Ethically Robust Comparisons of Bidimensional Distributions with an Ordinal Attribute*

Patrick MOYES

GREThA, CNRS, UMR 5113
Université de Bordeaux

&

Nicolas GRAVEL

Aix-Marseille University and AMSE (GREQAM)

Comparaisons Ethiquement Robustes de Distributions Bidimensionnelles avec un Attribut Ordinal

Résumé

Nous proposons des bases pour l'évaluation normative de distributions de deux attributs, le premier étant cardinalement mesurable et transférable entre les individus et le second ordinal et non-transférable. Le résultat que nous établissons constitue en quelque sorte l'analogue du théorème obtenu par Hardy, Littlewood et Pólya (1934) pour les distributions d'un attribut cardinal. Plus précisément, nous identifions les transformations des distributions qui garantissent que le bien-être social augmente pour l'unanimité utilitariste dès lors que la fonction d'utilité est concave dans l'attribut cardinal et que son utilité marginale par rapport au même attribut est décroissante par rapport à l'attribut ordinal. Nous établissons que ce classement à l'unanimité des distributions est équivalent au classement obtenu sur la base du critère des écarts de pauvreté ordonnés proposé par Bourguignon (1989). Enfin, nous montrons que, si une distribution domine une autre distribution selon le critère des écarts de pauvreté ordonnés, alors la première peut être obtenue à partir de la seconde au moyen d'une suite finie de telles transformations.

Mots-clés : *Analyse Normative, Utilitarianisme, Dominance Stochastique Bidimensionnelle, Transformations Égalisantes*

Ethically Robust Comparisons of Bidimensional Distributions with an Ordinal Attribute

Abstract

We provide foundations for robust normative evaluation of distributions of two attributes, one of which is cardinally measurable and transferable between individuals and the other is ordinal and non-transferable. The result that we establish takes the form of an analogue to the standard Hardy, Littlewood, and Pólya (1934) theorem for distributions of one cardinal attribute. More specifically, we identify the transformations of the distributions which guarantee that social welfare increases according to utilitarian unanimity provided that the utility function is concave in the cardinal attribute and that its marginal utility with respect to the same attribute is non-increasing in the ordinal attribute. We establish that this unanimity ranking of the distributions is equivalent to the Bourguignon (1989) ordered poverty gap quasi-ordering. Finally, we show that, if one distribution dominates another according to the ordered poverty gap criterion, then the former can be derived from the latter by means of an appropriate and finite sequence of such transformations

Keywords: Normative Analysis, Utilitarianism, Bidimensional Stochastic Dominance, Inequality Reducing Transformations

JEL: D31, D63, I32.

1. Introductory Remarks

The normative foundations of the comparison of distributions of a single attribute between a given number of individuals are by now well-established. They originate in the equivalence between three statements that are considered relevant answers to the question of when a distribution **x** can be considered normatively better than a distribution **y**. Given two distributions **x** and **y** with equal means, these statements, the equivalence of which was first established by Hardy, Littlewood, and Pólya (1934) and popularised later on among economists by Kolm (1969), Atkinson (1970), Dasgupta, Sen, and Starrett (1973), Sen (1973), Fields and Fei (1978) among others, are the following:

(a) Distribution **x** can be obtained from distribution **y** by means of a finite sequence of progressive – or equivalently Pigou-Dalton – transfers.

(b) All utilitarian ethical observers who assume that individuals convert the attribute into well-being by means of the same non-decreasing and concave utility function rank distribution **x** above distribution **y**.

(c) The Lorenz curve of distribution **x** lies nowhere above and somewhere below that of **y**, or equivalently, for all poverty lines, the poverty gap is no greater in distribution **x** than in distribution **y** and it is smaller for at least one poverty line.

This remarkable result, which can be generalised in a number of ways, points to three different aspects of the inequality measurement process.[1] The first statement aims at capturing the very *notion of inequality* reduction by associating it with elementary transformations of the distributions. The second statement is fundamentally normative and it assumes that society has an *aversion to inequality* which, in the utilitarian framework, is reflected by the concavity of the utility function. To some extent the first statement helps in clarifying the meaning of the restriction imposed on the utility function in the second statement. While these two conditions shed light on two different facets of the inequality concept, they do not prove very useful for deciding in practice when one distribution is more unequal than another. This is particularly true for the second statement which requires an infinite number of comparisons to be made before a distribution can be declared less unequal than another. The third statement resolves this problem by providing easily *implementable criteria* that allow one to recover the ranking of distributions implied by the first and second statements.

The last condition can also been interpreted as a means of identifying those pairs of distributions for which a consensus prevails among all inequality averse utilitarian ethical observers. If the Lorenz curves – or equivalently the poverty gap profiles – of two distributions do not intersect, then all the utilitarian ethical observers who have some aversion to inequality will rank these distributions in the same way. If, on the contrary, the Lorenz curves of two distributions

---

[1] As shown by Dasgupta *et al.* (1973), the equivalence continues to hold when more flexible social welfare functions are substituted for the utilitarian one. It is also possible to extend this result to the case where the distributions under comparison have differing means by adapting appropriately each of these three conditions (see Kolm (1969) and Shorrocks (1983)).

cross, then it is always possible to find two such ethical observers who will rank these distributions in the opposite way. This points to what may be considered an inherent limitation of this approach: it is very unlikely that the above conditions will allow one to rank conclusively *all* the distributions under consideration. Actually this pessimistic view is exaggerated and there is empirical evidence that the comparison of the Lorenz curves permits conclusive verdicts to be obtained in a significant number of cases.[2] Nevertheless, the Lorenz criterion is often considered to be a first round approach which must be supplemented by the use of ethically more demanding indices in a second stage. It is traditional practice to require that these summary indices be compatible with any of the conditions of the Hardy-Littlewood-Pólya result (see, e.g., Foster (1985)).

Important as they are, these foundations strictly concern distributions of a single attribute which is typically identified with individual income. However the ability of income alone to measure a person's well-being has been seriously challenged during the last thirty years and there has been an increasing concern for a more comprehensive approach. The focus on income is to a large extent justified by the assumption that it provides a good measure of the level of well-being achieved by an individual who behaves rationally in a comprehensive and fully competitive market environment. This neglects the fact that a number of commodities that contribute to a person's well-being cannot be given market values. Typical instances are amenities like recreational areas or publicly provided services like education or health care, for which markets are imperfect or even do not exist. Also, attributes such as family circumstances or health status affect a person's well-being but cannot be related to income exclusively. The recognition that a person's well-being cannot be fully summarised by income alone calls for a multidimensional approach to welfare and inequality measurement. While the literature has produced, in the last thirty years or so, a significant number of contributions that have formulated various criteria for comparing alternative distributions of multiple attributes, none of them has established an equivalence between elementary transformations, utilitarian unanimity over a class of utility functions and an empirically implementable criterion. We find it useful to describe the literature on multidimensional normative evaluation as having followed three different routes.

The first route is to impose particular conditions on the utility function that are assumed to capture the aversion to multidimensional inequality and to explore their implications for the ranking of the situations under comparison. The main purpose is to find implementable dominance tests that permit one to check if it is possible to reach a consensus among these well-defined classes of social welfare functions. Building on the results of Hadar and Russell (1974) in the multidimensional risk literature, Atkinson and Bourguignon (1982) prove that first and second order bidimensional stochastic dominance imply unanimity of judgements among all utilitarians for specific classes of utility functions. In particular, first order stochastic

---

[2] Sophisticated statistical inference techniques have been designed in order to test for the robustness of the ranking of distributions based on the Lorenz – and more generally on the stochastic dominance – quasi-orderings (see Beach and Davidson (1983), Bishop, Chakraborti, and Thistle (1989), Anderson (1996), Bishop and Formby (1999), Davidson and Duclos (2000), among others).

dominance is shown to be sufficient for utilitarian unanimity provided that the utility functions are non-decreasing in both attributes and submodular. They also suggest that an equivalence between their multidimensional stochastic dominance criteria and utilitarian unanimity might hold. Atkinson and Bourguignon (1987) provide an appealing interpretation of their stochastic dominance criteria in the particular case where one attribute is household income and the other one an ordinal index of needs such as the household size. Their criteria and equivalence results, developed originally for distributions of attributes with an identical marginal distribution of needs, have been subsequently extended by Jenkins and Lambert (1993) and Bazen and Moyes (2003) to more general situations.

Decancq (2011) proposes an extension of the first order stochastic dominance test of Atkinson and Bourguignon (1982) to the case where there are more than two attributes and where all attributes are only ordinally measurable. Generalising results obtained by Epstein and Tanny (1980) and Tchen (1980) for bivariate distributions to arbitrary multivariate ones, he establishes an equivalence between his criterion and expected utility unanimity over the class of submodular utility functions. He also shows that his criterion is equivalent to the possibility of going from the dominated distribution to the dominating one by a finite sequence of *multivariate rearrangements*. However, Decancq's results are derived in the context of decision making under uncertainty rather than in that of social choice. The properties of the utility function that guarantee that utilitarian unanimity is consistent with second order stochastic dominance are quite restrictive and in addition their meaning is not that clear. Bourguignon (1989) proposes an interesting empirically implementable criterion that lies in between the first and second order dominance criteria of Atkinson and Bourguignon (1982). He further proves that the ranking of distributions agreed to by all utilitarians coincides with that implied by his criterion provided that the utility functions are non-decreasing and concave in income and that the marginal utility of income is non-decreasing with needs.[3] However, he does not provide any indication about what the elementary transformations of the distribution that correspond to his dominance criterion are.

The second route, initiated by Kolm (1977), is built precisely with the aim of introducing elementary transformations of the distributions in a multidimensional context, and in identifying the properties of the individual utility function over which a utilitarian unanimity coincides with such transformations. Hence, Kolm (1977) considers the multiplication of the distribution by a bistochastic matrix as the appropriate generalisation of the Pigou-Dalton transfer to the multidimensional framework. He proves that, if a situation is obtained from another one by means of such a transformation, then the former situation is ranked above the latter by all the utilitarian ethical observers who evaluate the individuals' well-being by means of a concave utility function, and conversely. Kolm's approach avoids the difficulty inherent in Atkinson and Bourguignon (1982) by making clear what is meant by inequality reduction and then by identifying the normative judgements consistent with it. On the other hand, the elementary transformation he considers is somewhat specific: it consists in transferring an

---

[3] Fleurbaey, Hagneré, and Trannoy (2003) show that it is possible to refine the Bourguignon (1989) criterion by introducing equivalence scales and by allowing these to vary within predefined intervals.

identical fraction of each attribute from one individual to another. More importantly, he does not identify a criterion that would implement the utilitarian unanimity criterion for the class of concave utility functions.

If one allows the fractions transferred to vary across attributes and at the same time requires that the beneficiary of the transfers is poorer then the donor in all dimensions, then one is lead to the type of transformation considered by Müller and Scarsini (2011). These authors show that these intuitively equalising transformations can – under suitable conditions – be decomposed into two more basic transformations, where the first one decreases correlation without altering the marginal distributions of the attributes, while the second one reduces dispersion in one attribute without affecting the correlation between the attributes. As it turns out, the transformations proposed by Müller and Scarsini are generalisations – to the case of an arbitrary number of cardinally measurable attributes – of the elementary transformations considered in this paper and further discussed below. Müller and Scarsini (2011) (see also Meyer and Strulovici (2010)) show that, if one distribution is obtained from another by means of a finite sequence of their equalising transformations, then the former is preferred to the latter by all utilitarians whose utility functions are concave in all dimensions and submodular, and conversely.[4] However, while the equalising process considered by Müller and Scarsini (2011) is more attractive than the transformations of Kolm (1977), it must be admitted that an analogue to the stochastic dominance tests of Atkinson and Bourguignon (1982) is still missing.

The third route is followed by Koshevoy (1995), who suggests the use of the Lorenz zonotope as a generalisation of the Lorenz curve in the multidimensional framework. Among other things, he proves that a sufficient condition for one distribution to dominate a second according to his criterion – the Lorenz zonotope of the former distribution is included in the Lorenz zonotope of the latter – is that the former can be obtained from the latter through multiplication by a bistochastic matrix. However, this does not tell us a lot about the implicit equalisation process embedded in the Lorenz zonotope criterion since the converse statement does not hold. Furthermore, Koshevoy (1995) does not provide any indication about what properties of the utility function would guarantee that the ranking of distributions generated by unanimity among all utilitarian ethical observers coincide with that implied by the Lorenz zonotope criterion. Nor does he give evidence that the welfarist approach to inequality measurement is compatible with the Lorenz zonotope quasi-ordering.[5] While the welfare criterion considered by Kolm implies Koshevoy's quasi-ordering, the converse implication does not hold, and it is therefore difficult to justify the use of the Lorenz zonotope from a normative standpoint.

---

[4] We refer the reader to Marinacci and Montrucchio (2005) where the relationships between the notions of component-wise concave and submodular functions and those of inframodular functions are discussed.

[5] Rigorously speaking, the fact that the Lorenz zonotope cannot be given a utilitarian – and possibly a welfarist – justification does not mean that there exists no social welfare functions consistent with it. To the best of our knowledge, there does not seem to exist a proof of this claim and it is still an open question whether the Lorenz zonotope can be rationalised by means of a social welfare function.

This brief examination of the literature indicates that the best that has been achieved to date are equivalences between a criterion based on unanimity among those utilitarian ethical observers who subscribe to some common values and *either* (i) an implementable criterion (Atkinson and Bourguignon (1982, 1987), Decancq (2011), Bourguignon (1989)) *or* (ii) specific equalising transformations (Kolm (1977), Müller and Scarsini (2011), Meyer and Strulovici (2010, 2011)). While the literature has proposed different generalisations of the notion of a progressive transfer, it must be admitted that no decisive step has been made by proving that if one distribution is ranked above another distribution by some *easily* implementable criterion, then it is possible to obtain the dominating distribution from the dominated one by successive applications of such transformations. As a consequence, and with the exception of Decancq (2011), Epstein and Tanny (1980) and Tchen (1980) in the specific context of decision making under uncertainty, none of these studies has succeeded in providing an equivalence between the analogues of statements (a), (b) and (c) in a multidimensional setting up to now. Hence, it appears, that, despite the relative wealth of attempts made in this direction during the last thirty years, there are no multidimensional analogues to the Hardy-Littlewood-Pólya equivalence result.

In this paper, we aim at providing a step towards establishing such an equivalence for the case where there are only two attributes to be distributed, one of which has a cardinal nature while the other is only ordinally measurable, as in Atkinson and Bourguignon (1987) and Bourguignon (1989). As far as the normative evaluation is concerned, we follow the standard practice which consists in endowing each ethical observer with a utilitarian social welfare function and in requiring unanimity of judgements among all observers. Certainly, the utilitarian rule is a particular method for making distributive judgements that shows strictly-speaking no direct consideration for distributive justice. What matters for a utilitarian ethical observer is the sum of the individuals' utilities and not the way these utilities are distributed among the population. However, the utilitarian rule is a particularly convenient device and, contrary to what intuition suggests, it is more flexible than it might look at first sight.

Focusing on bidimensional distributions may look restrictive, but it provides a simple enough structure in which the basic problem we are interested in can be addressed. Our somewhat asymmetric treatment of the two attributes may seem more disputable, but there are a number of instances where the available information about several attributes of interest is only ordinal in nature. For instance, one may be interested in the evaluation of income distributions for households who differ in size and composition, as in Atkinson and Bourguignon (1987) or Bourguignon (1989). Even though precise information concerning household composition such as the number and ages of the family's members is available, one may be reluctant to assign a cardinal meaning to these figures. A similar situation arises when one has to compare the well-being of different populations on the basis of the distributions of the income and health status of their members. Indicators like infant mortality or life expectancy are routinely used for measuring a person's health status and, here again, one might be willing to assume that the only reliable information is of an ordinal nature (see for instance Allison and

Foster (2004)).[6] The consumption of local public goods such as the quality of schooling (see Gravel, Moyes, and Tarroux (2009)) is another example of an attribute to which one might hesitate to assign a cardinal meaning. On the other hand, there is wide agreement among economists about the fact that individual income per year – at least when used in a specific price configuration – is a cardinally meaningful attribute. It is therefore not unreasonable to consider two attributes that cannot be defined with the same degree of precision.

As for elementary transformations of the distributions, it seems quite natural to assume that, other things equal, a progressive transfer in the cardinal attribute between two individuals who have the same endowment of the ordinal attribute reduces inequality. For the sake of consistency with the unidimensional approach, a *within-type progressive transfer* – as we will call it henceforth – is the first equalising transformation we consider. The second transformation – which we refer to as a *favourable permutation* – is implicitly contained in Atkinson and Bourguignon (1982) and explicitly used by Müller and Scarsini (2011), Meyer and Strulovici (2010, 2011) and Decancq (2011). It consists in giving the cardinal attribute endowment of the better-off individual in both attributes to the worse-off individual in both attributes and *vice versa*. This permutation of the endowments in the cardinal attribute actually amounts to reducing the pairwise correlation – or positive association – existing between the two attributes. The favourable permutation is probably the most controversial elementary transformation we consider because it is not clear how inequality in the distribution of the two attributes is reduced as a result. Actually, it is the inequalities of well-being between the two individuals involved in the favourable permutation that are reduced. More precisely, it can be established that, for any increasing utility function, the well-being of the poorer individual increases while that of the richer individual decreases, thereby bringing them closer together on the utility scale. This consequence of a favourable permutation is reminiscent of the Hammond (1976) *equity condition* in an ordinal context and it is in this respect that it may be considered to be an equalising transformation. The *between-type progressive transfer*, which is a natural extension of a progressive transfer in our particular framework, is our third transformation. It consists in transferring an amount of the cardinal attribute from a better-off individual in both attributes to a worse-off individual in both attributes in such a way that the beneficiary of the transfer is not made richer than the donor in the cardinal attribute. Contrary to the bistochastic transformation of Kolm (1977), where the same equalising process is applied to both attributes, a between-type progressive transfer entails redistribution in only one dimension. Clearly, a between-type progressive transfer is a particular case of the equalising transformations considered by Müller and Scarsini (2011).

These three transformations are elementary in the sense that it seems difficult to conceive of simpler inequality-reducing operations into which they could be decomposed. In fact, this is not perfectly true as far as between-type progressive transfers are concerned and one may argue they are not as elementary as they might look at first glance. As will be seen below – and as was shown by Müller and Scarsini (2011) – it is possible to decompose a between-

---

[6] An important issue, particularly stressed by Allison and Foster (2004), concerns the implication for the ranking of the situations under comparisons of a change in the measurement scale of an attribute.

type progressive transfer into a within-type progressive transfer coupled with a favourable permutation. But, for such a decomposition to be made, an important qualification is needed: one has to add a phantom individual with appropriate endowments in both attributes. As will become clear later, this phantom individual is only instrumental and she has no impact on the ranking of the distributions.[7]

Turning finally to the question of the implementation of the unanimity of value judgements criterion, we focus on two dominance criteria that have been considered separately in the literature. The first criterion, proposed by Atkinson and Bourguignon (1982), is the one that has received more attention in the literature. It declares that one situation dominates another situation if the graph of the joint distribution function of the former situation lies nowhere above that of the latter. This condition is actually the bidimensional version of the poverty measurement approach based on the headcount ratio: poverty, as measured by the percentage of individuals who fall below predetermined levels of the two attributes, is less in the first situation than in the second, for all possible values of the attributes' poverty lines. The Bourguignon (1989) ordered poverty gap quasi-ordering is the second implementable criterion that we examine in this paper. The poverty gap is computed for the cardinal attribute by assigning to each individual a poverty line that depends negatively on her endowment of the ordinal attribute. Other things equal, it is more difficult for someone in good health or living in an area with good public facilities to be considered deprived in income than for someone handicapped or living in the slums.

The organisation of the paper is as follows. We introduce in Section 2 our bidimensional model along with our notation and preliminary definitions. In Section 3 we present the normative criterion, which will be used by the ethical observer in order to rank the distributions under comparison. Section 4 is concerned with the first criterion due to Atkinson and Bourguignon (1982) and it provides the analogue to the Hardy-Littlewood-Pólya result in this case. More precisely, we show that the domination of one situation over another is equivalent to requiring unanimity of value judgements among all utilitarian ethical observers whose marginal utility of the cardinal attribute is non-negative and decreasing in the level of the ordinal attribute. This implies in turn that the dominating distribution can be obtained from the dominated one by means of a finite sequence of favourable permutations of the cardinal attribute.

Our main result is the subject of Section 5, which examines the normative foundations of the Bourguignon (1989) ordered poverty gap dominance criterion. Introducing the additional requirement that the utility function is concave in the cardinal attribute for fixed levels of the ordinal attribute is enough to guarantee that the rankings of situations implied by the ordered poverty gap criterion and unanimity among all utilitarian ethical observers is identical. Exploiting the possibility of adding phantom individuals to the original population allows us to show that it is possible to derive the dominating situation from the dominated one by means of a finite sequence of within-type progressive transfers followed by a finite sequence of favourable permutations, and conversely. Our results have been derived under the assumptions

---

[7] Things would be different if, for instance, we were interested in the absolute and average welfare in the society, for in this case the introduction of this phantom individual would no longer be neutral.

that the distributions under comparison have the same mean income and that the marginal distribution of the ordinal attribute is fixed.

The utilitarian principle, upon which we drew when establishing our results, has been criticised on the ground that it is insensitive to the way utility is distributed among the society's members. We indicate in Section 6 how our results can be generalised when the distributions under comparison differ with respect to both mean income and the distribution of the ordinal variable, and when the ethical observers appeal to normative principles that might appear more acceptable than the utilitarian rule. Section 7 summarises the main results and hints at possible avenues for future research. In order to keep technicalities to a minimum in the course of the exposition, the proofs of the lemmas and propositions are relegated to Section 8.

