Inequality measurement with an ordinal and continuous variable — **Source link** ⬈

Nicolas Gravel, Brice Magdalou, Patrick Moyes

Related papers:

- Hammond Equity: a Generalization

- The distribution problem and Rawlsian reasoning

- A complete and strongly anonymous leximin relation on infinite streams

- Q-anonymous social welfare relations on infinite utility streams *

- Measuring health inequality using qualitative data.

Inequality measurement with an ordinal and continuous variable

Nicolas Gravel, Brice Magdalou, Patrick Moyes

Mesure des Inégalités avec une Variable Ordinale et Continue

Résumé

Quel serait l'analogue du quasi-ordre de Lorenz lorsque la variable d'intérêt est continue et ordinale? Nous soutenons qu'il est possible d'obtenir un tel critère dès lors qu'on accepte de substituer au transfert de Pigou-Dalton traditionnellement utilisé dans la littérature ce que nous appelons transfert progressif de Hammond. Selon ce critère, une distribution d'utilités est considérée comme moins inégale qu'une autre si elle est jugée meilleure par les extensions lexicographiques du maximin et du minimax, respectivement dénommées leximin et antileximax. Si l'on impose en outre que l'augmentation de l'utilité d'un individu est socialement désirable, alors le leximin s'impose, tandis que l'exigence selon laquelle le bien-être social augmente suite à une diminution d'utilité individuelle conduit au critère de l'antileximax. Incidemment, le papier offre une caractérisation alternative et particulièrement simple du principe du leximin largement utilisé dans la littérature concernant le choix social et le bien-être.

Inequality Measurement with an Ordinal and Continuous Variable

Abstract

What would be the analogue of the Lorenz quasi-ordering when the variable of interest is continuous and of a purely ordinal nature? We argue that it is possible to derive such a criterion by substituting for the Pigou-Dalton transfer used in the standard inequality literature what we refer to as a Hammond progressive transfer. According to this criterion, one distribution of utilities is considered to be less unequal than another if it is judged better by both the lexicographic extensions of the maximin and the minimax, henceforth referred to as the leximin and the antileximax, respectively. If one imposes in addition that an increase in someone's utility makes the society better off, then one is left with the leximin, while the requirement that society welfare increases as the result of a decrease of one person's utility results in the antileximax criterion. Incidentally, the paper provides an alternative and simple characterisation of the leximin principle widely used in the social choice and welfare literature.

1. Introduction[1]

While inequality is seen as a major concern in most societies and has given rise to a large body of literature, less attention has been paid to the implication of the measurability nature of the variable of interest for its appraisal. So far, the literature has mostly focused on the design of indices and criteria for measuring inequality when the attribute distributed among the population's members is (at least) of a cardinal nature. The measurement of income and wealth inequality is a typical example where one even goes on assuming that the variables are measurable on a ratio scale. There are however many other variables that contribute to a person's well-being whose distribution is deemed to be important from a normative point of view, but that, at the same time, cannot be given fully agreed cardinal values. This may be due to a variety of reasons ranging from the fact that, *either* some attributes are by essence not cardinally measurable, *or* there is too much imprecision in the measurement to accord significance to the cardinal values provided, *or* no consensus can be reached about the appropriate cardinal values taken by the attribute.

In this paper, we are interested in the measurement of inequality – and more generally of social welfare – when the information available is purely ordinal. The only restriction – which actually has important consequences – is that the variable of interest is continuous. For convenience, one can think of utility but the analysis carries over for other items like cognitive abilities as measured, for instance, by the PISA scores. The case of categorical data – individuals are allocated into a finite number of ordered categories – raises additional problems and is the subject of a companion paper (see Gravel, Magdalou, and Moyes (2018)). Is it legitimate, as some authors do, to import the conventional tools designed for assessing income inequality to this informationally less demanding domain? If it is not, then when can it be reasonably claimed that one distribution is less unequal than another?

In the standard (income) inequality literature, a large consensus prevails for assimilating inequality reduction with the operation consisting of transferring some income from a richer individual to a poorer one in such a way that the beneficiary of the transfer is not made richer than the donor. An important restriction inherent in the definition of a so-called progressive transfer is that the amount taken from the richer individual must equal the amount received by the poorer individual. While it modifies the magnitude of the income differences, a change of scale under cardinal measurability preserves the (weak) inequalities between such differences. This ensures, among other things, that the equality between the amount taken from the richer and that given to the poorer is not affected by a change of scale. This equality does not survive to a change of scale in an ordinal setting even though the relative positions of the individuals are preserved. There is however something that is immune to a change of scale in

both the ordinal and the cardinal frameworks: it is the fact that the better-off and the worse-off individuals taking part in the transfer become closer. This captures the most intuitive notion of inequality reduction one could have in mind and it is precisely this idea that was retained by Hammond (1976) in the formulation of his equity principle in a social choice framework (see also Hammond (1979)).[2]

Hammond's equity condition requires that, if a person's utility increases while at the same time the utility of a better-off person decreases and if in addition the positions on the utility scale of all individuals are preserved – what we refer to as a Hammond progressive transfer – then the society's welfare must improve. We are basically interested in comparing distributions of utilities with a particular focus on inequality, and to this aim we assume a social preference relation defined on the set of the distributions of utilities. Following Hammond's suggestion, we impose that the social preference relation has the property that, if one distribution is obtained from another by means of a Hammond progressive transfer, then it is ranked above the latter. Because there is no particular reason to prefer one social preference relation to another, we follow the dominance approach consisting in requiring unanimity over all those social preference relations obeying the equity condition. This leads us quite naturally to search for a practical means for testing whether one distribution is preferred to another by all the social preference relations that are compatible with the equity condition and – if we further believe that the identities of the utilities' receivers play no role – with the principle of anonymity.

A widely applied criterion for comparing distributions of utilities is the lexicographic extension of the maximin principle which we refer to as the leximin in what follows. The leximin ranks one distribution of utilities above another, if the worst-off person in the first distribution gets a higher utility than that of the worst-off person in the second distribution, or, in the case they enjoy the same utility in both distributions, if the second worst-off person in the first distribution gets a higher utility than what the second worst-off person is given in the second distribution, and so on. The lexicographic extension of the minimax – henceforth the antileximax – is defined analogously but starting the other way round and requiring that the utility of the best-off person, the second best-off person, and so on, is lower in the preferred distribution than in the other. It is remarkable that the simultaneous application of the leximin and the antileximax proves to be the appropriate procedure to check unanimity over all anonymous and equity regarding social preference relations. Adding a preference for more efficiently distributed utilities in addition to anonymity and equity precipitates the leximin, while the judgement that a decrease in someone's utility results in a better situation leads to the antileximax.

As far as the organisation of the paper is concerned, we proceed as follows. We present in Section 2 our general framework. We introduce in Section 3 our axioms and the different criteria for comparing situations. Section 4 contains the main results and their proofs. We discuss in Section 5 the relationships between our approach and other results in the inequality literature, while we provide in Section 6 a social choice (re)interpretation of our character-

---

[2] While this operation certainly reduces inequality in a two-person society, things are less obvious when the population consists of more than two individuals. There are indeed good reasons to consider that the fact of bringing two individuals closer in terms of their incomes has the effect of moving them farther apart from the other individuals in the society: for more on this, see, e.g., Magdalou and Moyes (2009).

isations of the leximin and antileximax criteria. Finally, Section 7 concludes the paper by summarising the results and suggesting avenues for further research.

2. The Framework

We are interested in the comparison of distributions of a continuous and ordinal attribute among n individuals ($n \geqslant 2$). A distribution for a population $N := \{1, 2, \ldots, n\}$ is a vector $\mathbf{u} := (u_1, \ldots, u_n)$, where $u_i \in \mathbb{R}$ may be viewed as the *utility* of individual i, but other interpretations are also possible. We henceforth refer to the vector $\mathbf{u}$ as a *profile* and we indicate by $\mathscr{U} := \mathbb{R}^n$ the set of profiles for the population N. In order to compare profiles, we have recourse to a *social preference relation* R on $\mathscr{U}$. We assume that the relation R is *reflexive* and *transitive*, i.e., (i) for all $\mathbf{u} \in \mathscr{U}$, one has $\mathbf{u}\,\mathrm{R}\,\mathbf{u}$ and (ii) for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathscr{U}$, $\mathbf{u}\,\mathrm{R}\,\mathbf{v}$ and $\mathbf{v}\,\mathrm{R}\,\mathbf{w}$ implies that $\mathbf{u}\,\mathrm{R}\,\mathbf{w}$, respectively. On the other hand, we do not impose that R is a complete relation on the set $\mathscr{U}$: the relation R can be an ordering or a quasi-ordering.[3] We indicate respectively by I and P the symmetric and asymmetric components of R defined in the usual way, and we denote by $\mathscr{R}$ the set of social preference relations. An important requirement throughout is that the ranking of the profiles under comparison provided by R is invariant to an increasing transformation of the individuals' utilities. Formally, this amounts to imposing:

ORDINAL SCALE INVARIANCE (OSI). For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$ and all increasing function $\varphi : \mathbb{R} \to \mathbb{R}$, we have $\mathbf{u}\,\mathrm{R}\,\mathbf{v}$ if and only if $\varphi(\mathbf{u})\,\mathrm{R}\,\varphi(\mathbf{v})$, where $\varphi(\mathbf{u}) := (\varphi(u_1), \ldots, \varphi(u_n))$ and $\varphi(\mathbf{v}) := (\varphi(v_1), \ldots, \varphi(v_n))$.

This restriction is reminiscent of the condition of *ordinal level comparability* (OLC) in the social choice literature (see, e.g., Sen (1977)). Finally, given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we write: (i) $\mathbf{u} \geq \mathbf{v}$ whenever $u_h \geqslant v_h$, for all $h \in N$; (ii) $\mathbf{u} > \mathbf{v}$ whenever $\mathbf{u} \geq \mathbf{v}$ and $\mathbf{u} \neq \mathbf{v}$.

