York, N.Y.: Published for the World Bank, Oxford University Press.

[27] Pazner E., D. Schmeidler 1978, "Egalitarian equivalent allocations: A new concept of economic equity," *Quarterly Journal of Economics* 92: 671–687.

[28] Ravallion M. 1996, "Issues in measuring and modelling poverty", *Economic Journal* 106: 1328-1343.

[29] Ravallion M. 2011, "On multidimensional indices of poverty", *Journal of Economic Inequality* 9: 235-248.

[30] Samuelson P.A. 1977, "Reaffirming the existence of "reasonable" Bergson-Samuelson social welfare functions," *Economica* 44: 81–88.

[31] Sen A. K. 1976, "Poverty: an ordinal approach to measurement", *Econometrica* 44: 219-231.

[32] Sen A. K. 1985, *Commodities and Capabilities*: Amsterdam, North-Holland.

[33] Shorrocks A. F. 1983, "Ranking Income Distributions", *Economica* 50: 3-17.

[34] Tsui K. Y. 2002, "Multidimensional poverty indices", *Social Choice and Welfare* 19: 69-93.

[35] van Doorslaer E. and A. M. Jones 2003, "Inequalities in self-reported health: validation of a new approach to measurement", *Journal of Health Economics* 22: 61-87

[36] Weymark J. A. 2006, "The normative approach to the measurement of multidimensional inequality", in F. Farina and E. Savaglio (eds.), "Inequality and Economic Integration", Routledge, London. 303-328.

[37] Zheng B. 1997, "Aggregate poverty measures", *Journal of Economic Surveys* 11: 123-162.

Appendix: Proofs

Proof. **(of Theorem 1)** We focus on the "only if" part. Let P be a poverty index satisfying the axioms. Let $N \subset \mathbb{N}_{++}$. By *Subgroup Consistency* applied by decomposing N into $\{i\}$ and $N \setminus \{i\}$, there exist functions $g : X^{N\setminus\{i\}} \times \mathcal{R}^{N\setminus\{i\}} \to \mathbb{R}$ and $f : \mathbb{R}^2 \to \mathbb{R}$ such that for all $(x_N, R_N) \in \mathcal{S}$,

$$P(x_N, R_N) = f(P(x_i, R_i), g(x_{N\setminus\{i\}}, R_{N\setminus\{i\}})).$$

By applying the same argument to decompose g and the following functions, we reach the conclusion that there exists a function $f : \mathbb{R}^{|N|} \to \mathbb{R}$ such that for all $(x_N, R_N) \in \mathcal{S}$,

$$P(x_N, R_N) = f((P(x_i, R_i))_{i \in N}).$$

Fix R_N for the moment. The line segment from 0 to $z(R_i)$ is a compact set. By *Focus*, the range of $P(x_i, R_i)$ for fixed R_i is the image of this line segment. By *Continuity*, the ranges of $P(x_i, R_i)$ is a compact intervals, of strictly positive length by *Pareto among the Poor*.

We claim that f is continuous on this domain. Indeed, take a sequence $p_N^k \to p_N^*$ in this domain and suppose $f\left(p_N^k\right)$ does not tend to $f(p_N^*)$. It then has a subsequence p_N^t such that $f\left(p_N^t\right)$ stays away from $f(p_N^*)$. There is a corresponding sequence x_N^t on the line segments from 0 to $(z(R_i))_{i \in N}$ such that $p_N^t = (P(x_i^t, R_i))_{i \in N}$ for all t. As the sequence is in a compact set, it has a converging subsequence $x_N^s \to x_N^*$. By *Continuity*, $P(x_N^s, R_N) \to P(x_N^*, R_N)$. Also by *Continuity*, $p_N^s \to p_N^*$ implies that $p_N^* = (P(x_i^*, R_i))_{i \in N}$. Therefore $f\left(p_N^*\right) = P(x_N^*, R_N)$ is the limit of $f\left(p_N^s\right)$. This contradicts the assumption that $f\left(p_N^s\right)$, a subsequence of $f\left(p_N^t\right)$, stays away from $f(p_N^*)$.

Now let us drop the restriction on R_N. By *Focus*, one can normalize $P(x_i, R_i) = 0$ whenever $x_i R_i z(R_i)$. Therefore the range of $P(x_i, R_i)$ is the union of the intervals for every possible R_i, all these intervals having 0 as their lower bound, so that their union is an interval. The f function being continuous on any of these (Cartesian products of) intervals, it is continuous on (the Cartesian product of) their union.