2. Notation and Preliminary Definitions

We consider a finite *population* or *society* consisting of n individuals ($n \geqq 2$), where each individual is endowed with two attributes, the first of which is assumed to be cardinal while the second is ordinal. In the case of a cardinal attribute, it makes sense to compare levels as well as differences of individual endowments, while only the first type of comparison is admissible when the attribute is ordinal. In other words, an affine transformation of the values taken by the attribute has no impact on the information provided by a cardinal variable, whereas arbitrary increasing transformations leave the information unchanged when the variable is ordinal. These assumptions can be viewed as either describing the intrinsic nature of the attribute or reflecting our perception of the attribute, in which case there may not be full agreement regarding the type of measurability attached to the attribute. It is important for our purpose to note that the type of measurability that one assumes has important consequence for the analysis as it restricts the range of operations that can be considered and as a consequence the list of properties of the functions which have the attribute as an argument. For instance, while it makes sense to transfer amounts of the cardinal attribute between individuals, we assume throughout that such operations are meaningless in the case of the ordinal attribute. To simplify the exposition, we identify the cardinal attribute with *income* and the ordinal attribute with *health*. Given the discussion above, the latter must not be taken too literally and it should rather be interpreted as a means of ordering individuals and defining their *types* in terms of health status.[8]

A *bidimensional distribution*, or more compactly a *situation*, for a population of n individ-

---

[8] Other examples of the ordinal attribute include the consumption of local public goods (Gravel *et al.* (2009)) or family size in the case of households (Atkinson and Bourguignon (1987), Bourguignon (1989), among others). In the latter case, family size is interpreted as an indicator of household needs and the ordinal attribute is a decreasing function of neediness.

uals is a $n \times 2$ matrix

$$(2.1) \qquad \mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) := \begin{bmatrix} x_1 & a_1 \\ \vdots & \vdots \\ x_i & a_i \\ \vdots & \vdots \\ x_n & a_n \end{bmatrix} = \begin{bmatrix} s_1 \\ \vdots \\ s_i \\ \vdots \\ s_n \end{bmatrix},$$

such that $s_i = (x_i, a_i)$ fully describes individual i, where $x_i \in \mathscr{D} := [\underline{v}, \overline{v}] \subset \mathbb{R}$ and $a_i \in \mathscr{A} := \{1, 2, \ldots, H\}$ are respectively the income and the health status of individual i. As can be noticed, the asymmetry in the treatment of the two attributes is reemphasised by our assumption that the cardinally measurable attribute takes continuously many different values while the ordinally measurable one only takes a finite number of values. Yet there is nothing fundamental in that distinction and we could equally well have presented our results by assuming either a finite number of possible values for both attributes or continuously many values for both. The (marginal) *distributions of income* and *health* in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a})$ are indicated respectively by $\mathbf{x} := (x_1, \ldots, x_n) \in \mathscr{D}^n$ and $\mathbf{a} := (a_1, \ldots, a_n) \in \mathscr{A}^n$.[9] The general set of situations for a population of n individuals is denoted as

$$(2.2) \qquad \mathscr{S}_n := \{ \mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \mid (x_i, a_i) \in \mathscr{D} \times \mathscr{A}, \ \forall \, i = 1, 2, \ldots, n \},$$

and we let $\mu(\mathbf{x})$ represent the *mean income* in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a})$.

We will make extensive use in what follows of the representation of the bidimensional distributions by means of their associated joint, marginal and conditional cumulative distribution functions. Given the situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \in \mathscr{S}_n$, we denote by $N(y, h) = \{ i \mid x_i = y$ and $a_i = h \}$ the set of individuals who receive an income equal to y and who have a health status equal to h. The *joint density function* of $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a})$ is given by

$$(2.3) \qquad f(y, h) = n(y, h)/n, \ \forall \, y \in \mathscr{D}, \ \forall \, h \in \mathscr{A},$$

where $n(y, h) = \#N(y, h)$. Similarly, we indicate by $Q(y, h) = \{ i \mid x_i \leq y$ and $a_i \leq h \}$ the set of individuals who receive an income no greater than y and whose health status is no greater than h. Letting $q(y, h) = \#Q(y, h)$, the *joint cumulative distribution function* is defined by

$$(2.4) \qquad F(y, h) = q(y, h)/n, \ \forall \, y \in \mathscr{D}, \ \forall \, h \in \mathscr{A}.$$

The set of individuals whose income in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a})$ is equal to y is denoted as $N_1(y) = \{ i \mid x_i = y \}$, while $N_2(h) = \{ i \mid a_i = h \}$ indicates the set of individuals whose health status is equal to h. The *marginal density functions of income* and *health* are respectively defined by

$$(2.5) \qquad f_1(y) = n_1(y)/n \ \text{ and } \ f_2(h) = n_2(h)/n, \ \forall \, y \in \mathscr{D}, \ \forall \, h \in \mathscr{A},$$

---

[9] To be rigourous we should write $\mathbf{x} := (x_1, \ldots, x_n)^T$ and $\mathbf{a} := (a_1, \ldots, a_n)^T$, where the superscript "T" denotes the transposed vector. Since there is little risk of confusion, we will omit throughout the transposition operator to lighten the notation.

where $n_1(y) = \#N_1(y)$ and $n_2(h) = \#N_2(h)$. The set of individuals whose income in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a})$ is no greater than y is indicated by $Q_1(y) = \{i \mid x_i \leqq y\}$, while the set of individuals whose health status is no greater than h is denoted as $Q_2(h) = \{i \mid a_i \leqq h\}$. The *marginal distribution functions of income* and *health* are respectively defined by

$$(2.6) \qquad F_1(y) = q_1(y)/n \ \text{ and } \ F_2(h) = q_2(h)/n, \ \forall \ y \in \mathscr{D}, \ \forall \ h \in \mathscr{A},$$

where $q_1(y) = \#Q_1(y)$ and $q_2(h) = \#Q_2(h)$. The *conditional density functions of income* and *health* are indicated by

$$(2.7) \qquad f_1(y \mid h) = n(y, h)/n_2(h) \ \text{ and } \ f_2(h \mid y) = n(y, h)/n_1(y), \ \forall \ y \in \mathscr{D}, \ \forall \ h \in \mathscr{A},$$

respectively. Let $Q_1(y \mid h) = \{i \mid x_i \leqq y \text{ and } a_i = h\}$ represent the set of individuals whose health status is equal to h and income no greater than y. Similarly, the set of individuals with income equal to y and health status no greater than h is denoted by $Q_2(h \mid y) = \{i \mid a_i \leqq h \text{ and } x_i = y\}$. The *conditional distribution functions of income* and *health* are defined by

$$(2.8) \qquad F_1(y \mid h) = q_1(y \mid h)/n_2(h) \ \text{ and } \ F_2(h \mid y) = q_2(h \mid y)/n_1(y), \ \forall \ y \in \mathscr{D}, \ \forall \ h \in \mathscr{A},$$

respectively, where $q_1(y \mid h) = \#Q_1(y \mid h)$ and $q_2(h \mid y) = \#Q_2(h \mid y)$.

Throughout this paper, we are interested in the comparison of situations $\mathbf{s}^\circ \equiv (\mathbf{x}^\circ; \mathbf{a}^\circ), \mathbf{s}^* \equiv (\mathbf{x}^*; \mathbf{a}^*) \in \mathscr{S}_n$. The associated joint, marginal and conditional density functions and distribution functions of $\mathbf{s}^*$ and $\mathbf{s}^\circ$ will be identified by means of the corresponding superscripts. Because the most difficult part of our results concerns the way the attributes are distributed among the individuals, we restrict our attention in the first part of the paper to the comparisons of situations whose marginal distribution functions of health are identical and mean incomes equal.

Before we examine the normative criteria that we will rely on for passing welfare judgements, we would like to insist on the implications of our informational constraints for these assessments which we alluded to above. Let the situations $\mathbf{s}^\circ \equiv (\mathbf{x}^\circ; \mathbf{a}^\circ)$, $\mathbf{s}^* \equiv (\mathbf{x}^*; \mathbf{a}^*)$, $\tilde{\mathbf{s}}^\circ \equiv (\tilde{\mathbf{x}}^\circ; \tilde{\mathbf{a}}^\circ)$ and $\tilde{\mathbf{s}}^* \equiv (\tilde{\mathbf{x}}^*; \tilde{\mathbf{a}}^*)$ be such that

$$(2.9) \qquad \tilde{x}_i^* = \alpha + \beta x_i^*, \ \tilde{x}_i^\circ = \alpha + \beta x_i^\circ, \ \tilde{a}_i^* = \psi(a_i^*) \ \text{ and } \ \tilde{a}_i^\circ = \psi(a_i^\circ), \ \forall \ i = 1, 2, \ldots, n,$$

where $\alpha \in \mathbb{R}$, $\beta > 0$ and ψ is increasing. Assuming that income is cardinally measurable and health ordinally measurable amounts to considering that the pairs $\{\mathbf{s}^*, \mathbf{s}^\circ\}$ and $\{\tilde{\mathbf{s}}^*, \tilde{\mathbf{s}}^\circ\}$ convey the same information. Then, all the normative criteria $\geq_J$ we will consider must have the property that

$$(2.10) \quad \forall \ \mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n : (\mathbf{x}^*; \mathbf{a}^*) \geq_J (\mathbf{x}^\circ; \mathbf{a}^\circ) \Longleftrightarrow (\alpha\mathbf{1} + \beta\mathbf{x}^*; \psi(\mathbf{a}^*)) \geq_J (\alpha\mathbf{1} + \beta\mathbf{x}^\circ; \psi(\mathbf{a}^\circ)),$$

whenever $\alpha \in \mathbb{R}$, $\beta > 0$ and ψ is increasing, and where $\mathbf{1} := (1, \ldots, 1) \in \mathbb{R}^n$. In other words the normative criteria $\geq_J$ are invariant with respect to particular modifications of the measurement scales of the attributes. This informational constraint is not innocuous: for instance, the standard relative Lorenz quasi-ordering, which is widely used for making inequality comparisons in the single attribute case, is not invariant with respect to affine transformations of the variable (see, e.g., Moyes (1994, Proposition 3.1)).

3. The Normative Criteria

Following the usual practice in the dominance approach, we assume that (i) all individuals transform their endowments of the two attributes into well-being by means of the same utility function and (ii) the distribution of the individual utilities in a situation provides all the relevant information for appraising this situation from a normative point of view. It is convenient to think of an *ethical observer* who is in charge of evaluating the different situations being compared on the basis of the distributions of utilities they generate. This way of proceeding is compatible with the *welfarist* approach according to which the utility function is the one actually used by the individuals in order to convert their endowments of the attributes into well-being.[10] Here, well-being is assumed to be a cardinally measurable and interpersonally comparable variable that summarises all the aspects of an individual's situation that are deemed relevant for normative evaluation. But the utility function can also be viewed as reflecting the assessment of the individual's situation by the ethical observer as it is typically assumed in the *non-welfarist* approach. Then, there is no presumption that utility is connected to the individual's actual well-being and one must rather interpret the utility function as a predefined *social norm*. It is immaterial for our purpose which of these two approaches one has in mind and our results apply equally to both cases.

The utility achieved by individual i in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a})$ as envisaged by the ethical observer is indicated by $U(s_i) = U(x_i, a_i)$. To simplify the exposition, we assume throughout that the *utility function* $U : D \times \mathscr{A} \to \mathbb{R}$ is twice differentiable in income and we denote by $\mathscr{U}$ the set of such functions.[11] We use $U(\mathbf{s}) \equiv U(\mathbf{x}; \mathbf{a}) := (U(x_1, a_1), \ldots, U(x_n, a_n))$ to indicate the *distribution of utility* generated by the situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \in \mathscr{S}_n$ when the utility function is $U \in \mathscr{U}$. The *utilitarian rule* ranks the situations being compared on the basis of the *sum of the utilities* they generate. More precisely, from the point of view of a utilitarian ethical observer endowed with the utility function $U \in \mathscr{U}$, situation $\mathbf{s}^*$ is considered to be no worse than situation $\mathbf{s}^\circ$ if and only if

$$(3.1) \qquad \sum_{i=1}^{n} U(x_i^*, a_i^*) \geqq \sum_{i=1}^{n} U(x_i^\circ, a_i^\circ).$$

The utility function U captures the utilitarian ethical observer's normative judgement and it is the only parameter by which such ethical observers can be distinguished. In order to rule out as much arbitrariness as possible, we require all utilitarian ethical observers whose utility functions U belong to a given class $\mathscr{U}^* \subset \mathscr{U}$ to agree on the ranking of the situations under comparison.

UTILITARIAN UNANIMITY RULE. We say that situation $\mathbf{s}^*$ is no worse than situation $\mathbf{s}^\circ$ for

---

[10] We refer the reader to Blackorby, Bossert, and Donaldson (2005) and Griffin (1986) for discussions of the welfarist approach in economics and philosophy, respectively.

[11] The differentiability assumption is not restrictive since it is always possible in our framework to approximate the discontinuous functions we might face in the proofs by suitable continuous and differentiable functions (see Fishburn and Vickson (1978)).

the (classical) utilitarian unanimity rule over the class $\mathscr{U}^* \subset \mathscr{U}$ if and only if

$$(3.2) \qquad \sum_{i=1}^{n} U\left(x_i^*, a_i^*\right) \geqq \sum_{i=1}^{n} U\left(x_i^\circ, a_i^\circ\right), \ \forall \, U \in \mathscr{U}^*.$$

In the course of the paper, we shall have to introduce particular permutations of income between individuals whose impact on the society's welfare will go in a definite direction. Whereas the question of knowing whether *partial* permutations lead to a welfare improvement is open to debate, it is generally admitted that a *complete* permutation of the situations of two individuals will leave social welfare and inequality unchanged.

PERMUTATION OF THE INDIVIDUALS' SITUATIONS. Given two situations $\mathbf{s}^\circ, \mathbf{s}^* \in \mathscr{S}_n$, we say that $\mathbf{s}^*$ is obtained from $\mathbf{s}^\circ$ by means of a *permutation of the individuals' situations* if there exists an $n \times n$ permutation matrix P such that $\mathbf{s}^* = P\mathbf{s}^\circ$.

It follows from our definitions that the utilitarian social welfare function – and as a consequence the utilitarian unanimity principle – is invariant to a permutation of the individuals' situations.

4. Bidimensional Headcount Poverty Dominance

While the primary interest of Atkinson and Bourguignon (1982) was not to investigate the transformations that reduce inequality in a bidimensional setting, the main condition they imposed on the utility function suggests they had implicitly in mind the following elementary transformation.

FAVOURABLE INCOME PERMUTATION. Given two situations $\mathbf{s}^\circ, \mathbf{s}^* \in \mathscr{S}_n$, we say that $\mathbf{s}^*$ is obtained from $\mathbf{s}^\circ$ by means of a *favourable income permutation* if there exist two individuals i and j ($i \neq j$) such that:

$$(4.1a) \qquad x_j^* = x_i^\circ < x_j^\circ = x_i^*; \ a_i^* = a_i^\circ < a_j^\circ = a_j^*; \ \text{and}$$

$$(4.1b) \qquad s_g^* = s_g^\circ, \ \forall \, g \neq i, j.$$

A favourable income permutation consists in exchanging the income endowment of the better-off individual in both attributes with that of the worse-off individual in both attributes.[12] We have represented in Figure 4.1 a favourable income permutation involving two individuals i and j whose initial situations are respectively $s_i^\circ = (x_i^\circ, a_i^\circ) = (u, h)$ and $s_j^\circ = (x_j^\circ, a_j^\circ) = (v, k)$ with $u < v$ and $h < k$. The permutation of the income endowments actually reduces the correlation – or equivalently the positive association – existing between the two attributes. The resulting situation may be considered less unequal than the original situation in the sense that

$$(4.2a) \qquad U(x_i^\circ, a_i^\circ) < U(x_i^*, a_i^*) \leqq U(x_j^*, a_j^*) < U(x_j^\circ, a_j^\circ), \ \text{or}$$

$$(4.2b) \qquad U(x_i^\circ, a_i^\circ) < U(x_j^*, a_j^*) < U(x_i^*, a_i^*) < U(x_j^\circ, a_j^\circ),$$

---

[12] Conditions (4.1a) and (4.1b) can equally be interpreted as defining a favourable health status permutation. There is indeed perfect symmetry, but we find it convenient to think of such transformations as income rather than health status changes.

Figure 4.1: A favourable income permutation
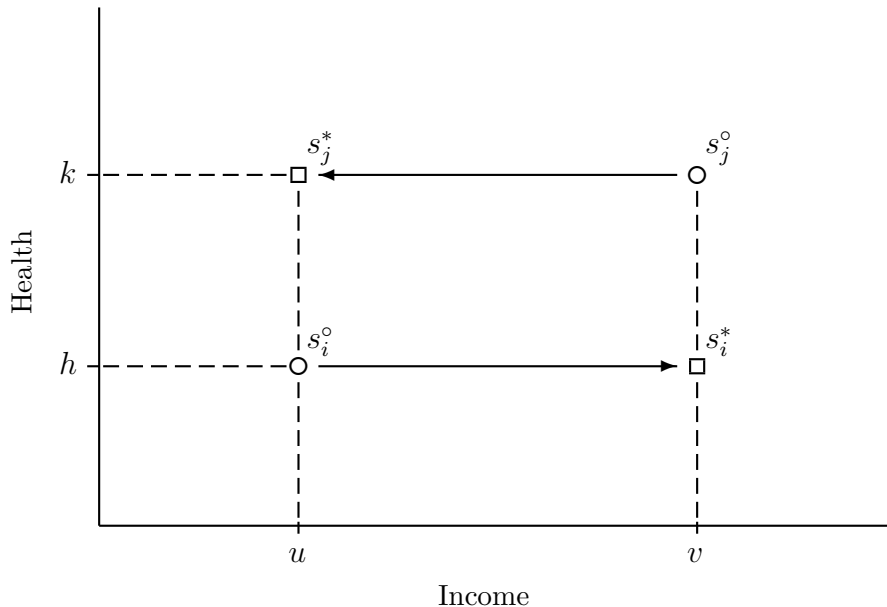
for all utility functions U increasing in both income and health. On the other hand, a favourable income permutation has no impact on the distribution of each attribute in the population: the marginal distribution functions of income and health are left unchanged and so is mean income. The next result identifies the conditions to be imposed on the utility function for social welfare as measured by the utilitarian rule to improve as the result of a favourable income permutation.

Lemma 4.1. *For all* $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$: $\sum_{i=1}^{n} U\left(x_i^*, a_i^*\right) \geqq \sum_{i=1}^{n} U\left(x_i^\circ, a_i^\circ\right)$ *whenever* $\mathbf{s}^*$ *is obtained from* $\mathbf{s}^\circ$ *by means of a favourable permutation if and only if*

C1 $$U_y(y, h) \geqq U_y(y, k), \ \forall \ y \in \mathscr{D}, \ \forall \ h, k \in \mathscr{A} \ (h < k),$$

where $U_y(y, h)$ *indicates the first derivative of* $U(y, h)$ *with respect to income.*

According to Lemma 4.1, a non-increasing marginal utility of income as health status increases is necessary and sufficient for a utilitarian ethical observer to consider that a favourable income permutation results in a weak welfare improvement. In order to save space and with a slight abuse of terminology we say that the utility functions that verify condition C1 are *submodular*.[13] We denote by

(4.3) $$\mathscr{U}_1 := \{U \in \mathscr{U} \mid U \text{ satisfies C1}\}$$

the class of utility functions with non-increasing marginal utility of income.

---

[13] Actually, a function $g : \mathbb{R}^2 \to \mathbb{R}$ is called *submodular* if $g(u + \varepsilon, v + \delta) - g(u, v + \delta) \leqq g(u + \varepsilon, v) - g(u, v)$, for all $(u, v) \in \mathbb{R}^2$ and all $\varepsilon, \delta > 0$. When the function g is differentiable, this reduces to the condition that the cross-derivative $g_{12}(u, v)$ is non-positive. Sometimes, one also says that g is *L-subadditive* (see, e.g., Marshall and Olkin (1979, Chapter 6, Section D).

Atkinson and Bourguignon (1982) have shown that the ranking of situations by the utilitarian unanimity rule over the class $\mathscr{U}_1$ is implied by the ranking obtained by comparing the graphs of their joint cumulative distribution functions by means of the first order stochastic dominance criterion. More precisely, we say that situation $\mathbf{s}^*$ *first order stochastic dominates* situation $\mathbf{s}^\circ$, which we write as $\mathbf{s}^* \geq_{FSD} \mathbf{s}^\circ$, if and only if:

$$(4.4) \qquad F^*(y, h) \leqq F^\circ(y, h), \ \forall \ y \in \mathscr{D}, \ \forall \ h \in \mathscr{A}.$$

Using (2.4), (2.5) and (2.8) and upon substitution, condition (4.4) can be equivalently rewritten as

$$(4.5) \qquad \sum_{g=1}^{h} f_2^*(g) \, H^*(z \mid g) \leqq \sum_{g=1}^{h} f_2^\circ(g) \, H^\circ(z \mid g), \ \forall \ z \in \mathscr{D}, \ \forall \ h \in \mathscr{A},$$

where

$$(4.6) \qquad H(z \mid g) := \frac{q_1(z \mid g)}{n_2(g)}$$

is the *headcount poverty* of the subpopulation of individuals having health status g in situation $\mathbf{s}$ for the income poverty line z. Condition (4.4) – or equivalently, condition (4.5) – expresses the fact that there is less poverty in situation $\mathbf{s}^*$ than in situation $\mathbf{s}^\circ$ for all possible values of the bidimensional poverty line (z, h), where poverty is measured by the percentage of individuals whose income and health status fall below z and h, respectively. More precisely, we say that situation $\mathbf{s}^*$ (weakly) *headcount poverty dominates* situation $\mathbf{s}^\circ$, which we write as $\mathbf{s}^* \geq_{HP} \mathbf{s}^\circ$, when condition (4.5) holds.

In conjunction with the result of Atkinson and Bourguignon (1982), Lemma 4.1 suggests that favourable income permutations might be the elementary transformations that lie behind first order stochastic – equivalently headcount poverty – dominance. The following result shows that this intuition is correct with certain qualifications.

