Equality among unequals*

Mathieu Faure† and Nicolas Gravel‡

September 25, 2018

Abstract

This paper establishes an equivalence between four incomplete rankings of distributions of income among agents who are vertically differentiated with respect to some other non-income characteristic (health, household size, etc.). The first ranking is that associated with the possibility of going from one distribution to the other by a finite sequence of income transfers from richer and more highly ranked agents to poorer and less highly ranked ones. The second ranking is the unanimity among utilitarian planners over all comparisons of two distributions assuming that agents' marginal utility of income is decreasing with respect to both income and the source of vertical differentiation. The third ranking is the Bourguignon (1989) ordered poverty gap dominance criterion. The fourth ranking is a new dominance criterion based on cumulative lowest incomes.

Keywords: Equalization, transfers, heterogeneous agents, poverty gap, dominance, cumulative sums of income, utilitarianism.

JEL Codes: D30, D63, D69, I32.

1 Introduction

When can a distribution of an attribute among a group of homogeneous agents be considered more equal than another? An important achievement of the modern theory of inequality measurement is the demonstration made by Hardy, Littlewood, and Polya (1952) - and popularized among economists by Dasgupta, Sen, and Starrett (1973), Kolm (1969), Sen (1973) and Fields and Fei (1978) - that the following four answers to this question are equivalent.

---

*With the usual disclaiming qualification, we are indebted to Sebastian Bervoets, Yann Bramoullé, Alain Chateauneuf and, especially, Patrick Moyes for useful comments and discussions. Financial support from the French ANR (Grant Contract No. ANR-16-CE41-0005-01) is also gratefully acknowledged.

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS. Email address: mathieu.faure@univ-amu.fr.

‡Centre de Sciences Humaines (Delhi) and Aix-Marseille University (Aix-Marseille School of Economics). Email address: nicolas.gravel@csh-delhi.com

(1) A is more equal than B if it can be obtained from B by means of a finite sequence of bilateral Pigou-Dalton transfers.

(2) A is more equal than B if all utilitarian planners who assume that individuals convert the attribute into well-being by the same concave utility function so agree.

(3) A is more equal than B if poverty, as measured by the poverty gap, is lower in A than in B for every definition of the poverty line.

(4) A is more equal than B if the cumulative income of the k poorest agents is greater in A than in B whatever k is (that is, if the distribution of income in A Lorenz dominates that in B).

This equivalence is of fundamental importance for (in)equality measurement because it shows the congruence of four distinct approaches to the question. The first one focuses on the *elementary operation* that intuitively captures the very notion of inequality reduction. The second approach links inequality measurement to a set of explicit normative principles and seeks consensus among all principles in this set. Finally, the third and fourth approaches provide *empirically implementable tests* - poverty gap or Lorenz dominance - to determine whether or not one distribution dominates another.

Remarkable as it is, this equivalence only concerns distributions of a single attribute, often identified with income, between otherwise perfectly homogeneous agents. Yet income is clearly not the only ethically relevant source of differentiation between economic agents. If these agents are collectivities such as households or jurisdictions, they differ not only by their total income but also by the number of members among whom the income must be shared. If the agents are individuals, they may also differ by non-income characteristics such as age, health, education or effort. What does "being more equal" mean when applied to distributions of an attribute among differentiated agents ? In short, how can one define equality among unequals ? This is the basic question addressed in this paper.

Specifically, we establish an equivalence between four notions of inequality reduction among unequals that are each is analogous, in nature, to one of the above four notions of equality among equals. The elementary operation that we propose to capture inequality reduction among unequals is like a Pigou-Dalton transfer, but with the stipulation that the donor in the transfer must be *both* richer *and* more highly ranked than the receiver. Moreover, contrary to what is usually required in a Pigou-Dalton transfer, we do not restrict the transfer to being lower than half the income difference between the giver and the receiver. The quantity transferred can be as large as the full income difference. The normative principles that we examine are those generated by comparisons of distributions by a utilitarian planner who assumes that agents convert income into utility by the same function exhibiting a marginal utility of income that is decreasing with respect to *both* income *and* the source of vertical differentiation. The empirically implementable criterion that we consider is the Bourguignon (1989) ordered poverty gap dominance criterion. This criterion requires that poverty, measured by the income poverty gap, be smaller in the dominating distribution than in the dominated distribution for any collection of poverty lines that are *negatively* related to the agent's vertical standing. Just as in the classical case with homogeneous agents, we also introduce a "partial cumulative sum of income" dominance criterion similar to the Lorenz one and prove that it is equivalent to ordered poverty gap dominance.

This paper contributes to the multidimensional - in fact two-dimensional - inequality measurement literature which has emerged in the last forty years or so. To the best of our knowledge, no contribution to this literature has succeeded in establishing an *equivalence* between an empirically implementable criterion (such as Lorenz or poverty gap dominance), a welfarist (or otherwise) unanimity over a class of functions that transform the attributes into achievement and an elementary operation that captures in an intuitive way the nature of the equalization sought.

For instance, Atkinson and Bourguignon (1982) (and before them Hadar and Russell (1974)) showed that first- and second-order multidimensional stochastic dominance imply utilitarian dominance over a class of individual utility functions that is specific to the order of dominance. They also suggested (without providing any proof) that there could be an equivalence between their multidimensional stochastic dominance criteria and utilitarian unanimity over their class of individual utility functions. But they did not identify an elementary operation that could be implied by their criteria or that could imply them. Atkinson and Bourguignon (1987) proposed a nice interpretation of one of the Atkinson and Bourguignon (1982) stochastic dominance criteria in the specific case of two attributes, one of which being interpreted as an ordinal index of needs (such as household size). Yet, they did not identify the elementary operation which, when performed a finite number of times, would coincide with the criterion. Their criterion and equivalence results, developed originally for distributions of attributes with identical (marginal) distribution of needs, were extended to more general situations by Jenkins and Lambert (1993) and Bazen and Moyes (2003), but without identifying the underlying elementary operations.

It was in the same two-dimensional context as that considered in Atkinson and Bourguignon (1987) that Bourguignon (1989) introduced his ordered poverty gap criterion. Bourguignon (1989) also identified the class of utility functions over which utilitarian unanimity was equivalent to his criterion. However, he did not identify the elementary operation that would be equivalent to it.

Elementary operations believed to lie behind the criteria proposed by Atkinson and Bourguignon (1982), Atkinson and Bourguignon (1987) and Bourguignon (1989) have been discussed by various authors, including Atkinson and Bourguignon (1982) themselves, Ebert (1997), Fleurbaey, Hagneré, and Trannoy (2003) and Moyes (2012) (among others). Yet none of these papers showed that performing these elementary operations a finite number of times was equivalent to the implementable criteria. In a related vein, Muller and Scarsini (2012) established an equivalence between a class of elementary transformations - multidimensional transfers and correlation-reducing permutations, to be discussed below - and a utilitarian unanimity over the class of increasing and submodular utility functions.[1] However, they did not succeed in identifying an implementable test - such as Lorenz or poverty gap dominance - that coincides with either their elementary

[1] See e.g. Marinacci and Montrucchio (2005) for a definition of these properties.

transformations or the utilitarian unanimity over their class of utility functions.

Progress towards establishing equivalence between an empirically implementable criterion, a utilitarian unanimity over a suitable class of individual utility functions and a finite sequence of elementary transformations has been made in two streams of the literature. One of them, initiated by Epstein and Tanny (1980) (see also Tchen (1980)), and significantly generalized by Decancq (2012), considers first-order stochastic dominance rankings of multivariate distributions in the context of decision making under uncertainty. In this setting, Decancq (2012) established an equivalence between first-order dominance among multivariate distributions with the same marginals and the possibility of going from the dominated to the dominating distribution by a finite sequence of Frechet rearrangements. By significantly generalizing results from Lehmann (1955) and Levhari, Paroush, and Peleg (1975), Osterdal (2010) also established an equivalence between utilitarian unanimity over the class of all increasing utility functions, the possibility of going from one distribution to another by a finite sequence of improving mass transfers, and a specific first-order stochastic dominance test that is less discriminant than the usual multivariate one considered in Hadar and Russell (1974) and Atkinson and Bourguignon (1982). None of these results, however, sheds light on the meaning of income equalization in a two-dimensional context.

Progress in this direction has been made by Gravel and Moyes (2012), who established a *form of equivalence* between the *three* following answers to the basic question of when a distribution A of income between vertically differentiated agents is normatively better than another distribution B:

(a) When A could be obtained from B by performing a finite sequence of *either* Pigou-Dalton transfers of income between agents of the same type *or* correlation-reducing permutations.

(b) When A is considered better than B by all utilitarian planners who assume that vertically differentiated agents convert income into well-being by the same utility function whose *marginal utility* of income is *decreasing* with respect to both income and the source of vertical differentiation.

(c) When A dominates B by the ordered poverty gap criterion.

Answer (a) combines two elementary operations. The first is the standard Pigou-Dalton transfer performed between agents of the same "type". The second is a *correlation-reducing* income permutation between two agents, one of them being both richer and more highly ranked than the other. A correlation-reducing permutation is an operation closely related to the notion of Frechet rearrangement used by Decancq (2012) (see also Tsui (1999), Atkinson and Bourguignon (1982) and Epstein and Tanny (1980)). Answers (b) and (c) are of course those considered here.

However, Gravel and Moyes (2012) did not succeed in proving that answer (c) (or answer (b)) implies answer (a) (clearly answer (a) implies answer (b) which in turn implies answer (c)). What they prove is that *if* distribution A dominates distribution B for the ordered poverty gap criterion, *then* it is possible to add dummy individuals - or *phantoms* - to both distributions A and B in such a way as to be able to go from phantom-augmented distribution B to phantom-augmented distribution A by first, performing a finite sequence of Pigou-Dalton transfers among agents of the same type and second, performing a finite sequence of favourable permutations. This inability to prove the equivalence of statements (a), (b) and (c) without resorting to phantoms is clearly disappointing. What matters, after all, are distributions of attributes between *actual* agents. The fact that these agents could make (receive) transfers to (from) non-existing phantoms can appear of secondary importance.

In this paper, we prove an equivalence between the above answers (b) and (c) and the possibility of going from dominated to dominating distributions by a finite sequence of elementary transfers of income from richer and more highly ranked agents to poorer and less highly ranked agents. We do so without resorting to phantoms or dummies. Our notion of transfer contains as particular cases both the correlation-reducing permutation and the within-type transfer considered in Gravel and Moyes (2012). An additional contribution of this paper is to establish an equivalence between the Bourguignon ordered poverty gap criterion and a Lorenz-like dominance criterion based on partial sums of incomes of the poorest agents. A difficulty involved in the latter criterion is identifying the m poorest agents when these agents are differentiated by a non-income characteristic. We therefore view this paper as providing, to the best of our knowledge, the first dominance foundation for income equalization among heterogeneous agents.

The organization of the remainder of the paper is as follows. In the next section, we introduce notations and give the definitions of the main criteria and elementary transformations considered. The main results are stated and discussed in Section 3. Section 4 discusses how the results of section 3 extend to the case where the number of agents and/or the total income to be distributed vary across distributions and Section 5 concludes.

2 The formal setting

2.1 Notations

We consider a finite *population* of n agents who are vertically differentiated into k *categories* or *types*, indexed by h. Agents in lower categories are assumed to be more needy (or worse off) *ceteris paribus* than agents in higher categories. These categories may refer to any non-pecuniary source of agent differentiation, such as health, number of members, education level, labour effort, etc. For any category h, we denote by $\mathcal{N}(h)$ the set of agents in category h and by $n(h) = \#\mathcal{N}(h)$ the number of those agents. Our objective is to provide a ranking of alternative distributions of income (or any other cardinally meaningful variable) between these differentiated agents on the basis of *equality*. Any such income distribution, $\mathbf{x}$ say, is depicted as a collection of k vectors $(x_1^h, ..., x_{n(h)}^h) \in \mathbb{R}^{n(h)}$ (for $h = 1, ..., k$). The criteria used in this paper for comparing alternative distributions are all *anonymous* conditional on the agent's type. Because of this, we find convenient to index the agents in category h (for $h = 1, ..., k$) according to their income and therefore to assume that $x_i^h \leq x_{i+1}^h$ for $i = 1, ..., n(h) - 1$. More compactly, we write $\mathbf{x} = \{(x_1^h, ..., x_{n(h)}^h)\}_{h=1}^{n(h)}$. Since we focus on pure equality considerations, we restrict our attention to income distributions $\mathbf{x}$ such that

$$\sum_{h=1}^{k} \sum_{i \in \mathcal{N}(h)} x_i^h = I \text{ for some real number } I.$$

We let $\mathcal{D}(I)$ denote the set of all such income distributions.
For any income poverty *threshold* $t \in \mathbb{R}$ and any distribution $\mathbf{x}$, we also denote by $\overline{\mathcal{P}}^{\mathbf{x}}(h,t)$ and $\mathcal{P}^{\mathbf{x}}(h,t)$ the (possibly empty) sets of agents of type h who are, respectively, weakly and strictly poor for threshold t in distribution $\mathbf{x}$. These sets are defined by:

$$\overline{\mathcal{P}}^{\mathbf{x}}(h,t) = \{i \in \mathcal{N}(h) : x_i \leq t\} \text{ and } \mathcal{P}^{\mathbf{x}}(h,t) = \{i \in \mathcal{N}(h) : x_i < t\}$$

while the number of poor that these sets contain are denoted respectively by $\overline{p}^{\mathbf{x}}(h,t) = \#\overline{\mathcal{P}}^{\mathbf{x}}(h,t)$ and $p^{\mathbf{x}}(h,t) = \#\mathcal{P}^{\mathbf{x}}(h,t)$. Given two distributions $\mathbf{x}, \mathbf{y} \in \mathcal{D}(I)$, we finally denote by $\underline{v}(\mathbf{x}, \mathbf{y})$ and $\overline{v}(\mathbf{x}, \mathbf{y})$, respectively, their lowest and highest income.

We now introduce the main concepts between which an equivalence will be established.

2.2 Elementary transformations

The main elementary transformation considered herein is the following notion of Between-Type transfers discussed in many papers, including Ebert (1997), Atkinson and Bourguignon (1982), Fleurbaey, Hagneré, and Trannoy (2003), Muller and Scarsini (2012) and Gravel and Moyes (2012).

Definition 1 (Between-Type Progressive Income Transfer) *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$. *We say that* $\mathbf{x}$ *is obtained from* $\mathbf{y}$ *by means of a Between-Type Progressive Income Transfer (BTPIT) if there are categories* g *and* h *for which* $g \leq h$, *two agents* $i^g \in \mathcal{N}(g)$ *and* $i^h \in \mathcal{N}(h)$ *for which* $y_{i^h}^h > y_{i^g}^g$ *and a number* $\alpha \in]0, \frac{y_{i^h}^h - y_{i^g}^g}{2}]$ *such that:*

(i) $x_i^g = y_{i+1}^g$ *for all* $i \in \mathcal{N}(g)$ *such that* $i^g \leq i < r_+^g(\alpha)$ *(if any).*

(ii) $x_{r_+^g}^g(\alpha) = y_{i^g}^g + \alpha$.

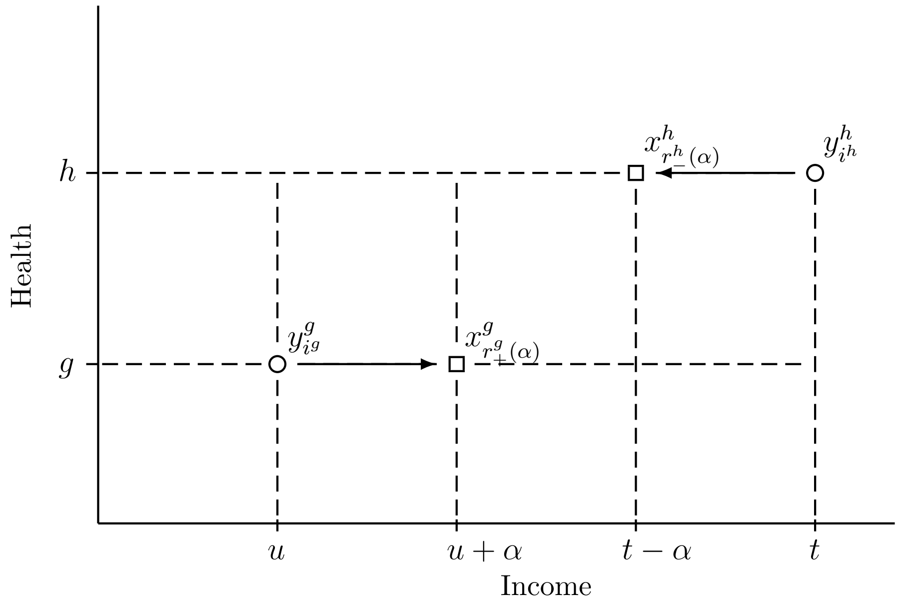
(iii) $x_i^h = y_{i-1}^h$ *for all* $i \in \mathcal{N}(h)$ *such that* $r_-^h(\alpha) < i \leq i^h$ *(if any)*

(iv) $x_{r_-^h(\alpha)}^h = y_{i^h}^h - \alpha$.

(v) $x_i^l = y_i^l$ *for any other pair* (i,l) *where* $l \in \{1, ..., k\}$ *and* $i \in \mathcal{N}(l)$.

where $r_+^g(\alpha) := \max\{i \in \mathcal{N}(g) : y_i^g < y_{i^g}^g + \alpha\}$, $r_-^h(\alpha) := \min\{i \in \mathcal{N}(h) : y_i^h > y_i^{i^h} - \alpha\}$.

Figure 1: A between-type progressive income transfer



A BTPIT resembles a standard one-dimensional Pigou-Dalton transfer. There is however a major difference: the beneficiary of the transfer must both be poorer than the donor and have a (weakly) lower status. Put differently, the transfer recipient must be deprived in both dimensions – income and status – compared to the donor. This kind of transfer is a particular case of the *equalizing transformation* considered by Muller and Scarsini (2012), where the transfers can be made in all dimensions. In the current setting, it would not make much sense to transfer the (ordinal) non-pecuniary variable by which agents differentiate themselves. A BTPIT is illustrated in Figure 1. Note that our definition of a BTPIT allows the donor to be of the same type as the receiver. Hence, the standard one-dimensional Pigou-Dalton transfer (conditional on type) is a particular case of BTPIT. Note also that our definition of BTPIT rules out the possibility of the amount transferred being more than half the income difference between giver and receiver.

This restriction can be eliminated by considering the following elementary transformation, called *Favourable Income Permutation* in Gravel and Moyes (2012).

Definition 2 (Favourable Income Permutation) *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$. *We say that* $\mathbf{x}$ *is obtained from* $\mathbf{y}$ *by means of a Favourable Income Permutation (FIP) if there are categories* g *and* h *for which* $g < h$ *and two agents* $i^g \in \mathcal{N}(g)$ *and* $i^h \in \mathcal{N}(h)$ *for which* $y^h_{i^h} > y^g_{i^g}$ *such that:*

(i) $x^g_i = y^g_{i+1}$ *for all* $i \in \mathcal{N}(g)$ *such that* $i^g \leq i < r^g(i^h)$ *(if any).*

(ii) $x^g_{r^g(i^h)} = y^h_{i^h}$.

(iii) $x^h_i = y^h_{i-1}$ *for all* $i \in \mathcal{N}(h)$ *such that* $r^h(i^g) < i \leq i^h$ *(if any).*

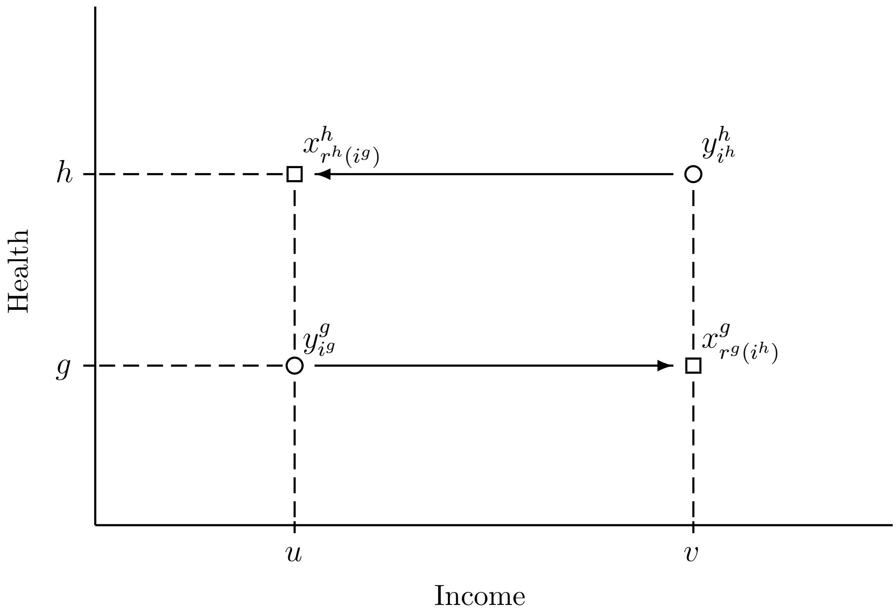
(iv) $x^h_{r^h(i^g)} = y^g_{i^g}$.

(v) $x^l_i = y^l_i$ *for any other pair* (i, l) *where* $l \in \{1, ..., k\}$ *and* $i \in \mathcal{N}(l)$.

where $r^g(i^h) := \max\{i \in \mathcal{N}(g) : y^g_i < y^h_{i^h}\}$ *and* $r^h(i^g) := \min\{i \in \mathcal{N}(h) : y^h_i > y^g_{i^g}\}$.

An FIP consists in exchanging the income endowment of a relatively rich agent belonging to a relatively high category with that of a poorer agent from a lower category. It can thus be viewed as an extreme form of BTPIT in which the total income difference between the two individuals is transferred. Figure 2 represents an FIP where agents i^h in category h and i^g in category $g < h$ exchange their incomes, respectively equal to v and u, with $v > u$.

Figure 2: A favourable income permutation



Gravel and Moyes (2012) showed that a BTPIT can always be decomposed into a (within-type) conventional Pigou-Dalton transfer followed by an FIP provided that a phantom individual is added. This individual must be endowed with the income of the beneficiary and the health status of the donor prior to the transfer. In this paper, we show that the possibility of going from a distribution $\mathbf{y}$ to a distribution $\mathbf{x}$ by a finite sequence of transfers that include FIP as a special (extreme) case is equivalent to the utilitarian dominance of $\mathbf{y}$ by $\mathbf{x}$ for a relatively large class of utility functions, which we now define.