3. Axioms and Definitions

Given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we say that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a *permutation* – or, shortly, that $\mathbf{u}$ is a permutation of $\mathbf{v}$ – if there exists a permutation matrix $Q := [q_{ij}]$ such that $\mathbf{u} = \mathbf{v}\,Q$. For later use, we indicate by $\tilde{\mathbf{u}} := (\tilde{u}_1, \ldots, \tilde{u}_n)$ the non-decreasing rearrangement of profile $\mathbf{u} := (u_1, \ldots, u_n)$ defined by $\tilde{u}_1 \leqslant \tilde{u}_2 \leqslant \cdots \leqslant \tilde{u}_n$. We note that, if $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a permutation, then $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$. Our first condition requires that the ranking of the profiles is not affected by a permutation of the individuals' utilities, which formally amounts to imposing:

ANONYMITY (A). For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have $\mathbf{u}\,\mathrm{I}\,\mathbf{v}$ whenever $\mathbf{u}$ is a permutation of $\mathbf{v}$.

It is typically assumed in the economic inequality literature that a transfer of a fixed amount of income from a richer individual to a poorer one that leaves their respective positions unchanged – a so-called progressive transfer – reduces inequality. In an ordinal framework, where the individuals' utilities are defined up to the same increasing transformation, the notion of a progressive transfer makes no sense. Indeed, the equality between the utility gain and the

---

[3] In this respect, our analysis framework is similar to that used by Tungodden (2000) but our approach differs significantly from his.

utility loss of the individuals involved in the transfer is likely to be challenged by a change of the scale of measurement. The following transformation, introduced by Hammond (1976), captures the very essence of the notion of inequality reduction without imposing the restriction that the utility gain of the receiver equals the utility loss of the donor. More precisely, given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we say that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a *Hammond progressive transfer*, if there exist two individuals $i, j \in N$ such that:

$$(3.1) \qquad v_i < u_i \leqslant u_j < v_j \text{ and } u_h = v_h, \text{ for all } h \neq i, j.$$

Following Hammond, we want that, if one profile is more equal than another, then it is ranked above the latter, hence the following condition:

HAMMOND'S EQUITY (HE). For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have $\mathbf{u} \, \mathrm{P} \, \mathbf{v}$ whenever $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a Hammond progressive transfer.

While inequality reduction is viewed as a main concern in social evaluation, it is also common to supplement the pursuit of more equality with efficiency considerations. Given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we say that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of an *increment*, if there exists an individual $i \in N$ such that:

$$(3.2) \qquad u_i > v_i \text{ and } u_h = v_h, \text{ for all } h \neq i.$$

Equivalently, we say that $\mathbf{v}$ is obtained from $\mathbf{u}$ by means of a *decrement*. Our next condition is standard and requires that a more efficient profile is always ranked above a less efficient one:

STRONG PARETO (SP). For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have $\mathbf{u} \, \mathrm{P} \, \mathbf{v}$ whenever $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of an increment.

The above condition implicitly assumes that utility is a desirable item: more utility is always preferred to less by any individual and also by the society. Depending on which interpretation of utility we have in mind, it may turn out that less utility might be preferable to more. So is the case when a waste has to be distributed among the population, for it is reasonable to assume that everybody would prefer to have less of it than more of it. In such cases, one is certainly willing to impose the following condition:

STRONG ANTIPARETO (SAP). For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have $\mathbf{u} \, \mathrm{P} \, \mathbf{v}$ whenever $\mathbf{u}$ is obtained $\mathbf{v}$ by means of a decrement.

For later use, it is convenient to introduce the following sets of social preference relations defined on the set of profiles $\mathscr{U}$:

$$(3.3a) \qquad \mathscr{R}_+^{\circ} := \{ \, \mathrm{R} \in \mathscr{R} \mid \text{conditions A and SP hold} \, \};$$

$$(3.3b) \qquad \mathscr{R}_-^{\circ} := \{ \, \mathrm{R} \in \mathscr{R} \mid \text{conditions A and SAP hold} \, \};$$

$$(3.3c) \qquad \mathscr{R}^* := \{ \, \mathrm{R} \in \mathscr{R} \mid \text{conditions A and HE hold} \, \};$$

$$(3.3d) \qquad \mathscr{R}_+^* := \{ \, \mathrm{R} \in \mathscr{R} \mid \text{conditions A, HE and SP hold} \, \}; \text{ and}$$

$$(3.3e) \qquad \mathscr{R}_-^* := \{ \, \mathrm{R} \in \mathscr{R} \mid \text{conditions A, HE and SAP hold} \, \}.$$

The following equivalence relation will be needed when we will define the different principles examined in the paper:

$$(3.4) \qquad \forall \, \mathbf{u}, \mathbf{v} \in \mathscr{U}; \mathbf{u} \, \mathrm{I}^* \, \mathbf{v} \Longleftrightarrow \tilde{u}_h = \tilde{v}_h, \forall \, h \in N.$$

The first principle we consider is due to Suppes (1966) and, although it does not incorporate any concern for equality, we will occasionally refer to it in subsequent discussion and proofs.

SUPPES QUASI-ORDERING. We say that $\mathbf{u}$ weakly Suppes dominates $\mathbf{v}$, which we write $\mathbf{u}\,R_{\text{SU}}\,\mathbf{v}$, if and only if, *either* $\mathbf{u}\,I_{\text{SU}}\,\mathbf{v}$, where $I_{\text{SU}} = I^*$, *or* $\mathbf{u}\,P_{\text{SU}}\,\mathbf{v}$ which intends to mean that:

$$(3.5) \qquad \tilde{u}_h \geqslant \tilde{v}_h,\ \forall\, h \in N \text{ and } \exists\, i \in N \mid \tilde{u}_i > \tilde{v}_i\,.$$

The second principle is the lexicographic extension of the maximin – known as the leximin ordering – introduced by Sen (1977) that gives priority to the worst-off persons in the society.

LEXIMIN. We say that $\mathbf{u}$ weakly leximin dominates $\mathbf{v}$, which we write $\mathbf{u}\,R_{\text{LM}}\,\mathbf{v}$, if and only if, *either* $\mathbf{u}\,I_{\text{LM}}\,\mathbf{v}$, where $I_{\text{LM}} = I^*$, *or* $\mathbf{u}\,P_{\text{LM}}\,\mathbf{v}$ which intends to mean that:

$$(3.6) \qquad \exists\, i \in N \mid \tilde{u}_h = \tilde{v}_h,\ \forall\, h \in \{1, 2, \ldots, i-1\} \text{ and } \tilde{u}_i > \tilde{v}_i\,.$$

According to the leximin, profile $\mathbf{u}$ is considered to be better than profile $\mathbf{v}$ if the person who is the most disadvantaged in profile $\mathbf{u}$ gets a higher utility than the most disadvantaged person in profile $\mathbf{v}$ or, if they both enjoy the same utility, then the second worst-off person in $\mathbf{u}$ gets a higher utility than her counterpart in $\mathbf{v}$, and so on. The third principle mirrors the preceding one by focusing on the situations of the best-off persons.

ANTILEXIMAX. We say that $\mathbf{u}$ weakly antileximax dominates $\mathbf{v}$, which we write $\mathbf{u}\,R_{\text{ALX}}\,\mathbf{v}$, if and only if, *either* $\mathbf{u}\,I_{\text{ALX}}\,\mathbf{v}$, where $I_{\text{ALX}} = I^*$, *or* $\mathbf{u}\,P_{\text{ALX}}\,\mathbf{v}$ which intends to mean that:

$$(3.7) \qquad \exists\, j \in N \mid \tilde{u}_j < \tilde{v}_j \text{ and } \tilde{u}_h = \tilde{v}_h,\ \forall\, h \in \{j+1, j+2, \ldots, n\}\,.$$

The antileximax ranks profile $\mathbf{u}$ above profile $\mathbf{v}$ as soon as the best-off person in profile $\mathbf{u}$ gets a lower utility than the best-off person in profile $\mathbf{v}$ or, if they have the same utility, whenever the second best-off person in $\mathbf{u}$ gets a lower utility than her counterpart in $\mathbf{v}$, and so on. The leximin and the antileximax pay attention exclusively to the worst-off and the best-off persons in the society, respectively. The next and last principle – which appears to be new to the best of our knowledge – can be seen as a compromise between the views expressed by the leximin and the antileximax.

LEXIMIN-ANTILEXIMAX. We say that $\mathbf{u}$ weakly leximin-antileximax dominates $\mathbf{v}$, which we write $\mathbf{u}\,R_{\text{LMX}}\,\mathbf{v}$, if and only if, *either* $\mathbf{u}\,I_{\text{LMX}}\,\mathbf{v}$, where $I_{\text{LMX}} = I^*$, *or* $\mathbf{u}\,P_{\text{LMX}}\,\mathbf{v}$ which intends to mean that:

$$(3.8a) \qquad \exists\, i, j \in N\, (i < j) \mid \tilde{u}_i > \tilde{v}_i;\ \tilde{u}_j < \tilde{v}_j;\ \text{and}$$

$$(3.8b) \qquad \tilde{u}_h = \tilde{v}_h,\ \forall\, h \in \{1, 2, \ldots, i-1\} \cup \{j+1, j+2, \ldots, n\}.$$

The leximin-antileximax expresses a concern for less unequally distributed utilities in the sense that the utilities are more concentrated in the preferred profile if we leave aside those individuals at both ends of the distributions who are not affected by the choice between the two profiles. It must be emphasised at this stage that the leximin-antileximax pays no attention to the utilities of those individuals who are ranked between i and j. A profile can be ranked above another by the leximin-antileximax while the utilities of these individuals are more unequal in the preferred profile than in the other, something that may be seen a limitation of this criterion. Consider for instance the profiles $\mathbf{u} = (1, 3, 4, 6, 7, 9)$ and $\mathbf{v} = (1, 2, 5, 5, 8, 9)$:

clearly, $\mathbf{u} \, \mathrm{P_{LMX}} \, \mathbf{v}$, while at the same time the utilities of the individuals with ranks 3 and 4 are more equal in $\mathbf{v}$ than in $\mathbf{u}$. Implicit in the definition of the leximin-antileximax is the idea that the reduction of the inequalities between individuals ranked i and j overcome all the possible – and whatever their number – inequalities between the individuals whose ranks lie between i and j. As we will see later on, it is an interesting – and somewhat surprising – property of the leximin-antileximax that the preferred profile can always been derived from the less preferred one by a succession of inequality reducing operations of the kind described above. It follows from the definitions above that these four criteria are nested in the way described below.