Proposition 4.1. *Let $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$ be such that $\mathbf{a}^* = \mathbf{a}^\circ$ and $\mathbf{s}^*$ is not a permutation of $\mathbf{s}^\circ$. Then, statements (a) and (b) below are equivalent:*

(a) *$\mathbf{s}^*$ is obtained from $\mathbf{s}^\circ$ by means of a finite sequence of favourable income permutations.*

(b-1) *$F^*(y, h) \leqq F^\circ(y, h)$, for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, H - 1$;*

(b-2) *$F_1^*(y) \equiv F^*(y, H) = F^\circ(y, H) \equiv F_1^\circ(y)$, for all $y \in \mathscr{D}$;*

(b-3) *$F_2^*(h) \equiv F^*(\overline{v}, h) = F^\circ(\overline{v}, h) \equiv F_2^\circ(h)$, for all $h \in \mathscr{A}$.*

Whereas condition (b-3) in Proposition 4.1 is but a restatement of the assumption that the situations under comparison have identical distributions of health statuses, the two other conditions are of a more fundamental nature. In particular, condition (b-2) implies that the distributions of income $\mathbf{x}^*$ and $\mathbf{x}^\circ$ are permutations of each other.

Making use of Lemma 4.1 and Proposition 4.1, we obtain the following result, which constitutes in our framework the analogue to the Hardy-Littlewood-Pólya result for first order stochastic – equivalently headcount poverty – dominance.

Theorem 4.1. *Let* $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$ *be such that* $\mathbf{a}^* = \mathbf{a}^\circ$ *and* $\mu(\mathbf{x}^*) = \mu(\mathbf{x}^\circ)$. *Then statements* (a)*,* (b) *and* (c) *below are equivalent:*

(a) $\mathbf{s}^*$ *is obtained from* $\mathbf{s}^\circ$ *by means of a finite sequence of permutations of the individuals' situations and/or favourable income permutations.*

(b) $\sum_{i=1}^{n} U(x_i^*, a_i^*) \geqq \sum_{i=1}^{n} U(x_i^\circ, a_i^\circ)$, *for all* $U \in \mathscr{U}_1$.

(c-1) $F^*(y, h) \leqq F^\circ(y, h)$, *for all* $y \in \mathscr{D}$ *and all* $h = 1, 2, \ldots, H - 1$;

(c-2) $F_1^*(y) \equiv F^*(y, H) = F^\circ(y, H) \equiv F_1^\circ(y)$, *for all* $y \in \mathscr{D}$;

(c-3) $F_2^*(h) \equiv F^*(\overline{v}, h) = F^\circ(\overline{v}, h) \equiv F_2^\circ(h)$, *for all* $h \in \mathscr{A}$.

PROOF.

(a) $\Longrightarrow$ (b). The fact that statement (a) implies statement (b) follows from repeated applications of Lemma 4.1 (sufficiency part) and the fact that the utilitarian rule is invariant with respect to permutations of the individuals' situations.

(b) $\Longrightarrow$ (c). Condition (c-3) is a consequence of the assumption that the marginal distributions of health status in situations $\mathbf{s}^*$ and $\mathbf{s}^\circ$ are identical. We first prove that condition (b) implies that

(4.7) $$F^*(y, h) \leqq F^\circ(y, h), \ \forall \ y \in \mathscr{D}, \ \forall \ h \in \mathscr{A}.$$

Suppose that (4.7) is not verified and let (v^*, h^*) be the smallest – in the lexicographic sense – couple (v, h), where $v \in \mathscr{D}$ and $h \in \mathscr{A}$, such that $F^*(v, h) > F^\circ(v, h)$. Consider then the function

(4.8) $$\varphi(y) := \begin{cases} 0 & \text{if} \quad \underline{v} \leqq y < v^*, \\[2mm] \vartheta & \text{if} \quad v^* \leqq y \leqq \overline{v}, \end{cases}$$

and let φ^* be a twice differentiable approximation of φ with positive first derivatives (see Fishburn and Vickson (1978, p. 75) for details). Now choose the utility function U such that $U(y, h) := \varphi^*(y)$, for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, h^* - 1$, and $U(y, h) := \vartheta$, for all $y \in \mathscr{D}$ and all $h = h^*, h^* + 1, \ldots, H$. By construction $U \in \mathscr{U}_1$, but one can easily check that

(4.9) $$\sum_{i=1}^{n} U(x_i^*, a_i^*) < \sum_{i=1}^{n} U(x_i^\circ, a_i^\circ).$$

Hence condition (b) is not verified. Thus, condition (4.7) holds, and so does condition (c-1). It remains to show that condition (c-2) is also fulfilled. Suppose it is not, which in conjunction with (4.7) implies that

(4.10) $$F^*(y, H) \leqq F^\circ(y, h), \ \forall \ y \in \mathscr{D}, \ \text{and} \ F^*(\tilde{y}, H) < F^\circ(\tilde{y}, h), \ \text{for some} \ \tilde{y} \in \mathscr{D}.$$

Upon integrating, this implies in turn that

(4.11) $$\mu(\mathbf{x}^*) = \overline{v} - \int_{\underline{v}}^{\overline{v}} F^*(y, H) \, dy > \overline{v} - \int_{\underline{v}}^{\overline{v}} F^\circ(y, H) \, dy = \mu(\mathbf{x}^\circ),$$

which contradicts our assumption that $\mu\left(\mathbf{x}^*\right) = \mu\left(\mathbf{x}^\circ\right)$.

(c) $\Longrightarrow$ (a). This follows from Proposition 4.1. $\qquad\square$

Theorem 4.1, in conjunction with the fact that headcount poverty dominance is a necessary condition for a situation to be ranked above another one by all utilitarian ethical observers whose marginal utility of income is non-increasing with health, supplements Atkinson and Bourguignon (1987, Proposition 2). It is fair to note that Epstein and Tanny (1980) and Tchen (1980) have proven a similar result in the context of decision making under uncertainty but using a different argument. Related results have been provided by Müller and Scarsini (2011) and Meyer and Strulovici (2010, 2011) when there are more than just two attributes. However, none of these articles provide an easy-to-check condition that plays a role similar to our statement (c). While there is a high (theoretical) presumption that the criterion defined by statement (c) does not allow us to discriminate a great deal among the situations under comparison, it is important to note that it is very robust to changes in the measurement units. First order stochastic – or equivalently, headcount poverty – dominance builds only on ordinal information and the results of the comparisons are not affected by changes in the measurement scales of the two attributes provided that they are increasing.

5. Bidimensional Ordered Poverty Gap Dominance

In a unidimensional framework, it is generally agreed that a transfer from a more endowed individual to a less endowed one reduces overall inequality provided that their relative positions are preserved. Invoking the consistency of the multidimensional approach to inequality with the unidimensional one, it is natural to require that such a transfer will – other things equal – reduce inequality in our bidimensional setting.[14]

WITHIN-TYPE PROGRESSIVE INCOME TRANSFER. Given two situations $\mathbf{s}^\circ, \mathbf{s}^* \in \mathscr{S}_n$, we say that $\mathbf{s}^*$ is obtained from $\mathbf{s}^\circ$ by means of a *within-type progressive income transfer* if there exist two individuals i and j $(i \neq j)$ such that:

(5.1a) $$x_i^\circ < x_i^* \leqq x_j^* < x_j^\circ; \ a_i^\circ = a_i^* = a_j^* = a_j^\circ;$$

(5.1b) $$x_i^* - x_i^\circ = x_j^\circ - x_j^*; \ \text{and}$$

(5.1c) $$s_g^* = s_g^\circ, \ \forall \ g \neq i, j.$$

We have represented in Figure 5.1 a within-type progressive income transfer involving two individuals i and j with the same health status h. By transferring the amount $\Delta = v - u = t - w > 0$ from the richer individual to the poorer one, we move individual i from $s_i^\circ = (x_i^\circ, a_i^\circ) = (u, h)$ to $s_i^* = (x_i^*, a_i^*) = (v, h)$ and individual j from $s_j^\circ = (x_j^\circ, a_j^\circ) = (t, h)$

---

[14] While this principle makes sense intuitively, it is subject to the same criticism as the one that has been addressed to the standard *principle of transfers* in the unidimensional framework. Namely, while the individuals involved in the transfer have become more equal, the inequalities between each of these two individuals and any other individual in the population has increased (see, e.g., Chateauneuf and Moyes (2006), Magdalou and Moyes (2009)).

to $s_j^* = (x_j^*, a_j^*) = (w, h)$, respectively. The marginal distribution function of health is left unchanged by a within-type progressive income transfer, while the distributions of income conditional on health are made more equal or at least no more unequal.

Figure 5.1: A within-type progressive income transfer
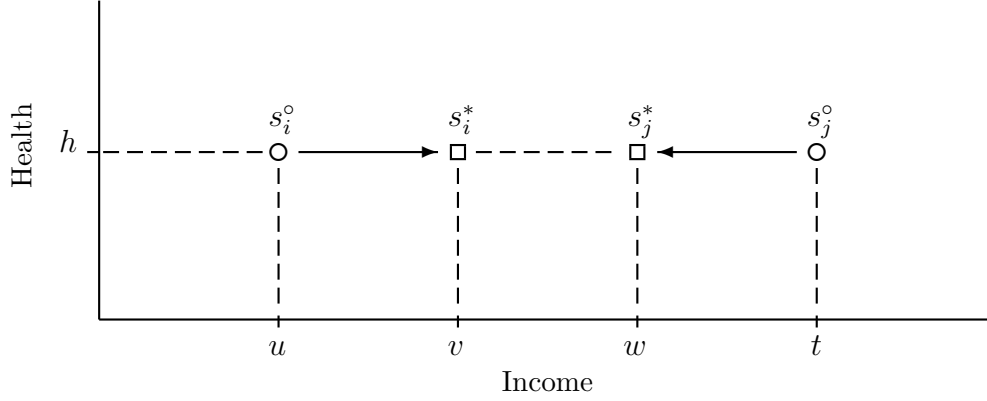
One can easily identify the property of the utility function that guarantees that social welfare as measured by the utilitarian rule improves as the result of a within-type progressive transfer:

Lemma 5.1. *For all* $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$: $\sum_{i=1}^n U(x_i^*, a_i^*) \geqq \sum_{i=1}^n U(x_i^\circ, a_i^\circ)$ *whenever* $\mathbf{s}^*$ *is obtained from* $\mathbf{s}^\circ$ *by means of a within-type progressive income transfer if and only if*

C2
$$U_{yy}(y, h) \leqq 0, \ \forall \ y \in \mathscr{D}, \ \forall \ h \in \mathscr{A},$$

where $U_{yy}(y, h)$ *indicates the second derivative of* $U(y, h)$ *with respect to income.*

The multidimensional nature of the situations under comparison actually plays no role in the concept of a within-type progressive income transfer, which is merely a restatement of the usual Pigou-Dalton transfer in our framework. The next transformation, which fully exploits the bidimensionality of a situation, constitutes in our model a very natural generalisation of a unidimensional progressive transfer.

BETWEEN-TYPE PROGRESSIVE INCOME TRANSFER. Given two situations $\mathbf{s}^\circ, \mathbf{s}^* \in \mathscr{S}_n$, we say that $\mathbf{s}^*$ is obtained from $\mathbf{s}^\circ$ by means of a *between-type progressive income transfer* if there exist two individuals i and j $(i \neq j)$ such that:

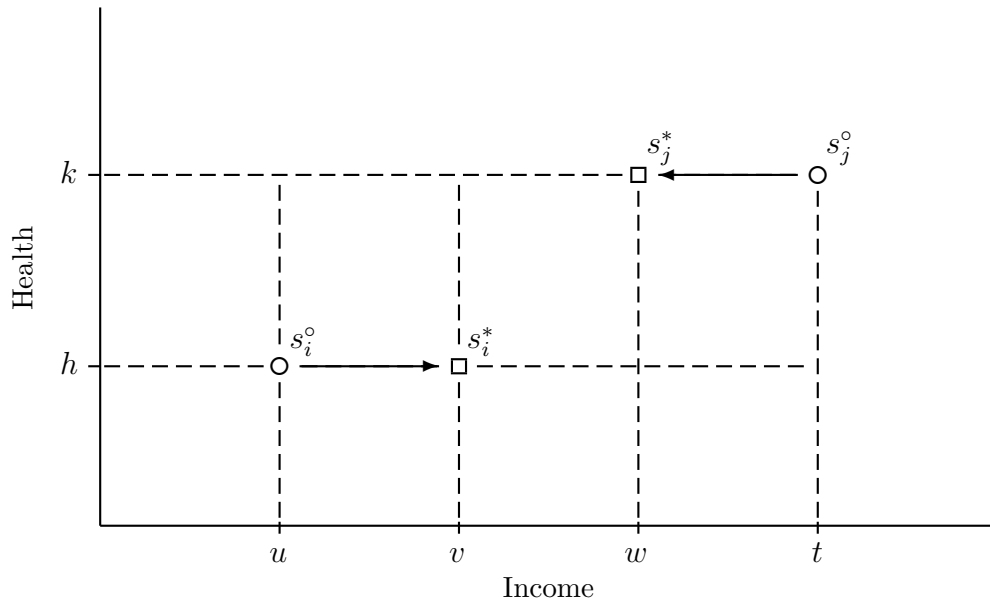(5.2a) $x_i^\circ < x_i^* \leqq x_j^* < x_j^\circ; \ a_i^\circ = a_i^* < a_j^* = a_j^\circ;$

(5.2b) $x_i^* - x_i^\circ = x_j^\circ - x_j^*; \ \text{and}$

(5.2c) $s_g^* = s_g^\circ, \ \forall \ g \neq i, j.$

A between-type progressive income transfer resembles a Pigou-Dalton transfer but there is a major difference: the beneficiary of the transfer must be poorer than the donor and she

Figure 5.2: A between-type progressive income transfer

must also have a lower health status. Put differently, the transfer recipient must be deprived in both dimensions – income and health – compared to the donor. This kind of transfer is a particular case of the *equalising transformation* considered by Müller and Scarsini (2011) where the transfers occur in *possibly* all dimensions. The fact that in our case health is an ordinal and non-transferable attribute rules out the possibility of making transfers along this dimension. We have represented in Figure 5.2 a between-type progressive income transfer where individual i with income u and health status h receives an additional income of $v - u$ while individual j with income t and health status k gives away an amount of income equal to $t - u$, where $u < v < w < t$, $v - u = t - w$ and $h < k$.

A between-type progressive income transfer can always be decomposed into a within-type progressive income transfer followed by a favourable permutation provided that one adds a phantom individual endowed with the income of the beneficiary and the health status of the donor prior to the transfer. This is illustrated in Figure 5.3 which describes the three steps involved in this decomposition process. To simplify things, suppose that the population consists of two individuals i and j and that $\mathbf{s}^* = (s_i^*, s_j^*)$ is obtained from $\mathbf{s}^\circ = (s_i^\circ, s_j^\circ)$ by means of a single between-type progressive transfer so that $s_i^\circ = (x_i^\circ, a_i^\circ) = (u, h)$, $s_j^\circ = (x_j^\circ, a_j^\circ) = (t, k)$, $s_i^* = (x_i^*, a_i^*) = (v, h)$ and $s_j^* = (x_j^*, a_j^*) = (w, k)$, where $u < v < w < t$, $v - u = t - w = \Delta$ and $h < k$. Consider now an individual g whose situation is given by $s_g^\circ = (x_g^\circ, a_g^\circ) = (u, k)$. Adding individual g to the initial population $\{i, j\}$, we obtain the augmented situation $(s_i^\circ, s_j^\circ; s_g^\circ)$. Individuals g and j have the same health status but individual j is richer than individual g. Taking an income amount $\Delta > 0$ from individual j and giving it to individual g we obtain the new situation $(s_i^\circ, s_j^*; s_g^*)$ which follows from $(s_i^\circ, s_j^\circ; s_g^\circ)$ by means of a within-type progressive income transfer. Observe that individual i is deprived in both income and health compared to individual g. Exchanging now the incomes of
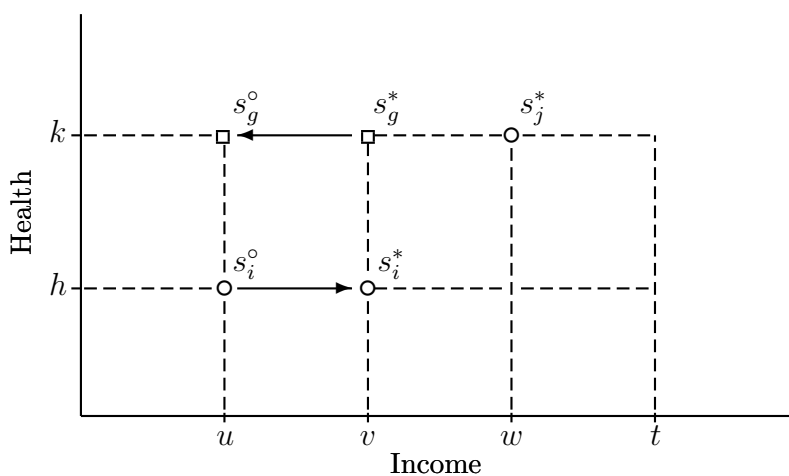
Figure 5.3: Decomposition of a between-type progressive income transfer
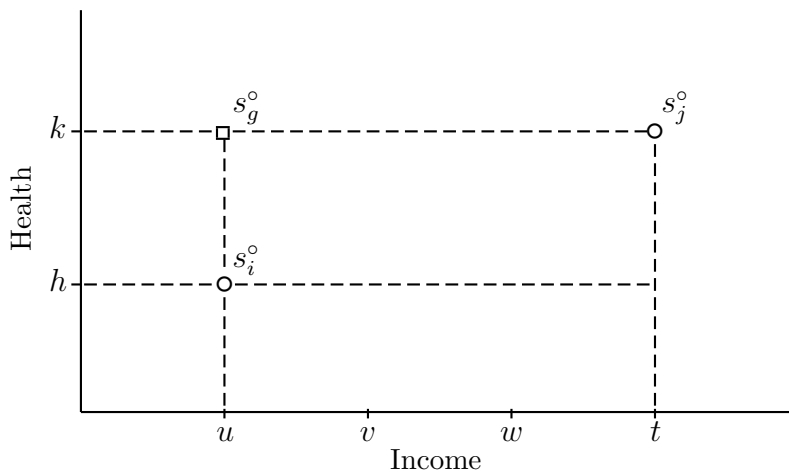
(a) Step 1. Addition of the phantom individual
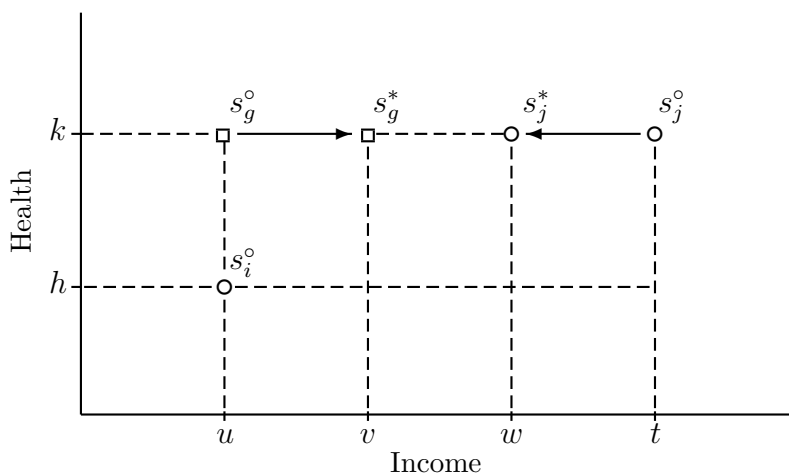
(b) Step 2. Within-type progressive income transfer

(c) Step 3. Favourable income permutation

individuals i and g we obtain the situation $(s_i^*, s_j^*; s_g^\circ)$ which follows from situation $(s_i^\circ, s_j^*; s_g^*)$ by means of a favourable income permutation. Individual g is back to her initial situation so that she actually played only an instrumental role in the decomposition. We could have equally proceeded by choosing $s_g^\circ = (x_g^\circ, a_g^\circ) = (t, h)$, which would have led to a within-type progressive income transfer from individual g to individual i followed by a favourable income permutation between individuals g and j.

Somewhat surprising is the fact that no condition in addition to the ones we have considered up to now need to be imposed on the utility function for a between-type progressive income transfer to imply a welfare improvement.

Lemma 5.2. *For all* $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$: $\sum_{i=1}^{n} U(x_i^*, a_i^*) \geqq \sum_{i=1}^{n} U(x_i^\circ, a_i^\circ)$ *whenever* $\mathbf{s}^*$ *is obtained from* $\mathbf{s}^\circ$ *by means of a between-type progressive income transfer if conditions* C1 *and* C2 *are satisfied.*

According to Lemma 5.2, it is sufficient for welfare as conceived by a utilitarian ethical observer to increase as the result of a between-type progressive income transfer that the marginal utility of income be non-increasing in income and in health status. The corresponding class of utility functions is indicated by

(5.3) $$\mathscr{U}_2 := \{U \in \mathscr{U} \mid U \text{ satisfies C1 and C2}\}.$$

Contrary to what happens with favourable income permutations and within-type progressive income transfers, it is not clear whether the conditions we obtain here are also necessary for a between-type progressive income transfer to result in a welfare improvement.[15]

Bourguignon (1989) introduces a dominance criterion that coincides with the one which commands unanimity over all utilitarian ethical observers who use a utility function that is non-decreasing and concave in income and whose marginal utility in income is non-increasing in health status. In order to define his criterion, we first need to introduce additional notation and technicalities. An *ordered poverty line* is a H-dimensional vector $\mathbf{z} := (z(1), z(2), \ldots, z(H))$ such that $z(1) \geqq z(2) \geqq \cdots \geqq z(H)$, where $z(h) \in \mathscr{D}$ is the poverty line assigned to all the individuals whose health status is equal to h. The poverty line faced by an individual is no longer exogenously given as in the standard unidimensional framework but it depends now on her personal situation. The *ordered poverty gap* in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \in \mathscr{S}_n$, when the ordered poverty line is $\mathbf{z}$ – denoted $\mathscr{P}(\mathbf{z}; \mathbf{s})$ – is defined by

(5.4) $$\mathscr{P}(\mathbf{z}; \mathbf{s}) = \sum_{h \in \mathscr{A}} f_2(h) \int_{\underline{v}}^{z(h)} F_1(t \mid h)\, dt = \sum_{h \in \mathscr{A}} f_2(h)\, P(z(h) \mid h),$$

where

(5.5) $$P(z(h) \mid h) = \frac{1}{n_2(h)} \sum_{i \in Q_1(z(h) \mid h)} (z(h) - x_i)$$

---

[15] Ebert (2000) makes use of a continuity argument to show that a decreasing in health status marginal utility of income and concavity with respect to income are also necessary for the sum of utilities to increase as the result of a between-type progressive income transfer.

is the *conditional poverty gap* of the group of individuals with health status equal to h when the poverty line is set to $z(h)$.