So, f is a real-valued function defined over the Cartesian product of connected and separable spaces (intervals). By *Pareto among the Poor*, each dimension of the domain of f is essential (f cannot be constant in any $P(x_i, R_i)$). Therefore, by *Continuity* and *Subgroup Consistency*, we can apply Debreu's theorem on additive representation of separable preferences:

37

there exists a continuous and strictly increasing function $G_N : \mathbb{R} \to \mathbb{R}$ and for all $i \in N$ a continuous function $f_i : \mathbb{R} \to \mathbb{R}$ such that

$$P(x_N, R_N) = G_N \left[\sum_{i=1}^{|N|} f_i(P(x_i, R_i)) \right].$$

By defining functions $\phi_i^N : X \times \mathcal{R} \to \mathbb{R}$ as follows: $\phi_i^N = f_i \circ P$, we get

$$P(x_N, R_N) = G_N \left[\sum_{i=1}^{|N|} \phi_i^N(x_i, R_i) \right].$$

By *Continuity*, ϕ_i^N are continuous (in their first argument). By *Pareto among the Poor*, $\phi_i^N(x_i, R_i) > \phi_i^N(x_i', R_i)$ whenever $z(R_i) \, R_i \, x_i' \, P_i \, x_i$. By *Focus*, we can normalize $\phi_i^N(x_i, R_i) = 0$ whenever $x_i \, R_i \, z(R_i)$.

In order to prove the second part of the Theorem, assume it is wrong. For all G and ϕ_i^N such that $P(x_N, R_N) = G \left[\sum_{i=1}^{|N|} \phi_i(x_i, R_i), N \right]$, there exist $(x_N, R_N), (x_{N'}', R_{N'}') \in \mathcal{S}$ such that

$$\frac{1}{|N|} \sum_{i=1}^{|N|} \phi_i(x_i, R_i) = \frac{1}{|N'|} \sum_{i=1}^{|N'|} \phi_i(x_i', R_i') \tag{2}$$

and

$$P(x_N, R_N) \neq P(x_{N'}', R_{N'}'). \tag{3}$$

Let $n = |N|$ and $n' = |N'|$. By *Replication Invariance*,

$$P(x_N, R_N) = P(x_N^{n'}, R_N^{n'}) = G \left[n' \sum_{i=1}^{|N|} \phi_i(x_i, R_i), nn' \right],$$

and

$$P(x_{N'}', R_{N'}') = P(x_{N'}'^{\,n}, R_{N'}'^{\,n}) = G \left[n \sum_{i=1}^{|N'|} \phi_i(x_i', R_i'), nn' \right].$$

Eq. 2 implies that

$$n' \sum_{i=1}^{|N|} \phi_i(x_i, R_i) = n \sum_{i=1}^{|N'|} \phi_i(x_i', R_i'),$$

38

in contradiction to Eq. 3. ∎

Proof. (of Theorem 2) We first prove that under *Nested Contours* and *Continuity, Nested Unions* is equivalent to the following property P*:

For all $N \subset \mathbb{N}_{++}$, $i \in N$, $(x_{N \setminus \{i\}}, R_{N \setminus \{i\}}) \in \mathcal{S}^{N \setminus \{i\}}$, $\overline{\mathcal{R}}^i \subseteq \mathcal{R}$, $p \in \mathbb{R}$, if for all $R_i \in \overline{\mathcal{R}}^i$ there exists $x_i(R_i) \in X$ such that

$$P\left((x_{N \setminus \{i\}}, x_i(R_i)), (R_{N \setminus \{i\}}, R_i)\right) = p,$$

then for all $R_i' \in \mathcal{R}$, $x_i' \in X$, if

$$L(x_i', R_i') \subseteq \cup_{R_i \in \overline{\mathcal{R}}^i} L(x_i(R_i), R_i),$$

then

$$P\left((x_{N \setminus \{i\}}, x_i'), (R_{N \setminus \{i\}}, R_i')\right) \geq p.$$