REMARK 3.1. For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have: (i) $\mathbf{u} \, \mathrm{R_{SU}} \, \mathbf{v}$ implies $\mathbf{u} \, \mathrm{R_{LM}} \, \mathbf{v}$; (ii) $\mathbf{v} \, \mathrm{R_{SU}} \, \mathbf{u}$ implies $\mathbf{u} \, \mathrm{R_{ALX}} \, \mathbf{v}$; and (iii) $\mathbf{u} \, \mathrm{R_{LM}} \, \mathbf{v}$ and $\mathbf{u} \, \mathrm{R_{ALX}} \, \mathbf{v}$ if and only if $\mathbf{u} \, \mathrm{R_{LMX}} \, \mathbf{v}$.

It must be noted that, while the Suppes criterion and the leximin-antileximax provide partial rankings of the feasible profiles, the leximin and the antileximax totally order the elements of $\mathscr{U}$. The following example illustrates the preceding definitions and show to which extent the different principles depart from each other when ranking particular profiles.

EXAMPLE 3.1. Let $n = 4$ and consider the four following profiles and their corresponding non-decreasing rearrangements:

$$\mathbf{u}^{(1)} = (1, 2, 5, 8); \quad \tilde{\mathbf{u}}^{(1)} = (1, 2, 5, 8);$$
$$\mathbf{u}^{(2)} = (4, 6, 5, 7); \quad \tilde{\mathbf{u}}^{(2)} = (4, 5, 6, 7);$$
$$\mathbf{u}^{(3)} = (4, 2, 3, 7); \quad \tilde{\mathbf{u}}^{(3)} = (2, 3, 4, 7);$$
$$\mathbf{u}^{(4)} = (4, 2, 5, 7); \quad \tilde{\mathbf{u}}^{(4)} = (2, 4, 5, 7).$$

Application of the four above criteria gives the results indicated in the tables below where the symbol "$+1$" at the intersection of row i and column j means that profile $\mathbf{u}^{(i)}$ is preferred to profile $\mathbf{u}^{(j)}$, the symbol "-1" that profile $\mathbf{u}^{(j)}$ is preferred to profile $\mathbf{u}^{(i)}$, and the symbol "$\#$" means that they are not comparable.

SUPPES

	$\mathbf{u}^{(2)}$	$\mathbf{u}^{(3)}$	$\mathbf{u}^{(4)}$
$\mathbf{u}^{(1)}$	#	#	#
$\mathbf{u}^{(2)}$		+1	+1
$\mathbf{u}^{(3)}$			−1

LEXIMIN

	$\mathbf{u}^{(2)}$	$\mathbf{u}^{(3)}$	$\mathbf{u}^{(4)}$
$\mathbf{u}^{(1)}$	−1	−1	−1
$\mathbf{u}^{(2)}$		+1	+1
$\mathbf{u}^{(3)}$			−1

ANTILEXIMAX

	$\mathbf{u}^{(2)}$	$\mathbf{u}^{(3)}$	$\mathbf{u}^{(4)}$
$\mathbf{u}^{(1)}$	−1	−1	−1
$\mathbf{u}^{(2)}$		−1	−1
$\mathbf{u}^{(3)}$			+1

LEXIMIN-ANTILEXIMAX

	$\mathbf{u}^{(2)}$	$\mathbf{u}^{(3)}$	$\mathbf{u}^{(4)}$
$\mathbf{u}^{(1)}$	−1	−1	−1
$\mathbf{u}^{(2)}$		#	#
$\mathbf{u}^{(3)}$			#

Figure 3.1 provides an illustration of the different criteria above for a population comprising two individuals. The points U and V correspond respectively to the profile $\mathbf{u} = (2, 5)$ and to its permutation $\mathbf{v} = (5, 2)$. For each criteria, the area in light grey (including its boundary)

corresponds to the profiles that are considered as least as good as **u**, while the area in dark grey (including its boundary) represents the profiles that are considered as least as bad as **u**. The white areas represent all the profiles that are not comparable with **u** and, thanks to anonymity, also with **v**. In the particular case where $n = 2$, the sets of the best elements for the leximin and the antileximax are identical to the sets of the best elements for the maximin and the minimax, respectively, hence the impression that the criteria are complete. In the case of the leximin-antileximax, it must be noted that the set of profiles that are better than **u** is constituted by the light grey area with the exclusion of its boundary coloured in black and of the points U and V. This corresponds to the intersection of the sets of profiles that are ranked above **u** by the leximin and the antileximax. Adding the points U and V gives the set of profiles that are considered to be as least as good as **u** by the leximin-antileximax.

Figure 3.1: Dominating profiles $(n = 2)$
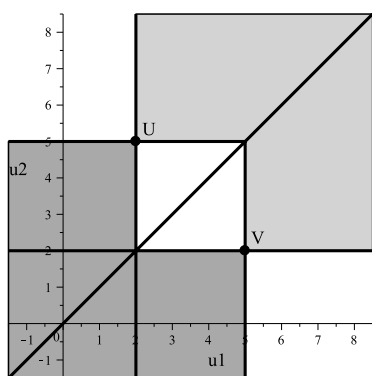
(a) Suppes
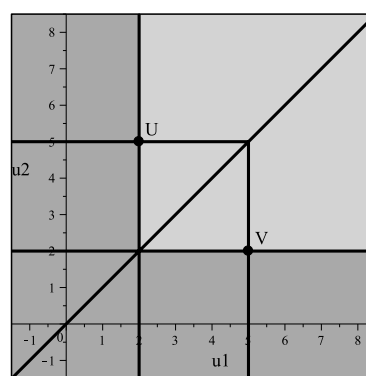
(b) Leximin
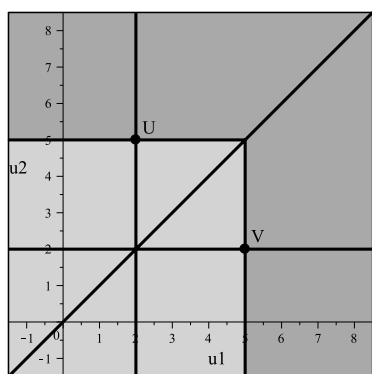
(c) Antileximax
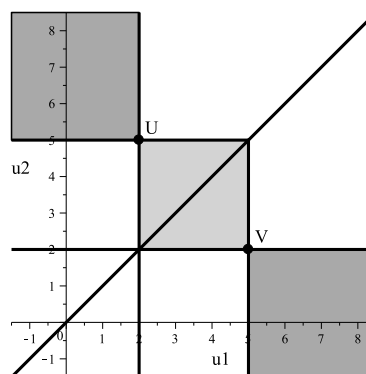
(d) Leximin-Antileximax

4. The Main Results

Our first result is but a mere restatement of a standard result in the inequality and risk literature and we mostly present it for the sake of completeness.

THEOREM 4.1. *For all* $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, *the following three statements are equivalent:*

(a) **u** *is obtained from* **v** *by means of a finite sequence of permutations and/or increments.*

(b) $\mathbf{u}\,\mathrm{R}\,\mathbf{v}$, *for all* $\mathrm{R} \in \mathscr{R}_+^\circ$.

(c) $\mathbf{u}\,\mathrm{R}_{\mathrm{SU}}\,\mathbf{v}$.

PROOF.

(a) $\Longrightarrow$ (c). If $\mathbf{u}$ is a permutation of $\mathbf{v}$, then $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$ and $\mathbf{u}\,\mathrm{I}_{\mathrm{SU}}\,\mathbf{v}$. If $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of increments, then it follows from Moyes (2013, Lemma 2.1) that $\tilde{\mathbf{u}} > \tilde{\mathbf{v}}$. To sum up, statement (a) implies that $\mathbf{u}\,\mathrm{R}_{\mathrm{SU}}\,\mathbf{v}$.

(c) $\Longrightarrow$ (a). Suppose that $\mathbf{u}\,\mathrm{R}_{\mathrm{SU}}\,\mathbf{v}$, in which case there are two possibilities: *either* $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$, *or* $\tilde{\mathbf{u}} > \tilde{\mathbf{v}}$. If $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$, then $\mathbf{u}$ is a permutation of $\mathbf{v}$. If $\tilde{\mathbf{u}} > \tilde{\mathbf{v}}$, then one obtains $\tilde{\mathbf{u}}$ starting from $\tilde{\mathbf{v}}$ by means of at most n increments. In the latter case, one will have recourse to permutations if it happens that $\tilde{\mathbf{u}} \neq \mathbf{u}$ and/or $\tilde{\mathbf{v}} \neq \mathbf{v}$.

(a) $\Longrightarrow$ (b). This follows from the definitions of conditions A and SP.

(b) $\Longrightarrow$ (c). The fact that (a) implies (c) guarantees that the Suppes criterion verifies condition A and SP, hence $\mathrm{R}_{\mathrm{SU}} \in \mathscr{R}_+^\circ$. Therefore, if statement (b) holds, then so does statement (c). $\square$

According to Theorem 4.1, the Suppes criterion represents the point of view of unanimity among all social preference relations that satisfy anonymity and strong Pareto. It further says that when this unanimous agreement holds, then the profile that is preferred can be derived from the other by making use of permutations and increments only.

The replacement of strong Pareto by Hammond equity gives the following result that may be considered the counterpart of the Hardy-Littlewood-Pólya-theorem in an ordinal framework.

THEOREM 4.2. *For all* $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, *the following three statements are equivalent:*

(a) $\mathbf{u}$ *is obtained from* $\mathbf{v}$ *by means of a finite sequence of permutations and/or Hammond progressive transfers.*

(b) $\mathbf{u}\,\mathrm{R}\,\mathbf{v}$, *for all* $\mathrm{R} \in \mathscr{R}^*$.

(c) $\mathbf{u}\,\mathrm{R}_{\mathrm{LMX}}\,\mathbf{v}$.

PROOF.