We now introduce the criterion that is at the origin of the present research and that will play a fundamental role in the following. We say that situation $\mathbf{s}^*$ *ordered poverty gap dominates* situation $\mathbf{s}^\circ$, which we write as $\mathbf{s}^* \geq_{OPG} \mathbf{s}^\circ$, if and only if:

$$(5.6) \qquad \mathscr{P}(\mathbf{z}; (\mathbf{x}^*; \mathbf{a}^*)) \leqq \mathscr{P}(\mathbf{z}; (\mathbf{x}^\circ; \mathbf{a}^\circ)), \ \forall \ \mathbf{z} \text{ such that } z(1) \geqq z(2) \geqq \cdots \geqq z(H).$$

Definition (5.4) makes clear that the ordered poverty gap is the weighted sum of the conditional poverty gaps, where the weights are equal to the marginal densities of health status. For later reference, we say that situation $\mathbf{s}^*$ *conditional poverty gap dominates* situation $\mathbf{s}^\circ$, which we write as $\mathbf{s}^* \geq_{CPG} \mathbf{s}^\circ$, if and only if:

$$(5.7) \qquad P^*(z \,|\, h) \leqq P^\circ(z \,|\, h), \ \forall \ z \in \mathscr{D}, \ \forall \ h \in \mathscr{A}.$$

When the situations under comparison have identical marginal distributions of health status, it is immediately clear that conditional poverty gap dominance implies ordered poverty gap dominance, while the converse implication does not hold. Straightforward computations indicate that

$$(5.8) \qquad \mathscr{P}(\mathbf{z}; (\tilde{\mathbf{s}}; \hat{\mathbf{s}})) = \mathscr{P}(\mathbf{z}; \tilde{\mathbf{s}}) + \mathscr{P}(\mathbf{z}; \hat{\mathbf{s}}), \ \forall \ \tilde{\mathbf{s}} \in \mathscr{S}_n, \ \forall \ \hat{\mathbf{s}} \in \mathscr{S}_q \ (n, q \geqq 1).$$

Hence, the ordered poverty gap $\mathscr{P}(\mathbf{z}; \mathbf{s})$ is additively separable.

While obvious, the next result, which is a direct consequence of the additive separability of the ordered poverty gap, will play a crucial role in subsequent developments.

Lemma 5.3. *For all* $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n \ (n \geq 2)$ *and all* $\tilde{\mathbf{s}} \in \mathscr{S}_q \ (q \geq 1)$, *we have:*

$$(5.9) \qquad \mathbf{s}^* \geq_{OPG} \mathbf{s}^\circ \iff (\mathbf{s}^*; \tilde{\mathbf{s}}) \geq_{OPG} (\mathbf{s}^\circ; \tilde{\mathbf{s}}).$$

The following proposition is nothing but a restatement of well-known results in the unidimensional case (see, e.g., Hardy *et al.* (1934), Berge (1963), Marshall and Olkin (1979)).

Proposition 5.1. *Let* $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$ *be such that* $\mathbf{a}^* = \mathbf{a}^\circ$ *and* $\mu(\mathbf{x}^*) = \mu(\mathbf{x}^\circ)$. *Then, statements* (a) *and* (b) *below are equivalent:*

(a) $P^*(z \,|\, h) \leqq P^\circ(z \,|\, h)$, *for all* $z \in \mathscr{D}$ *and all* $h \in \mathscr{A}$.

(b) $\mathbf{s}^*$ *is obtained from* $\mathbf{s}^\circ$ *by means of a finite sequence of within-type progressive income transfers.*

The next proposition will play a decisive role in the proof of our main result: it is a separation result, which states that ordered poverty gap domination can always be decomposed into headcount poverty domination and conditional poverty gap domination.[16]

---

[16] Related results are routinely used in the unidimensional framework for decomposing generalised Lorenz domination into first order stochastic domination and Lorenz domination (see, e.g., Marshall and Olkin (1979, 5A9)).

Proposition 5.2. *Let $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$ be such that $\mathbf{a}^* = \mathbf{a}^\circ$ and $\mu(\mathbf{x}^*) = \mu(\mathbf{x}^\circ)$. Then, statements (a) and (b) below are equivalent:*

(a-1) $\mathscr{P}(\mathbf{z}; \mathbf{s}^*) \leqq \mathscr{P}(\mathbf{z}; \mathbf{s}^\circ)$, *for all $\mathbf{z}$ such that $z(h)$ is non-increasing in h;*

(a-2) $F_2^*(h) = F_2^\circ(h)$, *for all $h \in \mathscr{A}$.*

   (b) *There exist $\tilde{\mathbf{s}} \in \mathscr{S}_q$ ($q \geqq 0$) and $\hat{\mathbf{s}}^\ddagger \in \mathscr{S}_n \times \mathscr{S}_q$ such that $\tilde{\mathbf{s}}^* \geq_{HP} \hat{\mathbf{s}}^\ddagger \geq_{CPG} \tilde{\mathbf{s}}^\circ$, where $\tilde{\mathbf{s}}^* := (\mathbf{s}^*; \tilde{\mathbf{s}})$ and $\tilde{\mathbf{s}}^\circ := (\mathbf{s}^\circ; \tilde{\mathbf{s}})$.*

We obtain the following theorem, which illuminates the normative foundations of the Bourguignon (1989) ordered poverty gap quasi-ordering.

Theorem 5.1. *Let $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$ be such that $\mathbf{a}^* = \mathbf{a}^\circ$ and $\mu(\mathbf{x}^*) = \mu(\mathbf{x}^\circ)$. Then, statements (a), (b) and (c) below are equivalent:*

   (a) *There exists $q \geqq 0$ and a situation $\tilde{\mathbf{s}} \in \mathscr{S}_q$ such that $(\mathbf{s}^*; \tilde{\mathbf{s}})$ is obtained from $(\mathbf{s}^\circ; \tilde{\mathbf{s}})$ by means of a finite sequence of permutations of the individuals' situations, favourable income permutations and/or within-type progressive income transfers.*

   (b) $\sum_{i=1}^n U(x_i^*, a_i^*) \geqq \sum_{i=1}^n U(x_i^\circ, a_i^\circ)$, *for all $U \in \mathscr{U}_2$.*

(c-1) $\mathscr{P}(\mathbf{z}; (\mathbf{x}^*; \mathbf{a}^*)) \leqq \mathscr{P}(\mathbf{z}; (\mathbf{x}^\circ; \mathbf{a}^\circ))$, *for all $\mathbf{z}$ such that $z(h)$ is non-increasing in h;*

(c-2) $F_2^*(h) = F_2^\circ(h)$, *for all $h \in \mathscr{A}$.*

PROOF.

(a) $\Longrightarrow$ (b). Repeated applications of Lemmas 4.1, 5.1 and 5.2 together with the fact that the utilitarian rule is invariant with respect to permutations of the individuals' situations implies that

$$(5.10) \qquad \sum_{i=1}^n U(x_i^*, a_i^*) + \sum_{i=1}^q U(x_i, a_i) \geq \sum_{i=1}^n U(x_i^\circ, a_i^\circ) + \sum_{i=1}^q U(x_i, a_i),$$

for all $U \in \mathscr{U}_2$, which implies statement (b).

(b) $\Longrightarrow$ (c). Condition (c-2) is a consequence of the assumption that the marginal distributions of health in situations $\mathbf{s}^*$ and $\mathbf{s}^\circ$ are identical. Therefore, we only have to prove that condition (b) implies condition (c-1). Suppose that there exists $\mathbf{z} := (z(1), z(2), \ldots, z(H))$ with $z(1) \geqq z(2) \geqq \cdots \geqq z(H)$ such that

$$(5.11) \qquad\qquad \mathscr{P}(\mathbf{z}; (\mathbf{x}^*; \mathbf{a}^*)) > \mathscr{P}(\mathbf{z}; (\mathbf{x}^\circ; \mathbf{a}^\circ)),$$

which, upon using (5.4), can equivalently be rewritten as

$$(5.12) \qquad\qquad \sum_{h=1}^H f_2^*(h) P^*(z(h) \mid h) > \sum_{h=1}^H f_2^\circ(h) P^\circ(z(h) \mid h).$$

Upon computation, condition (5.12) reduces to

$$(5.13) \qquad \sum_{h \in \mathscr{A}} \sum_{i \in Q_1^*(z(h) \,|\, h)} (z(h) - x_i^*) > \sum_{h \in \mathscr{A}} \sum_{i \in Q_1^\circ(z(h) \,|\, h)} (z(h) - x_i^\circ),$$

or equivalently,

$$(5.14) \qquad \sum_{h \in \mathscr{A}} \sum_{i \in N_2^*(h)} \max\{0, z(h) - x_i^*\} > \sum_{h \in \mathscr{A}} \sum_{i \in N_2^\circ(h)} \max\{0, z(h) - x_i^\circ\},$$

where by assumption $N_2^*(h) = N_2^\circ(h)$, for all $h = 1, 2, \ldots, H$. Choose

$$(5.15) \qquad \chi(y, h) := \begin{cases} y - z(h) & \text{if} \quad \underline{v} \leqq y < z(h), \\[2mm] 0 & \text{if} \quad z(h) \leqq y \leqq \overline{v}, \end{cases}$$

for all $h = 1, 2, \ldots, H$. Appealing to an argument in Fishburn and Vickson (1978, p. 76), it is possible to approximate $\chi(y, h)$ by a twice differentiable function $U(y, h)$ such that $U_y(y, h) \geqq 0$, $U_{yy}(y, h) \leqq 0$, and $U_y(y, h)$ is non-increasing in h, for all $y \in (\underline{v}, \overline{v})$ and all $h \in \mathscr{A}$. Therefore, $U \in \mathscr{U}_2$, but one can easily check that

$$(5.16) \qquad \sum_{i=1}^{n} U\left(x_i^*, a_i^*\right) < \sum_{i=1}^{n} U\left(x_i^\circ, a_i^\circ\right).$$

Hence, condition (b) is not verified.

(c) $\Longrightarrow$ (a). Assuming that statement (c) holds and invoking Proposition 5.2, we deduce that there exists a situation $\tilde{\mathbf{s}} \in \mathscr{S}_q$ $(q \geqq 0)$ and $\hat{\mathbf{s}}^{\ddagger} \in \mathscr{S}_n \times \mathscr{S}_q$ such that

$$(5.17) \qquad \tilde{\mathbf{s}}^* \geq_{HP} \hat{\mathbf{s}}^{\ddagger} \geq_{CPG} \tilde{\mathbf{s}}^\circ,$$

where $\tilde{\mathbf{s}}^* := (\mathbf{s}^*; \tilde{\mathbf{s}})$ and $\tilde{\mathbf{s}}^\circ := (\mathbf{s}^\circ; \tilde{\mathbf{s}})$. Finally, Propositions 4.1 and 5.1 establish that (5.17) implies statement (a), which completes the argument. $\qquad \square$

The equivalence between the unanimous agreement among all utilitarian ethical observers whose marginal utility of income is non-increasing in both attributes and ordered poverty gap dominance was first established by Bourguignon (1989). However, the normative meaning of this result is somewhat obscured by the fact that the nature of the underlying elementary transformations that are needed to transform the dominated situation into the dominating one was not revealed. Also, the ingenious technique of proof employed by Bourguignon (1989) relied on the introduction of auxiliary functions whose meaning is unclear. Theorem 5.1 is motivated by these critiques, but we have to admit that the answer it provides is not totally satisfactory. Indeed, in order to identify the elementary transformations, successive applications of which allow one to derive the dominating situation from the dominating one, it may be the case that we have to introduce a phantom situation. Thanks to the separability of the different normative criteria we appeal to, the phantom situation plays only an instrumental role in the derivation of our result. There is a strong presumption that this instrumental situation actually mirrors the auxiliary functions used by Bourguignon (1989) in his proof.

Leaving aside the fact that we do not provide a means for identifying with precision the phantom situation, the main limitation of Theorem 5.1 is that it is silent as far as between-type progressive income transfers are concerned. Indeed, we know from Lemma 5.2 that all utilitarian ethical observers whose marginal utility of income is non-increasing in income and health status will consider that a between-type progressive transfer results in improvement in social welfare. However, we have not succeeded in showing that a finite sequence of such transformations enables the dominated situation to be transformed into the dominating one, nor have we been able to prove that this is impossible. All we have shown is that it is possible to transform the dominated situation into the dominating situation – where *both situations are augmented by the phantom situation* – by means of favourable income permutations and within-type progressive income transfers.

6. Extensions

Variable population size   While we have kept the population size n fixed up to now, we note that our approach easily extends to the case of variable population sizes by invoking the *principle of population* due to Dalton (1920). According to this principle, the society's welfare *per head* is not affected by a replication of the distribution. All of our results generalise to this case provided one substitutes the *average utilitarian social welfare function* for the classical utilitarian one.

Different mean incomes and distributions of health status   So far we have restricted our attention to the case where the situations under comparison have identical marginal distributions of health status and equal mean incomes. These restrictions were to a large extent motivated by the fact that we were interested in the basic notion of inequality in our particular framework. It was therefore desirable to rule out any considerations for efficiency that can be associated with transformations of the distributions that result in modifications of the distributions of health status as well as changes in mean incomes. While these restrictions allowed us to focus on equity considerations, the practical relevance of the results we have obtained is limited. Theorem 5.1 can be extended to the general case where the marginal distributions are no longer fixed and mean incomes no longer equal provided that one is willing to introduce additional restrictions on the utility functions. Or equivalently, one has to think of those transformations of the distributions that modify the distribution of health status and mean income, and that at the same time can be considered improvements from a social welfare point of view.

The two following transformations are the most natural that one can think of when the situations under comparisons have different marginal distributions of income and health.

INCOME INCREMENT. Given two situations $\mathbf{s}^\circ, \mathbf{s}^* \in \mathscr{S}_n$, we say that situation $\mathbf{s}^*$ is obtained from situation $\mathbf{s}^\circ$ by means of an *income increment* if there exists one individual i such that:

(6.1)
$$x_i^* > x_i^\circ, \ a_i^* = a_i^\circ \text{ and } s_g^* = s_g^\circ, \ \forall \ g \neq i.$$

HEALTH STATUS INCREMENT. Given two situations $\mathbf{s}^\circ, \mathbf{s}^* \in \mathscr{S}_n$, we say that situation $\mathbf{s}^*$ is obtained from situation $\mathbf{s}^\circ$ by means of a *health status increment* if there exists one individual i such that:

$$(6.2) \qquad\qquad x_i^* = x_i^\circ, \; a_i^* > a_i^\circ \text{ and } s_g^* = s_g^\circ, \; \forall \, g \neq i.$$

It is immediately clear that social welfare as measured by the utilitarian rule will increase as the result of an income increment if and only if the utility function is non-decreasing in income, namely:

C3 $$\qquad\qquad U_y(y,h) \geqq 0, \; \forall \, y \in D, \; \forall \, h \in \mathscr{A}.$$

Similarly, for a health status increment to imply a social welfare improvement according to the same principle it is necessary and sufficient that the utility function is non-decreasing in health status:

C4 $$\qquad\qquad U(y,h) \leqq U(y,k), \; \forall \, y \in D, \; \forall \, h, k \in \mathscr{A} \; (h < k).$$

Consider now the following class of utility functions:

$$(6.3) \qquad\qquad \mathscr{U}_2^* := \{U \in \mathscr{U} \mid U \text{ satisfies C1, C2, C3 and C4}\}.$$

Then, dispensing with the restrictions that the situations under comparison have identical marginal distributions and equal means, we obtain:

Theorem 6.1. *Let $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$. Then, statements (a), (b) and (c) below are equivalent:*

(a) *There exists $q \geqq 0$ and a situation $\mathbf{s} \in \mathscr{S}_q$ such that $(\mathbf{s}^*; \mathbf{s})$ is obtained from $(\mathbf{s}^\circ; \mathbf{s})$ by means of a finite sequence of permutations of the individuals' situations, income increments, health status increments, favourable income permutations and/or within-type progressive income transfers.*

(b) $\sum_{i=1}^n U(x_i^*, a_i^*) \geqq \sum_{i=1}^n U(x_i^\circ, a_i^\circ)$, *for all $U \in \mathscr{U}_2^*$.*

(c-1) $\mathscr{P}(\mathbf{z}; (\mathbf{x}^*; \mathbf{a}^*)) \leqq \mathscr{P}(\mathbf{z}; (\mathbf{x}^\circ; \mathbf{a}^\circ))$, *for all $\mathbf{z}$ such that $z(h)$ is non-increasing in h;*

(c-2) $F_2^*(h) \leqq F_2^\circ(h)$, *for all $h \in \mathscr{A}$.*

From utilitarian unanimity to welfarist unanimity   We have followed the usual practice in the stochastic dominance literature that consists in appealing to the utilitarian rule when comparing alternative distributions. Despite this preeminence, utilitarianism has been criticised on different grounds, the most serious of which being its insensitivity to the distribution of the total sum of the individuals' well-beings (see, e.g., Sen (1973)). One might prefer to appeal to principles that express some aversion to the inequality in the distribution of individual utilities such as the *maximin* or the *leximin*. To conform with this view, let social welfare in situation $\mathbf{s} \in \mathscr{S}$ be given by $G(U(\mathbf{x}; \mathbf{a})) := G(U(x_1, a_1), \ldots, U(x_n, a_n))$, where $G \in \mathscr{G} := \{g : \mathbb{R}^n \to \mathbb{R}\}$ is the social welfare function. Then, situation $\mathbf{s}^*$ will be considered

no worse than situation $\mathbf{s}°$ by an ethical observer endowed with the utility function $U \in \mathscr{U}$ and the social welfare function G, if and only if

$$(6.4) \qquad G\left(U\left(\mathbf{x}^*; \mathbf{a}^*\right)\right) \geqq G\left(U\left(\mathbf{x}°; \mathbf{a}°\right)\right).$$

Aversion to inequality is typically captured by the property of Schur-concavity and we indicate by $\mathscr{G}_{SC}$ the set of social welfare functions that are monotone and Schur-concave (see Kolm (1969), Dasgupta *et al.* (1973)).[17] The utility function U and the social welfare function G capture the welfarist ethical observer's normative judgement and these are the only parameters that permit one to distinguish one such ethical observer from another. In order to rule out as much arbitrariness as possible we require that all welfarist ethical observers whose utility functions U belong to a given class $\mathscr{U}^* \subset \mathscr{U}$ and whose social welfare functions are members of a class $\mathscr{G}^* \subset \mathscr{G}$ to agree on the ranking of the situations under comparison.

WELFARIST UNANIMITY RULE. We say that situation $\mathbf{s}^*$ is no worse than situation $\mathbf{s}°$ for the welfarist unanimity rule over the class of utility functions $\mathscr{U}^* \subset \mathscr{U}$ and the class of social welfare functionals $\mathscr{G}^* \subset \mathscr{G}$, if and only if

$$(6.5) \qquad G\left(U\left(\mathbf{x}^*; \mathbf{a}^*\right)\right) \geqq G\left(U\left(\mathbf{x}°; \mathbf{a}°\right)\right), \ \forall \, U \in \mathscr{U}^*, \ \forall \, G \in \mathscr{G}^*.$$

Clearly welfarist unanimity implies utilitarian unanimity whatever the class of utility functions. Actually, since the converse statement does not generally hold, it is more demanding to require unanimous agreement among all welfarist ethical observers than among all utilitarian ones. Quite interesting for our purpose is the fact that the utilitarian and welfarist unanimity rules give the same ranking of the situations to be compared under suitable restrictions on the utility functions when $\mathscr{G}^* \subseteq \mathscr{G}_{SC}$. More precisely, Gravel and Moyes (2011) have shown that a sufficient condition for welfarist and utilitarian unanimities to rank social states in the same way is that the class of utility functions is closed under functional composition by all non-decreasing and concave functions, a property satisfied by *all* the classes of utility functions we have considered here. It follows that all the results of this paper for utilitarian unanimity are equally valid when one requires welfarist unanimity.

7. Conclusion

Our aim was to investigate the normative foundations of two implementable criteria – the so-called headcount poverty and ordered poverty gap quasi-orderings – designed for comparing distributions of two attributes, one of which – income – is cardinally measurable, while the other – health – is ordinal. More precisely we wanted (i) to identify the class of utility functions such that all utilitarian ethical observers whose value judgements belong to this class rank the situations under comparison in the same way as the dominance criterion, and (ii) to uncover the elementary transformations, finite applications of which, permit one to derive the dominating situation from the dominated one.

---

[17] The mapping $G : \mathbb{R}^n \to \mathbb{R}$ is *Schur-concave* if $G(B\mathbf{u}) \geqq G(\mathbf{u})$, for all $\mathbf{u} = (u_1, \ldots, u_n)$ and all bistochastic matrices B (see Marshall and Olkin (1979)). It is *monotone* if $G(\mathbf{u}) \geqq G(\mathbf{v})$, for all $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ such that $u_i \geqq v_i$, for all $i = 1, 2, \ldots, n$.

A marginal utility of income that is non-increasing in health ensures that the utilitarian unanimity rule and the headcount poverty quasi-ordering rank the situations in the same way, while a favourable income permutation is the transformation which, if applied a finite number of times, allows one to derive the dominating situation from the dominated one. As far as the headcount poverty – and equivalently, first order stochastic dominance – criterion is concerned, Theorem 4.1 provides the desired equivalence. This result supplements Atkinson and Bourguignon (1987, Proposition 2) by identifying the transformations of the distributions that permit us to derive by successive application the dominating distribution from the dominated one.