2.3 Utilitarian dominance.

This notion of dominance rides on the assumption that all agents of a given type transform their income into some type-dependant, ethically meaningful achievement (well-being, happiness, freedom, etc.) by means of the same (utility) function satisfying some minimal property. Specifically, the utility achieved by agent i of type h in distribution $\mathbf{x}$ is indicated by $U^h(x_i^h)$, where $U^h : \mathbb{R} \to \mathbb{R}$. The *utilitarian rule* ranks the distributions on the basis of the *sum of the utilities* they generate. More precisely, the utilitarian rule considers distribution $\mathbf{x}$ to be no worse than distribution $\mathbf{y}$ if and only if

$$\sum_{h=1}^{k} \sum_{i=1}^{n(h)} U^h \left(x_i^h\right) \geqq \sum_{h=1}^{k} \sum_{i=1}^{n(h)} U^h \left(y_i^h\right). \tag{1}$$

The list of type-dependent utility functions $U^1, ..., U^k$ used by the utilitarian rule reflects some normative evaluation of the contribution of income to every agent's achievement, conditional on the agent's type. For the sake of robust normative evaluation, the dominance approach commonly requires a consensus among a relatively large class, $\mathcal{U}^*$ say, of such lists of utility functions. This gives rise to the following general notion of utilitarian dominance.

Definition 3 *(Utilitarian Dominance). Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$*. We say that* $\mathbf{x}$ *utilitarian dominates* $\mathbf{y}$ *for the class* $\mathcal{U}^*$ *of collections of* k *utility functions if and only if*

$$\sum_{h=1}^{k} \sum_{i=1}^{n(h)} U^h(x_i^h) \geqq \sum_{h=1}^{k} \sum_{i=1}^{n(h)} U^h(y_i^h), \ \forall \ (U^1, ..., U^k) \in \mathcal{U}^*. \tag{2}$$

In this paper, we specifically consider the class $\mathcal{U}^*$ of type-dependent $U^1, ..., U^k$ such that that satisfy:

$$U^h(w + a) - U^h(w) \geq U^{h'}(w' + a) - U^{h'}(w') \tag{3}$$

for any non-negative real number a, any categories $h, h' \in \{1, ..., k\}$ with $h \leq h'$, and any income pair $(w, w^{'})$ such that $w \leq w'$. In words, $\mathcal{U}^*$ is the class of collections of utility functions U^h (for $h = 1, ..., k$) with the property that the contribution of an additional unit of income to the individual's advantage (as measured by the function U^h) is decreasing with respect to *both* income and type.

2.4 Ordered poverty gap dominance

The ordered poverty gap criterion has been proposed by Bourguignon (1989) for comparing income distributions between households of differing sizes. In order to discuss this criterion in the current context, we first define set $\mathcal{V} \subset \mathbb{R}^k$ by:

$$\mathcal{V} = \{(v_1, ..., v_k) \in \mathbb{R}^k : v_1 \geq v_2 \geq ... \geq v_k\} \tag{4}$$

Set $\mathcal{V}$ comprises all combinations of poverty lines (one such line for every type) that are weakly decreasing with respect to type. Given this set, we define the Ordered Poverty Gap (OPG) dominance criterion as follows.

Definition 4 *(Ordered Poverty Gap Dominance). Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\in \mathcal{D}(I)$. *We say that* $\mathbf{x}$ *dominates* $\mathbf{y}$ *for the Ordered Poverty Gap criterion, here denoted by* $\mathbf{x} \succsim^{OPG} \mathbf{y}$, *if the following holds:*

$$\sum_{h=1}^{k} \sum_{i \in \mathcal{N}(h)} \max(v_h - x_i^h, 0) \leq \sum_{h=1}^{k} \sum_{i \in \mathcal{N}(h)} \max(v_h - y_i^h, 0), \ \forall (v_1, ..., v_k) \in \mathcal{V} \tag{5}$$

In words, $\mathbf{x}$ dominates $\mathbf{y}$ for the OPG criterion if, for all possible poverty lines that are (weakly) decreasing with respect to agent's type, the minimum income required to eliminate poverty defined by these lines is lower in $\mathbf{x}$ than in $\mathbf{y}$.

While the OPG criterion requires a comparison of the poverty gap between two distributions for all lists of ordered poverty lines in $\mathcal{V}$ (a non-countable set), it is nonetheless easily implementable (see e.g. Decoster and Ooghe (2006) or Gravel, Moyes, and Tarroux (2009)). One way of implementation is to restrict attention to the (finite) subset of lists of poverty lines in $\mathcal{V}$ that are actually observed in the two distributions under comparison. Another is to use the ingenious alternative formulation of the OPG dominance criterion proposed by Bourguignon (1989) (p. 74, equation (12)) via an iterative procedure based on the largest difference in poverty gap between two distributions for all poverty lines above any arbitrary threshold. In the next subsection, we introduce an alternative dominance criterion that is somewhat analogous to Lorenz dominance, in that it is based on partial sums of the income of the m poorest agents and that is also easily implementable.

Before doing so, we introduce some additional notation pertaining to the OPG dominance criterion. Specifically, for any distribution $\mathbf{x} \in \mathcal{D}(I)$ and (ordered) poverty lines $(v_1, ..., v_k) \in \mathcal{V}$, we denote by $P^{\mathbf{x}}(v_1, ..., v_k)$ the Ordered Poverty Gap of this distribution for those poverty lines defined by:

$$P^{\mathbf{x}}(v_1, ..., v_k) = \sum_{h=1}^{k} \sum_{i \in \mathcal{N}^{\mathbf{x}}(h)} \max(v_t - x_i^h, 0) \tag{6}$$

2.5 Cumulative lowest incomes dominance

In the classical case of income distributions among homogeneous agents, it is well known that poverty gap dominance is equivalent to the requirement, known as Lorenz dominance, that the sum of the incomes of the m poorest agents is larger in the dominating than in the dominated distribution, whatever m is. In this subsection, we introduce a similar dominance criterion based on the sum of the m poorest incomes among vertically differentiated agents. This is challenging because there is no obvious way to define who the m poorest agents are when these agents are differentiated with respect to a non-income characteristic. For example, is an individual earning \$1000 a month and working 80 hours indisputably poorer than someone who earns \$1500 and works 160 hours?

In what follows, we adopt the view, consistent with that underlying the OPG dominance criterion, that whatever the definition of a poor agent is for a given type, any agent who is both of a lower type and poorer than this agent must also be considered as poor. This means that whether or not an agent is poor is itself a function of the agent's type. Thus we have the following definition of set $\Pi(\mathbf{x}, m)$ of m poorest agents in any distribution $\mathbf{x} \in \mathcal{D}(I)$ (for any $m \leq n$):

$$\Pi(\mathbf{x}, m) = \left\{ (r_1, ..., r_k) : 0 \leq r_h \leq n(h),\ \sum_h r_h = m,\ x^h_{r_h+1} > x^{h'}_{r_{h'}}\ \forall h' > h \right\}. \quad (7)$$

with the convention that $x^h_0 = \underline{v}(\mathbf{x}, \mathbf{y})$ and $x^h_{n_h+1} = \overline{v}(\mathbf{x}, \mathbf{y})$ for every h. In words, $\Pi(\mathbf{x}, m)$ is the set of all combinations $(r_1, ..., r_k)$ of *type-dependent ranks* of agents in their type's income distribution that have two characteristics. First, the ranks must sum to m so as to identify m "poorest" agents. Second, the ranks must be such that if an agent in a given category stands, in the income distribution of this category, precisely at the rank assigned to the category, then any agent who is *both* in a lower category *and* poorer than this agent must stand, in the income distribution of his/her category, at a lower rank than that of his/her category.
Note that for any distribution $\mathbf{x}$ and $1 \leq m \leq n$, the set $\Pi(\mathbf{x}, m)$ is *never* empty. Indeed, given m, define the index $h \in \{1, ..., k\}$ to be the category with the property that $n_h \geq 1$ and

$$m \in \left\{ \sum_{g=1}^{h-1} n_g + 1, \sum_{g=1}^{h-1} n_g + 2, ..., \sum_{g=1}^{h} n_g \right\}.$$

Then the list of ranks $(n_1, n_2, ..., n_{h-1}, m - \sum_{g=1}^{h-1} n_g, 0, ..., 0)$ belongs to $\Pi(\mathbf{x}, m)$. It should be noted that, for any given m, there will typically be many lists of type-dependent ranks in the set $\Pi(\mathbf{x}, m)$. To illustrate this point, consider again the example of income distributions between individuals who work different numbers of hours. Suppose there are three categories of earners in terms of monthly time devoted to work (ordered from worst to best): full-time, part-time, and no time at all (rentiers). Suppose there are three individuals, each falling into one of these three categories. Three possible distributions of their monthly income are depicted in the following table:

Distribution	Full-time	Part-time	Rentier
$\mathbf{x}$	1500	1000	2000
$\mathbf{y}$	1400	2200	900
$\mathbf{z}$	1600	1700	1200

Consider first distribution $\mathbf{x}$. There are two ways of defining *the* poorest individual as per set $\Pi(.)$ in this distribution. One is to consider the full-time earner as the poorest. This corresponds to the list of ranks $(1, 0, 0)$. The other is to consider the part-time earner to be *the* poorest. This corresponds to the list of ranks $(0, 1, 0)$. Note that this list of ranks satisfies the property that $x^1_1 = 1500 > 1000 = x^2_1 > x^3_0 = \underline{v}(\mathbf{x}, \mathbf{y})$. Any possibility of considering the rentier individual as the poorest, which would correspond to the list of ranks $(0, 0, 1)$, is ruled out by the definition of $\Pi(.)$.

While this definition of the set of m poorest agents seems appropriate when the agents are vertically differentiated with respect to a non-income characteristic, its use in a dominance criterion based on the sum of incomes of the m poorest agents raises an additional difficulty: the set of m poorest agents defined by (7) may vary across distributions. To illustrate this, consider distribution $\mathbf{y}$ in the table. In this distribution, while there are also two ways to define *the* poorest, these ways are different to what they are for distribution $\mathbf{x}$. In $\mathbf{y}$, either the full-time worker is the poorest, or the rentier.

Using this definition of $\Pi(\mathbf{x}, m)$, we now propose the following notion of cumulative lowest incomes dominance.

Definition 5 *Let $\mathbf{x}$ and $\mathbf{y}$ be two distributions in $\in \mathcal{D}(I)$. We say that $\mathbf{x}$ dominates $\mathbf{y}$ for the cumulative lowest incomes criterion, which we denote as $\mathbf{x} \succsim^{CLI} \mathbf{y}$ if, for any $m \in \{1, ..., n\}$, and any $(r_1, ..., r_k) \in \Pi(\mathbf{x}, m)$, there exists $(r'_1, ..., r'_k) \in \Pi(\mathbf{y}, m)$ such that:*

$$\sum_{g=1}^{h} r'_g \geq \sum_{g=1}^{h} r_g \tag{8}$$

for any $h = 1, ..., k$ and:

$$\sum_{h=1}^{k} \sum_{i \leq r_h} x_i^h \geq \sum_{h=1}^{k} \sum_{i \leq r'_h} y_i^h. \tag{9}$$

In words, income distribution $\mathbf{x}$ dominates income distribution $\mathbf{y}$ for the cumulative lowest incomes criterion if, for any m, and any admissible set of m poorest agents in $\mathbf{x}$, there is an admissible set of m poorest agents in $\mathbf{y}$ who jointly have a lower sum of income than the m poorest agents in $\mathbf{x}$ (Condition (9) *and* who are such that, for every category h, the sum, over all categories below h, of the category-specific ranks below which agents are considered poor is larger in $\mathbf{y}$ than in $\mathbf{x}$ (Condition (8)).

To illustrate this latter condition, which does not appear in the standard unidimensional Lorenz dominance criterion, consider again distributions $\mathbf{x}$ and $\mathbf{y}$. By considering $m = 1$, it is clear that $\mathbf{x}$ does not dominate $\mathbf{y}$ for the cumulative lowest incomes criterion. As discussed, $\Pi(\mathbf{x}, 1) = \{(1, 0, 0), (0, 1, 0)\}$ (the full-time and the part-time worker can both be *the* poorest in $\mathbf{x}$) and $\Pi(\mathbf{y}, 1) = \{(1, 0, 0), (0, 0, 1)\}$ (the full-time worker and the rentier can both be the poorest in $\mathbf{y}$). Consider the list of ranks $(0, 1, 0) \in \Pi(\mathbf{x}, 1)$. Then the two lists of ranks $(1, 0, 0)$ and $(0, 0, 1)$ in $\Pi(\mathbf{y}, 1)$ respectively violate conditions (9) and (8). In the other direction, it is clear that $\mathbf{y}$ can not dominate $\mathbf{x}$ by the cumulative lowest incomes criterion, because the full-time worker in $\mathbf{y}$ earns strictly less than the full-time worker in $\mathbf{x}$.

However, distribution $\mathbf{z}$ dominates distribution $\mathbf{x}$ (but not distribution $\mathbf{y}$) by the cumulative lowest incomes criterion. If we first focus on *the* poorest individual, there are two acceptable candidates for that position in $\mathbf{z}$: the full-time worker, or the rentier. If the full-time worker is chosen, then this worker can be chosen as the poorest in distribution $\mathbf{x}$ as well, and conditions (8) and (9) hold. On the other hand, if the rentier is chosen as the poorest individual in $\mathbf{z}$, then there is an acceptable candidate for the poorest in $\mathbf{x}$ - namely the part-time worker - who has a lower income in $\mathbf{x}$ than the rentier in $\mathbf{z}$. Suppose that we now focus on the *two* poorest individuals. There are two possible choices in $\mathbf{z}$: the full-time and part-time workers (ranks $(1,1,0)$), or the full-time worker and the rentier (ranks $(1,0,1)$). However there is only one acceptable way to define the two poorest individuals in $\mathbf{x}$, who must be the full-time and the part-time workers. Note that this only acceptable way satisfies Condition (8) with respect to either of the two acceptable choices of the two poorest in $\mathbf{z}$. Also, whichever pair is chosen as the two poorest agents in $\mathbf{z}$, they have a larger cumulative income than the two poorest in $\mathbf{x}$. Hence $\mathbf{z}$ does indeed dominate $\mathbf{x}$.

The cumulative lowest incomes dominance criterion is more difficult to implement empirically than its unidimensional Lorenz dominance cousin. For one thing, it can not be implemented by simply comparing two curves drawn for each of the two distributions in isolation. Yet the cumulative lowest incomes dominance criterion is easy to implement through the following procedure. For any two distributions $\mathbf{x}$ and $\mathbf{y}$, identify the lowest income in the lowest category and choose, as a possible candidate for the dominating distribution, the distribution where this lowest income in the lowest category is the largest. For any $m = 1, ..., n$, identify then all admissible (under Condition (7)) combinations of ranks that sum to m in the potentially dominating distribution. For any such combination, find all admissible (under Condition (7)) combinations of ranks that sum to m in the other distribution and that satisfy inequality (8). Select from these combinations of ranks those associated with the smallest sum of incomes, and compare this sum of incomes with that associated with the initial combination of ranks in the potentially dominating distribution. If the comparison of the two sums violates Condition (9), then there is no dominance and the procedure stops. Otherwise, the procedure continues for all m, in which case we conclude that there is dominance.

As shown below, this dominance criterion is closely associated with the notion of equalization linked to BTPIT and/or FIP.

3 Main result

The main theorem proved in this paper is the following.

Theorem 1 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$*. Then the following four statements are equivalent.*

(*i*) *It is possible to go from* $\mathbf{y}$ *to* $\mathbf{x}$ *by a finite sequence of BTPIT and/or FIP.*

(*ii*) $\mathbf{x}$ *utilitarian dominates* $\mathbf{y}$ *for the class* $\mathcal{U}^*$.

(*iii*) $\mathbf{x} \succsim^{OPG} \mathbf{y}$.

(*iv*) $\mathbf{x} \succsim^{CLI} \mathbf{y}$.

The proof of this theorem proceeds in several steps. The first, establishing that (i) implies (ii) and that (ii) implies (iii), is easy and quite well-known (see for example Ebert (1997) or Gravel and Moyes (2012)). It is formally stated in Proposition 4 (See section 6.1 in the appendix) which is proved for the sake of completeness. In section 6.2 in the appendix we state and prove the equivalence between (iii) and (iv). We can then turn to the proof of the most difficult implication that statement (iii) implies statement (i). Proving this implication amounts to constructing an algorithm for going from a distribution $\mathbf{y}$ to a distribution $\mathbf{x}$ by a finite sequence of either BTPIT and/or FIP based solely on the information that $\mathbf{x} \succsim^{OPG} \mathbf{y}$. In every step of the algorithm, either a BTPIT or an FIP must be performed in such a way that the result of this elementary operation remains dominated by distribution $\mathbf{x}$. Section 3.1 is devoted to technical lemmas while Section 3.2 establishes a very important first step in the construction of the algorithm. Specifically, we prove in Section 3.2 that if $\mathbf{x}$ strictly dominates $\mathbf{y}$ as per the OPG criterion, it is always possible to perform either an FIP or a BTPIT in a way that preserves the OPG dominance by $\mathbf{x}$ of the newly created distribution. We actually propose a *diagnostic tool* that can be used to determine whether an FIP or a BTPIT should be performed. Finally, in section 3.3, we define our algorithm and prove its finiteness, which concludes the proof of $(iii) \Rightarrow (i)$.

3.1 Some technical lemmas

We start tackling the most difficult implication of Theorem 1 (Statement (iii) implies Statement (i)) by proving several auxiliary results. We assume without loss of generality that, in the two distributions $\mathbf{x}$ and $\mathbf{y}$ under consideration, we have $x_i^h \neq y_j^h$ for every type $h = 1, ..., k$ and every $i, j \in \mathcal{N}(h)$. In effect, if this condition was not satisfied, i.e. if there was a type h for which $x_i^h = y_j^h$ for some $i, j \in \mathcal{N}(h)$, these two agents could be removed and we could proceed with the remaining population. Since the OPG criterion is additively separable, such a removal of agents with the same type and income from distributions $\mathbf{x}$ and $\mathbf{y}$ would not affect their ranking as per the OPG criterion.

The first auxiliary result of this section is the following lemma (proved, like all lemmas and formal claims in the Appendix) which says that if $\mathbf{x}$ is a distribution that dominates $\mathbf{y}$ for the OPG criterion, the poorest person in the worst category is weakly richer in $\mathbf{x}$ than in $\mathbf{y}$ and, conversely, the richest person in the best category is poorer in $\mathbf{x}$ than in $\mathbf{y}$.

Lemma 1 *Let $\mathbf{x}$ and $\mathbf{y}$ be distributions in $\mathcal{D}(I)$, for which $\mathbf{x} \succsim^{OPG} \mathbf{y}$. Then $y_1^1 < x_1^1$ and $y_{n(k)}^k > x_{n(k)}^k$.*

The next lemma states that, if a distribution $\mathbf{x}$ dominates a distribution $\mathbf{y}$ by the OPG criterion, then the sum of incomes held by agents in the $\overline{h}$ lowest categories must be larger or equal in $\mathbf{x}$ than in $\mathbf{y}$ for any $\overline{h}$.

Lemma 2 *Let $\mathbf{x}$ and $\mathbf{y}$ be two distributions in $\mathcal{D}(I)$ for which $\mathbf{x} \succsim^{OPG} \mathbf{y}$. Then $\sum_{h=1}^{\overline{h}} \sum_{i \in \mathcal{N}(h)} x_i^h \geq \sum_{h=1}^{\overline{h}} \sum_{i \in \mathcal{N}(h)} y_i^h$ for all $\overline{h} = 1, ..., k$.*

We next state an important lemma that provides a sufficient condition for performing an FIP from distribution $\mathbf{y}$ in such a way that the distribution obtained after making such an FIP remains dominated by $\mathbf{x}$ as per the OPG criterion.

Lemma 3 *Let $\mathbf{x}$ and $\mathbf{y}$ be two distributions in $\mathcal{D}(I)$ such that $\mathbf{x} \succsim^{OPG} \mathbf{y}$. Assume that*

- *there exist $\mathbf{w} \in \mathcal{V}$, $i_1 \in \mathcal{N}(1)$ and $h_0 \in \{2, ..., k\}$ such that*[2]

$$P^{\mathbf{y}}(\mathbf{w}) = P^{\mathbf{x}}(\mathbf{w}), \;\; y^1_{i_1} = w_1 = ... = w_{h_0} > w_{h_0+1};$$

- *there exists a category g_0 such that: $2 \leq g_0 \leq h_0$ and:*

$$\sum_{h=l+1}^{g_0} \overline{p}^{\mathbf{y}}(h, y^1_{i_1}) < \sum_{h=l+1}^{g_0} \overline{p}^{\mathbf{x}}(h, y^1_{i_1}) \; \forall \, l = 1, ..., g_0 - 1. \tag{10}$$

Then there exists a distribution $\overline{\mathbf{x}} \in \mathcal{D}(I)$ such that $\overline{\mathbf{x}}$ was obtained from $\mathbf{y}$ by an FIP and $\mathbf{x} \succsim^{OPG} \overline{\mathbf{x}}$.

Although this result is important, it is of limited immediate usefulness. There are actually no obvious ways to identify the of poverty lines vector $\mathbf{w}$ that is required by this lemma. We will nonetheless use Lemma 3 on two occasions in what follows.