(a) $\Longrightarrow$ (c). If $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of permutations, then $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$, and therefore $\mathbf{u}\,\mathrm{I}_{\mathrm{LMX}}\,\mathbf{v}$. Suppose now that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a single Hammond progressive transfer so that (3.1) holds for some $i, j \in N$. Consider the indices g, h, k and ℓ defined by

(4.1a) $$g := \max\left\{t \in \{1, 2, \ldots, n\} \,\middle|\, \tilde{v}_t \leqslant v_i\right\},$$

(4.1b) $$h := \max\left\{t \in \{1, 2, \ldots, n\} \,\middle|\, \tilde{u}_t \leqslant u_i\right\},$$

(4.1c) $$k := \min\left\{t \in \{1, 2, \ldots, n\} \,\middle|\, \tilde{u}_t \geqslant u_j\right\}, \text{ and}$$

(4.1d) $$\ell := \min\left\{t \in \{1, 2, \ldots, n\} \,\middle|\, \tilde{v}_t \geqslant v_j\right\},$$

that always exist. It follows from (3.1) and the definitions above that

(4.2a) $$\tilde{v}_g = v_i; \ \ \tilde{u}_h = u_i; \ \ \tilde{u}_k = u_j; \ \ \tilde{v}_\ell = v_j;$$

(4.2b) $$\tilde{v}_g < \tilde{u}_h \leqslant \tilde{u}_k < \tilde{v}_\ell; \text{ and}$$

(4.2c) $$1 \leqslant g \leqslant h < k \leqslant \ell \leqslant n;$$

hence the following situation:

1	$\cdots$	$g-1$	g	$g+1$	$\cdots$	$h-1$	h	$h+1$	$\cdots$	$k-1$	k	$k+1$	$\cdots$	$\ell-1$	ℓ	$\ell+1$	$\cdots$	n

$$
\begin{array}{c}
u_i \\ \| \\
\tilde{\mathbf{u}}: \tilde{u}_1 \leqslant \cdots \leqslant \tilde{u}_{g-1} \leqslant \tilde{u}_g \leqslant \tilde{u}_{g+1} \leqslant \cdots \leqslant \tilde{u}_{h-1} \leqslant \tilde{u}_h < \tilde{u}_{h+1} \leqslant \cdots \leqslant \tilde{u}_{k-1} < \tilde{u}_k \leqslant \tilde{u}_{k+1} \leqslant \cdots \leqslant \tilde{u}_{\ell-1} \leqslant \tilde{u}_\ell < \tilde{u}_{\ell+1} \leqslant \cdots \leqslant \tilde{u}_n
\end{array}
$$

with u_j over column k, and

$$
\tilde{\mathbf{v}}: \tilde{v}_1 \leqslant \cdots \leqslant \tilde{v}_{g-1} \leqslant \tilde{v}_g < \tilde{v}_{g+1} \leqslant \cdots \leqslant \tilde{v}_{h-1} \leqslant \tilde{v}_h \leqslant \tilde{v}_{h+1} \leqslant \cdots \leqslant \tilde{v}_{k-1} \leqslant \tilde{v}_k \leqslant \tilde{v}_{k+1} \leqslant \cdots \leqslant \tilde{v}_{\ell-1} < \tilde{v}_\ell \leqslant \tilde{v}_{\ell+1} \leqslant \cdots \leqslant \tilde{v}_n
$$

with v_i under column g and v_j under column ℓ.

Inspection of the table makes clear that, for $\mathbf{u}\,\mathrm{P_{LMX}}\,\mathbf{v}$, one needs to have $\tilde{u}_g > \tilde{v}_g$ and $\tilde{u}_\ell < \tilde{v}_\ell$. At the risk of providing too many details, we distinguish four cases.

CASE 1: $g = h$. Then, we have $\tilde{u}_t = \tilde{v}_t$, for all $t = 1, 2, \ldots, g-1$, and $\tilde{u}_g = u_i > v_i = \tilde{v}_g$.

CASE 2: $g < h$. By definition of the indices g and h, we have

$$
\tilde{u}_g \leqslant \tilde{u}_{g+1} \leqslant \cdots \leqslant \tilde{u}_{h-3} \leqslant \tilde{u}_{h-2} \leqslant \tilde{u}_{h-1} \leqslant \tilde{u}_h = u_i
$$
$$
v_i = \tilde{v}_g < \tilde{v}_{g+1} \leqslant \tilde{v}_{g+2} \leqslant \cdots \leqslant \tilde{v}_{h-2} \leqslant \tilde{v}_{h-1} \leqslant \tilde{v}_h
$$

from which we deduce that $\tilde{u}_t = \tilde{v}_t$, for all $t = 1, 2, \ldots, g-1$, and $\tilde{u}_g > \tilde{v}_g$.

CASE 3: $k = \ell$. Then, we have $\tilde{u}_\ell = u_j < v_j = \tilde{v}_\ell$ and $\tilde{u}_t = \tilde{v}_t$, for all $t = \ell+1, \ldots, n-1, n$.

CASE 4: $k < \ell$. By definition of the indices k and ℓ, we have

$$
u_j = \tilde{u}_k \leqslant \tilde{u}_{k+1} \leqslant \tilde{u}_{k+2} \leqslant \cdots \leqslant \tilde{u}_{\ell-2} \leqslant \tilde{u}_{\ell-1} \leqslant \tilde{u}_\ell
$$
$$
\tilde{v}_k \leqslant \tilde{v}_{k+1} \leqslant \cdots \leqslant \tilde{v}_{\ell-3} \leqslant \tilde{v}_{\ell-2} \leqslant \tilde{v}_{\ell-1} < \tilde{v}_\ell = v_j
$$

from which we deduce that $\tilde{u}_t = \tilde{v}_t$, for all $t = n, n-1, \ldots, \ell+1$, and $\tilde{u}_\ell < \tilde{v}_\ell$.

Therefore, we have $\tilde{u}_t = \tilde{v}_t$, for all $t \in \{1, 2, \ldots, g-1\} \cup \{\ell+1, \ell+2, \ldots, n\}$, $\tilde{u}_g > \tilde{v}_g$ and $\tilde{u}_\ell < \tilde{v}_\ell$, hence $\mathbf{u}\,\mathrm{P_{LMX}}\,\mathbf{v}$. When more than one Hammond progressive transfer is needed in order to transform $\mathbf{v}$ into $\mathbf{u}$, the result follows from the transitivity of $\mathrm{P_{LMX}}$. We therefore conclude that, if statement (a) holds, then $\mathbf{u}\,\mathrm{R_{LMX}}\,\mathbf{v}$.

(c) $\Longrightarrow$ (a). Suppose that $\mathbf{u}\,\mathrm{R_{LMX}}\,\mathbf{v}$, in which case there are two possibilities: *either* $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$, *or* $\tilde{\mathbf{u}} \neq \tilde{\mathbf{v}}$. If $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$, then $\mathbf{u}$ is a permutation of $\mathbf{v}$ and we are home. Now, if $\tilde{\mathbf{u}} \neq \tilde{\mathbf{v}}$, then $\mathbf{u}\,\mathrm{P_{LMX}}\,\mathbf{v}$, hence there exist two indices i and j ($1 \leqslant i < j \leqslant n$) such that:

(4.3a) $\qquad \tilde{u}_t = \tilde{v}_t, \forall t \in \{1, 2, \ldots, i-1\} \cup \{j+1, j+2, \ldots, n\};$

(4.3b) $\qquad \tilde{u}_i > \tilde{v}_i$ and $\tilde{u}_j < \tilde{v}_j.$

Consider now the indices g and h defined by

(4.4a) $\qquad g := \max\left\{t \in \{i, i+1, \ldots, j-2, j-1\} \,\middle|\, \tilde{u}_t > \tilde{v}_t\right\}$ and

(4.4b) $\qquad h := \min\left\{t \in \{g+1, g+2, \ldots, j-1, j\} \,\middle|\, \tilde{u}_t < \tilde{v}_t\right\}.$

Two such indices necessarily exist and it follows from their definitions that $i \leqslant g < h \leqslant j$ and

(4.5) $\qquad \tilde{u}_t = \tilde{v}_t, \forall t \in \{g+1, \ldots, h-1\},$

provided that $g \neq h - 1$.

The table below summarises the available information derived from the above definitions that is at our disposal for arguing. To prove the implication, we first show that it is possible to find a profile $\tilde{\mathbf{w}} := (\tilde{w}_1, \ldots, \tilde{w}_n)$ with $\tilde{w}_1 \leqslant \tilde{w}_2 \leqslant \cdots \leqslant \tilde{w}_n$ such that (i) $\tilde{\mathbf{w}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a Hammond progressive transfer, (ii) $\mathbf{u}\, \mathrm{R}_{\mathrm{LMX}}\, \mathbf{w}$, and (iii) $\tilde{w}_t = \tilde{u}_t$, for at least one $t \in \{g, h\}$. We find convenient to distinguish four cases.

	1	$\cdots$	$i-1$	i	$i+1$	$\cdots$	$g-1$	g	$g+1$	$\cdots$	$h-1$	h	$h+1$	$\cdots$	$j-1$	j	$j+1$	$\cdots$	n
$\tilde{\mathbf{u}}$:	$\tilde{u}_1 \leqslant$	$\cdots$	$\leqslant \tilde{u}_{i-1} <$	$\tilde{u}_i \leqslant$	$\tilde{u}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{g-1} \leqslant$	$\tilde{u}_g \leqslant$	$\tilde{u}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{h-1} \leqslant$	$\tilde{u}_h \leqslant$	$\tilde{u}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{j-1} \leqslant$	$\tilde{u}_j <$	$\tilde{u}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_n$
	$\|$		$\|$	$\vee$	$\veebar$		$\veebar$	$\vee$	$\|$		$\|$	$\wedge$	$\barwedge$		$\barwedge$	$\wedge$	$\|$		$\|$
$\tilde{\mathbf{v}}$:	$\tilde{v}_1 \leqslant$	$\cdots$	$\leqslant \tilde{v}_{i-1} \leqslant$	$\tilde{v}_i \leqslant$	$\tilde{v}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{g-1} \leqslant$	$\tilde{v}_g <$	$\tilde{v}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{h-1} <$	$\tilde{v}_h \leqslant$	$\tilde{v}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{j-1} \leqslant$	$\tilde{v}_j \leqslant$	$\tilde{v}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_n$