Our main contribution is to uncover the value judgements implicit in the Bourguignon (1989) ordered poverty gap criterion by showing that repeated applications of favourable income permutations and within-type progressive income transfers allow one to transform the dominated situation into the dominating one. This permitted us to obtain a result (Theorem 5.1) in the vein of the Hardy-Littlewood-Pólya theorem which complements Bourguignon's original result. However, things are to some extent less satisfactory than in the case of the headcount poverty criterion since we did not manage to establish this equivalence without resorting to the adjunction of a phantom situation. On the other hand, we know that all utilitarian ethical observers whose marginal utilities of income are non-increasing in both income and health status will record as welfare-improving a between-type progressive income transfer. However, it is still an open question whether it is possible to obtain the dominating situation from the dominated one by means of between-type progressive income transfers *without* resorting to phantom individuals.

Other limitations of the present analysis concern the number of attributes considered and also the assumptions relative to the measurability of the attributes we have made. Focusing on just two attributes is certainly restrictive and does not allow one to capture all the relevant dimensions of a person's well-being. The HDI is a good example of an aggregate measure focusing on three essential factors that contribute to a person's well-being: income, life expectancy and literacy. Increasing the number of dimensions is certainly one direction to go, but such an extension is likely to become very involved quite quickly. To give but one example, the meaning of correlation, which lies at the heart of the concept of a favourable income permutation, needs to be substantially reformulated when more than two attributes are considered. Suppose there are three attributes as in the HDI case: then the welfare impact of a favourable permutation involving the first and second attribute will depend on the quantity of the third attribute received by the individuals involved in the transformation.[18] This route is followed by Decancq (2011), who identifies the dominance criteria that generalise the headcount poverty quasi-ordering when there are more than two attributes. Also the question arises of knowing if it could be possible to adapt the Bourguignon (1989) approach in order

---

[18] A way of avoiding this difficulty is to make the extra value judgement that some attributes are separable from each other, which rules out any possibility of compensation between these attributes. This route is followed, for instance, by Muller and Trannoy (2011), who assume separability between life expectancy at birth and educational attainment, but still retain the possibility of using income to compensate for poor health and low education.

to rank situations involving only cardinal attributes, as it is done in Kolm (1977) or Atkinson and Bourguignon (1982).

Finally, it must be recognised that there is a general skepticism concerning the ability of the dominance approach to provide relevant information because of the incomplete nature of the quasi-orderings it relies on. The non-decisiveness of this approach is accentuated in the multi-attribute case and it is expected to be the more serious as the number of dimensions increases. The criteria we investigate in this paper are no exception and, to some extent, they may be considered as a first step that has to supplemented by the use of multidimensional cardinal indices, for instance, like those characterised in Ebert (1995). However, to conclude on a more positive note, it is worth mentioning that the criteria we have examined provide conclusive verdicts in a non-negligible number of cases, as the evidence in Gravel *et al.* (2009) shows.

8. Proofs of Lemmas and Propositions

PROOF OF LEMMA 4.1. See Moyes (2011a, Remark 3.1). □

Proposition 4.1 is actually a generalisation of Moyes (2011b, Theorem 3.1) and the proof follows closely the arguments used for establishing this result. Before we proceed to the proof of Proposition 4.1, we need to introduce additional notation and definitions. Since the notation is particularly heavy in what follows, we momentarily use $N(h)$ and $n(h)$ in place of $N_2(h)$ and $n_2(h)$ to indicate the set and number of individuals who have health statuses h in situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \in \mathscr{S}_n$. It is convenient to begin with to consider the case of distributions of income for a fixed and homogeneous population of m individuals ($m \geqq 2$). The non-decreasing rearrangement of an income distribution $\mathbf{u} := (u_1, \ldots, u_m) \in \mathscr{D}^m$ is indicated by $\tilde{\mathbf{u}} := (\tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_m)$, where $\tilde{u}_1 \leqq \tilde{u}_2 \leqq \cdots \leqq \tilde{u}_m$. Then, given two income distributions $\mathbf{u} := (u_1, \ldots, u_m), \mathbf{v} := (v_1, \ldots, v_m) \in \mathscr{D}^m$, we say that $\mathbf{u}$ *rank order dominates* $\mathbf{v}$, which we write $\mathbf{u} \geqq_{RO} \mathbf{v}$, if and only if $\tilde{u}_i \geqq \tilde{v}_i$, for all $i = 1, 2, \ldots, m$. We indicate respectively by $\sim_{RO}$ and $>_{RO}$ the symmetric and asymmetric components of $\geqq_{RO}$ defined in the usual way, and we note that $\mathbf{u} \sim_{RO} \mathbf{v}$ if and only if $\tilde{u}_i = \tilde{v}_i$, for all $i = 1, 2, \ldots, m$, in which case $\mathbf{u}$ is a permutation of $\mathbf{v}$.

We associate to the situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \in \mathscr{S}_n$ the *heterogeneous income distribution* $\mathbf{u} := (\mathbf{u}^1; \ldots; \mathbf{u}^H)$, where

$$(8.1) \qquad \mathbf{u}^h := (u_1^h, u_2^h, \ldots, u_{n(h)}^h),$$

is the income distribution of the subpopulation consisting of all individuals whose health statuses are equal to h. Without loss of generality, we assume that incomes within populations of individuals of the same type are non-decreasingly arranged so that

$$(8.2) \qquad u_1^h \leqq u_2^h \leqq \cdots \leqq u_{n(h)}^h.$$

We indicate by $\mathbf{u}^{(h)} := (\mathbf{u}^1; \mathbf{u}^2; \ldots; \mathbf{u}^h)$ the distribution of the incomes received by the indi-

viduals with health statuses equal to h or less and we note that $\mathbf{u}^{(H)} = \mathbf{u}$. We let

$$(8.3) \qquad \tilde{\mathbf{u}}^{(h)} := (\tilde{u}_1^{(h)}, \tilde{u}_2^{(h)}, \ldots, \tilde{u}_{\tilde{n}(h)-1}^{(h)}, \tilde{u}_{\tilde{n}(h)}^{(h)})$$

stand for the rearrangement of $\mathbf{u}^{(h)}$ such that

$$(8.4) \qquad \tilde{u}_1^{(h)} \leqq \tilde{u}_2^{(h)} \leqq \cdots \leqq \tilde{u}_{\tilde{n}(h)-1}^{(h)} \leqq \tilde{u}_{\tilde{n}(h)}^{(h)},$$

where $\tilde{n}(h) := \sum_{k=1}^h n(k)$.

Consider two situations $\mathbf{s}^* \equiv (\mathbf{x}^*; \mathbf{a}^*)$ and $\mathbf{s}^\circ \equiv (\mathbf{x}^\circ; \mathbf{a}^\circ)$ such that $\mathbf{a}^* = \mathbf{a}^\circ$, which implies that $n^*(h) = n^\circ(h) = n(h)$: the number of individuals with exactly health status h in the two situations is the same. Situations $\mathbf{s}^*$ and $\mathbf{s}^\circ$ only differ with respect to the way income is distributed among the health types. Let $\mathbf{u}$ and $\mathbf{v}$ be the heterogeneous income distributions associated with the situations $\mathbf{s}^*$ and $\mathbf{s}^\circ$, respectively. Then, we say that $\mathbf{u}$ *sequential rank order dominates* $\mathbf{v}$, which we write $\mathbf{u} \geq_{SRO} \mathbf{v}$, if and only if $\mathbf{u}^{(h)} \geq_{RO} \mathbf{v}^{(h)}$, for all $h = 1, 2, \ldots, H-1$, and $\mathbf{u}^{(H)} \sim_{RO} \mathbf{v}^{(H)}$. If in addition $\mathbf{u}^{(k)} >_{RO} \mathbf{v}^{(k)}$, for some $k < H$, then we say that $\mathbf{u}$ *sequential rank order strictly dominates* $\mathbf{v}$, which we write $\mathbf{u} >_{SRO} \mathbf{v}$. Important for subsequent developments is the fact that, if $\mathbf{u} \geq_{SRO} \mathbf{v}$, then $\mathbf{u}$ is a permutation of $\mathbf{v}$. It must be noted that conditions (b-1), (b-2) and (b-3) of Proposition 4.1 are actually equivalent to the fact that $\mathbf{u} \geq_{SRO} \mathbf{v}$.[19] Making use of (8.4), $\mathbf{u} \geq_{SRO} \mathbf{v}$ amounts to requiring that

$$(8.5a) \qquad \tilde{u}_g^{(h)} \geqq \tilde{v}_g^{(h)}, \ \forall\, g = 1, 2, \ldots, \tilde{n}(h), \ \forall\, h = 1, 2, \ldots, H-1, \ \text{and}$$

$$(8.5b) \qquad \tilde{u}_g^{(H)} = \tilde{v}_g^{(H)}, \ \forall\, g = 1, 2, \ldots, \tilde{n}(H).$$

Given the above conventions, the statement that $\mathbf{s}^* \equiv (\mathbf{x}^*; \mathbf{a}^*)$ is obtained from $\mathbf{s}^\circ \equiv (\mathbf{x}^\circ; \mathbf{a}^\circ)$ by means of a favourable permutation is equivalent to saying that there exist two health statuses $i, j \in \mathscr{A}$ with $i < j$ and two individuals s, t with $s \in N(i)$ and $t \in N(j)$ such that

$$(8.6a) \qquad v_s^i = u_t^j < u_s^i = v_t^j;$$

$$(8.6b) \qquad v_g^i = u_g^i, \ \forall\, g \neq s; \ \ v_g^j = u_g^j, \ \forall\, g \neq t; \ \text{and}$$

$$(8.6c) \qquad \mathbf{u}^{(g)} = \mathbf{v}^{(g)}, \ \forall\, g \neq i, j.$$

Finally, we introduce three technical results that will make the proof of Proposition 4.1 easier. The first lemma is well-known (see, e.g., Saposnik (1981)) and it simply states that the ranking of income distributions by utilitarian unanimity over the class of non-decreasing utility functions is equivalent to the ranking implied by the rank order criterion. More precisely, letting $\mathbf{\Phi} := \{\phi : \mathscr{D} \to \mathbb{R} \mid \phi \text{ is non-decreasing}\}$, we have:

Lemma 8.1. *Let* $\mathbf{u}, \mathbf{v} \in \mathscr{D}^m$ ($m \geqq 2$). *Then, statements* (a) *and* (b) *below are equivalent:*

(a) $\sum_{g=1}^m \phi(u_g) \geqq \sum_{g=1}^m \phi(v_g)$, *for all* $\phi \in \mathbf{\Phi}$.

(b) $\mathbf{u} \geq_{RO} \mathbf{v}$.

---

[19] This follows from the fact that rank order dominance and first order stochastic dominance generate the same ranking of distributions (see, e.g., Levy (1998, Chapter 4)).

The second lemma provides a similar equivalence for the case of heterogeneous income distributions by invoking the sequential rank order criterion. Before we state the result, we need to introduce the following class of H-tuples of functions:

$$(8.7) \qquad \boldsymbol{\Psi} := \left\{ \boldsymbol{\psi} := (\psi_1, \ldots, \psi_H) \mid \psi_h'(s) \geqq \psi_{h+1}'(s), \, \forall \, s \in \mathscr{D}, \, \forall \, h = 1, 2, \ldots, H-1 \right\}.$$

Then, we have:

Lemma 8.2. *Let* $\mathbf{u} := (\mathbf{u}^1; \ldots; \mathbf{u}^H), \mathbf{v} := (\mathbf{v}^1; \ldots; \mathbf{v}^H) \in \mathscr{D}^{n(1)} \times \cdots \times \mathscr{D}^{n(H)}$, *where* $n(h) \geqq 1$, *for all* $h \in \mathscr{A}$. *Then, statements* (a) *and* (b) *below are equivalent:*

(a) $\sum_{h=1}^{H} \sum_{i=1}^{n(h)} \psi_h(u_i^h) \geqq \sum_{h=1}^{H} \sum_{i=1}^{n(h)} \psi_h(v_i^h)$, *for all* $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_H) \in \boldsymbol{\Psi}$.

(b) $\mathbf{u} \geqq_{SRO} \mathbf{v}$.

PROOF.

(a) $\implies$ (b). Let $\boldsymbol{\lambda} := (\lambda_1, \ldots, \lambda_H)$ and consider the H-tuple $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_H)$ defined by $\psi_h(s) := \lambda_h \, \phi(s)$, for all $h = 1, 2, \ldots, H$. Clearly, $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_H) \in \boldsymbol{\Psi}$ provided that $\phi'(s) \geqq 0$ and $\lambda_1 \geqq \lambda_2 \geqq \cdots \geqq \lambda_H$, which we assume. Choosing successively $\boldsymbol{\lambda} := (1, 0, 0, \ldots, 0, 0)$, $\boldsymbol{\lambda} := (1, 1, 0, \ldots, 0, 0), \ldots, \boldsymbol{\lambda} := (1, 1, 1, \ldots, 1, 0)$, $\boldsymbol{\lambda} := (1, 1, 1, \ldots, 1, 1)$, and $\boldsymbol{\lambda} := (-1, -1, -1, \ldots, -1, -1)$, condition (a) reduces to

$$(8.8a) \quad \sum_{g=1}^{n(1)} \phi(u_g^1) + \cdots + \sum_{g=1}^{n(h)} \phi(u_g^h) \geqq \sum_{g=1}^{n(1)} \phi(v_g^1) + \cdots + \sum_{g=1}^{n(h)} \phi(v_g^h), \, \forall \, h = 1, 2, \ldots, H-1, \text{ and}$$

$$(8.8b) \quad \sum_{g=1}^{n(1)} \phi(u_g^1) + \cdots + \sum_{g=1}^{n(H)} \phi(u_g^H) = \sum_{g=1}^{n(1)} \phi(v_g^1) + \cdots + \sum_{g=1}^{n(h)} \phi(v_g^H),$$

which holds for all functions $\phi \in \boldsymbol{\Phi}$. Appealing to Lemma 8.1, we conclude that $\mathbf{u}^{(h)} \geqq_{RO} \mathbf{v}^{(h)}$, for all $h = 1, 2, \ldots, H-1$, and $\mathbf{u}^{(H)} \sim_{RO} \mathbf{v}^{(H)}$, hence $\mathbf{u} \geqq_{SRO} \mathbf{v}$.

(b) $\implies$ (a). Given any $\mathbf{u} := (\mathbf{u}^1; \ldots; \mathbf{u}^H) \in \mathscr{D}^{n(1)} \times \cdots \times \mathscr{D}^{n(H)}$, we have the following equality:

$$(8.9) \qquad \sum_{h=1}^{H} \sum_{i=1}^{n(h)} \psi_h(u_i^h) = \sum_{h=1}^{H} \left[\sum_{i=1}^{n(h)} \psi_h(u_i^h) + \sum_{k=1}^{h-1} \sum_{j=1}^{n(k)} \psi_h(u_j^k) - \sum_{k=1}^{h-1} \sum_{j=1}^{n(k)} \psi_h(u_j^k) \right]$$

$$= \sum_{h=1}^{H-1} \left[\sum_{k=1}^{h} \sum_{i=1}^{n(k)} \left(\psi_h(u_i^k) - \psi_{h+1}(u_i^k) \right) \right] + \sum_{k=1}^{H} \sum_{i=1}^{n(k)} \psi_H(u_i^k).$$

Letting $f_h(s) := \psi_h(s) - \psi_{h+1}(s)$, for $h = 1, 2, \ldots, H-1$, and $f_H(s) := \psi_H(s)$, and upon substituting into (8.9), we obtain

$$(8.10) \qquad \sum_{h=1}^{H} \sum_{i=1}^{n(h)} \psi_h(u_i^h) = \sum_{h=1}^{H-1} \left[\sum_{k=1}^{h} \sum_{i=1}^{n(k)} f_h(u_i^k) \right] + \sum_{k=1}^{H} \sum_{i=1}^{n(k)} f_H(u_i^k) = \sum_{h=1}^{H} \sum_{i=1}^{\tilde{n}(h)} f_h(\tilde{u}_i^{(h)}).$$

Because $f_h'(s) := \psi_h'(s) - \psi_{h+1}'(s)$, in the light of (8.10), condition (a) can be equivalently rewritten as

$$(8.11) \qquad \sum_{h=1}^{H} \sum_{i=1}^{n(h)} \left(\psi_h(u_i^h) - \psi_h(v_i^h) \right) = \sum_{h=1}^{H} \sum_{i=1}^{\tilde{n}(h)} \left(f_h(\tilde{u}_i^{(h)}) - f_h(\tilde{v}_i^{(h)}) \right) \geqq 0,$$

for all $f_h(s)$ that are non-decreasing in s. Making use of the *Mean Value Theorem*, condition (8.11) is equivalent to

$$(8.12) \qquad \sum_{h=1}^{H-1} \sum_{i=1}^{\tilde{n}(h)} f'_h(\xi_i^h) \left[\tilde{u}_i^{(h)} - \tilde{v}_i^{(h)} \right] + \sum_{i=1}^{\tilde{n}(H)} f'_H(\xi_i^H) \left[\tilde{u}_i^{(H)} - \tilde{v}_i^{(H)} \right] \geqq 0,$$

for some $\xi_i^h \in (\tilde{u}_i^{(h)}, \tilde{v}_i^{(h)})$, for all $i \in \{1, 2, \ldots, n(h)-1, n(h)\}$ and all $h \in \mathscr{A}$. Given (8.5) and since $f'_h(s) \geqq 0$, for all s and all $h = 1, 2, \ldots, H-1$, we conclude that it is sufficient for (8.12) to hold that $\mathbf{u} \geqq_{SRO} \mathbf{v}$. $\qquad \square$

The assumption that individuals, whatever their health statuses, have different incomes in the two situations under comparison makes things easier. In conjunction with Lemma 8.2, our last technical result confirms that there is no loss of generality when comparing heterogeneous income distributions by means of utilitarian unanimity to restrict attention to the subpopulation of individuals whose incomes differ in the two situations. Given $\mathbf{u} := (\mathbf{u}^1; \ldots; \mathbf{u}^H), \mathbf{v} := (\mathbf{v}^1; \ldots; \mathbf{v}^H) \in \mathscr{D}^{n(1)} \times \cdots \times \mathscr{D}^{n(H)}$, where $n(h) \geqq 1$, for all $h \in \mathscr{A}$, we define:

$$(8.13a) \qquad \mathscr{A}_0 := \left\{ h \in \mathscr{A} \ \middle| \ u_g^h = v_g^h, \ \forall \, g \in N(h) \right\};$$

$$(8.13b) \qquad \mathscr{A}_1 := \left\{ h \in \mathscr{A} \ \middle| \ u_g^h \neq v_g^h, \ \exists \, g \in N(h) \right\};$$

$$(8.13c) \qquad S(h) := \left\{ g \in N(h) \ \middle| \ u_g^h = v_g^h \right\} \ \text{for } h \in \mathscr{A}_1;$$

$$(8.13d) \qquad T(h) := \left\{ g \in N(h) \ \middle| \ u_g^h \neq v_g^h \right\} \ \text{for } h \in \mathscr{A}_1;$$

and we note that $\varnothing \subseteq S(h) \subset N(h)$, $\varnothing \subset T(h) \subseteq N(h)$, and $S(h) \cup T(h) = N(h)$, for all $h \in \mathscr{A}_1$. Given the H-tuple $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_H) \in \boldsymbol{\Psi}$, we denote by $\boldsymbol{\psi}(\mathscr{A}_1) := ((\psi_h)_{h \in \mathscr{A}_1})$ its restriction to $\mathscr{A}_1$ and by $\boldsymbol{\Psi}(\mathscr{A}_1)$ the set of such profiles. Then, we have the following obvious result:

Lemma 8.3. *Let* $\mathbf{u} := (\mathbf{u}^1; \ldots; \mathbf{u}^H), \mathbf{v} := (\mathbf{v}^1; \ldots; \mathbf{v}^H) \in \mathscr{D}^{n(1)} \times \cdots \times \mathscr{D}^{n(H)}$, *where* $n(h) \geqq 1$, *for all* $h \in \mathscr{A}$. *Then, statements* (a) *and* (b) *below are equivalent:*

(a) $\sum_{h=1}^{H} \sum_{i=1}^{n(h)} \psi_h(u_i^h) \geqq \sum_{h=1}^{H} \sum_{i=1}^{n(h)} \psi_h(v_i^h)$, *for all* $\boldsymbol{\psi} := (\psi_1, \ldots, \psi_H) \in \boldsymbol{\Psi}$.

(b) $\sum_{h \in \mathscr{A}_1} \sum_{i \in T(h)} \psi_h(u_i^h) \geqq \sum_{h \in \mathscr{A}_1} \sum_{i \in T(h)} \psi_h(v_i^h)$, *for all* $\boldsymbol{\psi}(\mathscr{A}_1) \in \boldsymbol{\Psi}(\mathscr{A}_1)$.