3.2 Identifying which elementary operation is possible: a diagnostic tool

An important prerequisite for performing any step of the algorithm that we want to construct is a "diagnostic tool" for determining which of the two elementary operations - FIP or BTPIT - can be performed at any given step of the algorithm. Our diagnostic tool is based on the critical value v_1^c that is defined as follows:

$$v_1^c := \inf \left\{ v_1 > y_1^1 : \exists v_2, ..., v_k \text{ s.t. } \mathbf{v} = (v_1, ..., v_k) \in \mathcal{V} \text{ and } P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \right\} \tag{11}$$

In words, v_1^c is the lowest poverty threshold *above* the smallest income in the lowest category in the dominated distribution $\mathbf{y}$ that can be part of a collection of ordered poverty thresholds for which the ordered poverty gap in distributions $\mathbf{x}$ and $\mathbf{y}$ is the same. It is clear that v_1^c is well-defined because the set:

$$\left\{ v_1 > y_1^1 : \exists v_2, ..., v_k \text{ s.t. } \mathbf{v} = (v_1, ..., v_k) \in \mathcal{V} \text{ and } P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \right\}$$

is not empty (it contains $\overline{v}(\mathbf{x}, \mathbf{y})$) and is also bounded from below (by $\underline{v}(\mathbf{x}, \mathbf{y})$). Two mutually exclusive cases are possible:

(A) $v_1^c > y_1^1$ and:

[2] With the convention that $w_{k+1} = \underline{v}(\mathbf{x}, \mathbf{y})$

(B) $v_1^c = y_1^1$.

As will now be shown, if case (A) holds, there is some margin to make a strict BTPIT to the poorest individual in category 1 (endowed with y_1^1) in such a way that the after-transfer distribution remains dominated by $\mathbf{x}$ as per the OPG criterion. This however does not mean that there cannot be an FIP involving an individual in category 1. If both an FIP and a BTPIT are possible, then our algorithmic procedure will always choose to perform the FIP. [3] As will also be shown, if on the other hand case (B) holds, then it is possible to involve the poorest individual of type 1 in an FIP while preserving the OPG dominance of $\mathbf{x}$ over the distribution obtained by doing so.

3.2.1 Case $(A) : v_1^c > y_1^1$.

In this case, v_h^c can recursively be defined (for any $h = 2, ..., k$) by:

$$v_h^c = \inf \left\{ v_h \geq \underline{v}(\mathbf{x}, \mathbf{y}) : \ \exists v_{h+1}, ..., v_k \text{ s.t. } \mathbf{v} = (v_1^c, .., v_{h-1}^c, v_h, .., v_k) \in \mathcal{V}, \ P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \right\} \tag{12}$$

For just for the same reason as for v_1^c, it is clear that v_h^c is well-defined for any $h = 2, ..., k$. By construction, we have $\mathbf{v}^c = (v_1^c, ..., v_k^c) \in \mathcal{V}$. We refer to vector $\mathbf{v}^c$ as to the critical vector.

We start by establishing the following important result: if an ordered list $\mathbf{v} \in \mathcal{V}$ of poverty lines is such that $v_1 > y_1^1$ and $v_{h_0} < v_{h_0}^c$ for some $h_0 \in \{2, ..., k\}$, then $P^{\mathbf{x}}(\mathbf{v}) < P^{\mathbf{y}}(\mathbf{v})$. Specifically, we prove the following result.

Lemma 4 *Let $\mathbf{x}$ and $\mathbf{y}$ be two distributions in $\mathcal{D}(I)$ such that $\mathbf{x} \succsim^{OPG} \mathbf{y}$. Suppose that $v_1^c > y_1^1$. Then, if $\mathbf{v} \in \mathcal{V}$ is such that $v_1 > y_1^1$ and $v_{h_0} < v_{h_0}^c$ for some $h_0 \in \{2, ..., k\}$, we will have $P^{\mathbf{x}}(\mathbf{v}) < P^{\mathbf{y}}(\mathbf{v})$.*

We now state as a corollary of Lemma 4 the following alternative definition of the critical vector $\mathbf{v}^c$.

Corollary 1 *Let $\mathbf{x}$ and $\mathbf{y}$ be two distributions in $\mathcal{D}(I)$ such that $\mathbf{x} \succsim^{OPG} \mathbf{y}$. Suppose that $v_1^c > y_1^1$. Then, for every $h = 2, ..., k$, we have:*

$$v_h^c = \min_{v_h} \left\{ \exists v_{-h} \in [\underline{v}(\mathbf{x}, \mathbf{y}), \overline{v}(\mathbf{x}, \mathbf{y})]^{k-1} : \ v_1 > y_1^1, \ \mathbf{v} = (v_h, v_{-h}) \in \mathcal{V} \text{ and } P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \right\}.$$

The next lemma establishes an important comparative statement about adjacent sets of strictly and weakly poor agents in $\mathbf{x}$ and $\mathbf{y}$ when these sets are defined with respect to the vector of ordered poverty lines $\mathbf{v}^c$ where the poverty lines assigned to the adjacent categories are the same. Specifically, the next lemma establishes the following.

[3] We explain in detail in Section 3.3 why we make this choice.

Lemma 5 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$*. Suppose that* $v_1^c > y_1^1$*. Then, for any* $h_0 \in \{1, ..., k\}$ *and* $\overline{h} \in \{0, ..., k - h_0\}$ *such that* $v_{h_0-1}^c > v_{h_0}^c = v_{h_0+\overline{h}}^c > v_{h_0+\overline{h}+1}^c$ [4]*, we have, for any* $l = 0, ..., \overline{h}$*:*

$$\sum_{h=h_0}^{h_0+l} \overline{p}^{\mathbf{x}}(h, v_h^c) \leq \sum_{h=h_0}^{h_0+l} \overline{p}^{\mathbf{y}}(h, v_h^c) \tag{13}$$

and:

$$\sum_{h=h_0+l}^{h_0+\overline{h}} p^{\mathbf{x}}(h, v_h^c) > \sum_{h=h_0+l}^{h_0+\overline{h}} p^{\mathbf{y}}(h, v_h^c). \tag{14}$$

Lemma 5 has the following important corollary, that will be quite useful in establishing the possibility of making a non-zero BTPIT to the poorest individual in the worst health category of distribution $\mathbf{y}$ when critical value v_1^c is strictly larger than the income (y_1^1) of this individual. This corollary in fact establishes the existence of (potential donors) individuals in a weakly larger category who have, in distribution $\mathbf{y}$, an income of v_1^c.

Corollary 2 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$*. Suppose that* $v_1^c > y_1^1$*. Let* $h_0 \in \{0, ..., k-1\}$ *be such that* $v_1^c = v_{h_0+1}^c > v_{h_0+2}^c$ *Then, there exists some* $j \in \{1, ..., h_0 + 1\}$ *for which* $y_i^j = v_1^c$*, for some* $i \in \mathcal{N}(j)$*.*

The next lemma shows that, when critical value v_1^c is strictly larger than y_1^1, we in fact have some leeway to perform a BTPIT while preserving OPG dominance. Specifically, the following lemma deals with ordered vectors of poverty lines that assign to the worst category a poverty line only marginally above the lowest income observed in that category. This lemma says, roughly, that for any such ordered vector of poverty lines, the poverty gap in the dominated distribution must exceed that of the dominating one by an even larger margin. The precise statement of this lemma is as follows.

Lemma 6 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$*. Suppose that* $v_1^c > y_1^1$*. Then, for some strictly positive but suitably small real number* ε_1*, we have:*

$$P^{\mathbf{y}}(y_1^1 + \varepsilon_1, v_2, ..., v_k) \geq P^{\mathbf{x}}(y_1^1 + \varepsilon_1, v_2, ..., v_k) + \varepsilon_1,$$

provided $(y_1^1 + \varepsilon_1, v_2, ..., v_k) \in \mathcal{V}$ *.*

We now establish in the following proposition that if $\mathbf{x} \succsim^{OPG} \mathbf{y}$, then it is possible to find a distribution $\widehat{\mathbf{x}}$ that is OPG dominated by $\mathbf{x}$ and that was obtained from $\mathbf{y}$ by a BTPIT whenever critical value v_1^c is strictly larger than y_1^1.

Proposition 1 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$*. Suppose that* $v_1^c > y_1^1$*. Then, there exists a distribution* $\widehat{\mathbf{x}} \in \mathcal{D}(I)$ *such that:*

[4] Using if necessary the convention that $v_0^c = \overline{v}(\mathbf{x}, \mathbf{y})$ and $v_{k+1}^c = \underline{v}(\mathbf{x}, \mathbf{y})$

- $\mathbf{x} \succsim^{OPG} \widehat{\mathbf{x}}$,
- $\widehat{\mathbf{x}}$ *was obtained from* $\mathbf{y}$ *by a BTPIT involving some agent* $i^j \in \mathcal{N}(j)$ *for some category* $j \in \{1, ..., k\}$ *and agent* $1 \in \mathcal{N}(1)$.

This proposition (and its proof) identifies a particular category $j \geq 1$ and a particular agent in that category (labeled i^j) that can transfer a strictly positive quantity of income to the poorest agent in category 1. Since we proved the proposition with the objective of constructing a finite sequence of such transfers, it is important for the sequence not to be unnecessarily long and, therefore, for each transfer to be as large as possible. This leads to the following notion of a *maximal transfer* .

Definition 6 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$. *Suppose that* $\hat{\mathbf{x}}$ *is obtained from* $\mathbf{y}$ *by means of a BTPIT where agent* $i^h \in \mathcal{N}(h)$ *transfers* $\alpha_0 \geq 0$ *to agent* $i^g \in \mathcal{N}(g)$ $(g \leq h)$ *and that* $\hat{\mathbf{x}} \precsim^{OPG} \mathbf{x}$. *We say that this transfer is* maximal *(with respect to* $\mathbf{x}$*) if any of the following conditions holds:*

(MT1) ***Equalizing transfer:*** *there exist* $i, i' \in \mathcal{N}(h)$ *such that* $\widehat{x}^h_{i'} = x^h_i$ *or there exists some* $i, i' \in \mathcal{N}(g)$ *such that* $\widehat{x}^g_{i'} = x^g_i$ *(that is, one of the two agents involved in the transfer obtains the income that they will have in final distribution* $\mathbf{x}$*).*[5]

(MT2) ***Breaking transfer:*** *for any* $\alpha < \alpha_0$, *the transfer of amount* α *from agent* $i^h \in \mathcal{N}(h)$ *to agent* $i^g \in \mathcal{N}(g)$ *is not equalizing. Additionally for any* $\alpha > \alpha_0$ *the distribution obtained by making a transfer of amount* α *from agent* i_h *to agent* i_g *is not dominated by* $\mathbf{x}$ *as per the OPG dominance criterion.*

(MT3) ***Half transfer:*** $\alpha_0 = (y^h_{i^h} - y^g_{i^g})/2$ *and, for any* $\alpha < \alpha_0$, *the transfer of amount* α *from agents* $i^h \in \mathcal{N}(h)$ *to agent* $i^g \in \mathcal{N}(g)$ *is not equalizing.*

Note that in the settings of Definition 6, given two agents $i^h \in \mathcal{N}(h)$ and $i^g \in \mathcal{N}(g)$, the amount of the maximal transfer between them is uniquely defined.

We illustrate this definition in the case where $k = 2$ by providing three examples of pairs of distributions $\mathbf{x}$ and $\mathbf{y}$ for which $v^c_1 > y^1_1$ and that give rise to the three different possibilities of maximal transfer.

Example 1 *As an example of an equalizing transfer, consider the distributions where* $\mathcal{N}(1) = \{1, 2\}$ *and* $\mathcal{N}(2) = \{1\}$ *and where:*

$$y^1_1 = 0,\ y^1_2 = 1,\ y^2_1 = 7,$$

$$x^1_1 = 2,\ x^1_2 = 4 \text{ and } x^2_1 = 2.$$

It is not hard to check that $\mathbf{x} \succsim^{OPG} \mathbf{y}$ *and that* $v^c_1 = 7 > y^1_1$.

[5]Recall that we always assume without loss of generality that, for every category h and every $i, i' \in \mathcal{N}(h)$, $y_{i'} \neq x_i$.

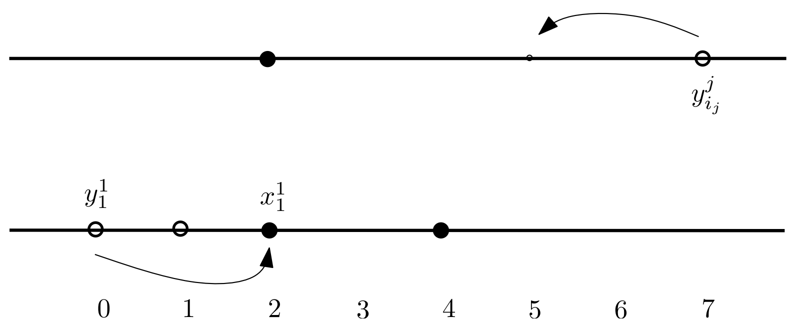


Figure 3: Equalizing transfer

It is then possible for the unique agent in $\mathcal{N}(2)$ *to transfer* 2 *units of income to agent 1 without breaking any OPG inequality, which corresponds to an equalizing transfer (see figure 3). Note that it would have been possible to transfer even more without breaking the OPG inequality. Yet we do not need to do so because, after receiving 2 units of income, agent 1 of category 1 has obtained the income of target distribution* $\mathbf{x}$*.*

Example 2 *As an example of a breaking transfer, consider the distributions where* $\mathcal{N}(1) = \{1, 2\} = \mathcal{N}(2)$ *and where:*

$$y_1^1 = 0 = y_1^2,\ y_2^1 = 7 = y_2^2,$$

$$x_1^1 = 5,\ x_2^1 = 6,\ x_1^2 = 1 \text{ and } x_2^2 = 2.$$

We have $\mathbf{x} \succsim^{OPG} \mathbf{y}$ *and, again, it turns out that* $v_1^c = 7 > y_1^1$*. It is possible for individual 2 in category 2 to transfer* 3 *units of income to individual 1 in category 1. Making this transfer changes the distribution from* $\mathbf{y}$ *to* $\widehat{\mathbf{x}}$*, where* $\widehat{\mathbf{x}}$ *is defined by:*

$$\widehat{x}_1^1 = 3,\ \widehat{x}_2^1 = 7,\ \widehat{x}_1^2 = 0 \text{ and } \widehat{x}_2^2 = 4.$$

As can be seen, $\mathbf{x} \succsim^{OPG} \widehat{\mathbf{x}}$*. Yet transferring* $3 + \varepsilon$ *(for any* $\varepsilon \in]0, 1/2]$*) would destroy this OPG dominance of the transformed distribution by target* $\mathbf{x}$*. See the appendix for a formal proof. This example is shown in Figure 4.*

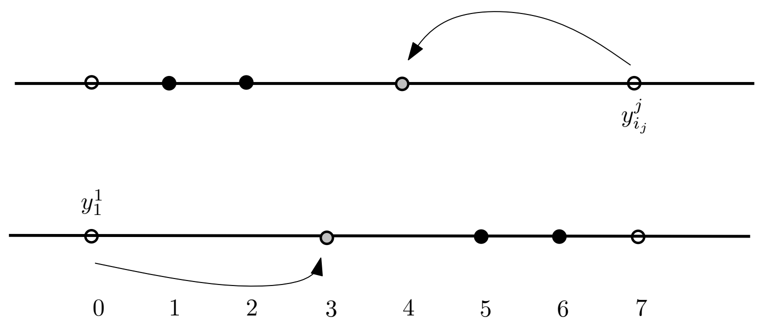


Figure 4: Breaking transfer

Example 3 *As an example of a half transfer (Figure 5), consider the distributions where* $\mathcal{N}(1) = \{1, 2\} = \mathcal{N}(2)$ *and:*

$$y_1^1 = 0 = y_1^2,\ y_2^1 = 6 = y_2^2,$$

$$x_1^1 = 4,\ x_2^1 = 5,\ x_1^2 = 1 \text{ and } x_2^2 = 2.$$

As can be seen, we have $\mathbf{x} \succsim^{OPG} \mathbf{y}$ *and* $v_1^c = 6 > x_1^1$*. It is possible for agent 2 in category 2 to transfer* 3 *to the poorest agent in category 1 - which is precisely half of their income difference - without breaking any of the inequalities that define OPG dominance.*

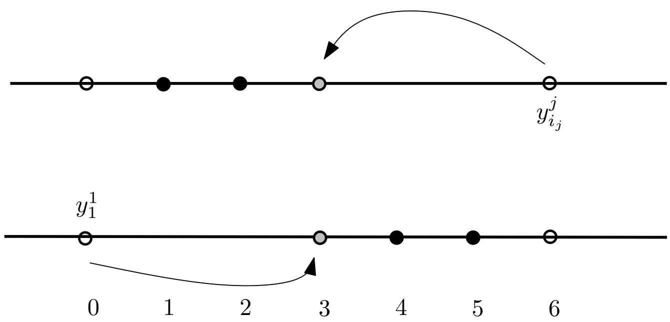


Figure 5: Half transfer

Proposition 1 shows that a BTPIT benefiting the poorest agent in category 1 in dominated distribution $\mathbf{y}$ can be performed in such a way that the distribution obtained after the transfer remains dominated by $\mathbf{x}$ as per the OPG criterion. However this proposition does not rule out the alternative possibility of performing an FIP between two individuals in such a way as to preserve the OPG dominance of the distribution obtained after performing this operation by $\mathbf{x}$. In the next proposition, we identify a circumstance which also entails the possibility of performing an FIP.

Proposition 2 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$*. Suppose that* $v_1^c > y_1^1$*. Let* $h_0 \in \{1, ..., k\}$ *be a category such that* $v_{h_0+1}^c < v_{h_0}^c = v_1^c$ *(with the convention that* $v_{k+1}^c = \underline{v}(\mathbf{x}, \mathbf{y})$*). Suppose also that*

- $\forall i \in \mathcal{N}(1)$*,* $x_i^1 > v_1^c$
- *For any category* h *such that* $h_0 \geq h \geq 2$*, we have* $y_i^h \neq v_1^c$*;*

Then there exists a distribution $\widehat{\mathbf{x}} \in \mathcal{D}(I)$ *such that* $\widehat{\mathbf{x}}$ *was obtained from* $\mathbf{y}$ *by an FIP and* $\mathbf{x} \succsim^{OPG} \widehat{\mathbf{x}}$*.*

We observe that the receiver of the FIP in Proposition 2 is not the poorest individual in category 1. It is another agent in category 1 whose income is equal to v_1^c. We now provide an example of a situation where both a BTPIT and an FIP are possible.

Example 4 *Let $k = 2$, $\mathcal{N}(1) = \{1,2,3\}$, $\mathcal{N}(2) = \{1,2,3,4\}$ and* $\mathbf{y}$ *and* $\mathbf{x}$ *be defined by*

$$y_1^1 = 2,\ y_2^1 = y_3^1 = 3,\ y_1^2 = 0,\ y_2^2 = y_3^2 = y_4^2 = 4$$

and:

$$x_1^1 = x_2^1 = x_3^1 = 4,\ x_1^2 = x_2^2 = x_3^2 = x_4^2 = 2.$$

As can be seen, $\mathbf{x} \succsim^{OPG} \mathbf{y}$, $v_1^c = v_2^c = 3$, and $y_1^1 = 2$. According to Proposition 1, agent 2 in category 2 can transfer some income to agent 1 in category 1. However, Proposition 2 states that an FIP is also possible. Indeed the conditions of this proposition are satisfied since $x_i^1 = 4 > v_1^c$ for all $i \in \mathcal{N}(1)$. We also have that $y_i^2 \neq 3$ for all $i \in \mathcal{N}(2)$. By virtue of Proposition 2, the distribution obtained after exchanging income 4 of agent 2 in category 2 with income 3 of agent 2 in category 1 remains dominated by distribution $\mathbf{x}$ *The situation is illustrated in the Figure below.*

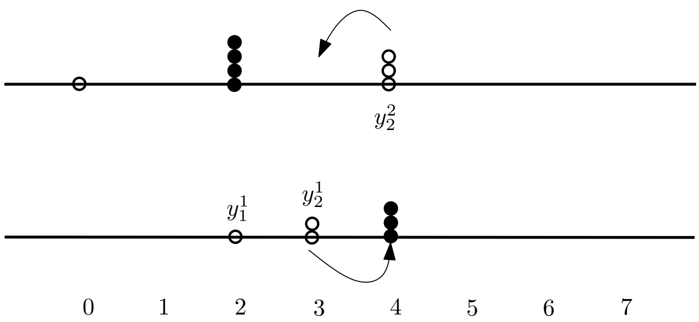


Figure 6: Agent 2 in category 2 can exchange income with agent 2 in category 1.

3.2.2 Case $(B) : v_1^c = y_1^1$

We start the analysis of this case by observing that requiring (5) to hold for all lists of poverty lines in the set $\mathcal{V}$ is equivalent to requiring this inequality to hold for the subset

$$\mathcal{V}' = \{(v_1, ..., v_k) \in \mathbb{R}^k : \overline{v}(\mathbf{x}, \mathbf{y}) \geq v_1 \geq v_2 \geq ... \geq v_k \geq \underline{v}(\mathbf{x}, \mathbf{y})\} \qquad (15)$$

of such lists of poverty lines, which is compact.

We also observe that, by the very definition of critical value v_1^c, there exists a sequence $\{\mathbf{w}^m\}$ of poverty lines vectors (with $\mathbf{w}^m \in \mathcal{V}'$ for every m) such that $P^{\mathbf{y}}(\mathbf{w}^m) - P^{\mathbf{x}}(\mathbf{w}^m) = 0$ and $w_1^m = y_1^1 + \varepsilon_1^m$, for $\varepsilon_1^m > 0$, and $\varepsilon_1^m \to 0$. By compactness of $\mathcal{V}'$ we can assume without loss of generality[6] that the sequence $\mathbf{w}^m$ of ordered poverty lines vectors converges to some limit $\overline{\mathbf{w}} \in \mathcal{V}'$. By continuity of the poverty gap function P, we have $P^{\mathbf{y}}(\overline{\mathbf{w}}) - P^{\mathbf{x}}(\overline{\mathbf{w}}) = 0$.

[6]Taking a subsequence if necessary.

Taking this limit vector $\overline{\mathbf{w}} \in \mathcal{V}'$ of ordered poverty lines, we first establish the existence, in initial distribution $\mathbf{y}$, of some agent in a category strictly larger than 1 with an income strictly larger than the lowest income observed in category 1. This agent will be a natural candidate for exchanging his/her higher income with that of the poorest agent in category 1. A crucial step for the identification of such an agent is the following lemma.