Implication: *Nested Unions* implies P* for the case $\left|\overline{\mathcal{R}}^i\right| = 2$. It is a straightforward induction argument to extend the property to any finite number. Now, let $L(x_i', R_i') \subseteq \cup_{R_i \in \overline{\mathcal{R}}^i} L(x_i(R_i), R_i)$ and consider a sequence $y^k \to x_i'$ such that $x_i' P_i' y^k$ for all k. Necessarily for all k, given our assumption that X is compact, there is a finite subset $\overline{\overline{\mathcal{R}}}^i \subset \overline{\mathcal{R}}^i$ such that $L(y^k, R_i') \subseteq \cup_{R_i \in \overline{\overline{\mathcal{R}}}^i} L(x_i(R_i), R_i)$. By the property for finite numbers, this implies

$$P\left((x_{N \setminus \{i\}}, y^k), (R_{N \setminus \{i\}}, R_i')\right) \geq p.$$

By *Continuity*,

$$P\left((x_{N \setminus \{i\}}, x_i'), (R_{N \setminus \{i\}}, R_i')\right) \geq p.$$

This proves the implication. The converse is omitted.

Part I: P* implies 2. By Theorem 1, we can restrict our attention to one agent situations. Let $p \in \mathbb{R}$. For all $R \in \mathcal{R}$, let $x_p(R) \in X$ be such that

$$P(x_p(R), R) = p.$$

Let $U^p \subseteq X$ be defined by

$$U^p = \bigcap_{R \in \mathcal{R}} U(x_p(R), R).$$

Because U^p is the (possibly infinite) intersection of compact and convex sets, it is compact and convex. Assume it is non-empty. Let $I^p \subseteq X$ be the

39

lower frontier of U^p. By the unrestricted domain assumption, there exists $R^p \in \mathcal{R}, x^p \in I^p$ such that $I(x^p, R^p) = I^p$. We claim that $P(x^p, R^p) = p$. By construction of U^p, $P(x^p, R^p) > p$ is impossible, because the intersection is taken over all $R \in \mathcal{R}$, including R^p. Let $\overline{\mathcal{R}} \in \mathcal{R}$ be defined by

$$\overline{\mathcal{R}} = \{R \in \mathcal{R} | L(x^p(R), R) \cap I^p \neq \emptyset\}.$$

Observe that

$$L(x^p, R^p) = \bigcup_{R \in \overline{\mathcal{R}}} L(x_p(R), R).$$

By P*, $P(x^p, R^p) \geq p$, which proves the claim.

Now, let $\Pi \subset \mathbb{R}$ be the set of $p \in \mathbb{R}$ such that U^p is non-empty. By continuity of P, Π is a closed and convex interval of $\mathbb{R}$. By construction, for all $p, p' \in \Pi$

$$p < p'^p \subset int U^{p'}.$$

Therefore, by the unrestricted domain assumption, there exists $R^\Pi \in \mathcal{R}$ such that for all $p \in \Pi, x^p \in I^p$, $I(x^p, R^\Pi) = I^p$. We now claim that P satisfies *Worst Preferences* for $R^w = R^\Pi$.

Let $x \in X, R \in \mathcal{R}$. We need to prove that $P(x, R) \leq P(x, R^\Pi)$. By construction, $P(x, R) = p$ for some $p \in \Pi$. Therefore, $U(x, R) \supseteq U^p$, so that $x \in L(x_p(R^\Pi), R^\Pi)$, which, by *Pareto Among the Poor*, implies $P(x, R^\Pi) \geq p$, the desired outcome.

Part II: 2 implies P*. Let $p \in \mathbb{R}$ and $\overline{\mathcal{R}} \in \mathcal{R}$ be such that for all $R \in \overline{\mathcal{R}}$ there exists $x(R) \in X$ such that $P(x, R) = p$. Let $x' \in X$ and $R' \in \mathcal{R}$ be such that

$$L(x', R') \subseteq \cap_{R \in \mathcal{R}} L(x(R), R) \equiv L^p.$$

We need to prove that $P(x', R') \geq p$.