PROOF OF PROPOSITION 4.1. One can easily check that if $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a finite sequence of favourable permutations, then $\mathbf{u} \geqq_{SRO} \mathbf{v}$. We therefore prove the converse implication and, thanks to Lemmas 8.2 and 8.3, we assume without loss of generality that

$$(8.14) \qquad u_g^h \neq v_g^h, \ \forall \, g \in N(h), \ \forall \, h \in \mathscr{A}.$$

Now consider the following indices:

$$(8.15a) \qquad i := \min \left\{ h \in \mathscr{A} \ \middle| \ \exists \, g \in N(h) : u_g^h \geqq u_\ell^p, \ \forall \, \ell \in N(p), \ \forall \, p \in \mathscr{A} \right\};$$

$$(8.15b) \qquad j := \min \left\{ h \in \mathscr{A} \ \middle| \ \exists \, g \in N(h) : v_g^h \geqq v_\ell^p, \ \forall \, \ell \in N(p), \ \forall \, p \in \mathscr{A} \right\};$$

(8.15c) $\qquad s := \min\left\{g \in N(i) \,\middle|\, u_g^i \geqq u_\ell^i,\, \forall\, \ell \in N(i)\right\};$

(8.15d) $\qquad t := \min\left\{g \in N(j) \,\middle|\, v_g^j \geqq v_\ell^j,\, \forall\, \ell \in N(j)\right\}.$

By definition of the indices i, j, s and t, we have

(8.16a) $\qquad u_1^i \leqq \cdots \leqq u_{s-1}^i < u_s^i = u_{s+1}^i = \cdots = u_{n(i)}^i;$

(8.16b) $\qquad u_g^h < u_s^i,\, \forall\, g \in N(h),\, \forall\, h = 1, 2, \ldots, i-1;$

(8.16c) $\qquad u_s^i \geqq u_g^h,\, \forall\, g \in N(h),\, \forall\, h = i+1, i+2, \ldots, H;$

(8.16d) $\qquad v_1^j \leqq \cdots \leqq v_{t-1}^j < v_t^j = v_{t+1}^j = \cdots = v_{n(j)}^j;$

(8.16e) $\qquad v_h^g < v_t^j,\, \forall\, g \in N(h),\, \forall\, h = 1, 2, \ldots, j-1;$

(8.16f) $\qquad v_t^j \geqq v_g^h,\, \forall\, g \in N(h),\, \forall\, h = j+1, i+2, \ldots, H.$

Furthermore, $u_s^i := \max\{u_g^h\} = \max\{v_g^h\} =: v_t^j$, because $\mathbf{u}$ is a permutation of $\mathbf{v}$. We note that by construction $i < j$. Indeed, we cannot have $i = j$, for, if this were the case, then $u_{n(i)}^i = v_{n(i)}^i$, which is ruled out by (8.14). Suppose next that $j < i$ and consider the distributions $\mathbf{u}^{(j)} := (\mathbf{u}^1; \mathbf{u}^2; \ldots; \mathbf{u}^j)$ and $\mathbf{v}^{(j)} := (\mathbf{v}^1; \mathbf{v}^2; \ldots; \mathbf{v}^j)$. By definition of v_t^j, we have

(8.17) $\qquad\qquad \tilde{u}_{\tilde{n}(j)}^{(j)} < u_s^i = v_t^j = \tilde{v}_{\tilde{n}(j)}^{(j)}.$

Hence $\neg\,[\mathbf{u}^{(j)} \geq_{RO} \mathbf{v}^{(j)}]$, which contradicts the fact that $\mathbf{u} \geq_{SRO} \mathbf{v}$ by assumption.

The idea of the proof is to operate favourable permutations within the subpopulation $N(i) \cup N(i+1) \cup \cdots \cup N(j-1) \cup N(j)$ so that, at the end of this process, one obtains a new distribution $\mathbf{z}$ such that (i) $u_g^i = z_g^i = v_t^j$, for some $g \in N(i)$, and (ii) $\mathbf{u} \geq_{SRO} \mathbf{z} >_{SRO} \mathbf{v}$. Let

(8.18a) $\quad k := \min\left\{h \in \{i, \ldots, j-1\} \,\middle|\, \exists\, g \in N(h) : v_g^h \geqq v_\ell^p,\, \forall\, \ell \in N(p),\, \forall\, p \in \{i, \ldots, j-1\}\right\};$

(8.18b) $\quad r := \min\left\{g \in N(k) \,\middle|\, v_g^k \geqq v_\ell^k,\, \forall\, \ell \in N(k)\right\}.$

We note that it is possible that $k = i$ and therefore that $r \in \{1, 2, \ldots, s-1\}$. By definition of the indices k and r, we have:

(8.19a) $\qquad v_r^k \leqq \cdots \leqq v_{r-1}^k \leqq v_r^k = v_{r+1}^k = \cdots = v_{n(k)}^k;$

(8.19b) $\qquad v_g^h < v_r^k,\, \forall\, g \in N(h),\, \forall\, h \in \{i, i+1, \ldots, k-1\};$

(8.19c) $\qquad v_r^k \geqq v_g^h,\, \forall\, g \in N(h),\, \forall\, h \in \{k+1, k+2, \ldots, j-1\};$

(8.19d) $\qquad v_r^k < v_t^j.$

Next we construct a new distribution $\mathbf{w} := (\mathbf{w}^1; \ldots; \mathbf{w}^H)$ starting from $\mathbf{v} := (\mathbf{v}^1; \ldots; \mathbf{v}^H)$ by means of a favourable permutation as indicated below

(8.20a) $\qquad \mathbf{w}^h = \mathbf{v}^h,\, \forall\, h \in \mathscr{A}\ (h \neq k, j);$

(8.20b) $\qquad w_g^k = v_g^k,\, \forall\, g \in N(k)\ (g \neq n(k));\ w_{n(k)}^k = v_t^j;$

(8.20c) $\qquad w_g^j = v_g^j,\, \forall\, g \in N(j)\ (g \neq t);\ w_t^j = v_r^k;$

Table 8.1: Construction of **w** starting from **v** by means of a favourable permutation

			Type i							Type k						Type j				
$\cdots$	1	$\cdots$	$s-1$	$\boxed{s}$	$s+1$	$\cdots$	$n(i)$	1	$\cdots$	$r-1$	$\boxed{r}$	$r+1$	$\cdots$	$\boxed{n(k)}$	1	$\cdots$	$t-1$	$\boxed{t}$	$t+1$	$\cdots$	$n(j)$
u $:$ $\cdots$	u^i_1	$\cdots$	u^i_{s-1}	$\boxed{u^i_s}$	u^i_{s+1}	$\cdots$	$u^i_{n(i)}$	u^k_1	$\cdots$	u^k_{r-1}	u^k_r	u^k_{r+1}	$\cdots$	$u^k_{n(k)}$	u^j_1	$\cdots$	u^j_{t-1}	u^j_t	u^j_{t+1}	$\cdots$	$u^j_{n(j)}$
w $:$ $\cdots$	v^i_1	$\cdots$	v^i_{s-1}	v^i_s	v^i_{s+1}	$\cdots$	$v^i_{n(i)}$	v^k_1	$\cdots$	v^k_{r-1}	$\boxed{v^k_r}$	v^k_{r+1}	$\cdots$	$\boxed{v^j_t}$	v^j_1	$\cdots$	v^j_{t-1}	$\boxed{v^k_{n(k)}}$	v^j_{t+1}	$\cdots$	$v^j_{n(j)}$
v $:$ $\cdots$	v^i_1	$\cdots$	v^i_{s-1}	v^i_s	v^i_{s+1}	$\cdots$	$v^i_{n(i)}$	v^k_1	$\cdots$	v^k_{r-1}	$\boxed{v^k_r}$	v^k_{r+1}	$\cdots$	$\boxed{v^k_{n(k)}}$	v^j_1	$\cdots$	v^j_{t-1}	$\boxed{v^j_t}$	v^j_{t+1}	$\cdots$	$v^j_{n(j)}$

and illustrated in Table 8.1.

We now examine the distributions $\mathbf{u}^{(h)}$, $\mathbf{w}^{(h)}$ and $\mathbf{v}^{(h)}$, for $h = 1, 2, \ldots, H$. We consider successively four cases.

CASE 1: $1 \leqq h \leqq i - 1$. By assumption $\mathbf{u}^{(h)} \geq_{RO} \mathbf{v}^{(h)}$ and by construction $\mathbf{w}^{(h)} = \mathbf{v}^{(h)}$. Hence, $\tilde{u}_g^{(h)} \geqq \tilde{w}_g^{(h)} = \tilde{v}_g^{(h)}$, for all $g \in N(h)$.

CASE 2: $i \leqq h \leqq j - 1$. If $k > h$, the argument is the same as in Case 1, so we can restrict attention to the case in which $k \in \{i, i + 1, \ldots, h - 1, h\}$. We indicate by $g^* = \chi(h, k)$ the largest rank with the income v_r^k in the ordered distribution $\tilde{\mathbf{v}}^{(h)}$, as shown in

$$(8.21) \quad \tilde{v}_1^{(h)} \leqq \cdots \leqq \tilde{v}_{g^* - (n(k) - r) - 1}^{(h)} \leqq \boxed{\tilde{v}_{g^* - (n(k) - r)}^{(h)} = \cdots = \boxed{\tilde{v}_{g^*}^{(h)} \equiv v_{n(k)}^k}} < \tilde{v}_{g^* + 1}^{(h)} \leqq \cdots \leqq \tilde{v}_{\tilde{n}(h)}^{(h)},$$

where we have made use of the fact that by definition $v_r^k = v_g^k$, for all $g = r + 1, r + 2, \ldots, n(k)$. We indicate by

$$(8.22) \qquad \rho^* := \# \left\{ g \in \{1, 2, \ldots, \tilde{n}(h) - 1, \tilde{n}(h)\} \mid \tilde{v}_g^{(h)} > \tilde{v}_{g^*}^{(h)} = v_r^k \right\} = \tilde{n}(h) - g^*$$

the number of individuals who in the distribution $\mathbf{v}^{(h)} := (\mathbf{v}^1; \mathbf{v}^2; \ldots; \mathbf{v}^h)$ have incomes greater than v_r^k. Because by definition $v_r^k \geqq v_g^h$, for all $g \in N(h)$ and all $h \in \{i, i + 1, \ldots, j - 1\}$, we note that

$$(8.23) \qquad n(i) + n(i + 1) + \cdots + n(h) \leqq g^* = \chi(h, k) \leqq \tilde{n}(h),$$

or equivalently that

$$(8.24) \qquad \rho^* \leqq \tilde{n}(h) - (n(i) + n(i + 1) + \cdots + n(h)) = n(1) + n(2) + \cdots + n(i - 1).$$

Suppose first that $g^* = \tilde{n}(h)$. Then, it follows from the definition of $\tilde{\mathbf{w}}^{(h)}$ and the fact that $\tilde{\mathbf{u}}^{(h)} \geq_{RO} \tilde{\mathbf{v}}^{(h)}$ by assumption that

$$(8.25a) \qquad \tilde{u}_g^{(h)} \geqq \tilde{w}_g^{(h)} = \tilde{v}_g^{(h)}, \ \forall \, g = 1, 2, \ldots, \tilde{n}(h) - 1, \ \text{and}$$

$$(8.25b) \qquad u_s^i = \tilde{u}_{\tilde{n}(h)}^{(h)} = \tilde{w}_{\tilde{n}(h)}^{(h)} = v_t^j > v_r^k = v_{n(k)}^k \equiv \tilde{v}_{\tilde{n}(h)}^{(h)}.$$

Consider next the case where $g^* \in \{1, 2, \ldots, \tilde{n}(h) - 1\}$. Invoking again the definition of $\tilde{\mathbf{w}}^{(h)}$ and the fact that $\tilde{\mathbf{u}}^{(h)} \geq_{RO} \tilde{\mathbf{v}}^{(h)}$ by assumption, we have

$$(8.26a) \qquad \tilde{u}_g^{(h)} \geqq \tilde{w}_g^{(h)} = \tilde{v}_g^{(h)}, \ \forall \, g = 1, 2, \ldots, g^* - 1, \ \text{and}$$

$$(8.26b) \qquad u_s^i = \tilde{u}_{\tilde{n}(h)}^{(h)} = \tilde{w}_{\tilde{n}(h)}^{(h)} = v_t^j > v_r^k = v_{n(k)}^k \equiv \tilde{v}_{\tilde{n}(h)}^{(h)}.$$

It remains to examine what happens when $g = g^*, g^* + 1, \ldots, \tilde{n}(h) - 1$.

By definition of the indices i and s, we have

$$(8.27) \qquad \tilde{u}_1^{(h)} \leqq \cdots \leqq \tilde{u}_{\tilde{n}(h-1)+1}^{(h)} \leqq \cdots \leqq \tilde{u}_{\tilde{n}(h-1)+s-1}^{(h)} < \boxed{\tilde{u}_{\tilde{n}(h-1)+s}^{(h)} = \cdots = \boxed{\tilde{u}_{\tilde{n}(h)}^{(h)} = u_s^i}}.$$

Given a non-empty and finite set $A := \{a_1, a_2, \ldots, a_m\}$, where $a_i \in \mathbb{R}$, for all $i = 1, 2, \ldots, m$ ($m \geqq 2$), we denote by $\max_\rho A$ the ρth-greatest element in A with $1 \leqq \rho \leqq m$. Since by definition $\tilde{u}^{(h)}_{\tilde{n}(h)} = u^i_s > \tilde{u}^{(i-1)}_g$, for all $g = 1, 2, \ldots, \tilde{n}(i-1) - 1, \tilde{n}(i-1)$, we deduce that

$$(8.28) \qquad \left\{\{\tilde{u}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} \subseteq \left\{\{\tilde{u}^{(h)}_g\}_{g=1,2,\ldots,\tilde{n}(h)-1}\right\}.$$

Thus we have:

$$\tilde{u}^{(h)}_{\tilde{n}(h)-1} = \max_1 \left\{\{u^{(h)}_g\}_{g=1,2,\ldots,\tilde{n}(h)-1}\right\} \geqq \max_1 \left\{\{u^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{u}^{(i-1)}_{\tilde{n}(i-1)-0}; \qquad \rho = 1$$

$$\tilde{u}^{(h)}_{\tilde{n}(h)-2} = \max_2 \left\{\{u^{(h)}_g\}_{g=1,2,\ldots,\tilde{n}(h)-1}\right\} \geqq \max_2 \left\{\{u^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{u}^{(i-1)}_{\tilde{n}(i-1)-1}; \qquad \rho = 2$$

$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots$$

$$\tilde{u}^{(h)}_{\tilde{n}(h)-\rho^*+1} = \max_{\rho^*-1} \left\{\{u^{(h)}_g\}_{g=1,2,\ldots,\tilde{n}(h)-1}\right\} \geqq \max_{\rho^*-1} \left\{\{u^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{u}^{(i-1)}_{\tilde{n}(i-1)-\rho^*+2}; \rho = \rho^* - 1$$

$$\tilde{u}^{(h)}_{\tilde{n}(h)-\rho^*} = \max_{\rho^*} \left\{\{u^{(h)}_g\}_{g=1,2,\ldots,\tilde{n}(h)-1}\right\} \geqq \max_{\rho^*} \left\{\{u^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{u}^{(i-1)}_{\tilde{n}(i-1)-\rho^*+1}; \qquad \rho = \rho^*.$$

More compactly:

$$(8.29) \quad \tilde{u}^{(h)}_{(\tilde{n}(h)-1)-\rho} = \max_\rho \left\{\{u^{(h)}_g\}_{g=1,2,\ldots,\tilde{n}(h)-1}\right\} \geqq \max_\rho \left\{\{u^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{u}^{(i-1)}_{\tilde{n}(i-1)-\rho+1},$$

for all $\rho = 1, 2, \ldots, \rho^* - 1, \rho^*$.

Similarly, since by definition $\tilde{v}^{(h)}_{g^*} = v^k_{n(k)} \geqq v^p_g$, for all $g \in N(p)$ and all $p = i, i+1, \ldots, h$, we deduce from (8.21) that

$$(8.30) \qquad \left\{\{\tilde{v}^{(h)}_g\}_{g=g^*+1,g^*+2,\ldots,\tilde{n}(h)}\right\} \subseteq \left\{\{\tilde{v}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\}.$$

This implies in turn that

$$\tilde{v}^{(h)}_{\tilde{n}(h)-0} = \max_1 \left\{\{\tilde{v}^{(h)}_g\}_{g=g^*+1,\ldots,\tilde{n}(h)}\right\} \leqq \max_1 \left\{\{\tilde{v}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{v}^{(i-1)}_{\tilde{n}(i-1)-0}; \qquad \rho = 1$$

$$\tilde{v}^{(h)}_{\tilde{n}(h)-1} = \max_2 \left\{\{\tilde{v}^{(h)}_g\}_{g=g^*+1,\ldots,\tilde{n}(h)}\right\} \leqq \max_2 \left\{\{\tilde{v}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{v}^{(i-1)}_{\tilde{n}(i-1)-1}; \qquad \rho = 2$$

$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots$$

$$\tilde{v}^{(h)}_{\tilde{n}(h)-\rho^*+2} = \max_{\rho^*-1} \left\{\{\tilde{v}^{(h)}_g\}_{g=g^*+1,\ldots,\tilde{n}(h)}\right\} \leqq \max_{\rho^*-1} \left\{\{\tilde{v}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{v}^{(i-1)}_{\tilde{n}(i-1)-\rho^*+2}; \rho = \rho^* - 1$$

$$\tilde{v}^{(h)}_{\tilde{n}(h)-\rho^*+1} = \max_{\rho^*} \left\{\{\tilde{v}^{(h)}_g\}_{g=g^*+1,\ldots,\tilde{n}(h)}\right\} \leqq \max_{\rho^*} \left\{\{\tilde{v}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{v}^{(i-1)}_{\tilde{n}(i-1)-\rho^*+1}; \qquad \rho = \rho^*.$$

More compactly:

$$(8.31) \quad \tilde{v}^{(h)}_{\tilde{n}(h)-\rho+1} = \max_\rho \left\{\{\tilde{v}^{(h)}_g\}_{g=g^*+1,g^*+2,\ldots,\tilde{n}(h)}\right\} \leqq \max_\rho \left\{\{\tilde{v}^{(i-1)}_g\}_{g=1,2,\ldots,\tilde{n}(i-1)}\right\} = \tilde{v}^{(i-1)}_{\tilde{n}(i-1)-\rho+1},$$

for all $\rho = 1, 2, \ldots, \rho^* - 1, \rho^*$.

Combining (8.29) and (8.31), making use of the definition of $\mathbf{w}$ and of the fact that $\mathbf{u}^{(i-1)} \geq_{RO} \mathbf{v}^{(i-1)}$, we obtain

$$\tilde{u}^{(h)}_{\tilde{n}(h)-1} \geqq \tilde{u}^{(i-1)}_{\tilde{n}(i-1)-0} \geqq \tilde{v}^{(i-1)}_{\tilde{n}(i-1)-0} \geqq \tilde{v}^{(h)}_{\tilde{n}(h)-0} = \tilde{w}^{(h)}_{\tilde{n}(h)-1} \geqq \tilde{v}^{(h)}_{\tilde{n}(h)-1}; \qquad \rho = 1$$

35

$$\tilde{u}_{\tilde{n}(h)-2}^{(h)} \geqq \tilde{u}_{\tilde{n}(i-1)-1}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(i-1)-1}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(h)-1}^{(h)} = \tilde{w}_{\tilde{n}(h)-2}^{(h)} \geqq \tilde{v}_{\tilde{n}(h)-2}^{(h)}; \qquad \rho = 2$$

$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots$$

$$\tilde{u}_{\tilde{n}(h)-\rho^*+1}^{(h)} \geqq \tilde{u}_{\tilde{n}(i-1)-\rho^*+2}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(i-1)-\rho^*+2}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(h)-\rho^*+2}^{(h)} = \tilde{w}_{\tilde{n}(h)-\rho^*+1}^{(h)} \geqq \tilde{v}_{\tilde{n}(h)-\rho^*+1}^{(h)}; \quad \rho = \rho^*-1$$

$$\tilde{u}_{\tilde{n}(h)-\rho^*}^{(h)} \geqq \tilde{u}_{\tilde{n}(i-1)-\rho^*+1}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(i-1)-\rho^*+1}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(h)-\rho^*+1}^{(h)} = \tilde{w}_{\tilde{n}(h)-\rho^*}^{(h)} \geqq \tilde{v}_{\tilde{n}(h)-\rho^*}^{(h)}; \qquad \rho = \rho^*.$$

More compactly:

$$(8.32) \qquad \tilde{u}_{\tilde{n}(h)-\rho}^{(h)} \geqq \tilde{u}_{\tilde{n}(i-1)-\rho+1}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(i-1)-\rho+1}^{(i-1)} \geqq \tilde{v}_{\tilde{n}(h)-\rho+1}^{(h)} = \tilde{w}_{\tilde{n}(h)-\rho}^{(h)} \geqq \tilde{v}_{\tilde{n}(h)-\rho}^{(h)},$$

for all $\rho = 1, 2, \ldots, \rho^* - 1, \rho^*$, or equivalently

$$(8.33) \qquad \tilde{u}_g^{(h)} \geqq \tilde{u}_{g+1}^{(i-1)} \geqq \tilde{v}_{g+1}^{(i-1)} \geqq \tilde{v}_g^{(h)} = \tilde{w}_g^{(h)} \geqq \tilde{v}_g^{(h)},$$

for all $g = g^*, g^* + 1, \ldots, \tilde{n}(h) - 2, \tilde{n}(h) - 1$. Thus, we conclude that

$$(8.34) \qquad \tilde{u}_g^{(h)} \geqq \tilde{w}_g^{(h)} \geqq \tilde{v}_g^{(h)}, \ \forall\, g = g^*, g^* + 1, \ldots, \tilde{n}(h) - 2, \tilde{n}(h) - 1.$$

CASE 3: $j \leqq h \leqq H - 1$. By assumption $\mathbf{u}^{(h)} \geq_{RO} \mathbf{v}^{(h)}$ and by construction $\mathbf{w}^{(h)}$ is a permutation of $\mathbf{v}^{(h)}$. Hence, $\tilde{u}_g^{(h)} \geqq \tilde{w}_g^{(h)} = \tilde{v}_g^{(h)}$, for all $g \in N(h)$.

CASE 4: $h = H$. By assumption $\mathbf{u}^{(H)} \sim_{RO} \mathbf{v}^{(H)}$ and by construction $\mathbf{w}^{(H)}$ is a permutation of $\mathbf{v}^{(H)}$. Hence, $\tilde{u}_g^{(H)} = \tilde{w}_g^{(H)} = \tilde{v}_g^{(H)}$, for all $g \in N(H)$.