CASE 1: $i < g < h < j$. Choosing $\tilde{\mathbf{w}}$ such that $\tilde{u}_g = \tilde{w}_g > \tilde{v}_g$, $\tilde{u}_h = \tilde{w}_h < \tilde{v}_h$, and $\tilde{w}_t = \tilde{v}_t$, for all $t \neq g, h$, we have the following situation:

	1	$\cdots$	$i-1$	i	$i+1$	$\cdots$	$g-1$	g	$g+1$	$\cdots$	$h-1$	h	$h+1$	$\cdots$	$j-1$	j	$j+1$	$\cdots$	n
$\tilde{\mathbf{u}}$:	$\tilde{u}_1 \leqslant$	$\cdots$	$\leqslant \tilde{u}_{i-1} <$	$\tilde{u}_i \leqslant$	$\tilde{u}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{g-1} \leqslant$	$\tilde{u}_g \leqslant$	$\tilde{u}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{h-1} \leqslant$	$\tilde{u}_h \leqslant$	$\tilde{u}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{j-1} \leqslant$	$\tilde{u}_j <$	$\tilde{u}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_n$
	$\|$		$\|$	$\vee$	$\veebar$		$\veebar$	$\|$	$\|$		$\|$	$\|$	$\barwedge$		$\barwedge$	$\wedge$	$\|$		$\|$
$\tilde{\mathbf{w}}$:	$\tilde{w}_1 \leqslant$	$\cdots$	$\leqslant \tilde{w}_{i-1} \leqslant$	$\tilde{w}_i \leqslant$	$\tilde{w}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{g-1} <$	$\tilde{w}_g \leqslant$	$\tilde{w}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{h-1} \leqslant$	$\tilde{w}_h <$	$\tilde{w}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{j-1} \leqslant$	$\tilde{w}_j \leqslant$	$\tilde{w}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_n$
	$\|$		$\|$	$\|$	$\|$		$\|$	$\vee$	$\|$		$\|$	$\wedge$	$\|$		$\|$	$\|$	$\|$		$\|$
$\tilde{\mathbf{v}}$:	$\tilde{v}_1 \leqslant$	$\cdots$	$\leqslant \tilde{v}_{i-1} \leqslant$	$\tilde{v}_i \leqslant$	$\tilde{v}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{g-1} \leqslant$	$\tilde{v}_g <$	$\tilde{v}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{h-1} <$	$\tilde{v}_h \leqslant$	$\tilde{v}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{j-1} \leqslant$	$\tilde{v}_j \leqslant$	$\tilde{v}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_n$

By construction $\tilde{\mathbf{w}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a Hammond progressive transfer, hence $\mathbf{w}\, \mathrm{I}_{\mathrm{LMX}}\, \tilde{\mathbf{w}}\, \mathrm{P}_{\mathrm{LMX}}\, \tilde{\mathbf{v}}\, \mathrm{I}_{\mathrm{LMX}}\, \mathbf{v}$. Inspection of the table above reveals that $\tilde{\mathbf{w}}$ is non-decreasingly arranged and that $\mathbf{u}\, \mathrm{I}_{\mathrm{LMX}}\, \tilde{\mathbf{u}}\, \mathrm{P}_{\mathrm{LMX}}\, \tilde{\mathbf{w}}\, \mathrm{I}_{\mathrm{LMX}}\, \mathbf{w}$. Furthermore $\tilde{u}_g = \tilde{w}_g$ and $\tilde{u}_h = \tilde{w}_h$.

CASE 2: $i = g < h < j$. Choosing $\tilde{\mathbf{w}}$ such that $\tilde{u}_g > \tilde{w}_g > \tilde{v}_g$, $\tilde{u}_h = \tilde{w}_h < \tilde{v}_h$, and $\tilde{w}_t = \tilde{v}_t$, for all $t \neq g, h$, we have the following situation:

	1	$\cdots$	$g-2$	$g-1$	$g=i$	$g+1$	$g+2$	$\cdots$	$h-2$	$h-1$	h	$h+1$	$\cdots$	$j-1$	j	$j+1$	$\cdots$	n
$\tilde{\mathbf{u}}$:	$\tilde{u}_1 \leqslant$	$\cdots$	$\leqslant \tilde{u}_{g-2} \leqslant$	$\tilde{u}_{g-1} \leqslant$	$\tilde{u}_g$	$\leqslant \tilde{u}_{g+1} \leqslant$	$\tilde{u}_{g+2} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{h-2} \leqslant$	$\tilde{u}_{h-1} \leqslant$	$\tilde{u}_h \leqslant$	$\tilde{u}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{j-1} \leqslant$	$\tilde{u}_j <$	$\tilde{u}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_n$
	$\|$		$\|$	$\|$	$\vee$	$\|$	$\|$		$\|$	$\|$	$\|$	$\barwedge$		$\barwedge$	$\wedge$	$\|$		$\|$
$\tilde{\mathbf{w}}$:	$\tilde{w}_1 \leqslant$	$\cdots$	$\leqslant \tilde{w}_{g-2} \leqslant$	$\tilde{w}_{g-1} <$	$\tilde{w}_g$	$\leqslant \tilde{w}_{g+1} \leqslant$	$\tilde{w}_{g+2} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{h-2} \leqslant$	$\tilde{w}_{h-1} \leqslant$	$\tilde{w}_h <$	$\tilde{w}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{j-1} \leqslant$	$\tilde{w}_j \leqslant$	$\tilde{w}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_n$
	$\|$		$\|$	$\|$	$\vee$	$\|$	$\|$		$\|$	$\|$	$\wedge$	$\|$		$\|$	$\|$	$\|$		$\|$
$\tilde{\mathbf{v}}$:	$\tilde{v}_1 \leqslant$	$\cdots$	$\leqslant \tilde{v}_{g-2} \leqslant$	$\tilde{v}_{g-1} \leqslant$	$\tilde{v}_g$	$< \tilde{v}_{g+1} \leqslant$	$\tilde{v}_{g+2} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{h-2} \leqslant$	$\tilde{v}_{h-1} <$	$\tilde{v}_h \leqslant$	$\tilde{v}_{h+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{j-1} \leqslant$	$\tilde{v}_j \leqslant$	$\tilde{v}_{j+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_n$

By construction $\tilde{\mathbf{w}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a Hammond progressive transfer, hence $\mathbf{w}\, \mathrm{I}_{\mathrm{LMX}}\, \tilde{\mathbf{w}}\, \mathrm{P}_{\mathrm{LMX}}\, \tilde{\mathbf{v}}\, \mathrm{I}_{\mathrm{LMX}}\, \mathbf{v}$. Inspection of the table above reveals that $\tilde{\mathbf{w}}$ is non-decreasingly arranged, that $\tilde{u}_h = \tilde{w}_h$, and that $\mathbf{u}\, \mathrm{I}_{\mathrm{LMX}}\, \tilde{\mathbf{u}}\, \mathrm{P}_{\mathrm{LMX}}\, \tilde{\mathbf{w}}\, \mathrm{I}_{\mathrm{LMX}}\, \mathbf{w}$.

CASE 3: $i < g < h = j$. Choosing $\tilde{\mathbf{w}}$ such that $\tilde{u}_g = \tilde{w}_g > \tilde{v}_g$, $\tilde{u}_h < \tilde{w}_h < \tilde{v}_h$, and $\tilde{w}_t = \tilde{v}_t$, for all $t \neq g, h$, we have the following situation:

	1	$\cdots$	$i-1$	i	$i+1$	$\cdots$	$g-1$	g	$g+1$	$\cdots$	$h-2$	$h-1$	$h=j$	$h+1$	$h+2$	$\cdots$	$n-1$	n
$\tilde{\mathbf{u}}$:	$\tilde{u}_1 \leqslant$	$\cdots$	$\leqslant \tilde{u}_{i-1} <$	$\tilde{u}_i \leqslant$	$\tilde{u}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{g-1} \leqslant$	$\tilde{u}_g \leqslant$	$\tilde{u}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{h-2} \leqslant$	$\tilde{u}_{h-1} \leqslant$	$\tilde{u}_h$	$\leqslant \tilde{u}_{h+1} \leqslant$	$\tilde{u}_{h+2} \leqslant$	$\cdots$	$\leqslant \tilde{u}_{n-1} \leqslant$	$\tilde{u}_n$
	$\|$		$\|$	$\vee$	$\veebar$		$\veebar$	$\|$	$\|$		$\|$	$\|$	$\wedge$	$\|$	$\|$		$\|$	$\|$
$\tilde{\mathbf{w}}$:	$\tilde{w}_1 \leqslant$	$\cdots$	$\leqslant \tilde{w}_{i-1} \leqslant$	$\tilde{w}_i \leqslant$	$\tilde{w}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{g-1} <$	$\tilde{w}_g \leqslant$	$\tilde{w}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{h-2} \leqslant$	$\tilde{w}_{h-1} \leqslant$	$\tilde{w}_h$	$< \tilde{w}_{h+1} \leqslant$	$\tilde{w}_{h+2} \leqslant$	$\cdots$	$\leqslant \tilde{w}_{n-1} \leqslant$	$\tilde{w}_n$
	$\|$		$\|$	$\|$	$\|$		$\|$	$\vee$	$\|$		$\|$	$\|$	$\wedge$	$\|$	$\|$		$\|$	$\|$
$\tilde{\mathbf{v}}$:	$\tilde{v}_1 \leqslant$	$\cdots$	$\leqslant \tilde{v}_{i-1} \leqslant$	$\tilde{v}_i \leqslant$	$\tilde{v}_{i+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{g-1} \leqslant$	$\tilde{v}_g <$	$\tilde{v}_{g+1} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{h-2} \leqslant$	$\tilde{v}_{h-1} <$	$\tilde{v}_h$	$\leqslant \tilde{v}_{h+1} \leqslant$	$\tilde{v}_{h+2} \leqslant$	$\cdots$	$\leqslant \tilde{v}_{n-1} \leqslant$	$\tilde{v}_n$

By construction $\tilde{\mathbf{w}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a Hammond progressive transfer, hence $\mathbf{w}\,\mathrm{I}_{\mathrm{LMX}}\,\tilde{\mathbf{w}}\,\mathrm{P}_{\mathrm{LMX}}\,\tilde{\mathbf{v}}\,\mathrm{I}_{\mathrm{LMX}}\,\mathbf{v}$. Inspection of the table above reveals that $\tilde{\mathbf{w}}$ is non-decreasingly arranged, that $\tilde{u}_g = \tilde{w}_g$, and that $\mathbf{u}\,\mathrm{I}_{\mathrm{LMX}}\,\tilde{\mathbf{u}}\,\mathrm{P}_{\mathrm{LMX}}\,\tilde{\mathbf{w}}\,\mathrm{I}_{\mathrm{LMX}}\,\mathbf{w}$.