Lemma 7 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$. *Suppose that* $v_1^c = y_1^1$. *Then, there exists* $h_0 \geq 2$ *such that* $y_1^1 = \overline{w}_1 = \overline{w}_2 = ... = \overline{w}_{h_0} > \overline{w}_{h_0+1}$. *Moreover there exists* $g_0 \leq h_0$ *such that* $g_0 \geq 2$,

$$\sum_{h=1}^{g_0} \overline{p}^{\mathbf{y}}(h, y_1^1) = \sum_{h=1}^{g_0} \overline{p}^{\mathbf{x}}(h, y_1^1),$$

and:

$$\sum_{h=1}^{l} \overline{p}^{\mathbf{y}}(h, y_1^1) > \sum_{h=1}^{l} \overline{p}^{\mathbf{x}}(h, y_1^1)$$

for all $l < g_0$.

This lemma indeed identifies a category g_0 strictly larger than 1 in which a "potential donor" to the poorest agent in the worst category can be selected. As we now establish, this donor can transfer to the poorest agent in category 1 the whole income difference, while maintaining the dominance of distribution $\mathbf{x}$ over the distribution created by the FIP.

Proposition 3 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$ *such that* $\mathbf{x} \succsim^{OPG} \mathbf{y}$. *Suppose that* $v_1^c = y_1^1$. *Then there exists a distribution* $\widehat{\mathbf{x}} \in \mathcal{D}(I)$ *such that* $\widehat{\mathbf{x}}$ *was obtained from* $\mathbf{y}$ *by an FIP and* $\mathbf{x} \succsim^{OPG} \widehat{\mathbf{x}}$.

In the next example, we illustrate Proposition 3

Example 5 *Assume that* $k = 2$ *and* $\mathcal{N}(1) = \mathcal{N}(2) = \{1, 2\}$ *and consider distributions* $\mathbf{y}$ *and* $\mathbf{x}$ *defined respectively by:*

$$y_1^1 = 3, \; y_2^1 = 7 \; y_1^2 = 0, \; y_2^2 = 4,$$

and:

$$x_1^1 = 5, \; x_2^1 = 6, \; x_1^2 = 1, \; x_2^2 = 2.$$

It can be verified that $\mathbf{x} \succsim^{OPG} \mathbf{y}$ *and* $v_1^c = 3 = y_1^1$. *Proposition 3 states that an FIP between an agent in category* 2 *and agent* 1 *in category 1 is possible without breaking the OPG inequality. Indeed agent* $i_2 = 2$ *can exchange his income with agent* 1 *in category 1.*

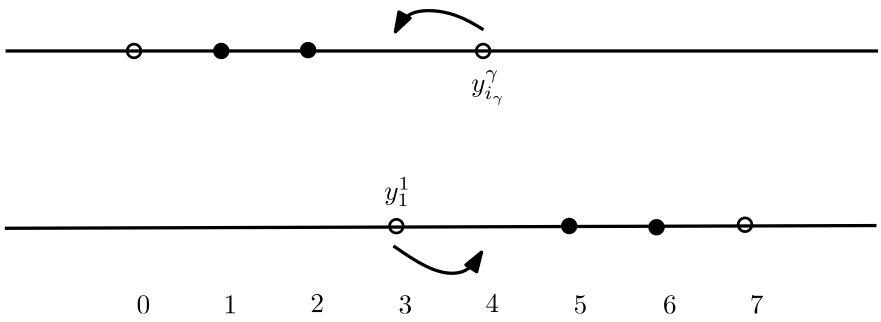


Figure 7: $v_1^c = y_1^1$; an FIP is possible

3.3 Proof of the main result $(iii) \Rightarrow (i)$

We now prove the last implication of Theorem 1.

Let $\mathbf{x}$ and $\mathbf{y}$ be as in Theorem 1. By a recursive argument on the finite set of agents, proving implication (iii) of Theorem 1 amounts to showing that $\mathbf{x} \succsim^{OPG} \mathbf{y}$ implies the possibility of going from $\mathbf{y}$ to some distribution $\overline{\mathbf{x}} \in \mathcal{D}(I)$ by a finite sequence of BTPIT and/or FIP in such a way that:

- $\mathbf{x} \succsim^{OPG} \overline{\mathbf{x}}$,
- there exists $h \in \{1, ..., k\}$ for which $x_i^h = \overline{x}_{\overline{i}}^h$ for some i and $\overline{i} \in \mathcal{N}(h)$.

Indeed, whenever one agent in one category has reached the income level that an agent of this category has in final distribution $\mathbf{x}$, we can remove that agent from that category and restart the procedure. Since the numbers of agents and categories are finite, this completes the proof. Let us therefore construct an algorithm for moving from $\mathbf{y}$ to some distribution $\overline{\mathbf{x}}$ as described above by a finite sequence of BTPIT and/or FIP. We construct this algorithm by first setting $\mathbf{x}(0) := \mathbf{y}$ and by recursively defining $\mathbf{x}(n+1)$ from $\mathbf{x}(n)$ in the following manner. Let $v_1^c(n)$ be the critical value defined as per (11) but applied to $\mathbf{x}(n)$ rather than to $\mathbf{y}$.

$(P1)$ If $v_1^c(n) = x_1^1(n)$ then perform an FIP, which is possible according to Proposition 3.

$(P2)$ If $v_1^c(n) > x_1^1(n)$ and if

- $\forall i \in \mathcal{N}(1)$, $x_i^1 > v_1^c(n)$,
- For any category h such that $h_0 \geq h \geq 2$, we have $x_i^h(n) \neq v_1^c(n)$,[7]

then perform an FIP, as described by Proposition 2 (remembering that the recipient of such an FIP is not the poorest individual in category 1 in that case).

(MT) otherwise perform the maximal transfer defined by Proposition 1 and Definition 6.

[7]Where h_0 is the category such that $v_1^c(n) = v_{h_0}^c(n) > v_{h_0+1}^c(n)$.

By construction, $\mathbf{x} \succsim^{OPG} \mathbf{x}(n)$ for any n. If there exists some $n^* \in \mathbb{N}_+$ such that, for some category $h \in \{1, ..., k\}$, we have:

$$x_i^h(n^*) = x_j^h$$

for some $i, j \in \mathcal{N}(h)$ then the algorithm ends and is said to be *finite*. If it does not end, then the algorithm generates an infinite (non-stationary) sequence. The only thing that remains to be proved is that the latter is impossible and that the algorithm is finite, as then $\overline{\mathbf{x}} := \mathbf{x}(n^*)$ satisfies the property stated above. We prove this by contradiction and therefore suppose that our algorithm generates an infinite sequence $(\mathbf{x}(n))_{n \in \mathbb{N}}$. We proceed by first establishing a series of claims (all proved in the Appendix).

Claim 1 *There exists some $n_0 \in \mathbb{N}$ such that, for all $n \geq n_0$, either case $(P2)$ or (MT) holds, so that $x_1^1(n) < v_1^c(n)$.*

An immediate consequence of this claim is that, for any $n \geq n_0$, we can also define the quantities $v_h^c(n)$ through expression (12).

Claim 2 *Let n_0 be the integer whose existence was established in Claim 1. Then, for all $n \geq n_0$ and all categories $h = 1, ..., k$, we have $v_h^c(n+1) \leq v_h^c(n)$.* [8]

In the next claim, we establish the existence of some step in the algorithm beyond which no FIP occurs.

Claim 3 *There exists $n_1 \in \mathbb{N}$ such that, for any $n \geq n_1$, the distribution $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ by means of a maximal transfer.*

We proved that, for any $n \geq n_1$, a maximal transfer of type (MT) occurs at time n. Since the algorithm is infinite, no transfer can be equalizing as per Definition 6. Hence, the maximal transfers at every step must be either a *breaking* or a *half* transfer of Definition 6. We next claim that at every step after n_1, if a breaking transfer is required by the algorithm, then the donor involved in the transfer will never be the donor again in a subsequent transfer.[9]

Claim 4 *There exists $n_2 \in \mathbb{N}$ such that, for any $n \geq n_2$, distribution $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ through a half transfer.*

[8]Note that the conclusion from this claim, that the critical value weakly decreases as n increases, is not true for operations of type (P1), which on the contrary necessarily weakly increase the critical value.

[9]While the proof of the claim is slightly cumbersome, the intuition behind it is relatively clear. Indeed, by its very definition, a breaking transfer is such that the donor cannot give more at this stage without breaking at least one of the OPG dominance inequalities. As n increases, the (ordered poverty gap) difference between distribution $\mathbf{x}(n)$ and distribution $\mathbf{x}$ gets smaller and smaller. Hence it becomes harder and harder to make a transfer without breaking some of the OPG inequalities.

We now establish that none of the donors involved in the half transfers that remain after all breaking transfers have been performed can be in category 1.

Claim 5 *For any $n \geq n_2$, the distribution $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ through a half transfer whose donor is not in category 1.*

We are now ready to establish a contradiction, and thus prove that the algorithm is finite. As proved in Claim 5, if the algorithm is infinite, there is some $n_2 \in \mathbb{N}$ such that, for $n \geq n_2$, $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ through a half transfer, the donor of which is not in category 1. Yet once an agent in category 1 has received a half transfer from an agent of a higher category, his/her income becomes equal to that of the donating agent . Hence, the donating agent cannot be selected again by the algorithm to donate to that same receiving agent. Since the number of agents is finite, this completes the proof. ■

There was a real risk of generating an infinite sequence of transfers starting from an OPG-dominated distribution $\mathbf{y}$ and going to a dominating distribution $\mathbf{x}$. This risk made us choose an FIP in the algorithm in the case labeled as $(P2)$ above, even though performing a BTPIT transfer would also be possible in that case under Proposition 1. The problem that might arise with a maximal BTPIT in case $(P2)$ is that of being trapped into performing an infinite sequence of maximal transfers, as illustrated by Example 4 above. If a BTPIT rather than an FIP were performed in this example, the maximal transfer would clearly be a half transfer of 1/2. Performing this transfer would yield the distribution $\widehat{\mathbf{x}}$ defined by:

$$\widehat{x}_1^1 = \widehat{x}_2^1 = 5/2,\ \widehat{x}_3^1 = 3,\ \widehat{x}_1^2 = 0,\ \widehat{x}_2^2 = \widehat{x}_3^2 = \widehat{x}_4^2 = 4$$

Note that the critical value $v_1^c(\widehat{\mathbf{x}})$ associated with $\widehat{\mathbf{x}}$ is still $3 > \widehat{x}_1^1 = 5/2$. Proposition 1 indicates that agent 3 in category 1 can make a transfer to one of the two poorest agents in that same category. The maximal transfer that agent 3 of category 1 can transfer to either one of the two poorest agents of category 1 is a half transfer of 1/4. If this transfer is performed, then distribution $\widehat{\widehat{\mathbf{x}}}$ is obtained, with $\widehat{\widehat{\mathbf{x}}}$ defined by:

$$\widehat{\widehat{x}}_1^1 = 5/2, \widehat{\widehat{x}}_2^1 = \widehat{\widehat{x}}_3^1 = 11/4,\ \widehat{\widehat{x}}_1^2 = 0,\ \widehat{\widehat{x}}_2^2 = \widehat{\widehat{x}}_3^2 = \widehat{\widehat{x}}_4^2 = 4$$

But from this distribution $\widehat{\widehat{\mathbf{x}}}$, the critical value $v_1^c(\widehat{\mathbf{x}})$ is 11/4 and this would have allowed a half transfer of 1/8 between either agent 2 or 3 in category 1 and the poorest agent 1 in this category and so on. Systematically resorting to the transfer allowed by Proposition 1 in that case would generate an infinite sequence of half transfers (with the "half" becoming smaller and smaller). It is to avoid this possibility that our algorithm imposes that the FIP allowed by Proposition 2 be performed every time the conditions of case $(P2)$ are verified.

4 Extensions

The analysis performed in the previous sections was restricted to distributions with the same number of agents in every category and the same total income. This restriction was motivated by our objective of defining pure equalization among unequals. If the number of agents in the various categories and/or the total income to be distributed vary from one distribution to the other, then inequality can not be the only criterion by which these distributions can be compared. However, the restriction of the analysis to distributions with the same number of agents in every category and the same total income is clearly limitative for practical applications. In what follow, we indicate how our analysis can be extended when the restrictions are removed. We develop in some detail the case where the number of agents in each category is allowed to vary across distributions, keeping constant the total number of agents (n), and the total income to be distributed. We then briefly indicate how these two last restrictions can be removed.

Denote by $\mathcal{N}_{\mathbf{x}}(h)$ the set of agents in category h for distribution $\mathbf{x}$ and by $n_{\mathbf{x}}(h) = \#\mathcal{N}_{\mathbf{x}}(h)$ the number of those agents. We write $\mathbf{x} = \{(x_1^h, ..., x_{n_{\mathbf{x}}(h)}^h)\}_{h=1}^{n_{\mathbf{x}}(h)}$. As mentioned, we maintain for the moment the assumption that

$$\sum_{h=1}^{k} n_{\mathbf{x}}(h) = n; \ \sum_{h=1}^{k} \sum_{i \in \mathcal{N}_{\mathbf{x}}(h)} x_i^h = I.$$

We call $\overline{\mathcal{D}}(I)$ the set of all such distributions. In order to allow for the number of agents in a given category to vary across distributions, we introduce the following elementary operation, that will be added to the FIP and BTPT discussed earlier.

Definition 7 (Categorical increments) *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\overline{\mathcal{D}}(I)$. *We say that* $\mathbf{x}$ *is obtained from* $\mathbf{y}$ *by means of a* categorical increment *if there are categories g and h for which $g < h$, $n_{\mathbf{x}}(g) = n_{\mathbf{y}}(g) - 1$, $n_{\mathbf{x}}(h) = n_{\mathbf{y}}(h) + 1$, as well as $i^g \in \mathcal{N}_{\mathbf{y}}(g)$ and $i^h \in \mathcal{N}_{\mathbf{x}}(h)$ such that:*

(i) $x_i^g = y_{i+1}^g$ *for all* $i \geq i^g$ *(if any).*

(ii) $x_{i^h}^h = y_{i^g}^g$.

(iii) $x_i^h = y_{i-1}^h$ *for all* $i \geq i^h + 1$ *(if any)*

(iv) $x_i^l = y_i^l$ *for any other pair* (i, l).

In words, $\mathbf{x}$ is obtained from $\mathbf{y}$ by means of a categorical increment if $\mathbf{x}$ and $\mathbf{y}$ only differ by the fact that one agent in some category in $\mathbf{y}$ reaches a superior category in $\mathbf{x}$ while keeping his/her income. The following result extends Theorem 1 to this new setting. Note that the different dominance criteria need to be adjusted in order to fit to the new notations. In particular we denote by $\mathcal{U}^{**}$ the class of type-dependent $(U^1, ..., U^k) \in \mathcal{U}^*$ satisfying the following additional property:

$$U^{h'}(w) \geq U^h(w) \ \ \forall h' \geq h. \tag{16}$$

Theorem 2 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\overline{\mathcal{D}}(I)$*. Then the following four statements are equivalent.*

(i) *It is possible to go from* $\mathbf{y}$ *to* $\mathbf{x}$ *by a finite sequence of categorical increments, BTPIT and/or FIP.*

(ii) $\mathbf{x}$ *utilitarian dominates* $\mathbf{y}$ *for the class* $\mathcal{U}^{**}$*,*

(iii) $\mathbf{x} \succsim^{OPG} \mathbf{y}$*:*

$$\sum_{h=1}^{k} \sum_{i \in \mathcal{N}_{\mathbf{x}}(h)} \max(v_h - x_i^h, 0) \leq \sum_{h=1}^{k} \sum_{i \in \mathcal{N}_{\mathbf{y}}(h)} \max(v_h - y_i^h, 0), \;\; \forall (v_1, ..., v_k) \in \mathcal{V} \tag{17}$$

(iv) $\mathbf{x} \succsim^{CLI} \mathbf{y}$*, where the set of eligible definitions of the* m *poorest agents now writes*

$$\Pi(\mathbf{x}, m) = \left\{ (r_1, ..., r_k) : 0 \leq r_h \leq n_{\mathbf{x}}(h), \; \sum_h r_h = m, \; x^h_{r_h+1} > x^{h'}_{r_{h'}} \; \forall h' > h \right\}.$$

Remark 1 *Observe that if* $\mathbf{x}$ *dominates* $\mathbf{y}$ *with respect to any of the four criteria then we necessarily have*

$$\sum_{h=1}^{g} n_{\mathbf{y}}(h) \geq \sum_{h=1}^{g} n_{\mathbf{x}}(h), \;\; \forall g = 1, ..., k.$$

Hence, first order stochastic dominance of its marginal distribution of needs is a necessary condition for $\mathbf{x}$ *to dominate* $\mathbf{y}$ *for any of the four criteria.*

Remark 2 *It is worth noticing that, when the number of agents in the different categories is allowed to vary across distributions, OPG dominance is not equivalent to inequality (17) for lists of ordered poverty lines of the subset* $\mathcal{V}'$ *defined in (15). Consider the following simple example involving 2 categories and 2 individuals (all in category 1 in distribution* $\mathbf{x}$ *and split between the two categories in distribution* $\mathbf{y}$*). Incomes are:* $x_1^1 = x_2^1 = 1$ *and* $y_1^1 = 0, y_1^2 = 2$*. Then one can easily check that* $\mathbf{x}$ *OPG-dominates* $\mathbf{y}$ *if we restrict the poverty lines to* $2 \geq v_1 \geq v_2 \geq 0$*. Nevertheless* $\mathbf{x}$ *does not dominate* $\mathbf{y}$ *for the criteria* $(i), (ii)$ *and* (iv)*. This non-dominance can be seen by considering, say, a poverty line of* 3 *in category* 1 *and* 0 *in category* 2*.*

We now briefly discuss how the results extend to the case where n (the total number of agents) is the same but where the total income to be distributed varies. For this sake, we need to introduce another elementary transformation, which we call *income increment.* This transformation is defined as follows.

Definition 8 (Income increments) *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\overline{\mathcal{D}}(I)$*. We say that* $\mathbf{x}$ *is obtained from* $\mathbf{y}$ *by means of an* income increment *if* $\mathcal{N}_{\mathbf{x}}(g) = \mathcal{N}_{\mathbf{y}}(g)$ *for any* g*, and there exists a category* h*,* $i^h \leq j^h \in \mathcal{N}(h)$ *and* $a > 0$ *such that:*

(i) $x_i^h = y_{i+1}^h$ *for* $i = i^h ... j^h - 1$ *(if any).*

(ii) $x_{j_h}^h = y_{i_h}^h + a$.

(iii) $x_i^l = y_i^l$ *for any other pair* (i, l).

In words, $\mathbf{x}$ is obtained from $\mathbf{y}$ by means of an income increment if $\mathbf{x}$ and $\mathbf{y}$ only differ by the fact that one agent in some category has an higher income in distribution $\mathbf{x}$ than in distribution $\mathbf{y}$.
We must also enlarge the class of utility functions over which utilitarian unanimity is looked for to the following class: let $\mathcal{U}^{***}$ be the class of type-dependent $(U^1, ..., U^k) \in \mathcal{U}^{**}$ satisfying the following additional property:

$$U^h(w + a) \geq U^h(w) \ \ \forall h, \ \ \forall a \geq 0. \tag{18}$$

We can then state the following theorem whose proof, which can easily be worked out using standard results in one-dimensional dominance analysis and the fact that the proof of the equivalence between (iii) and (iv) in Theorem 1 does not assume anything about the sum of income that is distributed, is omitted.

Theorem 3 *Let* $\mathbf{x} \in \overline{\mathcal{D}}(I)$ *and* $\mathbf{y} \in \overline{\mathcal{D}}(J)$. *Then the following four statements are equivalent.*

(i) *It is possible to go from* $\mathbf{y}$ *to* $\mathbf{x}$ *by a finite sequence of income increments, categorical increments, BTPIT and/or FIP.*

(ii) $\mathbf{x}$ *utilitarian dominates* $\mathbf{y}$ *for the class* $\mathcal{U}^{***}$.

(iii) $\mathbf{x} \succsim^{OPG} \mathbf{y}$.

(iv) $\mathbf{x} \succsim^{CLI} \mathbf{y}$.

Finally, the case where the total number of agents vary across distributions is handled by appealing to the so-called Dalton principle of population replication (see e.g. Dalton (1920)). In the present setting, this principle says that replicating any finite number of time a distribution of income among agents in different categories is a matter of social indifference.

5 Conclusion

In this paper, we provide a workable definition of "income equalization" when performed between agents who are vertically differentiated with respect to some other characteristic. The definition of such equalization is transferring from a richer and more highly ranked agent to a poorer and less highly ranked agent an amount of income that does not exceed the income difference between the two agents. If the transfer does not exceed half the income difference between the donor and the receiver, then such a transfer is called a BTPIT. If the transfer is larger than half the income difference, then the transfer can be viewed as a combination of a

BTPIT of less than half the income difference and an FIP. The paper has identified the normative foundations of this notion of equalization. Specifically, it has shown that the smallest transitive ranking of distributions consistent with this notion of equalization is the unanimity of all utilitarian planners' rankings considering that the marginal utility of income for every agent is decreasing with respect to both income and type. The paper has also identified two empirically implementable criteria - the Ordered Poverty gap criterion, and the cumulative lowest incomes criterion - each of which is equivalent to this notion of equalization. While Gravel and Moyes (2012) showed that the ordered poverty gap dominance of one distribution over another is equivalent to the possibility of going from a phantom-augmented dominated distribution to the phantom-augmented dominating one by a finite sequence of Pigou-Dalton transfers (between agents of a given type) and/or FIP, these authors could not establish the equivalence without resorting to dummy or phantom agents. The present paper is therefore, to the best of our knowledge, the first to provide an equivalence between a notion of normative dominance, an elementary principle of equalization, and two empirically implementable criteria that applies to distributions of a cardinally meaningful attribute among vertically differentiated agents.

It is our hope that the implementable criteria that we have justified in this fashion - the Bourguignon (1989) criterion and its equivalent formulation in the form of the cumulative smallest income dominance criterion - will be used with increasing confidence by practitioners evaluating inequalities among vertically differentiated agents.

6 Appendix

6.1 Proof of $(i) \Rightarrow (ii) \Rightarrow (iii)$

Proposition 4 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$*. Then, in Theorem 1, Statement (i) implies Statement (ii) and Statement (ii) implies statement (iii).*

Proof. $(i) \Rightarrow (ii)$ We must prove that both BTPIT and FIP increase the sum of utilities for any collection of utility functions $\{U^h\}_{h=1}^k \in \mathcal{U}^*$.