Let $x_p^w \in X$ be such that $x_w^p \in L^p$ and for all $x \in L^p : x_w^p R^w x$, that is, x_w^p is the best bundle for R^w in the lower contour of all R in $\overline{\mathcal{R}}$ through their $x(R)$. First, we claim that such a x_p^w exists. By the assumption that X is compact, L^p is compact. the claims follows from the assumption that all preferences are continuous. Second, we claim that $P(x_p^w, R^w) = p$. By construction, there is $R \in \overline{\mathcal{R}}$ such that $P(x_w^p, R) = p$. By *Worst Preferences*, $P(x_p^w, R^w) \geq p$. Assume $P(x_p^w, R^w) > p$. Let $x \in X$ be such that $x P^w x_p^w$ and $P(x, R^w) = p$. By construction of L^p and x_p^w, $U(x, R^w) \cap L^p = \emptyset$. By *Nested Contours*, $P(x, R^w) < P(x_p^w, R)$, a contradiction, which proves the claim.

Now, let $x'^w \in X$ be such that $x'^w \in L(x', R')$ and $x'^w R^w x$ for all $x \in L(x', R')$. By the same argument as to prove the claim above, $P(x', R') = P(x'^w, R^w)$. Finally, $L(x', R'^p$ implies that $P(x'^w, R^w) \geq P(x_p^w, R^w) = p$, which yields $P(x', R') \geq p$, the desired contradiction.

Part III: 2 implies 3. Let us construct function $\phi(\cdot, \cdot)$. By Theorem 1, we know that $\phi(z(R^w), R^w) = 0$. Without loss of generality, we can assume that $\phi(0^\ell, R^w) = 1$. By *Continuity*, $\phi(\cdot, \cdot)$ is continuous in its first argument. Let $R_i \in \mathcal{R}$, $x_i, x_i' \in X$ be such that $L(x_i, R_i) \cap U(x_i', R^w) \neq \emptyset$ and $intL(x_i, R_i) \cap intU(x_i', R^w) = \emptyset$ (that is, $I(x_i, R_i)$ is tangent to $I(x_i', R^w)$ from below). We need to prove that $\phi(x_i, R_i) = \phi(x_i', R^w)$.

Assume $\phi(x_i, R_i) > \phi(x_i', R^w)$. Let $y \in L(x_i, R_i) \cap U(x_i', R^w)$. By continuity of ϕ, in a neighborhood of y there exists $x_i'' \in X$ such that $x_i'' P^w x_i'$ and $\phi(x_i'', R_i) > \phi(x_i', R^w)$. By *Pareto among the Poor*, $\phi(x_i', R^w) > \phi(x_i'', R^w)$ and therefore $\phi(x_i'', R_i) > \phi(x_i'', R^w)$, contradicting *Worst Preferences*.

Assume $\phi(x_i, R_i) < \phi(x_i', R^w)$. By continuity of ϕ, there exists $x_i'' \in X$ such that $x_i P_i x_i''$ and yet

$$\phi(x_i'', R_i) < \phi(x_i', R^w). \tag{4}$$

Let $R_i' \in \mathcal{R}$ satisfy: $I(x_i'', R_i) \in \mathcal{I}(R_i')$ and $I(x_i', R^w) \in \mathcal{I}(R^w)$. By *Nested Contours*, $\phi(x_i'', R_i') = \phi(x_i'', R_i)$ and $\phi(x_i', R_i') = \phi(x_i', R^w)$, which, in view of Eq. (4), violates *Pareto among the Poor*, because $x_i' P_i' x_i''$.

Part IV: It is easy to see that 3 implies 2. ∎

Proof. (of Theorem 3) Let $T \in \mathcal{T}$ be the set of bundles for which $\mathcal{T}$-*Poverty Sensitivity* applies. Let $R^w \in \mathcal{R}$ be the preference relation for which *Worst Preferences* is satisfied. Let $z \in T$ be defined by: $z I^w z(R^w)$. Let $\lambda(\cdot, \cdot)$ be defined by $\lambda(x_i, R_i) = \lambda$ if and only if $x_i I_i \lambda z$.

Assume R^w is not the Leontief preference ordering defined in the proposition (the construction of the key ingredients of the proof is illustrated in Figure 9, where the convexity of T is used by focusing on the line segment from 0 to z). Then, there exist $R_k \in \mathcal{R}$, $x_k^* \in X$ such that

$$x_k^* \in L(x_k^*, R_k) \quad \cap \quad U(x_k^*, R^w), \tag{5}$$

$$intL(x_k^*, R_k) \quad \cap \quad intU(x_k^*, R^w) = \emptyset, \tag{6}$$

$$\lambda(x_k^*, R_k) \quad < \quad \lambda(x_k^*, R^w). \tag{7}$$

Moreover, there exists a neighborhood $U_{x_k^*}$ of x_k^* such that for all $x_k \in U_{x_k^*}$, Eqs. (5)–(7) apply. Therefore, we can assume $\lambda(x_k^*, R^w) < 1$. By Theorem 2:

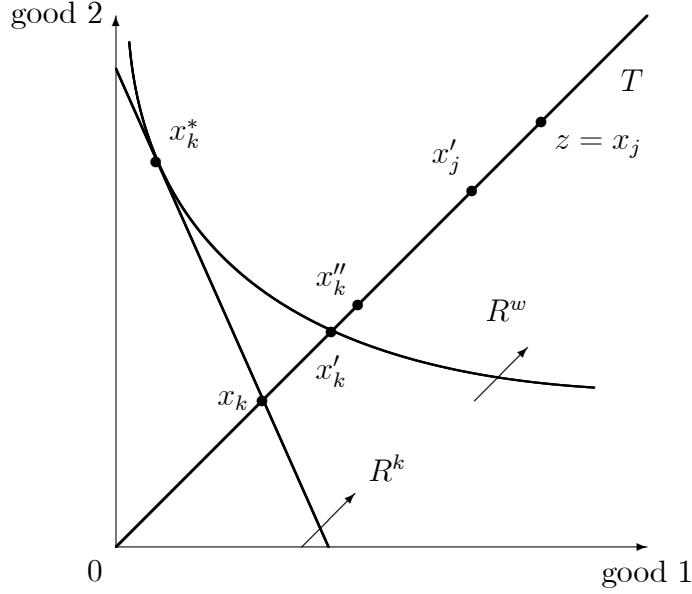$$\phi(x_k^*, R_k) = \phi(x_k^*, R^w). \tag{8}$$

41

Figure 9: Illustration of the proof of Theorem 3

Let $x_k, x'_k, x_j, x'_j \in T$ be defined by

$$
\begin{aligned}
x_k &\in I(x^*_k, R_k), \\
x'_k &\in I(x^*_k, R^w), \\
x_j &= z, \\
x'_j &= x_j - (x'_k - x_k).
\end{aligned}
$$

By construction, $\phi(x_j, R^w) < \phi(x'_j, R^w)$, and $\phi(x_k, R_k) = \phi(x'_k, R^w)$.

By *Continuity* of P, Theorem 2 implies that there exists $x''_k \in T$ such that $x''_k P^w x'_k$ and

$$
\phi(x_k, R_k) + \phi(x_j, R^w) < \phi(x''_k, R^w) + \phi(x'_j, R^w). \tag{9}
$$

Let $N = \{j, k\}$. Let $R'_j = R'_k \in \mathcal{R}$ be such that

$$
\begin{aligned}
I(x_k, R'_k) &= I(x_k, R_k), \\
I(x''_k, R'_k) &= I(x''_k, R^w), \\
I(x'_j, R'_j) &= I(x'_j, R^w), \\
I(x_j, R'_j) &= I(x_j, R^w).
\end{aligned}
$$

42

By *Pareto among the Poor*, $P(x_j, R'_j) < P(x_k, R'_k)$. Therefore, by $\mathcal{T}$-*Poverty Sensitivity*, given that $x_j + x_k = x'_j + x'_k$,

$$P((x'_j, x'_k), (R'_j, R'_k)) < P((x_j, x_k), (R'_j, R'_k)).$$

By *Pareto among the Poor*,

$$P((x'_j, x''_k), (R'_j, R'_k)) < P((x'_j, x'_k), (R'_j, R'_k)).$$

By transitivity,

$$P((x'_j, x''_k), (R'_j, R'_k)) < P((x_j, x_k), (R'_j, R'_k)). \tag{10}$$

By Theorem 2 and Eq. (8),

$$
\begin{aligned}
P((x'_j, x''_k), (R'_j, R'_k)) &= G(\frac{1}{2}(\phi(x'_j, R^w) + \phi(x''_k, R^w))), \\
P((x_j, x_k), (R'_j, R'_k)) &= G(\frac{1}{2}(\phi(x_j, R^w) + \phi(x_k, R^w))).
\end{aligned}
$$

Consequently, by Eq. (9),

$$P((x'_j, x''_k), (R'_j, R'_k)) > P((x_j, x_k), (R'_j, R'_k)),$$

in contradiction to Eq. (10). ■