CASE 4: $i = g < h = j$. It is identical to case 1 with the particularity that now we have the simple situation depicted below:

	1	2	$\cdots$	$g-2$	$g-1$	$g=i$	$g+1$	$g+2$	$\cdots$	$h-2$	$h-1$	$h=j$	$h+1$	$h+2$	$\cdots$	$n-1$	n
$\tilde{\mathbf{u}}$:	$\tilde{u}_1 \leqslant$	$\tilde{u}_2 \leqslant$	$\cdots \leqslant$	$\tilde{u}_{g-2} \leqslant$	$\tilde{u}_{g-1} <$	$\tilde{u}_g$	$\leqslant \tilde{u}_{g+1} \leqslant$	$\tilde{u}_{g+2} \leqslant$	$\cdots \leqslant$	$\tilde{u}_{h-2} \leqslant$	$\tilde{u}_{h-1} \leqslant$	$\tilde{u}_h$	$< \tilde{u}_{h+1} \leqslant$	$\tilde{u}_{h+2} \leqslant$	$\cdots \leqslant$	$\tilde{u}_{n-1} \leqslant$	$\tilde{u}_n$
	$\parallel$	$\parallel$		$\parallel$	$\parallel$	$\parallel$	$\parallel$	$\parallel$		$\parallel$	$\parallel$	$\parallel$	$\parallel$	$\parallel$		$\parallel$	$\parallel$
$\tilde{\mathbf{w}}$:	$\tilde{w}_1 \leqslant$	$\tilde{w}_2 \leqslant$	$\cdots \leqslant$	$\tilde{w}_{g-2} \leqslant$	$\tilde{w}_{g-1} <$	$\tilde{w}_g$	$\leqslant \tilde{w}_{g+1} \leqslant$	$\tilde{w}_{g+2} \leqslant$	$\cdots \leqslant$	$\tilde{w}_{h-2} \leqslant$	$\tilde{w}_{h-1} \leqslant$	$\tilde{w}_h$	$< \tilde{w}_{h+1} \leqslant$	$\tilde{w}_{h+2} \leqslant$	$\cdots \leqslant$	$\tilde{w}_{n-1} \leqslant$	$\tilde{w}_n$
	$\parallel$	$\parallel$		$\parallel$	$\parallel$	$\vee$	$\parallel$	$\parallel$		$\parallel$	$\parallel$	$\wedge$	$\parallel$	$\parallel$		$\parallel$	$\parallel$
$\tilde{\mathbf{v}}$:	$\tilde{v}_1 \leqslant$	$\tilde{v}_2 \leqslant$	$\cdots \leqslant$	$\tilde{v}_{g-2} \leqslant$	$\tilde{v}_{g-1} \leqslant$	$\tilde{v}_g$	$< \tilde{v}_{g+1} \leqslant$	$\tilde{v}_{g+2} \leqslant$	$\cdots \leqslant$	$\tilde{v}_{h-2} \leqslant$	$\tilde{v}_{h-1} <$	$\tilde{v}_h$	$\leqslant \tilde{v}_{h+1} \leqslant$	$\tilde{v}_{h+2} \leqslant$	$\cdots \leqslant$	$\tilde{v}_{n-1} \leqslant$	$\tilde{v}_n$

By construction $\tilde{\mathbf{w}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a Hammond progressive transfer, hence $\mathbf{w}\,\mathrm{I}_{\mathrm{LMX}}\,\tilde{\mathbf{w}}\,\mathrm{P}_{\mathrm{LMX}}\,\tilde{\mathbf{v}}\,\mathrm{I}_{\mathrm{LMX}}\,\mathbf{v}$. Inspection of the table above reveals that $\tilde{\mathbf{w}}$ is non-decreasingly arranged and $\mathbf{u}\,\mathrm{I}_{\mathrm{LMX}}\,\tilde{\mathbf{u}} = \tilde{\mathbf{w}}\,\mathrm{I}_{\mathrm{LMX}}\,\mathbf{w}$.

Let us first denote by $d(\mathbf{u},\mathbf{v}) := \#\{t \in \{1,2,\ldots,n\} \mid u_t \neq v_t\}$ the number of distinct components in $\mathbf{u}$ and $\mathbf{v}$. To sum up, we have shown that, if $\tilde{\mathbf{u}}\,\mathrm{P}_{\mathrm{LMX}}\,\tilde{\mathbf{v}}$, then it is possible to find a profile $\tilde{\mathbf{w}}$ such that (i) $\tilde{\mathbf{w}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a Hammond progressive transfer, (ii) $\tilde{\mathbf{u}}\,\mathrm{R}_{\mathrm{LMX}}\,\tilde{\mathbf{w}}$, and (iii) $\tilde{w}_t = \tilde{u}_t$, for at least one $t \in \{g,h\}$, which implies that $d(\tilde{\mathbf{u}},\tilde{\mathbf{w}}) \leqslant n-1$. Repeating the same argument as above, we obtain a sequence of profiles $\{\tilde{\mathbf{w}}^s\}$ such that at each step s one has: (i) $\tilde{\mathbf{w}}^s$ is obtained from $\tilde{\mathbf{w}}^{s-1}$ by means of a Hammond progressive transfer, (ii) $\tilde{\mathbf{u}}\,\mathrm{R}_{\mathrm{LMX}}\,\tilde{\mathbf{w}}^s$, and (iii) $\tilde{w}_t^s = \tilde{u}_t$, for at least one t. Letting $\tilde{\mathbf{w}}^1 \equiv \tilde{\mathbf{w}}$, it follows that

$$(4.6) \qquad 0 \leqslant d(\tilde{\mathbf{u}},\tilde{\mathbf{w}}^s) < d(\tilde{\mathbf{u}},\tilde{\mathbf{w}}^{s-1}) < \cdots < d(\tilde{\mathbf{u}},\tilde{\mathbf{w}}^2) < d(\tilde{\mathbf{u}},\tilde{\mathbf{w}}^1) < d(\tilde{\mathbf{u}},\tilde{\mathbf{v}}) \leqslant n,$$

for all s. The sequence $\{d(\tilde{\mathbf{u}},\tilde{\mathbf{w}}^s)\}$ is bounded – from above and below – and strictly decreasing. We therefore conclude that profile $\tilde{\mathbf{u}}$ can be obtained from profile $\tilde{\mathbf{v}}$ by making use of at most $n-1$ Hammond progressive transfers. Permutations will be used to complete the argument in the case where $\tilde{\mathbf{u}} \neq \mathbf{u}$ and/or $\tilde{\mathbf{v}} \neq \mathbf{v}$.

(a) $\Longrightarrow$ (b). This follows from the definitions of conditions A and HE.

(b) $\Longrightarrow$ (c). The fact that (a) implies (c) guarantees that the leximin-antileximax criterion verifies conditions A and HE, hence $\mathrm{R}_{\mathrm{LMX}} \in \mathscr{R}^*$. Therefore, if statement (b) holds, then so does statement (c). $\qquad \square$

According to Theorem 4.2, the only way to make sure that one profile is ranked above another by all the social preference relations that satisfy anonymity and Hammond equity is to subject the two profiles to the verdicts of the leximin and of the antileximax. More precisely, one profile will be preferred to another by unanimity among all these social preference relations if and only if both criteria agree. When this happens, Theorem 4.2 goes on saying that the dominating profile can always be derived from the dominated profile by making use of only permutations and Hammond progressive transfers.

So far we have focused on inequality and we have therefore prevented ourselves from taking into account any consideration for efficiency. Adding strong Pareto to anonymity and Hammond equity precipitates the leximin as the following result demonstrates.

THEOREM 4.3. *For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, the following three statements are equivalent:*

(a) **u** *is obtained from* **v** *by means of a finite sequence of permutations, Hammond progressive transfers and/or increments.*

(b) **u** R **v**, *for all* R $\in \mathscr{R}_+^*$.

(c) **u** R$_{\mathrm{LM}}$ **v**.

PROOF.

(a) $\implies$ (c). We know from Theorem 4.1 that, if **u** is obtained from **v** by means of permutations and/or increments, then **u** R$_{\mathrm{SU}}$ **v**, and it follows from Remark 3.1 that **u** R$_{\mathrm{LM}}$ **v**. Similarly, Theorem 4.2 tells us that **u** R$_{\mathrm{LMX}}$ **v** whenever **u** results from **v** through a finite sequence of Hammond progressive transfers, and Remark 3.1 again ensures that **u** R$_{\mathrm{LM}}$ **v**.

(c) $\implies$ (a). Suppose that **u** R$_{\mathrm{LM}}$ **v**, in which case there are two possibilities. If **u** I$_{\mathrm{LM}}$ **v**, then $\tilde{\mathbf{u}} = \tilde{\mathbf{v}}$ and **u** can be obtained from **v** by means of permutations only. Now, if **u** P$_{\mathrm{LM}}$ **v**, then there exists $i \in N$ such that $\tilde{u}_h = \tilde{v}_h$, for all $h \in \{1, 2, \ldots, i-1\}$, and $\tilde{u}_i > \tilde{v}_i$, and there are two cases to consider.

CASE 1: $\tilde{u}_h \geqslant \tilde{v}_h$, for all $h \in \{i+1, i+2, \ldots, n\}$. Then **u** P$_{\mathrm{SU}}$ **v** and it follows from Theorem 4.1 that **u** can be obtained from **v** by means of a finite sequence of permutations and/or increments.

CASE 2: $\tilde{u}_h < \tilde{v}_h$, for some $h \in \{i+1, \ldots, n\}$. Letting $g := \max \{h \in \{i+1, \ldots, n\} \mid \tilde{u}_h < \tilde{v}_h\}$, we have $\tilde{u}_h = \tilde{v}_h$, for all $h \in \{1, 2, \ldots, i-1\} \cup \{j+1, j+2, \ldots, n\}$, $\tilde{u}_i > \tilde{v}_i$ and $\tilde{u}_j < \tilde{v}_j$. Thus **u** P$_{\mathrm{LMX}}$ **v**, and it follows from Theorem 4.2 that **u** can be obtained from **v** by means of a finite sequence of permutations and/or Hammond progressive transfers.