BTPIT: Assume that $\mathbf{x}$ has been obtained from $\mathbf{y}$ by a BTPIT. Then, using Definition 1, there are categories g and h satisfying $g \leq h$, agents $i^g \in \mathcal{N}(g)$ and $i^h \in \mathcal{N}(h)$ satisfying $y_{i^g}^g < y_{i^h}^h$ and a number $\alpha \in [0, (y_{i^h}^h - y_{i^g}^g)/2]$ for which we have:

$$\begin{aligned}
& \sum_{j=1}^{k}\sum_{i=1}^{n(j)}(U^j(x_i^j) - U^j(y_i^j)) \\
= \quad & U^g(x_{r_+^g(\alpha)}^g) - U^g(y_{i^g}^g) + U^h(x_{r_-^h(\alpha)}^h) - U^h(y_{i^h}^h) \\
= \quad & U^g(y_{i^g}^g + \alpha) - U^g(y_{i^g}^g) - [U^h(y_{i^h}^h) - U^h(y_{i^h}^h - \alpha)] \\
\geq \quad & 0 \text{ (if the functions } U^1, ..., U^k \text{ belong to } \mathcal{U}^*)
\end{aligned}$$

FIP: Assume that $\mathbf{x}$ has been obtained from $\mathbf{y}$ by an FIP. Then, using Definition 2, there are categories g and h satisfying $g < h$, agents $i^g \in \mathcal{N}(g)$ and $i^h \in \mathcal{N}(h)$ satisfying $y^g_{i^g} < y^h_{i^h}$ for which we have:

$$\begin{aligned}
\sum_{j=1}^{k}\sum_{i=1}^{n(j)}(U^j(x_i^j) - U^j(y_i^j)) &= U^g(x^g_{r^g(i^h)}) - U^g(y^g_{i^g}) + U^h(x^h_{r^h(i^g)}) - U^h(y^h_{i^h}) \\
&= U^g(y^h_{i^h}) - U^g(y^g_{i^g}) + U^h(y^g_{i^g}) - U^h(y^h_{i^h}) \\
&= U^g(y^h_{i^h}) - U^g(y^g_{i^g}) - [U^h(y^h_{i^h}) - U^h(y^g_{i^g})] \\
&\geq 0 \text{ (if the functions } U^1, ..., U^k \text{ belong to } \mathcal{U}^*)
\end{aligned}$$

Repeating the arguments (for the FIP and/or the BTPIT) for any finite sequence of distributions of income completes the proof of the first implication for the theorem.

$(ii) \Rightarrow (iii)$. Let $\mathbf{x}$ and $\mathbf{y}$ be two distributions in $\mathcal{D}(I)$ for which the inequality:

$$\sum_{h=1}^{k}\sum_{i=1}^{n(t)} U^h(x_i^h) - \sum_{h=1}^{k}\sum_{i=1}^{n(t)} U^h(y_i^h) \geq 0 \tag{19}$$

holds for all lists of utility functions $\{U^h\}_{h=1}^k$ in $\mathcal{U}^*$. Choose any vector $\mathbf{v} = (v_1, ..., v_k)$ in the set $\mathcal{V}$ and define the k functions $U^h : \mathbb{R} \longrightarrow \mathbb{R}$ (for $h = 1, ..., k$) by $U^h(w) = \min[w - v_h, 0]$. Let us show that the collection of k functions $\{U^h\}_{h=1,...,k}$ satisfies inequality (3) for any vector $\mathbf{v} = (v_1, ..., v_k)$ in $\mathcal{V}$, and therefore belongs to $\mathcal{U}^*$. Consider any $u \geq 0$, $w \leq w'$ and $h \leq h'$. First note that the quantities $U^h(w + u) - U^h(w)$ and $U^{h'}(w' + u) - U^{h'}(w')$ belong to $[0, u]$.

If $w \geq v_h$ then $w + u \geq v_h$ and $w' + u \geq w' \geq v_{h'}$. Thus (3) holds with both sides equal to zero.

If $w \leq v_h$ then $U^h(w) = w - v_h$ and $U^h(w + u) - U^h(w) = \min(u, v_h - w)$. Note also that:

$$U^{h'}(w' + u) - U^{h'}(w') \leq -U^{h'}(w') \leq v_{h'} - w' \leq v_h - w.$$

Hence $U^{h'}(w' + u) - U^{h'}(w') \leq \min(u, v_h - w)$ and inequality (3) holds for that case also. We have therefore proved that the collection of functions $\{U^h\}_{h=1,...,k}$ belongs to the class $\mathcal{U}^*$ for all $\mathbf{v} = (v_1, ..., v_k) \in \mathcal{V}$. Since inequality (19) holds for all such functions so that we have:

$$\begin{aligned}
\sum_{h=1}^{k}\sum_{i=1}^{n(h)} \min[x_i^h - v_h, 0] &\geq \sum_{h=1}^{k}\sum_{i=1}^{n(h)} \min[y_i^h - v_h, 0] \\
&\Longleftrightarrow \\
\sum_{h=1}^{k}\sum_{i=1}^{n(h)} \max[v_h - x_i^h, 0] &\leq \sum_{h=1}^{k}\sum_{i=1}^{n(h)} \max[v_h - y_i^h, 0]
\end{aligned}$$

for all $\mathbf{v} = (v_1, ..., v_k) \in \mathcal{V}$, as required by the OPG criterion. ■

6.2 Proof of $(iii) \Leftrightarrow (iv)$

Proposition 5 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be two distributions in* $\mathcal{D}(I)$. *Then* $\mathbf{x} \succsim^{OPG} \mathbf{y}$ *if and only if* $\mathbf{x} \succsim^{CLI} \mathbf{y}$.

Proof. Assume first that $\mathbf{x} \succsim^{CLI} \mathbf{y}$, and consider any $\mathbf{v} \in \mathcal{V}$. We must show that $P^{\mathbf{x}}(\mathbf{v}) \leq P^{\mathbf{y}}(\mathbf{v})$. Let the vector of ranks $(r_1, ..., r_k)$ be defined, for any $h \in \{1, ..., k\}$ by:

$$\begin{aligned} r_h &= \max\{i \in \mathcal{N}(h) : x_i^h \leq v_h\}, \text{ if } \{i \in \mathcal{N}(h) : x_i^h \leq v_h\} \neq \varnothing \\ &= 0 \text{ otherwise} \end{aligned}$$

Clearly, $r_h \in \{0, ..., n_h\}$ for any h. Let $m = \sum_h r_h$. We first show that $(r_1, ..., r_k) \in \Pi(\mathbf{x}, m)$. By contradiction, assume that $(r_1, ..., r_k) \notin \Pi(\mathbf{x}, m)$. By definition of $\Pi(\mathbf{x}, m)$ (Expression (7)), there exist categories h and h' satisfying $h' > h$ such that $x_{r_h+1}^h \leq x_{r_{h'}}^{h'}$. By definition of r_h, we have that $x_{r_{h'}}^{h'} \geq x_{r_h+1}^h > x_{r_h}^h \geq \underline{v}(\mathbf{x}, \mathbf{y})$ (at least under the convention that $x_0^h = \underline{v}(\mathbf{x}, \mathbf{y})$). This means that $x_{r_{h'}}^{h'} = v^{h'} > v^h = x_{r_h}^h$ (if $r_h \geq 1$) and $x_{r_{h'}}^{h'} = v^{h'} > v^h$ (if $r_h = 0$ and $v^h < x_1^h$). But this contradicts the fact that $\mathbf{v} \in \mathcal{V}$. Observe that:

$$P^{\mathbf{x}}(\mathbf{v}) = \sum_{h=1}^{k} r_h v_h - \sum_{h=1}^{k} \sum_{i \leq r_h} x_i^h$$

Let $(r'_1, ..., r'_k) \in \Pi(\mathbf{y}, m)$ be any vector of ranks such that $\sum_{g=1}^{h} r'_g \geq \sum_{g=1}^{h} r_g$ for any h and $\sum_{h=1}^{k} \sum_{i \leq r_h} x_i^h \geq \sum_{h=1}^{k} \sum_{i \leq r'_h} y_i^h$. Such a vector of ranks exists because $\mathbf{x} \succsim^{CLI} \mathbf{y}$. We then have:

$$\begin{aligned} P^{\mathbf{y}}(\mathbf{v}) &= \sum_{h=1}^{k} \sum_{i \in N(h)} \max\{0, v_h - y_i^h\} \geq \sum_{h=1}^{k} \sum_{i \leq r'_h} \max\{0, v_h - y_i^h\} \\ &\geq \sum_{h=1}^{k} \sum_{i \leq r'_h} (v_h - y_i^h) = \sum_{h=1}^{k} r'_h v_h - \sum_{h=1}^{k} \sum_{i \leq r'_h} y_i^h. \end{aligned}$$

Note that

$$\sum_{h=1}^{k} r'_h v_h = \sum_{h=1}^{k} r_h v_h + \sum_{h=1}^{k-1} (v_h - v_{h+1}) (\sum_{g=1}^{h} r'_g - \sum_{g=1}^{h} r_g) \geq \sum_{h=1}^{k} r_h v_h.$$

(because $\sum_{g=1}^{h} r'_g \geq \sum_{g=1}^{h} r_g$ and $(v_h - v_{h+1}) \geq 0$ for any h). It follows that $P^{\mathbf{y}}(\mathbf{v}) \geq P^{\mathbf{x}}(\mathbf{v})$, as required.

Assume now that $\mathbf{x} \succsim_{OPG} \mathbf{y}$. Take any integer $m \in \{1, ..., n\}$ and any $(r_1, ..., r_k) \in \mathbf{\Pi}(\mathbf{x}, m)$. Define the set $F(\mathbf{y}, r_1, ..., r_k)$ by:

$$F(\mathbf{y}, r_1, ..., r_k) = \left\{ (i_1, ..., i_k) \in \Pi(\mathbf{y}, m) : \sum_{g=1}^{h} i_g \geq \sum_{g=1}^{h} r_g \, \forall h, \; \sum_{g=1}^{k} i_g = m \right\}$$

This set is non-empty. Indeed, if $m \in \{\sum_{g=1}^{h-1} n_g + 1, ..., \sum_{g=1}^{h} n_g\}$ for some $h = 1, ..., k$, then the combination of ranks $(i_1, ..., i_k) := (n_1, n_2, ..., n_{h-1}, m - \sum_{g=1}^{h-1} n_g, 0, ..., 0)$ belongs to $\Pi(\mathbf{y}, m)$ by construction. Moreover, we have

$$\sum_{g=1}^{\widehat{h}} i_g = \sum_{g=1}^{\widehat{h}} n_g \geq \sum_{g=1}^{\widehat{h}} r_g \;\; \forall \, \widehat{h} = 1, ..., h-1$$

and

$$\sum_{g=1}^{\widehat{\widehat{h}}} i_g = m \geq \sum_{g=1}^{\widehat{\widehat{h}}} r_g \; \forall \; \widehat{\widehat{h}} = h, ..., k.$$

Hence $(i_1, ..., i_k) \in F(\mathbf{y}, r_1, ..., r_k)$. The non-empty set $F(\mathbf{y}, r_1, ..., r_k)$ thus contains all admissible (as per set Π) lists of ranks for distribution $\mathbf{y}$ that satisfy Inequality (9) of Definition 5 *vis-à-vis* the list of ranks $(r_1, ..., r_k)$. Define now the vector of ranks $(r'_1, ..., r'_k) \in \{0, ..., n_1\} \times ... \times \{0, ..., n_k\}$ by:

$$(r'_1, ..., r'_k) \in \underset{(i_1,...,i_k) \in F(\mathbf{y}, r_1,...,r_k)}{\arg\min} \sum_{h=1}^{k} \sum_{i \leq i_h} y_i^h, \tag{20}$$

Hence, the list of ranks $(r'_1, ..., r'_k)$ are those that define a set of m "poorest" agents in situation $\mathbf{y}$, under the constraint that $\sum_{g=1}^{h} r'_g \geq \sum_{g=1}^{h} r_g$ for any h. Consider now the vector of poverty lines $\mathbf{v}$ defined as follows (for any $h = 1, ..., k$)[10]:

$$v_h := \max \left\{ \{y_{r_l}^l : l \geq h\} \cup \{y_{r_g}^g : g < h \text{ and } \sum_{e=1}^{j} r'_e > \sum_{e=1}^{j} r_e \; \forall j = g, ..., h-1\} \right\}.$$

This vector $\mathbf{v}$ is well-defined because, for every h, at least one of sets $\{y_{r_l}^l : l \geq h\}$ or $\{y_{r_g}^g : g < h \text{ and } \sum_{e=1}^{j} r'_e > \sum_{e=1}^{j} r_e \; \forall j = g, ..., h-1\}$ is not empty. Let us first check that $\mathbf{v} \in \mathcal{V}$. This amounts to showing that if h and h' are two categories such that $h < h'$, then $v_h \geq v_{h'}$. But this is an immediate consequence of the fact that, for any h and h' such that $h < h'$, set $\{y_{r_l}^l : l \geq h\} \cup \{y_{r_l}^g : g < h \text{ and } \sum_{e=1}^{j} r'_e > \sum_{e=1}^{j} r_e \; \forall j = g, ..., h-1\}$ contains set $\{y_{r_l}^l : l \geq h'\} \cup \{y_{r_l}^g : g < h' \text{ and } \sum_{e=1}^{j} r'_e > \sum_{e=1}^{j} r_e \; \forall j = g, ..., h'-1\}$ as a subset. As a result, the maximum taken over the larger set cannot be smaller than the maximum taken over the subset. We now prove that, for any $h = 1, ..., k$, we have $v_h \in [y_{r_h}^h, y_{r_h+1}^h]$. By definition of v_h, the only inequality that needs to be established is $v_h \leq y_{r_h+1}^h$. By contradiction, assume that $v_h > y_{r_h+1}^h$. Since $(r'_1, ..., r'_k) \in \Pi(\mathbf{y}, m)$, we have $y_{r'_j}^j < y_{r'_g+1}^g$ for every two categories g and j such that $g < j$. Hence the fact that $v_h > y_{r_h+1}^h$ may only be due to the existence of a category $g < h$ such that $\sum_{e=1}^{l} r'_e > \sum_{e=1}^{l} r_e \; \forall l = g, ..., h-1$ and $y_{r'_g}^g > y_{r'_h+1}^h$. Consider then the list of ranks $(r''_1, ..., r''_k) \in \{0, ..., n_1\} \times ... \times \{0, ..., n_k\}$ defined by $r''_e = r'_e$ for $e \neq g, h$, $r''_h = r'_h + 1$ and $r''_g = r'_g - 1$ Observe that $(r''_1, ..., r''_k) \in F(y, r_1, ..., r_k)$. Now:

$$\sum_{h=1}^{k} \sum_{i \leq r''_h} y_i^h = \sum_{h=1}^{k} \sum_{i \leq r'_h} y_i^h + y_{r'_h+1}^h - y_{r'_g}^g < \sum_{h=1}^{k} \sum_{i \leq r'_h} y_i^h \tag{21}$$

(because $y_{r'_g}^g > y_{r'_h+1}^h$). But this contradicts the definition of $(r'_1, ..., r'_k)$ provided by (20). Since $y_{r'_h+1}^h \geq v_h \geq y_{r'_h}^h$ for all h, we have:

$$P^{\mathbf{y}}(\mathbf{v}) = \sum_{h=1}^{k} \sum_{i \leq r'_h} (v_h - y_i^h) = \sum_{h=1}^{k} r'_h v_h - \sum_{h=1}^{k} \sum_{i \leq r'_h} y_i^h.$$

[10]Again with the convention that, if $r'_h = 0$, then $y_{r'_h}^h = \underline{v}(\mathbf{x}, \mathbf{y})$.

and:

$$\begin{aligned} P^{\mathbf{x}}(\mathbf{v}) &= \sum_{h=1}^{k}\sum_{i\in\mathcal{N}(h)} \max\{0, v_h - x_i^h\} \geq \sum_{h=1}^{k}\sum_{i\leq r_h} \max\{0, v_h - x_i^h\} \\ &\geq \sum_{h=1}^{k}\sum_{i\leq r_h}(v_h - x_i^h) = \sum_{h=1}^{k} r_h v_h - \sum_{h=1}^{k}\sum_{i\leq i_h} x_i^h. \end{aligned}$$

Now, as established above:

$$\sum_{h=1}^{k} r'_h v_h = \sum_{h=1}^{k} r_h v_h + \sum_{h=1}^{k-1}(v_h - v_{h+1})(\sum_{g=1}^{h} r'_g - \sum_{g=1}^{h} r_g).$$

Let $h \leq k-1$. By construction of the vector of poverty line $\mathbf{v}$, if $\sum_{g=1}^{h} r'_g > \sum_{g=1}^{h} r_g$ then $v_h = v_{h+1}$. Indeed let g_0 be the category such that $\sum_{e=1}^{g} r'_e > \sum_{e=1}^{g} r_e$ for $g = g_0, ..., h$ and $\sum_{e=1}^{g_0-1} r'_e = \sum_{e=1}^{g_0-1} r_e$. Then $v_h = v_{h+1} = \max_{g\geq g_0} y_{i_g}^g$. It follows that $\sum_{h=1}^{k} r'_h v_h = \sum_{h=1}^{k} r_h v_h$. By assumption, $P^{\mathbf{y}}(\mathbf{v}) \geq P^{\mathbf{x}}(\mathbf{v})$. Hence

$$\sum_{h=1}^{k}\sum_{i\leq r'_h} y_i^h = \sum_{h=1}^{k} r'_h v_h - P^{\mathbf{y}}(\mathbf{v}) \leq \sum_{h=1}^{k} r_h v_h - P^{\mathbf{x}}(\mathbf{v}) \leq \sum_{h=1}^{k}\sum_{i\leq r_h} x_i^h,$$

which proves the result. ■

6.3 Proof of the results of section 3.1

6.3.1 Lemma 1.

For the first statement, assume by contraposition that $y_1^1 > x_1^1$. Consider then the vector of poverty lines $(y_1^1, \underline{v}(\mathbf{x},\mathbf{y}), ..., \underline{v}(\mathbf{x},\mathbf{y})) \in \mathcal{V}$. One has:

$$\begin{aligned} P^{\mathbf{y}}(y_1^1, \underline{v}(\mathbf{x},\mathbf{y}), ..., \underline{v}(\mathbf{x},\mathbf{y})) &= 0 \text{ and:} \\ P^{\mathbf{x}}(y_1^1, \underline{v}(\mathbf{x},\mathbf{y}), ..., \underline{v}(\mathbf{x},\mathbf{y})) &\geq y_1^1 - x_1^1 > 0 \end{aligned}$$

so that $x \succsim^{OPG} y$ does not hold, as required. The second statement holds by a mirror argument. ■

6.3.2 Lemma 2.

For any type $\overline{h} = 1, ..., k$, the vector of poverty lines $v^{\overline{h}} \in D^k$ defined by:

$$\mathbf{v}^{\overline{h}} = (\underbrace{\overline{v}(\mathbf{x},\mathbf{y}), ..., \overline{v}(\mathbf{x},\mathbf{y})}_{\overline{h}}, \underbrace{\underline{v}(\mathbf{x},\mathbf{y}), ..., \underline{v}(\mathbf{x},\mathbf{y})}_{k-\overline{h}})$$

clearly belongs to $\mathcal{V}$. Hence, since $x \succsim^{OPG} y$, we have:

$$
\begin{array}{ccc}
P^{\mathbf{x}}(\mathbf{v}^{\overline{h}}) & \leq & P^{\mathbf{y}}(\mathbf{v}^{\overline{h}}) \\
& \Longleftrightarrow & \\
\sum_{h=1}^{\overline{h}} n(h)\overline{v}(\mathbf{x},\mathbf{y}) - \sum_{h=1}^{\overline{h}} \sum_{i\in\mathcal{N}(h)} x_i^h & \leq & \sum_{h=1}^{\overline{h}} n(h)\overline{v}(\mathbf{x},\mathbf{y}) - \sum_{h=1}^{\overline{h}} \sum_{i\in\mathcal{N}(h)} y_i^h \\
& \Longleftrightarrow & \\
\sum_{h=1}^{\overline{h}} \sum_{i\in\mathcal{N}(h)} x_i^h & \geq & \sum_{h=1}^{\overline{h}} \sum_{i\in\mathcal{N}(h)} y_i^h \quad \blacksquare
\end{array}
$$

6.3.3 Lemma 3.

By inequality (10) there exists at least one agent with income strictly larger than $y_{i_1}^1$ in one of the categories $\{2, ..., g_0\}$. Define $\gamma \in \{2, ..., g_0\}$ and $i_\gamma \in \mathcal{N}(\gamma)$ by

$$\gamma = \min\{g \geq 2 : \exists i \in \mathcal{N}(g) \text{ such that } y_i^g > y_{i_1}^1\},\ i_\gamma = \min\{i \in \mathcal{N}(\gamma) : y_i^\gamma > y_{i_1}^1\}$$

We now prove that for any $\mathbf{v} \in \mathcal{V}$, we have:

$$P^{\mathbf{y}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) \geq \min\{y_{i_\gamma}^\gamma, v_1\} - \max\{y_{i_1}^1, v_\gamma\}; \tag{22}$$

If either $v_1 \leq y_{i_1}^1$ or $v_\gamma \geq y_{i_\gamma}^\gamma$, inequality (22) trivially holds (because $\min\{y_{i_\gamma}^\gamma, v_1\} - \max\{y_{i_1}^1, v_\gamma\} \leq 0$ in this case). Hence we suppose that $v_1 > y_{i_1}^1$ and $v_\gamma < y_{i_\gamma}^\gamma$. We establish the result by considering three different cases.