To summarise, we have $\mathbf{u}^{(h)} \geq_{RO} \mathbf{w}^{(h)} \geq_{RO} \mathbf{v}^{(h)}$, for all $h = 1, 2, \ldots, H - 1$, with $\mathbf{w}^{(h)} >_{RO} \mathbf{v}^{(h)}$, for at least one h, and $\mathbf{u}^{(H)} \sim_{RO} \mathbf{w}^{(H)} \sim_{RO} \mathbf{v}^{(H)}$. Therefore $\mathbf{u} \geq_{SRO} \mathbf{w} >_{SRO} \mathbf{v}$. If $\mathbf{u} = \mathbf{w}$, then distribution $\mathbf{u}$ results from distribution $\mathbf{v}$ by means of a single favourable permutation and the proof is complete. If $\mathbf{u} \neq \mathbf{w}$, then we apply the above reasoning to the distributions $\hat{\mathbf{u}}$ and $\hat{\mathbf{w}}$ obtained by deleting all indices h and g such that $u_g^h = w_g^h$. It is possible that no such indices exist in which case $\hat{\mathbf{u}} = \mathbf{u}$ and $\hat{\mathbf{w}} = \mathbf{w}$. By successive permutations of the kind described above, we finally obtain a distribution $\mathbf{z}$ such that $u_{n(i)}^i = u_s^i = v_t^j = z_{n(i)}^i$. Since $i < j$, we need at most $(j - i)$ favourable permutations in order to obtain $\mathbf{z}$ starting from $\mathbf{v}$. This gives a maximum of $(H - 1)$ permutations in the case where $i = 1$ and $j = H$ with the consequence that $\mathbf{u}$ and $\mathbf{z}$ differ in at most $(\tilde{n}(H) - 1)$ elements. Repeating the above operation, we obtain distribution $\mathbf{u}$ from distribution $\mathbf{v}$ by means of a finite sequence of no more than $\tilde{n}(H)(H - 1)$ favourable permutations. $\qquad \square$

PROOF OF LEMMA 5.1. See Moyes (2011a, Remark 3.2). $\qquad \square$

PROOF OF LEMMA 5.2. Suppose that situation $\mathbf{s}^*$ is obtained from situation $\mathbf{s}^\circ$ by means of a between-type progressive income transfer as it is represented in Figure 5.2. For social welfare to increase as the result, we must have

$$(8.35) \qquad \sum_{g \neq i,j} U(x_g^\circ, a_g^\circ) + U(v, h) + U(w, k) \geqq \sum_{g \neq i,j} U(x_g^\circ, a_g^\circ) + U(u, h) + U(t, k).$$

Making use of the decomposition of a between-type progressive income transfer depicted in Figure 5.3, this is equivalent to

$$(8.36) \qquad U(v,h) + U(u,k) + U(w,k) \geqq U(u,h) + U(u,k) + U(t,k).$$

Invoking condition C1, we have

$$(8.37) \qquad U(v,h) + U(u,k) + U(w,k) \geqq U(u,h) + U(v,k) + U(w,k),$$

while condition C2 guarantees that

$$(8.38) \qquad U(u,h) + U(v,k) + U(w,k) \geqq U(u,h) + U(u,k) + U(t,k).$$

Combining the inequalities (8.37) and (8.38), we obtain (8.36), which proves that conditions C1 and C2 are sufficient for social welfare to increase as the result of a between-type progressive income transfer. $\square$

PROOF OF LEMMA 5.3. It is a direct consequence of the additive separability of the ordered poverty gap (see equation (5.8)). $\square$

PROOF OF PROPOSITION 5.1. This follows from standard results in the literature (see, e.g., Hardy *et al.* (1934), Berge (1963), Marshall and Olkin (1979)). $\square$

Before we proceed to the proof of Proposition 5.2, we need introduce some additional piece of notation. First, given a situation $\mathbf{s} \in \mathscr{S}_n$, we let

$$(8.39) \qquad M_1(\mathbf{s}) := \{v_j \in \mathscr{D} \mid \exists i : x_i = v_j\}$$

and we label the distinct elements in $M_1(\mathbf{s})$ so that $v_1 < v_2 < \cdots < v_{m_1(\mathbf{s})}$, where $m_1(\mathbf{s}) := \#M_1(\mathbf{s})$. Similarly, given $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$, we let

$$(8.40) \qquad M_1(\mathbf{s}^*, \mathbf{s}^\circ) := \{v_j \in \mathscr{D} \mid \exists i : x_i^* = v_j \text{ or } x_i^\circ = v_j\},$$

where by convention the elements in $M_1(\mathbf{s}^*, \mathbf{s}^\circ)$ are labelled in such a way that $v_1 < v_2 < \cdots < v_{m_1(\mathbf{s}^*, \mathbf{s}^\circ)}$, with $m_1(\mathbf{s}^*, \mathbf{s}^\circ) := \#M_1(\mathbf{s}^*, \mathbf{s}^\circ)$. Given the situation $\mathbf{s} \equiv (\mathbf{x}; \mathbf{a}) \in \mathscr{S}_n$, we find convenient to let

$$(8.41) \qquad \mathscr{P}(y \mid h) := f_2(h) P(y \mid h),$$

for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, H$. Thus, $\mathscr{P}(y \mid h)$ is the (average) conditional poverty gap of the group of individuals with health status h multiplied by the proportion of these individuals in the whole population. For further reference, we note that (8.41) can be rewritten as

$$(8.42) \qquad \mathscr{P}(y \mid h) = f_2(h) \int_{\underline{v}}^{y} F(t \mid h) \, dt = \int_{\underline{v}}^{y} \int_{\underline{v}}^{t} f(\xi, h) \, d\xi \, dt.$$

On the other hand, we have

$$(8.43) \qquad \int_{\underline{v}}^{y} F(t, h) \, dt = \sum_{k=1}^{h} \int_{\underline{v}}^{y} \int_{\underline{v}}^{t} f(\xi, k) \, d\xi \, dt, \ \forall \, y \in \mathscr{D}, \ \forall \, h = 1, 2, \ldots, H.$$

Combining (8.42) and (8.43), we get

$$(8.44) \qquad \mathscr{P}(y \,|\, h) = \int_{\underline{v}}^{y} F(t, h) \, dt - \int_{\underline{v}}^{y} F(t, h-1) \, dt, \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H,$$

where $F(y, 0) := 0$, for all $y \in \mathscr{D}$. Summing over $h = 1, 2, \ldots, H$, we obtain

$$(8.45) \qquad \int_{\underline{v}}^{y} F(t, h) \, dt = \sum_{k=1}^{h} \mathscr{P}(y \,|\, k), \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H.$$

Clearly, $\mathscr{P}(y \,|\, h)$ inherits all the properties of $P(y \,|\, h)$. In particular: $\mathscr{P}(y \,|\, h) = 0$, for all $y \leqq \min\{x_i\}$ and all $h = 1, 2, \ldots, H$, $\mathscr{P}(y \,|\, h)$ is an increasing and convex piecewise linear function over $[\min\{x_i\}, \max\{x_i\}]$ with kinks occurring for values of y in $M_1(\mathbf{s})$, and

$$(8.46) \qquad \mathscr{P}(y \,|\, h) = f_2(h) \, [y - \mu(\mathbf{x}(h))], \ \forall \ y \geqq \max\{x_i\}, \ \forall \ h = 1, 2, \ldots, H.$$

Differentiating (8.42), we obtain the right and left derivatives of $\mathscr{P}(y \,|\, h)$ with respect to income, denoted respectively as

$$(8.47a) \qquad \mathscr{P}_y(y \,|\, h)_+ = f_2(h) \, F(y \,|\, h) = \frac{n_2(h)}{n} \frac{q(y \,|\, h)}{n_2(h)} = \frac{q(y \,|\, h)}{n} \ \text{ and}$$

$$(8.47b) \qquad \mathscr{P}_y(y \,|\, h)_- = f_2(h) \, F_-(y \,|\, h) = \frac{n_2(h)}{n} \frac{q_-(y \,|\, h)}{n_2(h)} = \frac{q_-(y \,|\, h)}{n},$$

where $q(y \,|\, h) = \#\{i \,|\, x_i \leqq y \text{ and } a_i = h\}$ and $q_-(y \,|\, h) = \#\{i \,|\, x_i < y \text{ and } a_i = h\}$. Using (8.47a) and (8.47b), we define the second derivative

$$(8.48) \qquad \mathscr{P}_{yy}(y \,|\, h) = \mathscr{P}_y(y \,|\, h)_+ - \mathscr{P}_y(y \,|\, h)_- = f_2(h) \, [F(y \,|\, h) - F_-(y \,|\, h)]$$

$$= f_2(h) \, f(y \,|\, h) = f(y, h) = \frac{n(y, h)}{n}.$$

We note that the first derivatives of $\mathscr{P}(y \,|\, h)$ are related to the conditional distribution function of income, while the second derivative is equal to the joint density function. Given two situations $\mathbf{s}^*, \mathbf{s}^\circ \in \mathscr{S}_n$, we let $v_- := \min\{x_i^*, x_i^\circ\}$ and $v_+ := \max\{x_i^*, x_i^\circ\}$ to simplify notation.

PROOF OF PROPOSITION 5.2. That condition (b) is sufficient for $\mathbf{s}^*$ to ordered poverty gap dominate $\mathbf{s}^\circ$ follows from Lemma 5.3 and the fact that the classes of utility functions that are respectively submodular and concave in income for every type are subsets of the class $\mathscr{U}_2$. The proof of the converse statement is rather lengthy and it will involve four steps. In the *first step*, assuming that $\mathbf{s}^*$ ordered poverty gap dominates $\mathbf{s}^\circ$, we follow Bourguignon (1989) and we derive a function $\Psi(y \,|\, h)$ that will play a crucial role in our separation result. This function is used in the *second step* to construct *pseudo conditional poverty curves* with the property that (i) they are nowhere above the conditional poverty curves of $\mathbf{s}^\circ$ and (ii) they generate a *pseudo joint distribution function* that lies nowhere below the joint distribution function of $\mathbf{s}^*$. However, the corresponding densities may well be negative for some $(y, h) \in \mathscr{D} \times \mathscr{A}$, which implies that these conditional poverty curves can be *neither* increasing *nor* convex on the domain $\mathscr{D}$. In the *third step*, starting with the function $\Psi(y \,|\, h)$, we construct two situations

$\tilde{\mathbf{s}}$ and $\hat{\mathbf{s}}^{\ddagger}$, where the first one will play the role of the phantom situation, while the second will be our separating situation. Finally, we verify in the *fourth step* that by augmenting the initial situations with the phantom situation it is possible to obtain the desired separation. To some extent, the second step of the proof provides an indication about the reasons why it is impossible to separate the dominating situation from the dominated one when the population is fixed.

STEP 1. Suppose that statement (a) holds or equivalently that

$$(8.49) \qquad \sum_{h=1}^{H} \left[\mathscr{P}^{\circ}(z(h) \,|\, h) - \mathscr{P}^{*}(z(h) \,|\, h) \right] \geqq 0, \ \forall \ z(1) \geqq z(2) \geqq \cdots \geqq z(H).$$

Consider the function $\Psi : \mathscr{D} \times \{0, 1, 2, \ldots, H\}$ defined recursively by $\Psi(y \,|\, 0) := 0$, for all $y \in \mathscr{D}$, and

$$(8.50) \qquad \Psi(y \,|\, h) := \min \left\{ \mathscr{P}^{\circ}(t_h \,|\, h) - \mathscr{P}^{*}(t_h \,|\, h) + \Psi(t_h \,|\, h-1) \mid t_h \geqq y \right\},$$

for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, H$. By construction, $\Psi(y \,|\, h)$ is a piecewise linear function where kinks occur possibly for values of y in $M_1\,(\mathbf{s}^*, \mathbf{s}^{\circ})$. We indicate respectively by

$$(8.51\text{a}) \qquad \Psi_y(y \,|\, h)_+ =: \phi_2(h)\,\psi(y \,|\, h) \text{ and}$$

$$(8.51\text{b}) \qquad \Psi_y(y \,|\, h)_- =: \phi_2(h)\,\psi(y \,|\, h)_-$$

the right and left derivatives of $\Psi(y \,|\, h)$ with respect to income, and we note that they are rational numbers. The second derivative is

$$(8.52) \qquad \Psi_{yy}(y \,|\, h) = \Psi_y(y \,|\, h)_+ - \Psi_y(y \,|\, h)_- := \phi_2(h) \left[\psi(y \,|\, h) - \psi(y \,|\, h)_- \right]$$

$$= \phi_2(h)\,\phi(y \,|\, h) = \phi(y, h),$$

and it is also a rational number being the difference of two rational numbers.

CLAIM 1. The function Ψ has the following properties:

A1 $\Psi(y \,|\, h) \geqq 0, \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H$;

A2 $\Psi(y \,|\, h)$ is non-decreasing in $y, \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H$;

A3 $\Psi(y \,|\, h) - \Psi(y \,|\, h-1) \leqq \mathscr{P}^{\circ}(y \,|\, h) - \mathscr{P}^{*}(y \,|\, h), \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H$;

A4 $\Psi(y \,|\, h) \leqq \sum_{k=1}^{h} \left[\mathscr{P}^{\circ}(y \,|\, k) - \mathscr{P}^{*}(y \,|\, k) \right], \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H$;

A5 $\Psi(\overline{v} \,|\, h) - \Psi(\overline{v} \,|\, h-1) = \mathscr{P}^{\circ}(\overline{v} \,|\, h) - \mathscr{P}^{*}(\overline{v} \,|\, h), \ \forall \ h = 1, 2, \ldots, H$;

A6 $\Psi(\overline{v} \,|\, h) = \sum_{k=1}^{h} \left[\mathscr{P}^{\circ}(\overline{v} \,|\, k) - \mathscr{P}^{*}(\overline{v} \,|\, k) \right], \ \forall \ h = 1, 2, \ldots, H$;

A7 $\Psi(y \,|\, H) = 0, \ \forall \ y \in \mathscr{D}$.

PROOF.

<u>A1</u>. Given $h \in \{1, 2, \ldots, H\}$ and $y \in \mathscr{D}$, we denote by $t_h(y)$ the smallest $t_h \geqq y$ such that

$$(8.53) \quad \mathscr{P}^\circ(t_h \,|\, h) - \mathscr{P}^*(t_h \,|\, h) + \Psi(t_h \,|\, h-1) \leqq \mathscr{P}^\circ(v_h \,|\, h) - \mathscr{P}^*(v_h \,|\, h) + \Psi(v_h \,|\, h-1),$$

for all $v_h \geqq y$.

When $h = 1$, the result follows from choosing $\mathbf{z} = (y, \underline{v}, \underline{v}, \ldots, \underline{v})$, where $y \geqq \underline{v}$. Then, condition (8.49) implies that

$$(8.54) \qquad\qquad \mathscr{P}^\circ(y \,|\, 1) - \mathscr{P}^*(y \,|\, 1) \geqq 0, \ \forall \ y \geqq \underline{v},$$

which implies that

$$(8.55) \qquad\qquad \Psi(y \,|\, 1) = \mathscr{P}^\circ(t_1(y) \,|\, 1) - \mathscr{P}^*(t_1(y) \,|\, 1) \geqq 0,$$

for all $y \in \mathscr{D}$.

When $h = 2$, we choose $\mathbf{z} = (t_1(y), y, \underline{v}, \ldots, \underline{v})$, where $t_1(y) \geqq y \geqq \underline{v}$, and it follows from condition (8.49) that

$$(8.56) \qquad \mathscr{P}^\circ(y \,|\, 2) - \mathscr{P}^*(y \,|\, 2) + \mathscr{P}^\circ(t_1(y) \,|\, 1) - \mathscr{P}^*(t_1(y) \,|\, 1) \geqq 0, \ \forall \ y \geqq \underline{v},$$

which implies in turn that

$$(8.57) \quad \Psi(y \,|\, 2) = \mathscr{P}^\circ(t_2(y) \,|\, 2) - \mathscr{P}^*(t_2(y) \,|\, 2) + \mathscr{P}^\circ(t_1 \circ t_2(y) \,|\, 1) - \mathscr{P}^*(t_1 \circ t_2(y) \,|\, 1) \geqq 0,$$

for all $y \in \mathscr{D}$.

In the general case where $h > 2$, choose

$$(8.58) \qquad \mathbf{z} = (t_1 \circ t_2 \circ \cdots \circ t_{h-1} \circ t_h(y), t_2 \circ \cdots \circ t_{h-1} \circ t_h(y), \ldots, t_{h-1}(y), y, \underline{v}, \ldots, \underline{v}),$$

and note that

$$(8.59) \qquad t_1 \circ t_2 \circ \cdots \circ t_{h-1} \circ t_h(y) \geqq t_2 \circ \cdots \circ t_{h-1} \circ t_h(y) \geqq \ldots \geqq t_{h-1}(y) \geqq y.$$

Then, condition (8.49) implies that

$$(8.60) \qquad \mathscr{P}^\circ(y \,|\, h) - \mathscr{P}^*(y \,|\, h) +$$

$$\mathscr{P}^\circ(t_{h-1}(y) \,|\, h-1) - \mathscr{P}^*(t_{h-1}(y) \,|\, h-1) +$$

$$\vdots$$

$$\mathscr{P}^\circ(t_2 \circ \cdots \circ t_{h-1} \circ t_h(y) \,|\, 2) - \mathscr{P}^*(t_2 \circ \cdots \circ t_{h-1} \circ t_h(y) \,|\, 2) +$$

$$\mathscr{P}^\circ(t_1 \circ t_2 \circ \cdots \circ t_{h-1} \circ t_h(y) \,|\, 1) - \mathscr{P}^*(t_1 \circ t_2 \circ \cdots \circ t_{h-1} \circ t_h(y) \,|\, 1) \geqq 0,$$

for all $y \geqq \underline{v}$. This implies in turn that

(8.61) $\Psi(y \mid h) =$

$\mathscr{P}^{\circ}(t_h(y) \mid h) - \mathscr{P}^*(t_h(y) \mid h) +$

$\mathscr{P}^{\circ}(t_{h-1} \circ t_h(y) \mid h-1) - \mathscr{P}^*(t_{h-1} \circ t_h(y) \mid h-1) +$

$\vdots$

$\mathscr{P}^{\circ}(t_2 \circ t_3 \circ \cdots \circ t_{h-1} \circ t_h(y) \mid 2) - \mathscr{P}^*(t_2 \circ t_3 \circ \cdots \circ t_{h-1} \circ t_h(y) \mid 2) +$

$\mathscr{P}^{\circ}(t_1 \circ t_2 \circ t_3 \circ \cdots \circ t_{h-1} \circ t_h(y) \mid 1) - \mathscr{P}^*(t_1 \circ t_2 \circ t_3 \circ \cdots \circ t_{h-1} \circ t_h(y) \mid 1) \geqq 0,$

for all $y \geqq \underline{v}$.

A2. Consider a function $\Xi : \mathbb{R} \times \{1, 2, \ldots, H\} \to \mathbb{R}$. By the definition of the min function, we have

(8.62) $\min \{ \Xi(t, h) \mid t \geqq y^{\circ} \} \leqq \min \{ \Xi(t, h) \mid t \geqq y^* \} , \ \forall \ y^* > y^{\circ}, \ \forall \ h = 1, 2, \ldots, H.$

The result follows by choosing $\Xi(t, h) = \mathscr{P}^{\circ}(t \mid h) - \mathscr{P}^*(t \mid h) + \Psi(t \mid h-1)$.

A3. By definition of $\Psi(y \mid h)$, we have

(8.63) $\Psi(y \mid h) \leqq \mathscr{P}^{\circ}(t_h \mid h) - \mathscr{P}^*(t_h \mid h) + \Psi(t_h \mid h-1), \ \forall \ t_h \geqq y, \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H.$

In particular, for $t_h = y$, we get

(8.64) $\Psi(y \mid h) \leqq \mathscr{P}^{\circ}(y \mid h) - \mathscr{P}^*(y \mid h) + \Psi(y \mid h-1), \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H,$

and A3 follows.

A4. This result follows from A3 by summation from 1 to h, where $h = 1, 2, \ldots, H$.

A5. By the definition of $\Psi(y \mid h)$, we obtain

(8.65) $\Psi(\overline{v} \mid h) = \mathscr{P}^{\circ}(\overline{v} \mid h) - \mathscr{P}^*(\overline{v} \mid h) + \Psi(\overline{v} \mid h-1), \ \forall \ h = 1, 2, \ldots, H,$

when $y = \overline{v}$, which is precisely A5.

A6. This result follows from A5 by summation from 1 to h, where $h = 1, 2, \ldots, H$.

A7. It follows from A4 that

(8.66) $$\Psi(\overline{v} \mid H) = \sum_{h=1}^{H} \left[\mathscr{P}^{\circ}(\overline{v} \mid h) - \mathscr{P}^*(\overline{v} \mid h) \right].$$

Making use of (8.46) and upon substitution into (8.66), we obtain

(8.67)
$$\Psi(\overline{v} \mid H) = \sum_{h=1}^{H} f_2^{\circ}(h)\left[\overline{v} - \mu(\mathbf{x}^{\circ}(h))\right] - \sum_{h=1}^{H} f_2^{*}(h)\left[\overline{v} - \mu(\mathbf{x}^{*}(h))\right]$$

$$= \left(\overline{v} - \sum_{h=1}^{H} f_2^{\circ}(h)\,\mu(\mathbf{x}^{\circ}(h))\right) - \left(\overline{v} - \sum_{h=1}^{H} f_2^{*}(h)\,\mu(\mathbf{x}^{*}(h))\right)$$

$$= \sum_{h=1}^{H} f_2^{*}(h)\,\mu\left(\mathbf{x}^{*}(h)\right) - \sum_{h=1}^{H} f_2^{\circ}(h)\,\mu\left(\mathbf{x}^{\circ}(h)\right)$$

$$= \mu\left(\mathbf{x}^{*}\right) - \mu\left(\mathbf{x}^{\circ}\right) = 0.$$

We know from A2 that $\Psi(y \mid H)$ is non-decreasing in y. This, in conjunction with (8.67) implies that $\Psi(y \mid H) = 0$, for all $y \in \mathscr{D}$. $\qquad\square$

STEP 2. Now define

(8.68)
$$\mathscr{P}^{\ddagger}(y \mid h) := \mathscr{P}^{*}(y \mid h) + \left[\Psi(y \mid h) - \Psi(y \mid h-1)\right], \ \forall\, y \in \mathscr{D}, \ \forall\, h = 1, 2, \ldots, H.$$

Then, we have:

CLAIM 2. The function $\mathscr{P}^{\ddagger}(y \mid h)$ has the following properties:

B1 $\mathscr{P}^{\ddagger}(y \mid h) = 0, \ \forall\, y \leqq v_{-}, \ \forall\, h = 1, 2, \ldots, H$;

B2 $\mathscr{P}^{\ddagger}(y \mid h) \leqq \mathscr{P}^{\circ}(y \mid h), \ \forall\, y < v_{+}, \ \forall\, h = 1, 2, \ldots, H$;

B3 $\mathscr{P}^{\ddagger}(y \mid h) = \mathscr{P}^{\circ}(y \mid h), \ \forall\, y \geqq v_{+}, \ \forall\, h = 1, 2, \ldots, H$;

B4 $F^{\ddagger}(y, h) \geqq F^{*}(y, h), \ \forall\, y \in \mathscr{D}, \ \forall\, h = 1, 2, \ldots, H-1$; and

B5 $F^{\ddagger}(y, H) = F^{*}(y, H), \ \forall\, y \in \mathscr{D}$.