(a) $\implies$ (b). This follows from the definitions of conditions A, HE and SP.

(b) $\implies$ (c). The fact that (a) implies (c) guarantees that the leximin verifies conditions A, HE and SP, hence R$_{\mathrm{LM}} \in \mathscr{R}_+^*$. We therefore conclude that, if statement (b) holds, then so does statement (c). □

The search for unanimity of point of views among the social preference relations that satisfy anonymity, Hammond equity and strong Pareto leads to the leximin. In addition, Theorem 4.3 confirms that only permutations, Hammond progressive transfers and increments are needed to convert the dominated profile into the dominating one.

Not surprisingly, substitution of strong antiPareto for strong Pareto, in conjunction with anonymity and Hammond equity, gives the next result which emphasises the decisive role played by the antileximax in the search for unanimity over the corresponding class of social preference relations.

THEOREM 4.4. *For all* **u**, **v** $\in \mathscr{U}$, *the following three statements are equivalent:*

(a) **u** *is obtained from* **v** *by means of a finite sequence of permutations, Hammond progressive transfers and/or decrements.*

(b) **u** R **v**, *for all* R $\in \mathscr{R}_-^*$.

(c) **u** R$_{\mathrm{ALX}}$ **v**.

PROOF. It is a straightforward adaptation of the arguments used when proving Theorem 4.3 above and it is therefore omitted. □

According to Theorem 4.4, the antileximax is the right criterion to use in order to check that one profile is ranked above another by all the social preference relations that satisfy anonymity, Hammond equity and strong antiPareto. If the antileximax succeeds to decide between two profiles, then Theorem 4.4 guarantees that the dominating profile can always be derived from the dominated profile by making use of only permutations, Hammond progressive transfers and decrements.

While Theorems 4.1 to 4.4 possess the same structure, close attention reveals a notable difference between Theorems 4.1 and 4.2 and Theorems 4.3 and 4.4 that may look surprising at first sight. It originates in statement (c) of the different theorems that introduces the Suppes quasi-ordering, the leximin-antileximax, the leximin, and the antileximax. Because the leximin and the antileximax are orderings – they are reflexive, transitive and complete – it follows that the classes of social preferences considered in statement (b) of Theorems 4.3 and 4.4, respectively, contains only one element, contrary to what happens in the case of Theorems 4.1 and 4.2. We will elaborate on this particular point in the next section.

5. Relationship With Other Results in the Inequality Literature

Comparison with the Lorenz approach Theorems 4.1 to 4.4 may be considered the natural adaptations of well-known results in the inequality and welfare literature when the variable that is distributed is continuous and ordinal. The Suppes quasi-ordering is actually identical to the first stochastic dominance criterion – or, equivalently, the quantile dominance criterion – extensively used in the risk and inequality literature (see, e.g., Fishburn and Vickson (1978)). In this respect, Theorem 4.1 is but a restatement of well-known equivalences with the difference that we had no recourse to representations of the social preference ordering, be they of the expected or non-expected utility types (see, e.g., Levy (1998) for a survey of the literature). Actually, we even do not require the social preference to be an ordering and only impose that it is a reflexive and transitive relation.

The leximin-antileximax leaves aside efficiency considerations and focuses exclusively on the idea of inequality reduction: in an ordinal framework, it can therefore be seen as the counterpart of the Lorenz criterion, also known as *majorisation* in the mathematical literature (see Marshall and Olkin (1979)). Like the Lorenz criterion, the leximin-antileximax considers that bringing two individuals closer on the utility scale without reversing their respective positions reduces inequality. But, contrary to the Lorenz criterion, it does not impose the additional constraint that gains and losses balance, which makes no sense in an ordinal framework. Theorem 4.2 may therefore be considered the ordinal version of the celebrated Hardy-Littlewood-Pólya theorem that provides the normative foundation of the Lorenz criterion by uncovering its connections with progressive transfers and the unanimous agreement among all utilitarians endowed with a concave utility function. It clarifies the way Hammond progressive transfers, unanimity over the class of relations that obey the Hammond equity condition and the leximin-antileximax criterion are related.

While the Lorenz criterion is exclusively concerned with inequality, the generalised Lorenz criterion introduced by Shorrocks (1983) allows one to compare distributions from a welfare point of view. More precisely, the generalised Lorenz criterion – also referred to as *weak submajorisation* (see again Marshall and Olkin (1979)) – insists that increments and progressive

transfers give rise to a social welfare improvement. The leximin plays a similar role in an ordinal framework by substituting Hammond progressive transfers for standard progressive transfers. On the other hand, it is not uncommon that one has to compare situations where an undesirable item is distributed among a population of individuals. For instance, one may think of a decision-maker who has to allocate a waste and who considers that (i) the smaller is the quantity that a person receives, the better off she is, and (ii) the more equally the waste is shared among the population – everybody contributes – the better off the society is. In such a case, *weak supermajorisation* seems to be an appropriate criterion for comparing the alternative distributions or, at least, to rule out those distributions that are deemed unacceptable with no ambiguity. Theorem 4.4 provides good reasons for appealing to the antileximax when such comparisons are made in an ordinal framework.

A common critique addressed to the Lorenz criterion – as well as to its various extensions – is that it fails to provide a complete ranking of the distributions to be compared. While a somewhat similar critique could be addressed to the leximin-antileximax, it must be emphasised that the leximin and the antileximax are exempt from this deficiency: they are able to rank *all* the profiles under consideration. This does not occur in the standard framework where the variable of interest is of a cardinal nature: like the Lorenz criterion, weak submajorisation and weak supermajorisation may not rank all distributions. Because the leximin-antileximax is by definition the intersection of two orderings – the leximin and the antileximax – it fails to be complete: for more on the relation between orderings and quasi-orderings, see Donaldson and Weymark (1998).[4] What is to some extent surprising is the ability of the leximin and the antileximax to totally order the profiles under comparison.[5]

Finally, while we argued that the leximin, antileximax and leximin-antileximax are the right criteria to use in an ordinal framework when one is concerned with inequality reduction, we do not see any objection to appeal to these criteria in richer informational settings. But it must be insisted on the fact that, by doing so, one implicitly adopts a somewhat extreme view of inequality reduction according to which bringing two individuals' utilities closer is always worth from a social point of view, whatever magnitude of the gains and losses in utility. In particular, the inequality judgement made by the Lorenz criterion is compatible with that captured by the leximin-antileximax: if one distribution is ranked above another by the Lorenz criterion, then it will be the same with the leximin-antileximax. In other words, in the standard income inequality setting, domination in terms of the leximin-antileximax is a necessary condition for Lorenz domination.[6] Similarly, the fact that one profile is ranked above another by the leximin is necessary for it to generalised Lorenz dominate the other profile.

---

[4] Actually, the same holds true for the Lorenz criterion that is the intersection of weak supermajorisation – or, equivalently, generalised Lorenz dominance – and weak submajorisation.

[5] The adjunction of strong Pareto to the anonymity and Hammond equity conditions suffices for rendering the social preference relation complete. This has to be contrasted to what happens in the standard framework: imposing that the social welfare function is monotone increasing in addition to being Schur-concave does not make the corresponding dominance criterion complete.

[6] This is readily inferred from the fact that, in a cardinal framework, a progressive transfer is but a particular case of a Hammond progressive transfer.

Inclusion quasi-ordering and unanimity From now on, we assume without loss of generality that the profiles under consideration are non-decreasingly arranged. A recurrent limitation of a progressive transfer – be it a Hammond transfer or a standard one – is the fact that, while it reduces inequality between the individuals involved, the transfer may generate an increase in the inequalities between each of these two individuals and the rest of the population.[7] Consider for instance the profiles $\mathbf{u}^{(5)} = (1, 2, 5, 8, 9)$ and $\mathbf{u}^{(6)} = (1, 4, 5, 7, 9)$, where $\mathbf{u}^{(6)}$ is obtained from $\mathbf{u}^{(5)}$ by means of a Hammond progressive transfer involving individuals 2 and 4. While $u_2^{(5)} < u_2^{(6)} < u_4^{(6)} < u_4^{(5)}$, we observe that $u_1^{(6)} = u_1^{(5)} < u_2^{(5)} < u_2^{(6)}$ and $u_4^{(6)} < u_4^{(5)} < u_5^{(5)} = u_5^{(6)}$.[8] Whereas individuals 2 and 4 came closer to each other in terms of utility, individuals of 1 and 2, as well as individuals of 4 and 5, moved away from each other. Thus, the reduction in inequality between individuals 2 and 4 attributable to the Hammond progressive transfer is achieved at the cost of an increase in inequality for the pairs of individuals $\{1, 2\}$ and $\{4, 5\}$.

The implicit concept of inequality mentioned above refers to the notion of *pairwise inclusion* that suggests a slightly more general way of appraising inequality in an ordinal framework (see Kolm (1999, Section 3.2)). Let us say that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a *pairwise inclusion* if there exist two individuals i and j such that

$$(5.1) \qquad [u_i, u_j] \subseteq [v_i, v_j] \text{ and } u_h = v_h, \text{ for all } h \neq i, j.$$

We find convenient to write $(u_i, u_j) \succsim (v_i, v_j)$ when $[u_i, u_j] \subseteq [v_i, v_j]$ and we denote respectively by $\succ$ and $\sim$ the asymmetric and symmetric components of $\succsim$. The inclusion is *semi-strict* if (5.1) holds and, *either* $u_i \neq v_i$, *or* $u_j \neq v_j$. If $u_i \neq v_i$ and $u_j \neq v_j$, then the inclusion is *strict* and it reduces to a Hammond progressive transfer. Table 5.1 indicates the corresponding

Table 5.1: Pairwise inclusion ranking in the case of profiles $\mathbf{u}^{(5)}$ and $\mathbf{u}^{(6)}$

	2	3	4	5
1	$(1, 4) \prec (1, 2)$	$(1, 5) \sim (1, 5)$	$(1, 7) \succ (1, 8)$	$(1, 9) \sim (1, 9)$
2		$(4, 5) \succ (2, 5)$	$(4, 7) \succ (2, 8)$	$(4, 9) \succ (2, 9)$
3			$(5, 7) \succ (5, 8)$	$(5, 9) \sim (5, 9)$
4				$(7, 9) \prec (8, 9)$

ranking of the couples $(u_i^{(5)}, u_j^{(5)})$ and $(u_i^{(6)}, u_j^{(6)})$ for each pair of individuals $\{i, j\}$, and confirms that a Hammond progressive transfer fails to reduce inequality over all pairs of individuals.