Case (i): $y_{i_1}^1 \leq v_{g_0} \leq v_1 \leq y_{i_\gamma}^\gamma$.
By definition of $y_{i_\gamma}^\gamma$, we have that that:

$$\overline{p}^{\mathbf{y}}(h, w) = \overline{p}^{\mathbf{y}}(h, w')$$

for $h = 2, ..., g_0$ and any w and $w' \in [y_{i_1}^1, y_{i_\gamma}^\gamma[$. Indeed, the number of poor in categories $2, .., g_0$ at distribution y does not change when we move the poverty line applicable to all these categories from $y_{i_1}^1$ to $y_{i_\gamma}^\gamma$). Combining this with inequality (10) and the fact that $\overline{p}^{\mathbf{x}}(h,w)$ is non-decreasing with respect to w, we have:

$$\begin{aligned}
& \sum_{g=2}^{g_0}\sum_{i\in\mathcal{N}(g)}[\max(v_1-y_i^g,0)-\max(v_1-x_i^g,0)] \\
\leq & \sum_{g=2}^{g_0}\sum_{i\in\mathcal{N}(g)}[\max(v_g-y_i^g,0)-\max(v_g-x_i^g,0)] \\
+ & \sum_{g=2}^{g_0}[\overline{p}^{\mathbf{y}}(g,v_g)-\overline{p}^{\mathbf{x}}(g,v_g)](v_1-v_g) \\
= & \sum_{g=2}^{g_0}\sum_{i\in\mathcal{N}(g)}[\max(v_g-y_i^g,0)-\max(v_g-x_i^g,0] \\
+ & \sum_{h=2}^{g_0}\sum_{g=h}^{g_0}[\overline{p}^{\mathbf{y}}(g,v_g)-\overline{p}^{\mathbf{x}}(g,v_g)](v_{h-1}-v_h) \\
\leq & \sum_{g=2}^{g_0}\sum_{i\in\mathcal{N}(g)}[\max(v_g-y_i,0)-\max(v_g-x_i,0]+v_{g_0}-v_1
\end{aligned}$$

Hence:

$$\begin{aligned}
P^{\mathbf{y}}(\mathbf{v})-P^{\mathbf{x}}(\mathbf{v}) &= P^{\mathbf{y}}(v_1,...,v_1,v_{g_0+1},...,v_k)-P^{\mathbf{x}}(v_1,...,v_1,v_{g_0+1},...,v_k) \\
&\quad -\sum_{g=2}^{g_0}\sum_{i\in\mathcal{N}(g)}[\max(v_1-y_i^g,0)-\max(v_1-x_i^g,0)] \\
&\quad +\sum_{g=2}^{g_0}\sum_{i\in\mathcal{N}(g)}[\max(v_g-y_i^g,0)-\max(v_g-x_i^g,0)] \\
&\geq P^{\mathbf{y}}(v_1,...,v_1,v_{g_0+1},...,v_k) \\
&\quad -P^{\mathbf{x}}(v_1,...,v_1,v_{g_0+1},...,v_k)+v_1-v_{g_0} \\
&\geq v_1-v_{g_0} \text{ (because } (v_1,...,v_1,v_{s_0+1},...,v_k)\in\mathcal{V}) \\
&\geq v_1-v_\gamma \\
&= \min\{y_{i_\gamma}^\gamma,v_1\}-\max\{y_{i_1}^1,v_\gamma\}.
\end{aligned}$$

as required.

Case (ii): $v_1>y_{i_\gamma}^\gamma$ and $v_{g_0}\geq y_{i_1}^1$.
In this case, there exists some $\underline{h}\in\{1,...,\gamma-1\}$ such that $v_1\geq...\geq v_{\underline{h}}>y_{i_\gamma}^\gamma\geq v_{\underline{h}+1}\geq ...\geq v_k$. Let $\tilde{v}=(y_{i_\gamma}^\gamma,...,y_{i_\gamma}^\gamma,v_{\underline{h}+1},...,v_k)$. Then $\tilde{v}$ belongs to case (i) and, consequently:

$$P^{\mathbf{y}}(\tilde{\mathbf{v}})-P^{\mathbf{x}}(\tilde{\mathbf{v}})\geq\tilde{v}_1-\tilde{v}_\gamma=y_{i_1}^1-v_\gamma.$$

Moreover denoting $\hat{v}:=(v_1,...,v_{\underline{h}},v_{\underline{h}+1}^c,...,v_k^c)$, we have, by definition of v^c :

$$\begin{aligned}
(P^{\mathbf{y}}(\tilde{\mathbf{v}})-P^{\mathbf{x}}(\tilde{\mathbf{v}}))-(P^{\mathbf{y}}(\mathbf{v})-P^{\mathbf{x}}(\mathbf{v})) &= (P^{\mathbf{y}}(\mathbf{v}^c)-P^{\mathbf{x}}(\mathbf{v}^c))-(P^{\mathbf{y}}(\hat{\mathbf{v}})-P^{\mathbf{x}}(\hat{\mathbf{v}})) \\
&= P^{\mathbf{x}}(\hat{\mathbf{v}})-P^{\mathbf{x}}(\hat{\mathbf{v}}) \\
&\leq 0
\end{aligned}$$

Hence we have:

$$P^{\mathbf{y}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) \geq P^{\mathbf{y}}(\tilde{\mathbf{v}}) - P^{\mathbf{x}}(\tilde{\mathbf{v}}) \geq y^{\gamma}_{i_\gamma} - v_\gamma = \min\{y^{\gamma}_{i_\gamma}, v_1\} - \max\{y^1_{i_1}, v_\gamma\}$$

as required.

Case (iii): $v_{g_0} < y_{i_1}$.(without any assumption on the relative standing of v_1 vis-à-vis $y^{\gamma}_{i_\gamma}$)
In this case, there exists some $\overline{h} < \{1, ..., g_0 - 1\}$ such that $v_{\overline{h}} \geq y^1_{i_1} > v_{\overline{h}+1} \geq ... \geq v_k$. We first note that:

$$\begin{aligned} & \sum_{g=\overline{h}+1}^{k} \sum_{i \in \mathcal{N}(g)} [\max(v_g - y^g_i, 0) - \max(v_1 - x^g_i, 0)] \\ \geq & \sum_{g=\overline{h}+1}^{k} \sum_{i \in \mathcal{N}(g)} [\max(w_g - y^g_i, 0) - \max(w_1 - x^g_i, 0)] \end{aligned}$$

because assuming otherwise would imply that:

$$\begin{aligned} P^{\mathbf{y}}(w_1, ..., w_{\overline{h}}, v_{\overline{h}+1}, ..., v_k) - P^{\mathbf{x}}(w_1, ..., w_{\overline{h}}, v_{\overline{h}+1}, ..., v_k) &< P^{\mathbf{y}}(\mathbf{w}) - P^{\mathbf{x}}(\mathbf{w}) \\ &= 0, \end{aligned}$$

and this inequality contradicts the statement that $x \succsim^{OPG} y$ (since the vector of poverty lines $(w_1, ..., w_{\overline{h}}, v_{\overline{h}+1}, ..., v_k)$ belongs to V).
Let $\tilde{v} := (v_1, ..., v_{\overline{h}}, w_{\overline{h}+1}, ..., w_k) \in \mathcal{V}$. Observe with care that the vector $\tilde{v}$ so defined corresponds either to case (i) (if $v_1 \leq y^{\gamma}_{i_\gamma}$) or to case (ii) (if $v_1 > y^{\gamma}_{i_\gamma}$). Observe also that $\max\{y^1_{i_1}, \tilde{v}_\gamma\} \leq \max\{y^1_{i_1}, v_\gamma\}$. Indeed if $\tilde{v}_\gamma \leq y^1_{i_1}$ there is nothing to prove. If on the other hand $\tilde{v}_\gamma > y^1_{i_1}$, then $\tilde{v}_\gamma = v_\gamma$ by definition of $\overline{h}$ and the inequality $\max\{y^1_{i_1}, \tilde{v}_\gamma\} \leq \max\{y^1_{i_1}, v_\gamma\}$ also holds. Collecting these observations, we obtain that

$$\begin{aligned} P^{\mathbf{y}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) &\geq P^{\mathbf{y}}(\tilde{\mathbf{v}}) - P^{\mathbf{x}}(\tilde{\mathbf{v}}) \\ &\geq \min\{y^{\gamma}_{i_\gamma}, v_1\} - \max\{\tilde{v}_\gamma, y^1_{i_1}\} \text{ (by cases (i) or (ii))} \\ &\geq \min\{y^{\gamma}_{i_\gamma}, v_1\} - \max\{y^1_{i_1}, v_\gamma\} \end{aligned}$$

which proves (22) in that last case.

Let us now establish the existence of a distribution $\overline{\mathbf{x}} \in D(I)$ that has been obtained from y by an FIP and that is such that $x \succsim^{OPG} \overline{\mathbf{x}}$. Let $\overline{\mathbf{x}}$ be the distribution obtained from y by means of an FIP from agent $i^\gamma \in N(\gamma)$ to agent $i^1 \in N(1)$. Let us show that $x \succsim^{OPG} \overline{\mathbf{x}}$. Consider any vector $v \in \mathcal{V}$ of ordered poverty lines. If $v_\gamma \geq y^{\gamma}_{i^\gamma}$ or $v_1 \leq y^1_{i^1}$, it is clear that $P^{\overline{\mathbf{x}}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \geq P^{\mathbf{x}}(\mathbf{v})$. If on the other hand $v_\gamma < y^{\gamma}_{i^\gamma}$ and $v_1 > y^1_{i^1}$, by straightforward computations[11]:

$$\begin{aligned} P^{\overline{\mathbf{x}}}(\mathbf{v}) &= P^{\mathbf{y}}(\mathbf{v}) - \max\{y^{\gamma}_{i^\gamma} - v_\gamma, 0\} - \max\{y^1_{i^1} - v_1, 0\} \\ &\quad + \max\{y^1_{i^1} - v_\gamma, 0\} + \max\{y^{\gamma}_{i^\gamma} - v_1, 0\} \\ &= P^{\mathbf{y}}(\mathbf{v}) - y^{\gamma}_{i^\gamma} + v_\gamma + \max\{y^1_{i^1} - v_\gamma, 0\} + \max\{y^{\gamma}_{i^\gamma} - v_1, 0\} \\ &= P^{\mathbf{y}}(\mathbf{v}) - (y^{\gamma}_{i^\gamma} - \max\{y^{\gamma}_{i^\gamma} - v_1, 0\}) + (v_\gamma - \max\{y^1_{i^1} - v_\gamma, 0\}) \\ &= P^{\mathbf{y}}(\mathbf{v}) - \min\{y^{\gamma}_{i^\gamma}, v_1\} + \max\{y^1_{i^1}, v_\gamma\}. \end{aligned}$$

[11] Some of them using the fact that $\max\{a, b\} = b + \max\{a - b, 0\}$ and $\min\{c, d\} = c - \max\{c - d, 0\}$

Using the inequality (22) proved above, this implies that:

$$P^{\overline{\mathbf{x}}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) \geq P^{\mathbf{y}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) - \min\{y^{\gamma}_{i^{\gamma}}, v_1\} + \max\{y^{1}_{i^{\gamma}}, v_{\gamma}\} \geq 0.$$

which proves the result. ■

6.4 Proof of the lemmas, claims and corollaries of Section 3.2

6.4.1 Lemma 4.

Define the two vectors of poverty lines v^- and v^+ by:

$$\begin{aligned} v^-_h &= \min(v_h, v^c_h) \text{ and,} \\ v^+_h &= \max(v_h, v^c_h). \end{aligned}$$

It is clear that v^- and v^+ both belong to $\mathcal{V}$. By definition of v^- and v^+, we have:

$$\begin{aligned} & P^{\mathbf{x}}(\mathbf{v}^+) - P^{\mathbf{y}}(\mathbf{v}^+) + P^{\mathbf{x}}(\mathbf{v}^-) - P^{\mathbf{y}}(\mathbf{v}^-) \\ = \; & P^{\mathbf{x}}(\mathbf{v}^c) - P^{\mathbf{y}}(\mathbf{v}^c) + P^{\mathbf{x}}(\mathbf{v}) - P^{\mathbf{y}}(\mathbf{v}) \end{aligned} \tag{23}$$

By definition of v^c, we have $P^{\mathbf{x}}(v^c) - P^{\mathbf{y}}(v^c) = 0$. Assume therefore by contradiction that $P^{\mathbf{x}}(\mathbf{v}) - P^{\mathbf{y}}(\mathbf{v}) = 0$ so that, using equality (23), we have:

$$P^{\mathbf{x}}(\mathbf{v}^+) - P^{\mathbf{y}}(\mathbf{v}^+) + P^{\mathbf{x}}(\mathbf{v}^-) - P^{\mathbf{y}}(\mathbf{v}^-) = 0$$

As there exists h_0 such that $v_{h_0} < v^c_{h_0}$, we must have $v^-_{h_0} = v_{h_0} < v^c_{h_0}$. Moreover $v^-_1 = \min(v_1, v^c_1) > y^1_1$. Consequently, by the recursive definition of v^c, we must have that:

$$P^{\mathbf{x}}(\mathbf{v}^-) - P^{\mathbf{y}}(\mathbf{v}^-) < 0$$

But this implies that:

$$P^{\mathbf{x}}(\mathbf{v}^+) - P^{\mathbf{y}}(\mathbf{v}^+) > 0$$

a contradiction of the fact that $x \succsim^{OPG} y$ and that v^+ belongs to $\mathcal{V}$. ■

6.4.2 Corollary 1.

Using the recursive definition of the vector v^c provided by (4), it is clear that:

$$v^c_h \geq \min_{v_h} \left\{ \exists v_{-h} : \; v_1 > y^1_1, \; \mathbf{v} = (v_h, v_{-h}) \in \mathcal{V} \text{ and } P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \right\}.$$

for all h. To prove that:

$$v^c_h \leq \min_{v_h} \left\{ \exists v_{-h} : \; v_1 > y^1_1, \; \mathbf{v} = (v_h, v_{-h}) \in \mathcal{V} \text{ and } P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(\mathbf{v}) \right\}.$$

we simply note that, thanks to Lemma 4, any vector $v \in \mathcal{V}$ such that $v_1 > y^1_1$ and $P^{\mathbf{x}}(\mathbf{v}) - P^{\mathbf{y}}(\mathbf{v}) = 0$ must also satisfy $v_h \geq v^c_h$ for all $h \in \{2, ..., k\}$. ■

6.4.3 Lemma 5.

We first note that, for any sufficiently small strictly positive number ε, the vector of poverty lines:

$$(v_1^c, ..., v_{h_0-1}^c, \underbrace{v_{h_0}^c + \varepsilon, ..., v_{h_0}^c + \varepsilon}_{l-h_0}, \underbrace{v_{h_0}^c, ..., v_{h_0}^c}_{\overline{h}-l}, v_{h_0+\overline{h}+1}^c, ..., v_k^c)$$

belongs to the set $\mathcal{V}$. Hence, since $x \succsim^{OPG} y$, we have:

$$\begin{aligned}
& P^{\mathbf{x}}(\mathbf{v}^c) + \varepsilon[\sum_{h=h_0}^{h_0+l} \overline{p}^{\mathbf{x}}(h, v_h^c)] \\
= \; & P^{\mathbf{x}}(v_1^c, ..., v_{h_0-1}^c, v_{h_0}^c + \varepsilon, ..., v_{h_0}^c + \varepsilon, v_{h_0}^c, ..., v_{h_0}^c, v_{h_0+\overline{h}+1}^c, ..., v_k^c) \\
\leq \; & P^{\mathbf{y}}(v_1^c, ..., v_{h_0-1}^c, v_{h_0}^c + \varepsilon, ..., v_{h_0}^c + \varepsilon, v_{h_0}^c, ..., v_{h_0}^c, v_{h_0+\overline{h}+1}^c, ..., v_k^c) \\
= \; & P^{\mathbf{y}}(\mathbf{v}^c) + \varepsilon[\sum_{h=h_0}^{h_0+l} \overline{p}^{\mathbf{y}}(h, v_h^c)]
\end{aligned}$$

which, when combined with the fact that $P^{\mathbf{x}}(v^c) - P^{\mathbf{y}}(v^c) = 0$ by definition of v^c, implies inequality (13). Similarly the vector of poverty lines:

$$(v_1^c, ..., v_{h_0-1}^c, \underbrace{v_{h_0}^c, ..., v_{h_0}^c}_{l-h_0}, \underbrace{v_{h_0}^c - \varepsilon, ..., v_{h_0}^c - \varepsilon}_{\overline{h}-l}, v_{h_0+\overline{h}+1}^c, ..., v_k^c)$$

belongs to the set $\mathcal{V}$ for a small enough ε. By the recursive definition of v^c, we have:

$$\begin{aligned}
= \; & P^{\mathbf{x}}(v_1^c, ..., v_{h_0-1}^c, v_{h_0}^c, ..., v_{h_0}^c, v_{h_0}^c - \varepsilon, ..., v_{h_0}^c - \varepsilon, v_{h_0+\overline{h}+1}^c, ..., v_k^c) \\
< \; & P^{\mathbf{y}}(v_1^c, ..., v_{h_0-1}^c, v_{h_0}^c, ..., v_{h_0}^c, v_{h_0}^c - \varepsilon, ..., v_{h_0}^c - \varepsilon, v_{h_0+\overline{h}+1}^c, ..., v_k^c)
\end{aligned}$$

and, therefore:

$$\begin{aligned}
& P^{\mathbf{x}}(\mathbf{v}^c) - \varepsilon[\sum_{h=h_0+l}^{h_0+\overline{h}} p^{\mathbf{x}}(h, v_h^c)] \\
= \; & P^{\mathbf{x}}(v_1^c, ..., v_{h_0-1}^c, v_{h_0}^c, ..., v_{h_0}^c, v_{h_0}^c - \varepsilon, ..., v_{h_0}^c - \varepsilon, v_{h_0+\overline{h}+1}^c, ..., v_k^c) \\
< \; & P^{\mathbf{y}}(v_1^c, ..., v_{h_0-1}^c, v_{h_0}^c, ..., v_{h_0}^c, v_{h_0}^c - \varepsilon, ..., v_{h_0}^c - \varepsilon, v_{h_0+\overline{h}+1}^c, ..., v_k^c) \\
= \; & P^{\mathbf{y}}(\mathbf{v}^c) - \varepsilon[\sum_{h=h_0+l}^{h_0+\overline{h}} p^{\mathbf{y}}(h, v_h^c)]
\end{aligned}$$

which, when combined with $P^{\mathbf{x}}(v^c) - P^{\mathbf{y}}(v^c) = 0$, implies inequality (14). ■

6.4.4 Corollary 2.

If $v_1^c > v_2^c$, we can apply Lemma 5 to the case where $h_0 = 1$ and $\overline{h} = 0$. In this case, Inequalities (13) and (14) write:

$$\overline{p}^{\mathbf{x}}(1, v_1^c) \leq \overline{p}^{\mathbf{y}}(1, v_1^c)$$

and:

$$p^{\mathbf{x}}(1, v_1^c) > p^{\mathbf{y}}(1, v_1^c).$$

Hence, there must exist an agent $i \in N(1)$ such that $y_i^1 = v_1^c$. More generally, if $v_1^c = v_2^c = ... = v_{k+1}^c > v_{k+2}^c$, one applies Lemma 5 to the case where $h_0 = 1$ (taking $l = \overline{h}$ in (13) and $l = 0$ in (14)) which gives

$$\sum_{h=1}^{\overline{h}+1} \overline{p}^{\mathbf{x}}(1, v_1^c) \leq \sum_{h=1}^{\overline{h}+1} \overline{p}^{\mathbf{y}}(1, v_1^c)$$

and:

$$\sum_{h=1}^{\overline{h}+1} p^{\mathbf{x}}(1, v_1^c) > \sum_{h=1}^{\overline{h}+1} p^{\mathbf{y}}(1, v_1^c).$$

One then obtains the existence of some $j \in \{1, ..., \overline{h} + 1\}$ and some $i \in N(j)$ such that $y_i^j = v_j^c = v_1^c$. ■

6.4.5 Lemma 6.

Given $\mathbf{x}$ and $\mathbf{y}$ two distributions, define the income support of these two distributions:

$$\mathcal{I}(\mathbf{x}, \mathbf{y}) = \{a : \exists h \in \{1, ..., k\},\ \ i \in \{1, ..., n(h)\} \text{ such that } x_i^h = a \text{ or } y_i^h = a\}.$$

Then choose the strictly positive number small enough as to satisfy:

$$\varepsilon_1 < \min_{a,b \in \mathcal{I}(\mathbf{x},\mathbf{y}), a \neq b} |a - b| \tag{24}$$

and

$$\varepsilon_1 < v_1^c - y_1^1. \tag{25}$$

Consider then any numbers $v_2, ..., v_k$ such that $v = (y_1^1 + \varepsilon_1, v_2, ..., v_k) \in \mathcal{V}$ and let $h_0 \in \{1, ..., k\}$ be such that $v_h > y_1^1$ for all $h \in \{1, ..., h_0\}$ and $v_h \leq y_1^1$ for all $j \in \{h_0 + 1, ..., k\}$ (if there are such j). One can then write the vector v as:

$$\mathbf{v} = (y_1^1 + \varepsilon_1, y_1^1 + \varepsilon_2, ..., y_1^1 + \varepsilon_{h_0}, v_{h_0+1}, v_{h_0+2}, ..., v_k)$$

for some (possibly empty) list $\varepsilon_2, ..., \varepsilon_{h_0}$ satisfying $\varepsilon_1 \geq \varepsilon_2 ... \geq \varepsilon_{h_0} > 0$. Let us prove that:

$$P^{\mathbf{y}}(\mathbf{v}) \geq P^{\mathbf{x}}(\mathbf{v}) + \varepsilon_1$$

Clearly, for ε_1 satisfying (24) and (25), we have:

$$P^{\mathbf{y}}(\mathbf{v}) = P^{\mathbf{y}}(y_1^1, ..., y_1^1, v_{h_0+1}, v_{h_0+2}, ..., v_k) + \sum_{h=1}^{h_0} \varepsilon_h \overline{p}^{\mathbf{y}}(h, y_1^1),$$

and:

$$P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{x}}(y_1^1, ..., y_1^1, v_{h_0+1}, v_{h_0+2}, ..., v_k) + \sum_{h=1}^{h_0} \varepsilon_h \overline{p}^{\mathbf{x}}(h, y_1^1),$$

and, therefore:

$$P^{\mathbf{y}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) = P^{\mathbf{y}}(y_1^1, ..., y_1^1, v_{h_0+1}, ..., v_k) - P^{\mathbf{x}}(y_1^1, ..., y_1^1, v_{h_0+1}, ..., v_k)$$

$$+\sum_{h=1}^{h_0}\varepsilon_h[\overline{p}^{\mathbf{y}}(h,y_1^1)-\overline{p}^{\mathbf{x}}(h,y_1^1)] \tag{26}$$