PROOF.

B1. This follows from the fact that $\mathscr{P}^{\circ}(y \mid h) = 0$ and $\Psi(y \mid h) = 0$, for all $y \leqq v_{-}$ and all $h = 1, 2, \ldots, H$.

B2 and B3. Using (8.50) and invoking respectively A3 and A5, we get

(8.69a)
$$\mathscr{P}^{\ddagger}(y \mid h) \leqq \mathscr{P}^{*}(y \mid h) + \left[\mathscr{P}^{\circ}(y \mid h) - \mathscr{P}^{*}(y \mid h)\right] = \mathscr{P}^{\circ}(y \mid h), \ \forall\, y \in \mathscr{D}, \text{ and}$$

(8.69b)
$$\mathscr{P}^{\ddagger}(y \mid h) = \mathscr{P}^{*}(y \mid h) + \left[\mathscr{P}^{\circ}(y \mid h) - \mathscr{P}^{*}(y \mid h)\right] = \mathscr{P}^{\circ}(y \mid h), \ \forall\, y \geqq v_{+},$$

for all $h = 1, 2, \ldots, H$.

B4 and B5. Taking the right derivative of (8.43), we get

(8.70)
$$F(y, h) = \sum_{k=1}^{h} \mathscr{P}_y(y \mid k)_{+}, \ \forall\, y \in \mathscr{D}, \ \forall\, h = 1, 2, \ldots, H.$$

Making use of (8.68), we obtain

$$(8.71) \quad F^{\ddagger}(y,h) - F^*(y,h) = \sum_{k=1}^{h} \left[\mathscr{P}^*_y(y \mid k)_+ + \left(\Psi_y(y \mid k)_+ - \Psi_y(y \mid k-1)_+ \right) - \mathscr{P}^*_y(y \mid k)_+ \right]$$

which simplifies to

$$(8.72) \quad F^{\ddagger}(y,h) - F^*(y,h) = \sum_{k=1}^{h} \left[\Psi_y(y \mid k)_+ - \Psi_y(y \mid k-1)_+ \right] = \Psi_y(y \mid h)_+,$$

for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, H$. Invoking A2 we deduce that

$$(8.73) \quad F^{\ddagger}(y,h) \geqq F^*(y,h), \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H.$$

We know from A7 that $\Psi(y \mid H) = 0$, for all $y \in \mathscr{D}$. This implies that $\Psi_y(y \mid H)_+ = 0$, for all $y \in \mathscr{D}$, and we therefore conclude that

$$(8.74) \quad F^{\ddagger}(y,H) = F^*(y,H), \ \forall \ y \in \mathscr{D}$$

which makes the proof complete. $\qquad \square$

To sum up, starting with $\mathscr{P}^*(y \mid h)$, $\mathscr{P}^{\circ}(y \mid h)$ and the assumption that $\mathscr{P}(\mathbf{z};(\mathbf{x}^*;\mathbf{a}^*)) \leqq \mathscr{P}(\mathbf{z};(\mathbf{x}^{\circ};\mathbf{a}^{\circ}))$, for all $\mathbf{z}$ such that $z(h)$ is non-increasing in h, we have constructed $\mathscr{P}^{\ddagger}(y \mid h)$ verifying properties B1 to B5. However, nothing guarantees that $\mathscr{P}^{\ddagger}_{yy}(y \mid h) =: f^{\ddagger}(y \mid h)$ is a density function, i.e:

$$(8.75) \quad f^{\ddagger}(y,h) = f^{\ddagger}_2(h) \left[F^{\ddagger}(y \mid h) - F^{\ddagger}_-(y \mid h) \right] \geqq 0, \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H; \text{ and}$$

$$(8.76) \quad \sum_{h=1}^{H} \int_{\underline{v}}^{\overline{v}} f^{\ddagger}(y,h) \, dy = \sum_{h=1}^{H} \int_{\underline{v}}^{\overline{v}} f^{\ddagger}_2(h) \left[F^{\ddagger}(y \mid h) - F^{\ddagger}_-(y \mid h) \right] \, dy = 1.$$

This implies in particular that $\mathscr{P}^{\ddagger}(y \mid h)$ may fail to be increasing and convex over the domain $[v_-, v_+]$. More importantly, $f^{\ddagger}(y,h)$ cannot be considered a density function and no situation $\mathbf{s}^{\ddagger}$ can be associated with it. Therefore, we find convenient to think of $\mathscr{P}^{\ddagger}(y \mid h)$ and $F^{\ddagger}(y,h)$ as a *pseudo conditional poverty curve* and a *pseudo joint distribution function*, respectively.

STEP 3. Differentiating (8.68) two times with respect to income, we obtain

$$(8.77) \quad f^{\ddagger}(y,h) = f^*(y,h) + \left[\phi(y,h) - \phi(y,h-1) \right], \ \forall \ y \in \mathscr{D}, \ \forall \ h = 1, 2, \ldots, H,$$

where by definition $f^*(y,h)$ and $\phi(y,h)$ are rational numbers. By assumption, the situations $\mathbf{s}^*$ and $\mathbf{s}^{\circ}$ have the same dimension equal to n, and by definition $f^*(y,h) = n^*(y,h)/n$, for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, H$. Then, it is always possible to find integers $n^{\ddagger}(y,h)$ and $r(y,h)$ such that

$$(8.78) \quad \frac{n^{\ddagger}(y,h)}{n} = \frac{n^*(y,h)}{n} + \left[\frac{r(y,h)}{n} - \frac{r(y,h-1)}{n} \right],$$

which implies in turn that

$$(8.79) \quad n^{\ddagger}(y,h) = n^*(y,h) + \left[r(y,h) - r(y,h-1) \right],$$

for all $y \in \mathscr{D}$ and all $h = 1, 2, \ldots, H$. Given any $h \in \{1, 2, \ldots, H\}$, there are two possibilities: *either* $n^{\ddagger}(y, h) \geqq 0$, for all $y \in \mathscr{D}$, *or* $n^{\ddagger}(y, h) < 0$, for some $y \in \mathscr{D}$.

Consider the function $\tilde{n} : \mathscr{D} \times \{1, 2, \ldots, H\} \to \mathbb{N}$ defined by

$$(8.80) \qquad \tilde{n}(y, h) := \max\left\{0, -n^{\ddagger}(y, h)\right\} \geqq 0, \ \forall \, y \in \mathscr{D}, \ \forall \, h = 1, 2, \ldots, H.$$

Then define

$$(8.81a) \qquad \tilde{n}^{*}(y, h) := n^{*}(y, h) \qquad\qquad\qquad + \tilde{n}(y, h),$$

$$(8.81b) \qquad \tilde{n}^{\circ}(y, h) := n^{\circ}(y, h) \qquad\qquad\qquad + \tilde{n}(y, h),$$

$$(8.81c) \qquad \hat{n}^{\ddagger}(y, h) := n^{*}(y, h) + \left[r(y, h) - r(y, h-1)\right] + \tilde{n}(y, h),$$

for all $y \in \mathscr{D}$ and $h = 1, 2, \ldots, H$. We note that $n^{*}(y, h)$, $n^{\circ}(y, h)$, $r(y, h)$, and $\tilde{n}(y, h)$ are all integers. Upon summation, we have

$$(8.82) \qquad \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{*}, \mathbf{s}^{\circ})} \tilde{n}^{*}(y, h) = \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{*}, \mathbf{s}^{\circ})} n^{*}(y, h) + \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{*}, \mathbf{s}^{\circ})} \tilde{n}(y, h) = n + \tilde{n},$$

where $\tilde{n} := \sum_{h=1}^{H} \sum_{v \in M_1(\mathbf{s}^{*}, \mathbf{s}^{\circ})} \tilde{n}(y, h)$. Similarly, we have

$$(8.83) \qquad \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} \tilde{n}^{\circ}(y, h) = \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} n^{\circ}(y, h) + \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} \tilde{n}(y, h) = n + \tilde{n}.$$

Now, we deduce from (8.81c), (8.82), and (8.83) that a necessary and sufficient condition for

$$(8.84) \qquad \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} \hat{n}^{\ddagger}(y, h) = n + \tilde{n},$$

is that

$$(8.85) \qquad \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} \left[r(y, h) - r(y, h-1)\right] = 0.$$

This is equivalent to

$$(8.86) \qquad \sum_{h=1}^{H} \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} \left[\Psi_{yy}(y \mid h) - \Psi_{yy}(y \mid h-1)\right] = \sum_{y \in M_1(\mathbf{s}^{\circ}, \mathbf{s}^{\circ})} \Psi_{yy}(y \mid H) = 0,$$

where we have made use of condition A7 according to which $\Psi(y \mid H) = 0$, for all $y \in \mathscr{D}$. The density functions of $\tilde{\mathbf{s}}^{*} := (\mathbf{s}^{*}; \tilde{\mathbf{s}})$, $\tilde{\mathbf{s}}^{\circ} := (\mathbf{s}^{\circ}; \tilde{\mathbf{s}})$, and $\hat{\mathbf{s}}^{\ddagger}$ are respectively given by

$$(8.87a) \qquad \tilde{f}^{*}(y, h) = (1 - \lambda) f^{*}(y, h) + \lambda \tilde{f}(y, h),$$

$$(8.87b) \qquad \tilde{f}^{\circ}(y, h) = (1 - \lambda) f^{\circ}(y, h) + \lambda \tilde{f}(y, h), \text{ and}$$

$$(8.87c) \qquad \hat{f}^{\ddagger}(y, h) = (1 - \lambda)\left[f^{*}(y, h) + (\phi(y, h) - \phi(y, h-1))\right] + \lambda \tilde{f}(y, h),$$

where

(8.88a) $\tilde{f}(y, h)$ is the density function of the phantom distribution $\tilde{\mathbf{s}} \in \mathscr{S}_q,$

(8.88b) $\lambda = \dfrac{\tilde{n}}{n + \tilde{n}},$ and

(8.88c) $q = \tilde{n}.$

STEP 4. It remains to check that conditions B2, B3, B4 and B5 hold with

(i) $\widetilde{\mathscr{P}^{\ddagger}}(y \mid h)$ and $\widetilde{\mathscr{P}^{\circ}}(y \mid h)$ in place of $\mathscr{P}^{\ddagger}(y \mid h)$ and $\mathscr{P}^{\circ}(y \mid h)$, respectively, and

(ii) $\widehat{F}^{\ddagger}(y, h)$ and $\widetilde{F}^{*}(y, h)$ in place of $F^{\ddagger}(y, h)$ and $F^{*}(y, h)$, respectively.

Using (8.42), we obtain

(8.89a) $\displaystyle \widetilde{\mathscr{P}^{\circ}}(y \mid h) = \int_{\underline{v}}^{y} \int_{\underline{v}}^{t} \left[(1 - \lambda) f^{\circ}(\xi, h) + \lambda \tilde{f}(\xi, h) \right] d\xi \, dt,$ and

(8.89b) $\displaystyle \widetilde{\mathscr{P}^{\ddagger}}(y \mid h) = \int_{\underline{v}}^{y} \int_{\underline{v}}^{t} \left[(1 - \lambda) \left(f^{*}(\xi, h) + (\phi(\xi, h) - \phi(\xi, h - 1)) \right) + \lambda \tilde{f}(\xi, h) \right] d\xi \, dt,$

for all $y \in \mathscr{D}$ and $h = 1, 2, \ldots, H$. Then, we have

(8.90) $\widetilde{\mathscr{P}^{\circ}}(y \mid h) - \widetilde{\mathscr{P}^{\ddagger}}(y \mid h) =$

$$\int_{\underline{v}}^{y} \int_{\underline{v}}^{t} (1 - \lambda) \left[f^{\circ}(\xi, h) - f^{*}(\xi, h) \right] d\xi \, dt - \int_{\underline{v}}^{y} \int_{\underline{v}}^{t} (1 - \lambda) \left[\phi(\xi, h) - \phi(\xi, h - 1) \right] d\xi \, dt =$$

$$(1 - \lambda) \left[\left(\mathscr{P}^{\circ}(y \mid h) - \mathscr{P}^{\ddagger}(y \mid h) \right) - \left(\Psi(y \mid h) - \Psi(y \mid h - 1) \right) \right],$$

for all $y \in \mathscr{D}$ and $h = 1, 2, \ldots, H$. Invoking A3 and A5, we conclude that

(8.91a) $\widetilde{\mathscr{P}^{\circ}}(y \mid h) \geqq \widetilde{\mathscr{P}^{\ddagger}}(y \mid h), \ \forall \, y \in \mathscr{D}, \ \forall \, h = 1, 2, \ldots, H - 1;$ and

(8.91b) $\widetilde{\mathscr{P}^{\circ}}(y \mid H) \geqq \widetilde{\mathscr{P}^{\ddagger}}(y \mid H), \ \forall \, y \in \mathscr{D}.$

Furthermore, we have

(8.92a) $\widetilde{F}^{*}(y, h) = (1 - \lambda) F^{*}(y, h) + \lambda F^{\circ}(y, h),$ and

(8.92b) $\widehat{F}^{\ddagger}(y, h) = (1 - \lambda) F^{*}(y, h) + \lambda F^{\circ}(y, h) + (1 - \lambda) \displaystyle\sum_{k=1}^{h} \int_{\underline{v}}^{y} \left[\phi(t, k) - \phi(t, k - 1) \right] dt,$

for all $y \in \mathscr{D}$ and $h = 1, 2, \ldots, H$. Making use of (8.52), we get

(8.93) $\displaystyle \sum_{k=1}^{h} \int_{\underline{v}}^{y} \left[\phi(t, k) - \phi(t, k - 1) \right] dt = \sum_{k=1}^{h} \left[\Psi_{y}(y \mid k) - \Psi_{y}(y \mid k - 1) \right] = \Psi_{y}(y \mid h),$

for all $y \in \mathscr{D}$ and $h = 1, 2, \ldots, H$. Invoking respectively A2 and A7, we deduce that

(8.94a) $\widehat{F}^{\ddagger}(y, h) \geqq \widetilde{F}^{*}(y, h), \ \forall \, y \in \mathscr{D}, \ \forall \, h = 1, 2, \ldots, H - 1;$ and

(8.94b) $\widehat{F}^{\ddagger}(y, H) = \widetilde{F}^{*}(y, H), \ \forall \, y \in \mathscr{D}.$

Therefore, we have identified a phantom situation $\tilde{\mathbf{s}} \in \mathscr{S}_q$ and a separating situation $\hat{\mathbf{s}}^{\ddagger} \in \mathscr{S}_n \times \mathscr{S}_q$ such that $\tilde{\mathbf{s}}^{*} \geq_{HP} \hat{\mathbf{s}}^{\ddagger} \geq_{CPG} \tilde{\mathbf{s}}^{\circ}$, where $\tilde{\mathbf{s}}^{*} := (\mathbf{s}^{*}; \tilde{\mathbf{s}})$ and $\tilde{\mathbf{s}}^{\circ} := (\mathbf{s}^{\circ}; \tilde{\mathbf{s}})$. $\square$

References

Allison, A. R. and Foster, J. E. (2004). Measuring health inequality using qualitative data. *Journal of Health Economics*, **23**, 503–524.

Anderson, G. (1996). Nonparametric tests of stochastic dominance in income distributions. *Econometrica*, **64**, 1183–1193.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, **2**, 244–263.

Atkinson, A. B. and Bourguignon, F. (1982). The comparison of multidimensioned distributions of economic status. *Review of Economic Studies*, **49**, 183–201.

Atkinson, A. B. and Bourguignon, F. (1987). Income distributions and differences in needs. In G. Feiwel, editor, *Arrow and the Foundations of the Theory of Economic Policy*, pages 350–370. MacMillan, New York.

Bazen, S. and Moyes, P. (2003). International comparisons of income distributions. *Research on Economic Inequality*, **9**, 85–111.

Beach, C. M. and Davidson, R. (1983). Distribution-free statistical inference with Lorenz curves and income shares. *Review of Economic Studies*, **50**, 723–735.

Berge, C. (1963). *Topological Spaces, Including a Treatment of Multi-Valued Functions, Vector Spaces and Convexity*. Oliver and Boyd, Edimburgh.

Bishop, J. A. and Formby, J. P. (1999). Tests of significance for Lorenz partial orders. In J. Silber, editor, *Handbook of Income Inequality Measurement*, pages 315–336. Kluwer Academic Publishers, Boston/Dordrecht/London.

Bishop, J. A., Chakraborti, S., and Thistle, P. D. (1989). Asymptotically distribution-free statistical inference for generalized Lorenz curves. *Review of Economics and Statistics*, **71**, 725–727.

Blackorby, C., Bossert, W., and Donaldson, D. (2005). *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge University Press, New York.

Bourguignon, F. (1989). Family size and social utility: Income distribution dominance criteria. *Journal of Econometrics*, **42**, 67–80.

Chateauneuf, A. and Moyes, P. (2006). Measuring inequality without the Pigou-Dalton condition. In M. McGillivray, editor, *Inequality, Poverty and Well-Being*, pages 22–65. Palgrave MacMillan, Basingtoke/New York.

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

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

Davidson, R. and Duclos, J.-Y. (2000). Statistical inference for stochastic dominance and for

the measurement of poverty and inequality. *Econometrica*, **68**, 1435–1464.

Decancq, K. (2011). Elementary multivariate rearrangements and stochastic dominance on a Fréchet class. *Journal of Economic Theory*. Forthcoming.

Ebert, U. (1995). Income inequality and differences in household size. *Mathematical Social Sciences*, **30**, 37–53.

Ebert, U. (2000). Sequential generalized Lorenz dominance and transfer principles. *Bulletin of Economic Research*, **52**, 113–123.

Epstein, L. G. and Tanny, S. (1980). Increasing generalized correlation: A definition and some economic consequences. *Canadian Journal of Economics*, **13**, 16–34.

Fields, G. S. and Fei, J. C. (1978). On inequality comparisons. *Econometrica*, **46**, 305–316.

Fishburn, P. C. and Vickson, R. G. (1978). Theoretical foundations of stochastic dominance. In G. Whitmore and M. Findlay, editors, *Stochastic Dominance*, pages 39–113. Lexington Books, Lexington, MA.

Fleurbaey, M., Hagneré, C., and Trannoy, A. (2003). Welfare comparisons with bounded equivalence scales. *Journal of Economic Theory*, **110**, 309–336.

Foster, J. E. (1985). Inequality measurement. In H. P. Young, editor, *Fair Allocation*, pages 38–61. American Mathematical Society, Providence, RI.

Gravel, N. and Moyes, P. (2011). Utilitarianism or welfarism: Does it make a difference? *Social Choice and Welfare*. Forthcoming.

Gravel, N., Moyes, P., and Tarroux, B. (2009). Robust international comparisons of distributions of disposable income and regional public goods. *Economica*, **76**, 432–461.

Griffin, J. (1986). *Well-Being, Its Meaning, Measurement and Moral Importance*. Clarendon Press, Oxford.

Hadar, J. and Russell, W. (1974). Stochastic dominance in choice under uncertainty. In M. Balch, D. McFadden, and S. Wu, editors, *Essays on Economics Behavior under Uncertainty*, pages 135–150. North-Holland, Amsterdam.

Hammond, P. J. (1976). Equity in two person situations: Some consequences. *Econometrica*, **47**, 1127–1135.

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1934). *Inequalities*. Cambridge University Press, Cambridge.

Jenkins, S. P. and Lambert, P. J. (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 J. Margolis and H. Guitton, editors, *Public Economics*, pages 145–200. MacMillan, London.

Kolm, S.-C. (1977). Multidimensional egalitarianisms. *Quarterly Journal of Economics*, **91**, 1–13.

Koshevoy, G. (1995). Multivariate Lorenz majorization. *Social Choice and Welfare*, **12**, 93–102.

Levy, H. (1998). *Stochastic Dominance. Investment Decision Making Under Uncertainty.* Kluwer Academic Publishers, Dordrecht.

Magdalou, B. and Moyes, P. (2009). Deprivation, welfare and inequality. *Social Choice and Welfare*, **32**, 253–273.

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

Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

Meyer, M. and Strulovici, B. (2010). The supermodular stochastic ordering. Unpublished Manuscript, Nuffield College, Oxford University.

Meyer, M. and Strulovici, B. (2011). Increasing interdependence of multivariate distributions. *Journal of Economic Theory*. Forthcoming.

Moyes, P. (1994). Inequality reducing and inequality preserving transformations of incomes: Symmetric and individualistic transformations. *Journal of Economic Theory*, **63**, 271–298.

Moyes, P. (2011a). Comparisons of heterogeneous distributions and dominance criteria. *Journal of Economic Theory*. Forthcoming.

Moyes, P. (2011b). Rearrangements and sequential rank order dominance: A result with economic applications. GREThA Discussion Paper 2011-35, Université Montesquieu Bordeaux IV.

Müller, A. and Scarsini, M. (2011). Fear of loss, inframodularity, and transfers. *Journal of Economic Theory*. Forthcoming.

Muller, C. and Trannoy, A. (2011). A dominance approach to the appraisal of the distribution of well being across countries. *Journal of Public Economics*, **95**, 239–246.

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

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

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

Tchen, A. H. (1980). Inequalities for distributions with given marginals. *The Annals of Probability*, **8**, 814–827.

Cahiers du GREThA
Working papers of GREThA

GREThA UMR CNRS 5113

Université Montesquieu Bordeaux IV
Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

http://gretha.u-bordeaux4.fr/

Cahiers du GREThA (derniers numéros – last issues)