If

$$P^{\mathbf{y}}(y_1^1,...,y_1^1,v_{h_0+1},...,v_k)>P^{\mathbf{x}}(y_1^1,...,y_1^1,v_{h_0+1},...,v_k)$$

then there is nothing to prove. Indeed, from Lemma 1 and the assumption that $y_i^1\neq x_i^1$ for all $i\in\mathcal{N}(1)$, we have that $\overline{p}^{\mathbf{y}}(1,y_1^1)\geq 1$ and $\overline{p}^{\mathbf{x}}(1,y_1^1)=0$. Hence:

$$\begin{aligned}&P^{\mathbf{y}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)-P^{\mathbf{x}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)\\&+\varepsilon_1[\overline{p}^{\mathbf{y}}(1,y_1^1)-\overline{p}^{\mathbf{x}}(1,y_1^1)]\\&\qquad>\varepsilon_1\end{aligned}$$

for any ε_1 satisfying (24) and (25). Because of this, we can choose ε_1 sufficiently small so as make the numbers $\varepsilon_2,...,\varepsilon_{h_0}$ sufficiently small for the inequality:

$$\begin{aligned}&P^{\mathbf{y}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)-P^{\mathbf{x}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)\\&+\varepsilon_1[\overline{p}^{\mathbf{y}}(1,y_1^1)-\overline{p}^{\mathbf{x}}(1,y_1^1)]+\sum_{h=2}^{h_0}\varepsilon_h[\overline{p}^{\mathbf{y}}(h,y_1^1)-\overline{p}^{\mathbf{x}}(h,y_1^1)]\\\geq\ &\varepsilon_1\end{aligned}$$

to hold. Suppose now that:

$$P^{\mathbf{y}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)=P^{\mathbf{x}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)$$

In that case, it follows from (26) that:

$$P^{\mathbf{y}}(\mathbf{v})-P^{\mathbf{x}}(\mathbf{v})=\sum_{h=1}^{h_0}\varepsilon_h[\overline{p}^{\mathbf{y}}(h,y_1^1)-\overline{p}^{\mathbf{x}}(h,y_1^1)]$$

This equality can equivalently be written as:

$$P^{\mathbf{y}}(\mathbf{v})-P^{\mathbf{x}}(\mathbf{v})=\sum_{h=1}^{h_0}[\varepsilon_h-\varepsilon_{h+1}]\sum_{g=1}^{h}[\overline{p}^{\mathbf{y}}(g,y_1^1)-\overline{p}^{\mathbf{x}}(g,y_1^1)] \tag{27}$$

using the convention that $\varepsilon_{h_0+1}=0$. Note that, by definition of v_1^c, we must have $P^{\mathbf{y}}(\mathbf{v})-P^{\mathbf{x}}(\mathbf{v})>0$ if ε_1 satisfies (25). Note also that, for all $h\in\{1,...,h_0\}$, we have:

$$\sum_{g=1}^{h}[\overline{p}^{\mathbf{y}}(g,y_1^1)-\overline{p}^{\mathbf{x}}(g,y_1^1)]\geq 1 \tag{28}$$

Indeed, by definition of v_1^c, we have for every strictly positive $\delta\leq\varepsilon_1$:

$$\begin{aligned}&P^{\mathbf{y}}(\underbrace{y_1^1+\delta,...,y_1^1+\delta}_{h},\underbrace{y_1^1,...,y_1^1}_{h_0-h},v_{h_0+1},v_{h_0+2},...,v_k)\\>\ &P^{\mathbf{x}}(\underbrace{y_1^1+\delta,...,y_1^1+\delta}_{h},\underbrace{y_1^1,...,y_1^1}_{h_0-h},v_{h_0+1},v_{h_0+2},...,v_k)\end{aligned}$$

Yet,

$$\begin{aligned}
& P^{\mathbf{y}}(\underbrace{y_1^1+\delta,...,y_1^1+\delta}_{h},\underbrace{y_1^1,...,y_1^1}_{h_0-h},v_{h_0+1},v_{h_0+2},...,v_k) \\
= \; & P^{\mathbf{y}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)+\delta\sum_{g=1}^{h}\overline{p}^{\mathbf{y}}(g,y_1^1)
\end{aligned}$$

and:

$$\begin{aligned}
& P^{\mathbf{x}}(\underbrace{y_1^1+\delta,...,y_1^1+\delta}_{h},\underbrace{y_1^1,...,y_1^1}_{h_0-h},v_{h_0+1},v_{h_0+2},...,v_k) \\
= \; & P^{\mathbf{x}}(x_1^1,...,x_1^1,v_{h_0+1},v_{h_0+2},...,v_k)+\delta\sum_{g=1}^{h}\overline{p}^{\mathbf{x}}(g,y_1^1)
\end{aligned}$$

Hence, under the assumption that:

$$P^{\mathbf{y}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)=P^{\mathbf{x}}(y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k)$$

we have:

$$\begin{aligned}
& P^{\mathbf{y}}(y_1^1+\delta,...,y_1^1+\delta,y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k) \\
& -P^{\mathbf{x}}(y_1^1+\delta,...,y_1^1+\delta,y_1^1,...,y_1^1,v_{h_0+1},v_{h_0+2},...,v_k) \\
= \; & \delta\sum_{g=1}^{h}[\overline{p}^{\mathbf{y}}(g,y_1^1)-\overline{p}^{\mathbf{x}}(g,y_1^1)]>0
\end{aligned}$$

which establishes Inequality (28). Together with (27), this leads to the conclusion that:

$$P^{\mathbf{y}}(\mathbf{v})-P^{\mathbf{x}}(\mathbf{v})\geq\varepsilon_1-\epsilon_{h_0+1}=\varepsilon_1.$$

as required. ■

6.4.6 Proposition 1

There exists some $h_0\in\{1,...,k\}$ such that $v_1^c=v_2^c=...v_{h_0}^c>v_{h_0+1}^c$ (using, if necessary, the convention that $v_{k+1}^c=\underline{v}(\mathbf{x},\mathbf{y})$). Then, using Corollary 2, we conclude that there is some category $j\in\{1,...,h_0\}$ and some individual $i^j\in\mathcal{N}(j)$ such that $y_{i^j}^j=v_1^c$ and $\forall h\in\{j,j+1,...,h_0\}$, $i\in\mathcal{N}(h)$, $y_i^h\neq v_1^c$ (that is, j is the highest category in the set $\{1,...,h_0\}$ for which there is an individual in distribution $\mathbf{y}$ whose income is equal to v_1^c.). It is important to note that we do not preclude the possibility that $j=1$. Let us show that there exists a distribution $\widehat{\mathbf{x}}\in\mathcal{D}(I)$ such that $\mathbf{x}\succsim^{OPG}\widehat{\mathbf{x}}$ and $\widehat{\mathbf{x}}$ was obtained from $\mathbf{y}$ by a BTPIT. For any strictly positive integer m, let $\widehat{\mathbf{x}}^m$ be the distribution obtained from distribution $\mathbf{y}$ by performing a BTPIT of an amount of $1/m$ from agent $i^j\in\mathcal{N}(j)$ to agent $1\in\mathcal{N}(1)$. We claim that there exists some m sufficiently large that $\mathbf{x}\succsim^{OPG}\widehat{\mathbf{x}}^m$. Assume by contradiction that no such m exists. This implies the existence of a sequence of ordered poverty lines vectors $\mathbf{v}^m\in\mathcal{V}'$ such that $P^{\widehat{\mathbf{x}}^m}(\mathbf{v}^m)<P^{\mathbf{x}}(\mathbf{v}^m)$.

Note that, for every strictly positive real integer m, and whatever the ordered vector of poverty lines $\mathbf{v} \in \mathcal{V}'$ is, we have:

$$\begin{aligned} P^{\widehat{\mathbf{x}}^m}(\mathbf{v}) &= P^{\mathbf{y}}(\mathbf{v}) - 1/m \text{ if } v_1 \geq y_1^1 + 1/m \text{ and } v_j \leq y_{ij}^j - 1/m, & (29)\\ &= P^{\mathbf{y}}(\mathbf{v}) - \max(v_1 - y_1^1, 0) \text{ if } v_1 < y_1^1 + \frac{1}{m},\, v_j \leq y_{ij}^j - \frac{1}{m} & (30)\\ &= P^{\mathbf{y}}(\mathbf{v}) + \min(v_j - y_{ij}^j, 0) \text{ if } v_1 \geq y_1^1 + \frac{1}{m},\, v_j > y_{ij}^j - \frac{1}{m}. & (31)\end{aligned}$$

Because of this, we can assume without loss of generality that $v_1^m \geq y_1^1 + 1/m$ and $v_j^m \leq v_1^c - 1/m$. By compactness of $\mathcal{V}'$, $\mathbf{v}^m$ admits a subsequence that converges to some vector of ordered poverty lines $\mathbf{v} \in \mathcal{V}'$. By continuity, we must have $P^{\mathbf{y}}(\mathbf{v}) = P^{\mathbf{x}}(\mathbf{v})$. Hence, by definition of critical value v_1^c, either:
(i) $v_1 = y_1^1$ or
(ii) $v_1 \geq v_1^c$

• If case (i) holds, then we have:

$$\begin{aligned} P^{\widehat{\mathbf{x}}^m}(\mathbf{v}^m) &\geq P^{\mathbf{y}}(\mathbf{v}^m) - 1/m \text{ by (29)-(31)}\\ &\geq P^{\mathbf{x}}(\mathbf{v}^m) - 1/m + v_1^m - y_1^1 \text{ by Lemma 6, taking } v_1^m - y_1^1 = \varepsilon_1\\ &\geq P^{\mathbf{x}}(\mathbf{v}^m)\end{aligned}$$

which is a contradiction.

• Now, if case (ii) holds and $v_1 \geq v_1^c$ then by Corollary 1, we must have $v_h \geq v_h^c$ for $h = 2, ..., k$. In particular, since $v_j^m \leq v_1^c - 1/m$ and $\mathbf{v}^m$ admits a subsequence that converges to $\mathbf{v}$, we must have $v_j = v_j^c = v_1^c$. We can actually assume without loss of generality that, for every $h = 1, ..., h_0$, $v_h^m \in \{v_1^c - 1/m, v_1^c\}$ (for large enough m, $v_1^c - 1/m$ and v_1^c are the only two incomes observed in distributions $\mathbf{y}$, $\widehat{\mathbf{x}}^m$ and $\mathbf{x}$ for poverty lines v_h^m relevant for categories $h = 1, ..., h_0$). Hence, for some $g \in \{1, ..., j\}$, we have: $v_g^m = ... = v_j^m = ...v_{h^0}^m = v_1^c - 1/m$, $v_1^m = v_2^m = ... = v_{g-1}^m = v_1^c$. Since:

$$\sum_{h=g}^{h_0} p^{\mathbf{x}}(h, v_1^c) > \sum_{h=g}^{h_0} p^{\mathbf{y}}(h, v_1^c)$$

we have:

$$\begin{aligned} &\sum_{h=g}^{h_0} \sum_{i \in \mathcal{P}^y(h, v_h^m)} [v_h^m - y_i^h] - \sum_{h=g}^{h_0} \sum_{i \in \mathcal{P}^x(h, v_h^m)} [v_h^m - x_i^h] - 1/m\\ \geq &\sum_{h=g}^{h_0} \sum_{i \in \mathcal{P}^y(h, v_1^c)} [v_1^c - y_i^h] - \sum_{h=g}^{h_0} \sum_{i \in \mathcal{P}^x(h, v_1^c)} [v_1^c - x_i^h]\end{aligned}$$

and, therefore:

$$\begin{aligned} &P^{\mathbf{y}}(\mathbf{v}^m) - P^{\mathbf{x}}(\mathbf{v}^m) - 1/m\\ \geq\ &P^{\mathbf{y}}(v_1^c, ..., v_1^c, v_{h_0+1}^c, ..., v_k^c) - P^{\mathbf{x}}(v_1^c, ..., v_1^c, v_{h_0+1}^c, ..., v_k^c) \geq 0\end{aligned}$$

Finally $P^{\widehat{\mathbf{x}}^m}(\mathbf{v}^m) \geq P^{\mathbf{y}}(\mathbf{v}^m) - 1/m \geq P^{\mathbf{x}}(\mathbf{v}^m)$, a contradiction.

6.4.7 Example 2

Consider distribution $\widehat{\mathbf{x}}^\varepsilon$ defined by:

$$\widehat{x}_1^{\varepsilon 1} = 3+\varepsilon,\ \ \widehat{x}_2^{\varepsilon 2} = 4-\varepsilon,\ \ \widehat{x}_2^{\varepsilon 1} = \widehat{x}_2^1 = 7 \text{ and } \widehat{x}_1^{\varepsilon 2} = \widehat{x}_1^2 = 0.$$

Then, using the ordered vector $\mathbf{v} = (3+\varepsilon, 3+\varepsilon)$ of poverty lines, we have:

$$\begin{aligned} P^{\widehat{\mathbf{x}}^\varepsilon}(3+\varepsilon, 3+\varepsilon) &= \max(3+\varepsilon-(3+\varepsilon),0)+\max(3+\varepsilon-7,0) \\ &\quad +\max(3+\varepsilon-0,0)+\max(3+\varepsilon-(4-\varepsilon),0) \\ &= 3+\varepsilon \text{ (if } \varepsilon \in]0,1/2]) \\ &< P^{\mathbf{x}}(3+\varepsilon,3+\varepsilon) \\ &= \max(3+\varepsilon-5,0)+\max(3+\varepsilon-6,0) \\ &\quad +\max(3+\varepsilon-1,0)+\max(3+\varepsilon-2,0) \\ &= 3+2\varepsilon \end{aligned}$$

6.4.8 Proposition 2

We prove this proposition using Lemma 3. We note first that, since $x_i^1 > v_1^c\ \forall i \in \mathcal{N}(1)$, we must have $v_1^c = v_2^c$. Suppose by contradiction that $v_1^c > v_2^c$. This means that, for any number ε such that $v_1^c - v_2^c > \varepsilon > 0$, we have that $(v_1^c - \varepsilon, v_2^c, ..., v_k^c) \in \mathcal{V}$ and

$$\begin{aligned} & P^{\mathbf{y}}(v_1^c-\varepsilon, v_2^c, ..., v_k^c) - P^{\mathbf{x}}(v_1^c-\varepsilon, v_2^c, ..., v_k^c) \\ =\ & -\varepsilon p^{\mathbf{y}}(1, v_1^c) + \varepsilon p^{\mathbf{x}}(1, v_1^c) + P^{\mathbf{y}}(v_1^c, v_2^c, ..., v_k^c) - P^{\mathbf{x}}(v_1^c, v_2^c, ..., v_k^c) \\ =\ & -\varepsilon p^{\mathbf{y}}(1, v_1^c) \\ <\ & 0 \end{aligned}$$

since $p^{\mathbf{x}}(1, v_1^c) = 0 = P^{\mathbf{y}}(v_1^c, v_2^c, ..., v_k^c) - P^{\mathbf{x}}(v_1^c, v_2^c, ..., v_k^c)$ (by definition of $\mathbf{v}^c$) and $p^{\mathbf{y}}(1, v_1^c) \geq 1$. But this is a contradiction of the fact that $\mathbf{x} \succsim^{OPG} \mathbf{y}$. It then follows that $h_0 \geq 2$ and that the second bullet statement of the proposition (e.g. "For any category h such that $h_0 \geq h \geq 2$, we have $y_i^h \neq v_1^c$") is not empty. This fact and Corollary 2 establish the existence of an agent $i_1 \in \mathcal{N}(1)$ for which $y_{i_1}^1 = v_1^c$. Hence vector of poverty lines $\mathbf{v}^c$ is just like vector $\mathbf{w}$ in the antecedent clause of Lemma 3. It then follows from Lemma 5 that:

$$\sum_{h=l+1}^{h_0} \overline{p}^{\mathbf{y}}(h, y_{i_1}^1) = \sum_{h=l+1}^{h_0} p^{\mathbf{y}}(h, y_{i_1}^1) < \sum_{h=l+1}^{h_0} p^{\mathbf{x}}(h, y_{i_1}^1) \leq \sum_{h=l+1}^{h_0} \overline{p}^{\mathbf{x}}(h, y_{i_1}^1)$$

for $l = 1, ..., h_0 - 1$, which implies that Inequality (10) in the antecedent clause of Lemma 3 holds. The existence of a distribution $\widehat{\mathbf{x}} \in \mathcal{D}(I)$ such that $\widehat{\mathbf{x}}$ was obtained from $\mathbf{y}$ by an FIP and $\mathbf{x} \succsim^{OPG} \widehat{\mathbf{x}}$ then immediately follows from this lemma.

6.4.9 Lemma 7.

Let $\overline{\mathbf{w}}$ be the vector of ordered poverty lines that is the limit of the sequence $\{w^m\}$ of ordered poverty lines vectors satisfying $w_1^m = y_1^1 + \varepsilon_1^m$ with $\varepsilon_1^m \to 0$ and $P^{\mathbf{x}}(w^m) - P^{\mathbf{y}}(w^m) = 0$ for all m that was mentioned in Case (B) of section 3.2. We first show that

$\overline{w}_1 = \overline{w}_2$. By contradiction, suppose that $\overline{w}_2 < \overline{w}_1$. Then, there exists a large enough m for which $w_2^m < \overline{w}_1 = y_1^1$. Also, for a large enough m, we have that:

$$w_1^m \overline{p}^{\mathbf{x}}(1, w_1^m) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{x}}(1, w_1^m)} x_i^1 = w_1 \overline{p}^{\mathbf{x}}(1, y_1^1) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{x}}(1, y_1^1)} x_i^1 = 0$$

thanks to Lemma 1 and the fact that $y_1^1 \neq x_1^1$. Moreover we have:

$$\begin{aligned} w_1^m \overline{p}^{\mathbf{y}}(1, w_1^m) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{y}}(1, w_1^m)} y_i^1 &= (\varepsilon_1^m + y_1^1)\overline{p}^{\mathbf{y}}(1, y_1^1) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{y}}(1, w_1)} y_i^1 \\ &> y_1^1 \overline{p}^{\mathbf{y}}(1, y_1^1) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{y}}(1, w_1)} y_i^1 = 0 \end{aligned}$$

because $\overline{p}^{\mathbf{y}}(1, y_1^1) \geq 1$. Hence:

$$\begin{aligned} P^{\mathbf{y}}(\mathbf{w}^m) - P^{\mathbf{x}}(\mathbf{w}^m) &= w_1^m \overline{p}^{\mathbf{y}}(1, w_1^m) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{y}}(1, w_1^m)} y_i^1 + \sum_{h=2}^{k} P^{\mathbf{y}}(h, w_h^m) \\ &\quad -[w_1^m \overline{p}^{\mathbf{x}}(1, w_1^m) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{x}}(1, w_1^m)} x_i^1 + \sum_{h=2}^{k} P^{\mathbf{x}}(h, w_h^m)] \\ &= 0 \\ &> y_1^1 \overline{p}^{\mathbf{y}}(1, y_1^1) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{y}}(1, w_1)} y_i^1 + \sum_{h=2}^{k} P^{\mathbf{y}}(h, w_h^m) \\ &\quad -[w_1 \overline{p}^{\mathbf{x}}(1, y_1^1) - \sum_{i \in \overline{\mathcal{P}}^{\mathbf{x}}(1, y_1^1)} x_i^1 + \sum_{h=2}^{k} P^{\mathbf{x}}(h, w_h^m)] \\ &= P^{\mathbf{y}}(y_1^1, w_2^m, ..., w_k^m) - P^{\mathbf{x}}(y_1^1, w_2^m, ..., w_k^m) \end{aligned}$$

a contradiction. We now show that $w_2^m > y_1^1$. Indeed, for m large enough, we have:

$$P^{\mathbf{y}}(\mathbf{w}^m) = P^{\mathbf{y}}(y_1^1, w_2^m, ..., w_k^m) + \varepsilon_1^m p^{\mathbf{y}}(1, y_1^1)$$

and:

$$P^{\mathbf{x}}(\mathbf{w}^m) = P^{\mathbf{x}}(y_1^1, w_2^m, ..., w_k^m) + \varepsilon_1^m p^{\mathbf{x}}(1, y_1^1)$$

Moreover, by Lemma 1 and because,$y_i^1 \neq x_i^1$,we have that

$$\overline{p}^{\mathbf{y}}(1, y_1^1) \geq 1 > 0 = p^{\mathbf{x}}(1, y_1^1).$$

Hence, we have:

$$\begin{aligned} P^{\mathbf{y}}(\mathbf{w}^m) - P^{\mathbf{x}}(\mathbf{w}^m) &= 0 \\ &\geq P^{\mathbf{y}}(y_1^1, w_2^m, ..., w_k^m) - P^{\mathbf{x}}(y_1^1, w_2^m, ..., w_k^m) + \varepsilon_1^m \end{aligned}$$

Because of this (and the fact that $\varepsilon_1^m > 0$), assuming that $w_2^m \leq y_1^1$ and, therefore, that the vector of poverty lines $(y_1^1, w_2^m, ..., w_k^m)$ belongs to $\mathcal{V}$ would be contradictory with the fact that $x \succsim^{OPG} y$. We also know that $w_{h_0+1}^m < y_1^1$. Let $l := \min\{h \geq 1 : w_h^m \leq y_1^1\}$. As we have just shown $3 \leq l \leq h_0 + 1$. For $h = 1, ..., l$, we have $w_h^m = y_1^1 + \varepsilon_h^m$ with

$\varepsilon_1^m \geq \varepsilon_2^m \geq ... \geq \varepsilon_l^m > 0$. We know already that $\overline{p}^{\mathbf{y}}(1, y_1^1) \geq 1 > 0 = \overline{p}^{\mathbf{x}}(1, y_1^1)$. Suppose that the main claim of the lemma was false. In that case, we would have:

$$\sum_{g=1}^{h} p^{\mathbf{y}}(g, y_1^1) > \sum_{g=1}^{h} p^{\mathbf{x}}(g, y_1^1)$$

for all $h = 1, ..., h_0$. Yet:

$$\begin{aligned}
0 &= P^{\mathbf{y}}(\mathbf{w}^m) - P^{\mathbf{x}}(\mathbf{w}^m) \\
&= P^{\mathbf{y}}(y_1^1, ..., y_1^1, w_{l+1}^m, ..., w_k^m) - P^{\mathbf{x}}(y_1^1, ..., y_1^1, w_{l+1}^m, ..., w_k^m) \\
&+ \sum_{h=1}^{l} [\overline{p}^{\mathbf{y}}(h, y_1^1) - \overline{p}^{\mathbf{x}}(h, y_1^1)]\varepsilon_h^m \\
&= P^{\mathbf{y}}(y_1^1, ..., y_1^1, w_{s+1}^m, ..., w_k^m) - P^{\mathbf{x}}(y_1^1, ..., y_1^1, w_{l+1}^m, ..., w_k^m) \\
&+ \sum_{h=1}^{l-1} (\varepsilon_h^m - \varepsilon_{h+1}^m) \sum_{g=1}^{h} [\overline{p}^{\mathbf{y}}(g, y_1^1) - \overline{p}^{\mathbf{x}}(g, y_1^1)] + \varepsilon_l^m \sum_{g=1}^{l} [p^{\mathbf{y}}(g, y_1^1) - p^{\mathbf{x}}(g, y_1^1)] \\
&\geq \sum_{h=1}^{s-1} (\varepsilon_h^m - \varepsilon_{h+1}^m) \sum_{g=1}^{h} [\overline{p}^{\mathbf{y}}(g, y_1^1) - \overline{p}^{\mathbf{x}}(g, y_1^1)] + \varepsilon_l^m \sum_{g=1}^{l} [\overline{p}^{\mathbf{y}}(g, y_1^1) - \overline{p}^{\mathbf{x}}(g, y_1^1)] \\
&\geq \varepsilon_l^m > 0
\end{aligned}$$

a contraction. ■

6.4.10 Proposition 3

We base the argument on Lemma 3. We must therefore prove that the limit vector of poverty lines $\overline{\mathbf{w}}$ satisfies the conditions imposed on vector $\mathbf{w}$ of this lemma. From Lemma 7, we have that $y_1^1 = \overline{w}_1 = \overline{w}_2 = ... = \overline{w}_{h_0} > \overline{w}_{h_0} + 1$ for some $h_0 \geq 2$. We also know from Lemma 7 that there is a category $g_0 \leq h_0$ satisfying $g_0 \geq 2$ for which we have:

$$\sum_{h=1}^{g_0} \overline{p}^{\mathbf{y}}(h, y_1^1) \leq \sum_{h=1}^{g_0} \overline{p}^{\mathbf{x}}(h, y_1^1),$$

and

$$\sum_{h=1}^{l} \overline{p}^{\mathbf{y}}(h, y_1^1) > \sum_{h=1}^{l} \overline{p}^{\mathbf{x}}(h, y_1^1)$$

for all $l = 1, ..., g_0 - 1$. As a consequence we have:

$$\sum_{h=l+1}^{g_0} \overline{p}^{\mathbf{y}}(h, y_1^1) < \sum_{h=l+1}^{g_0} \overline{p}^{\mathbf{x}}(h, y_1^1)$$

for $l = 1, ..., g_0 - 1$. and the conclusion of the proposition follows from Lemma 3.

6.4.11 Claim 1

Once an agent of a category higher than 1 has been involved in an FIP defined as in Proposition 3, his/her income becomes weakly smaller than that of the poorest income observed in category 1. Since the income of an agent in any category $h \geq 2$ never increases through the algorithm described above, and since the number of agents in categories higher than 1 is finite, it follows that the the number of FIP of type $(P1)$ in the algorithm is bounded above by the number of agents in categories $2, ..., k$. ■

6.4.12 Claim 2

By definition of the algorithm and the critical vector $\mathbf{v^c(n)}$, we have $P^{\mathbf{x}(n)}(\mathbf{v^c(n)}) = P^{\mathbf{x}}(\mathbf{v^c(n)})$ which directly implies that :

$$P^{\mathbf{x}(n)}(\mathbf{v^c(n)}) = P^{\mathbf{x}(n+1)}(\mathbf{v^c(n)}) = P^{\mathbf{x}}(\mathbf{v^c(n)}) \tag{32}$$

We first observe that if the distribution $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ through a maximal transfer (MT), the donor's income is equal to $v_1^c(n)$ and therefore the recipient being the poorest agent in category 1, we necessarily have $x_1^1(n+1) < v_1^c(n)$. On the other hand, if distribution $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ through an FIP of type $(P2)$, the recipient has an income equal to $v_1^c(n) > x_1^1(n)$, so that $x_1^1(n) = x_1^1(n+1) < v_1^c(n)$. Hence, in either case, we have $x_1^1(n+1) < v_1^c(n)$. Now by definition of $v_1^c(n+1)$ as an infimum, identity (32) and the fact that $x_1^1(n+1) < v_1^c(n)$, we have $v_1^c(n+1) \leq v_1^c(n)$.

Now combining $P^{\mathbf{x}}(v^c(n)) = P^{\mathbf{x}(n+1)}(v^c(n))$ and $P^{\mathbf{x}}(v^c(n+1)) = P^{\mathbf{x}(n+1)}(v^c(n+1))$ on the one hand and Corollary 1 on the other, it follows that $v_h^c(n+1) \leq v_h^c(n)$ holds for all h as well.

6.4.13 Claim 3

We first observe that $\overline{p}^{\mathbf{x}(n)}(1, v_1^c(n))$ is weakly decreasing for $n \geq n_0$, where n_0 is the integer whose existence was established in Claim 1. Indeed $v_1^c(n)$ is weakly decreasing for $n \geq n_0$ and an agent in category 1 can be designated as the donor at step n only if the algorithm prescribes a maximal transfer and his/her income is equal to $v_1^c(n)$. This proves that the number of agents in category 1 of distribution $\mathbf{x}(n)$ whose income is weakly smaller than $v_1^c(n)$ necessarily weakly decreases as n increases.

Assume now that at some stage $n \geq n_0$ we are in case (P2). In that case, the receiving agent's income is equal to $v_1^c(n)$. Hence $\overline{p}^{\mathbf{x}(n+1)}(1, v_1^c(n+1)) < \overline{p}^{\mathbf{x}(n)}(v_1^c(n))$. As a result, there can be at most $n(1)$ operations of type (P2) in the algorithm after step n_0.

6.4.14 Claim 4

Let n_1 be as in Claim 3. We need to establish that there can only be finitely many breaking transfers after stage n_1. Consider any $n \geq n_1$ and suppose that $\mathbf{x}(n+1)$ is obtained from $\mathbf{x}(n)$ through a breaking transfer of amount $\alpha > 0$ from agent $j_h \in \mathcal{N}(h)$ (with $h \geq 1$) to agent $1 \in \mathcal{N}(1)$. Let $r_+^1(\alpha) \in \mathcal{N}(1)$ and $r_-^h(\alpha) \in \mathcal{N}(h)$ be as in Definition 1: that is $x^1_{r_+^1(\alpha)}(n+1) = x_1^1(n) + \alpha$ and $x^h_{r_-^h(\alpha)}(n+1) = x^h_{j_h}(n) - \alpha$.

Let $\delta > 0$. By definition of a breaking transfer, there exists $\mathbf{v}(\delta) \in \mathcal{V}'$ such that:

$$P^{\mathbf{x}(n+1)^\delta}(\mathbf{v}(\delta)) < P^{\mathbf{x}}(\mathbf{v}(\delta)),$$

where $\mathbf{x}(n+1)^\delta$ denotes the distribution that would be obtained if the transfer at time n was equal to $\alpha + \delta$ rather than α. By compactness of $\mathcal{V}'$, we may assume without loss of generality that $\lim_{\delta \to 0} \mathbf{v}(\delta) = \mathbf{v}^* \in \mathcal{V}'$. By continuity, we then have $P^{\mathbf{x}(n+1)}(\mathbf{v}^*) = P^{\mathbf{x}}(\mathbf{v}^*)$. Note that without loss of generality, we can assume that $v_1(\delta) \geq x_1^1(n) + \alpha$. This implies that $v_1^* \geq x_1^1(n) + \alpha$.

We now show that $v_1^* > x_1^1(n+1)$. By contradiction assume that $v_1^* \leq x_1^1(n+1)$. Then, since $x_1^1(n+1) \leq x_1^1(n) + \alpha$, we necessarily have $v_1^* = x_1^1(n) + \alpha = x_1^1(n+1)$, that is, the poorest agent in category 1 remains the poorest agent after receiving α at step n.

Thus at next step (step $n+1$) the algorithm identifies him as the recipient again. Let $h' \geq 1$ and $j_{h'} \in \mathcal{N}(h')$ be the donor at next step $n+1$ and suppose he/she transfers $\delta > 0$. Since $v^*_{h'} \leq v^*_1 = x^1_1(n+1) < x^{h'}_{j_{h'}}(n+1)$, we have

$$P^{\mathbf{x}(n+2)}(\mathbf{v}(\delta)) = P^{\mathbf{x}(n+1)^\delta}(\mathbf{v}(\delta)) < P^{\mathbf{x}}(\mathbf{v}(\delta)),$$

a contradiction.

Since $v^*_1 > x^1_1(n+1)$, we must have $v^c_h(n+1) \leq v^*_h$ for any h by Corollary 1. By Claim 2, this implies that $v^c_h(m) \leq v^*_h$ for any $m \geq n+1$.
We now claim that for $m > n$, if the donor at step m is in category h, his/her income can not be equal to $x^h_{r^h_-(\alpha)}(n+1)$. Given the finiteness of the population, this will conclude the proof, because it will exclude the donor at stage n from donating again at a future step. Suppose, to the contrary, that there exists $m \geq n+1$ and $l_h \in \mathcal{N}(h)$ such that $v^c_1(m) = v^c_h(m) = x^h_{l_h}(m) = x^h_{r^h_-}(n+1)$, and that agent l_h transfers $\delta_0 > 0$ to agent $1 \in \mathcal{N}(1)$ at stage m. We then have

$$x^1_1(m) < x^h_{l_h}(m) = x^h_{r^h_-(\alpha)}(n+1) = v^c_1(m) \leq v^*_1.$$

Assume without loss of generality that δ_0 is small enough so that $x^1_1(m) \leq v_1(\delta_0) - \delta_0$. Then

$$P^{\mathbf{x}(m+1)}(\mathbf{v}(\delta_0)) - P^{\mathbf{x}(m)}(\mathbf{v}(\delta_0)) = P^{\mathbf{x}(n+1)^{\delta_0}}(\mathbf{v}(\delta_0)) - P^{\mathbf{x}(n+1)}(\mathbf{v}(\delta_0))$$

(both quantities are equal to $-\delta_0 + \max\{0, v_h(\delta_0) - (x^h_{r^h_-(\alpha)}(n+1) - \delta_0)\} - \max\{0, v_h(\delta_0) - x^h_{r^h_-(\alpha))}(n+1)\}$).
Since $P^{\mathbf{x}(m)}(\mathbf{v}(\delta_0)) \leq P^{\mathbf{x}(n+1)}(\mathbf{v}(\delta_0))$, we have

$$P^{\mathbf{x}(m+1)}(\mathbf{v}(\delta_0)) \leq P^{\mathbf{x}(n+1)^{\delta_0}}(\mathbf{v}(\delta_0)) < P^{\mathbf{x}}(v(\delta_0)),$$

a contradiction.

6.4.15 Claim 5

Let $n \geq n_2$. We proved already that the operation at stage n is necessarily a half transfer. Let h_0 be the category such that $v^c_1(n) = v^c_{h_0}(n) > v^c_{h_0+1}(n)$. Suppose by contradiction that the algorithm designates $i \in \mathcal{N}(1)$ to be the donor at stage n. By Proposition 1, it implies that $\forall h \in \{2, ..., h_0\}$, $\forall i \in \mathcal{N}(h)$, we must have $x^h_i(n) \neq v^c_1(n)$ because, otherwise, an agent in category $h > 1$ would be the donor. Note also that, by the very definition of a maximal transfer, we must have $x^1_i > v^c_1(n)$ for any $i \in \mathcal{N}(1)$ because assuming otherwise would make the transfer equalizing, which it can not be. Consequently the conditions of (P2) hold, which is a contradiction.

6.5 Proof of Theorem 2

$(i) \Rightarrow (ii)$ follows from the proof in the homogeneous settings, along with the fact that $U^{h'} \geq U^h$ for $h' \geq h$ to handle the categorical increments.
$(ii) \Rightarrow (iii)$ the proof of the corresponding implication in Section 6.1 can be applied here without any change.

$(iii) \Rightarrow (i)$ This implication is not trivial. Suppose that (iii) holds:

$$\Delta(\mathbf{v}) \leq 0 \;\; \forall \mathbf{v} \in \mathcal{V},$$

where $\Delta(\mathbf{v}) := \sum\limits_{h=1}^{k} \sum\limits_{i \in \mathcal{N}_{\mathbf{x}}(h)} \max(v_h - x_i^h, 0) - \sum\limits_{h=1}^{k} \sum\limits_{i \in \mathcal{N}_{\mathbf{y}}(h)} \max(v_h - y_i^h, 0)$. Then we claim that

$$\sum_{h=1}^{g} n_{\mathbf{y}}(h) \geq \sum_{h=1}^{g} n_{\mathbf{x}}(h), \;\; \forall g = 1, ..., k.$$

Let us prove this claim. Suppose that there exists some g^* such that $\sum_{h=1}^{g^*} n_{\mathbf{y}}(h) < \sum_{h=1}^{g^*} n_{\mathbf{x}}(h)$. For any $\alpha \geq \overline{v}(\mathbf{x}, \mathbf{y})$ let $\mathbf{v}(\alpha) \in \mathcal{V}$ be given by $v(\alpha)_h := \underline{v}(\mathbf{x}, \mathbf{y})$ for $h = g^* + 1, ..., k$ and $v(\alpha)_h = \alpha$ for $h = 1, ..., g^*$. We then have

$$\sum_{h=1}^{k} \sum_{i \in \mathcal{N}_{\mathbf{x}}(h)} \max(v(\alpha)_h - x_i^h, 0) = \sum_{h=1}^{g^*} \sum_{i \in \mathcal{N}_{\mathbf{x}}(h)} (\alpha - x_i^h) = \alpha \sum_{h=1}^{g^*} n_{\mathbf{x}}(h) - \sum_{h=1}^{g^*} \sum_{i \in \mathcal{N}_{\mathbf{x}}(h)} x_i^h$$

As a consequence

$$\Delta(\mathbf{v}(\alpha)) = \alpha \left(\sum_{h=1}^{g^*} n_{\mathbf{x}}(h) - \sum_{h=1}^{g^*} n_{\mathbf{y}}(h) \right) + \sum_{h=1}^{g^*} \left(\sum_{i \in \mathcal{N}_{\mathbf{y}}(h)} y_i^h - \sum_{i \in \mathcal{N}_{\mathbf{x}}(h)} x_i^h \right).$$

As α goes to $+\infty$, the right-hand side term goes to $+\infty$, which contradicts (iii), and proves the claim.

If $\sum_{h=1}^{g} n_{\mathbf{y}}(h) = \sum_{h=1}^{g} n_{\mathbf{x}}(h), \;\; \forall g = 1, ..., k$ then there is nothing to prove. Let us assume that this is not the case and let $h^* \in \{1, ..., k-1\}$ be the lowest category such that $\sum_{h=1}^{g} n_{\mathbf{y}}(h) > \sum_{h=1}^{g} n_{\mathbf{x}}(h)$. Let $\overline{\mathbf{y}}$ be the distribution obtained from $\mathbf{y}$ through an increment of agent $y_{n_{\mathbf{y}}(h^*)}^{h^*}$ to category $h^* + 1$. We claim that $\overline{\mathbf{y}} \precsim_{OPG} \mathbf{x}$. This will conclude the proof by induction. Let $\mathbf{v} \in \mathcal{V}$. If $v_{h^*} \leq y_{n_{\mathbf{y}}(h^*)}^{h^*}$ then $P^{\mathbf{x}}(\mathbf{v}) - P^{\overline{\mathbf{y}}}(\mathbf{v}) = P^{\mathbf{x}}(\mathbf{v}) - P^{\mathbf{y}}(\mathbf{v}) \leq 0$. If $v_{h^*} > y_{n_{\mathbf{y}}(h^*)}^{h^*}$ then

$$P^{\mathbf{y}}(\mathbf{v}) - P^{\overline{\mathbf{y}}}(\mathbf{v}) = v_{h^*} - y_{n_{\mathbf{y}}(h^*)}^{h^*} - \max\{0, v_{h^*+1} - y_{n_{\mathbf{y}}(h^*)}^{h^*}\}.$$

On the other hand, let $\tilde{\mathbf{v}} := (v_1, ..., \max\{y_{n_{\mathbf{y}}(h^*)}^{h^*}, v_{h^*+1}\}, v_{h^*+1}, ...v_k)$. Clearly $\tilde{\mathbf{v}} \in \mathcal{V}$ and

$$\begin{aligned}
P^{\mathbf{y}}(\mathbf{v}) - P^{\mathbf{x}}(\mathbf{v}) &\geq P^{\mathbf{y}}(\tilde{\mathbf{v}}) - P^{\mathbf{x}}(\tilde{\mathbf{v}}) + (n_{\mathbf{y}}(h^*) - n_{\mathbf{x}}(h^*))(v_{h^*} - \max\{y_{n_{\mathbf{y}}(h^*)}^{h^*}, v_{h^*+1}\}) \\
&\geq v_{h^*} - \max\{y_{n_{\mathbf{y}}(h^*)}^{h^*}, v_{h^*+1}\} \\
&= v_{h^*} - y_{n_{\mathbf{y}}(h^*)}^{h^*} - \max\{0, v_{h^*+1} - y_{n_{\mathbf{y}}(h^*)}^{h^*}\} \\
&= P^{\mathbf{y}}(\mathbf{v}) - P^{\overline{\mathbf{y}}}(\mathbf{v})
\end{aligned}$$

Consequently $P^{\overline{\mathbf{y}}}(\mathbf{v}) \geq P^{\mathbf{x}}(\mathbf{v})$, which concludes the proof.
$(iii) \Leftrightarrow (iv)$: the proof of the corresponding implication in Section 3.2 can be applied here without any change. ■

References

Atkinson, A. B., and F. Bourguignon (1982): "The Comparison of Multi-dimensioned Distribution of Economic Status," *Review of Economic Studies*, 49, 183–201.

——— (1987): "Income Distributions and Differences in Needs," in *Arrow and the Foundation of the Theory of Economic Policy*, ed. by G. R. Feiwel. Macmillan, London.

Bazen, S., and P. Moyes (2003): "Comparisons of income distributions accross heterogenous populations," *Research on Economic Inequality*, 9, 85–115.

Bourguignon, F. (1989): "Family Size and Social Utility: Income Distribution Dominance Criteria," *Journal of Econometrics*, 42, 67–80.

Dalton, H. (1920): "The Measurement of the inequality of incomes," *Economic Journal*, 30, 348–361.

Dasgupta, P., A. K. Sen, and D. Starrett (1973): "Notes on the Measurement of Inequality," *Journal of Economic Theory*, 6, 180–187.

Decancq, K. (2012): "Elementary Multivariate Frechet Rearrangements and Stochastic Dominance on a Frechet Class," *Journal of Economic Theory*, 147, 1450–1459.

Decoster, A., and E. Ooghe (2006): "A Bounded Index Test to Make Robust Heterogenous Welfare Comparisons," *Review of Income and Wealth*, 52, 361–376.

Ebert, U. (1997): "Social Welfare when Needs Differ: An Axiomatic Approach," *Economica*, 64, 233–244.

Epstein, L., and S. M. Tanny (1980): "Increasing Generalized Correlation: A definition and some economic consequences," *Canadian Journal of Economics*, 13, 16–34.

Fields, G., and J. Fei (1978): "On inequality Comparisons," *Econometrica*, 46, 305–316.

Fleurbaey, M., C. Hagneré, and A. Trannoy (2003): "Welfare Comparisons with Bounded Equivalence Scales," *Journal of Economic Theory*, 110, 309–336.

Gravel, N., and P. Moyes (2012): "Ethically Robust Comparisons of bi-dimensionnal Distributions with an Ordinal Attribute," *Journal of Economic Theory*, 147, 1384–1426.

Gravel, N., P. Moyes, and B. Tarroux (2009): "Robust International Comparisons of Distributions of Disposable Income and Access to Regional Public Goods," *Economica*, 76, 432–461.

Hadar, J., and W. Russell (1974): "Stochastic Dominance in Choice under Uncertainty," in *Essays on Economics Behavior under Uncertainty*, ed. by M. S. Balch, D. L. McFadden, and S. Wu. North Holland, Amsterdam, UK.

Hardy, G. H., J. E. Littlewood, and G. Polya (1952): *Inequalities, 2nd edition.* Cambridge University Press, Cambridge, UK.

Jenkins, S. P., and P. J. Lambert (1993): "Ranking Income Distributions when Needs Differ," *Review of Income and Wealth*, 39, 337–356.

Kolm, S. C. (1969): "The Optimal Production of Social Justice," in *Public Economics*, ed. by H. Guitton, and J. Margolis. Macmillan, London.

Lehmann, E. L. (1955): "Ordered Family of Distributions," *The Annals of Mathematics and Statistics*, 26, 399–419.

Levhari, D., J. Paroush, and B. Peleg (1975): "Efficiency Analysis for Multivariate Distributions," *Review of Economic Studies*, 42, 87–91.

Marinacci, M., and L. Montrucchio (2005): "Ultramodular functions," *Mathematics of Operation Research*, 30, 311–332.

Moyes, P. (2012): "Comparaisons of Heterogeneous Distributions and Dominance Criteria," *Journal of Economic Theory*, 147, 1351–1383.

Muller, A., and M. Scarsini (2012): "Fear of Loss, Inframodularity and Transfers," *Journal of Economic Theory*, 147, 1490–1500.

Osterdal, L. P. (2010): "The Mass Transfer Approach to Multivariate Discrete First Order Stochastic Dominance: Direct Proof and Implications," *Mathematical Social Sciences*, 46, 1222–1228.

Sen, A. K. (1973): *On Economic Inequality.* Oxford, Clarendon.

Tchen, A. H. (1980): "Inequalities for Distributions with Given Marginals," *Annals of Probability*, 8, 814–827.

Tsui, K. Y. (1999): "Multidimensional Inequality and Multidimensionnal Generalized Entropy Measures: an Axiomatic Derivation," *Social Choice and Welfare*, 16, 145–157.