An obvious solution in order to avoid such a situation would be to declare that inequality decreases if and only if *all* pairwise inequalities are simultaneously reduced. The following quasi-ordering, that is due to Kolm (1997, Chapter III-C), precisely aims at capturing this basic idea. Given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we say that $\mathbf{u}$ *unanimously pairwise inclusion*

---

[7] Questionnaire studies have highlighted the fact that a large proportion of respondents do not subscribe to the view according to which a progressive transfer reduces inequality (see for instance Amiel and Cowell (1999)). In this respect, a notion like that of a uniform on the right progressive transfer proposed by Magdalou and Moyes (2009), which is closely related to the reduction of deprivation, is more likely to fit the views expressed by the interviewed public.

[8] It must be noted however that, for all the other pairs of individuals $\{i, j\}$ such that $i < j$, we have, *either* $u_i^{(5)} < u_i^{(6)} < u_j^{(6)} < u_j^{(5)}$, *or* $u_i^{(5)} = u_i^{(6)} < u_j^{(6)} = u_j^{(5)}$.

dominates $\mathbf{v}$, which we write $\mathbf{u} \geq_{\mathrm{UPI}} \mathbf{v}$, if:

$$(5.2) \qquad [u_i, u_j] \subseteq [v_i, v_j], \quad \text{for all } i, j \in N \, (i \neq j).$$

In order to decide whether $\mathbf{u}$ dominates $\mathbf{v}$ according to unanimous pairwise inclusion, one has to perform $n(n-1)/2$ pairwise comparisons, something that rapidly involves a huge number of computations in practice. Interestingly, Kolm (1997) claims that unanimous pairwise inclusion dominance is equivalent to what he called bi-truncation dominance that actually only requires n pairwise comparisons to be made. Given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we say that $\mathbf{u}$ *bi-truncation dominates* $\mathbf{v}$, which we write $\mathbf{u} \geq_{\mathrm{BTR}} \mathbf{v}$, if there exist two individuals i, j and two reals a, b such that $v_i \leqslant a < b \leqslant v_j$ and:

$$(5.3a) \qquad u_h = a, \quad \text{for } h = 1, 2, \ldots, i;$$

$$(5.3b) \qquad u_h = v_h, \text{ for } h = i+1, \ldots, j-1; \text{ and}$$

$$(5.3c) \qquad u_h = b, \quad \text{for } h = j, j+1, \ldots, n.$$

It is almost immediate that $\mathbf{u} \geq_{\mathrm{BTR}} \mathbf{v}$ implies that $\mathbf{u} \geq_{\mathrm{UPI}} \mathbf{v}$. Conversely, one can show that, if $\mathbf{u} \geq_{\mathrm{UPI}} \mathbf{v}$, then one has

$$(5.4) \qquad v_1 \leqslant u_1 = a \leqslant u_2 = v_2 \leqslant \cdots \leqslant v_{n-1} = u_{n-1} \leqslant b = u_n \leqslant v_n,$$

which implies that (5.3) holds with $i = 1$ and $j = n$, hence $\mathbf{u} \geq_{\mathrm{BTR}} \mathbf{v}$. One might be tempted to apply the same technique to Hammond progressive transfers and say that the profile $\mathbf{u}$ is *unanimously Hammond dominates* the profile $\mathbf{v}$, which we write $\mathbf{u} \geq_{\mathrm{UH}} \mathbf{v}$, if $v_i < u_i \leqslant u_j < v_j$, for all $i \neq j$. For instance, in the case where $n = 3$, the above definition requires that the following inequalities are verified:

$$(5.5a) \qquad \{1, 2\} : \ v_1 < u_1 \leqslant u_2 < v_2;$$

$$(5.5b) \qquad \{1, 3\} : \ v_1 < u_1 \leqslant u_3 < v_3; \text{ and}$$

$$(5.5c) \qquad \{2, 3\} : \ v_2 < u_2 \leqslant u_3 < v_3.$$

Clearly, conditions (5.5a) and (5.5c) cannot be met simultaneously, which shows that the unanimous Hammond dominance criterion fails to exist whenever $n \geqslant 3$.

Contrary to the notion of a Hammond progressive transfer, the increment and the decrement need not be matched in a pairwise inclusion. According to the latter, all increments are considered inequality reducing transformations as long as there exists some individual whose initial endowment is not smaller than that of the individual who benefited from the increment. A similar remark applies *mutatis mutandis* in the case of decrements. Note that, in the approach of Allison and Foster (2004), increments and decrements reduce inequality subject to the proviso that (i) increments take place below the median, (ii) decrements occur above the median, and (iii) neither of them jump over the median. Therefore, the implicit notion of *inequality reducing transformation* considered by these authors is a particular case of the pairwise inclusion of Kolm. A possibility in order to disentangle inequality reduction and efficiency improvements in the pairwise inclusion approach would be to say that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of an increment if $u_i = v_i + \Delta > v_i$ for some i and there exists no j such that $u_i \leqslant v_j$. According to this definition, the only admissible increment is such that $u_i = v_i + \Delta > v_n$, hence $\Delta > v_n - v_i$: all other increases in person i' utility would result in pairwise inclusions.

Variable population size The criteria discussed in the paper can be easily adapted to the general case where the population size is allowed to vary through the introduction of the *principle of population* (PP) of Dalton (1920). Let us say that distribution $\mathbf{u} := (u_1, \ldots, u_n)$ is a *replication* of distribution $\mathbf{v} := (v_1, \ldots, v_m)$ if there exists $q \geqslant 2$ such that $\mathbf{u} = (\mathbf{v}; \ldots; \mathbf{v}) \in \mathbb{R}^{qm}$. A social preference relation R defined over $\bigcup_{n=2}^{\infty} \mathbb{R}^n$ verifies the principle of population if $\mathbf{u} \, \mathrm{I} \, \mathbf{v}$ whenever $\mathbf{u}$ is a replication of $\mathbf{v}$. Then, given two profiles $\mathbf{u}, \mathbf{v} \in \bigcup_{n=2}^{\infty} \mathbb{R}^n$, we would say that $\mathbf{u}$ (weakly) leximin-antileximax dominates $\mathbf{v}$, which we write $\mathbf{u} \, \mathrm{R}_{\mathrm{LMX}}^* \, \mathbf{v}$, if and only if, *either* $U(p) = V(p)$, for all $p \in [0, 1]$, *or*

$$(5.6a) \qquad \exists \, 0 < s < t < 1 \mid U(s) > V(s); \; U(t) < V(t); \text{ and}$$

$$(5.6b) \qquad U(p) = V(p), \, \forall \, p \in [0, s) \cup (t, 1],$$

where $U(p)$ and $V(p)$ are the *quantile* functions corresponding to $\mathbf{u}$ and $\mathbf{v}$ defined in the usual way (see, e.g., Moyes (1999, Section 2.2). We indicate by $\mathrm{I}_{\mathrm{LMX}}^*$ and $\mathrm{P}_{\mathrm{LMX}}^*$ the symmetric and asymmetric components of $\mathrm{R}_{\mathrm{LMX}}^*$. To illustrate things, consider the profiles $\mathbf{u} = (1, 2, 3, 4)$ and $\mathbf{v} = (1, b, 4)$. Indicate respectively by $\tilde{\mathbf{u}} = (1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4)$ and $\tilde{\mathbf{v}} = (1, 1, 1, 1, b, b, b, b, 4, 4, 4, 4)$ the 3-fold and 4-fold replicates of $\mathbf{u}$ and $\mathbf{v}$. Then, we have

$$(5.7) \qquad \tilde{u}_4 > \tilde{v}_4; \; \tilde{u}_9 < \tilde{v}_9 \text{ and } \tilde{u}_i = \tilde{v}_i, \, \forall \, i \in \{1, 2, 3\} \cup \{10, 11, 12\}.$$

Equivalently

$$(5.8) \qquad U\left(\frac{1}{4}\right) > V\left(\frac{1}{4}\right); \; U\left(\frac{3}{4}\right) < V\left(\frac{3}{4}\right) \text{ and } U(p) = V(p), \, \forall \, p \in \left[0, \frac{1}{4}\right) \cup \left(\frac{3}{4}, 1\right],$$

and therefore $\mathbf{u} \, \mathrm{I}_{\mathrm{LMX}}^* \, \tilde{\mathbf{u}} \, \mathrm{P}_{\mathrm{LMX}}^* \, \tilde{\mathbf{v}} \, \mathrm{I}_{\mathrm{LMX}}^* \, \mathbf{v}$ as long as $1 < b < 4$. One adapts in a similar way the definitions of the Suppes, leximin and antileximax criteria to accommodate the possibility that the populations are of different sizes.

6. Some Implications for Social Choice

Invoking the *welfarism theorem* makes clear the connection between our Theorems 4.1 to 4.4 and standard results in the social choice literature. The welfarism theorem allows one to translate the ordering of the social states into an ordering of the utility profiles associated with these social states provided that the social welfare ordering (SWO) satisfies the conditions of *universal domain* (UD), *independence with respect to irrelevant alternatives* (IIA), and *Pareto indifference* (PI) (see, e.g., Bossert and Weymark (2004, Theorem 4.2)). Assuming *ordinal level comparability* (OLC) and UD, it has been shown that, if a social welfare ordering satisfies IIA, A and SP, then it is *positional dictatorship* (see Gevers (1979) and Roberts (1980)). On the one hand, this result is somewhat related to our Theorem 4.1, where it is said that profile $\mathbf{u}$ is preferred to profile $\mathbf{v}$ by the Suppes criterion if and only if *all* positional dictators agree to rank $\mathbf{u}$ above $\mathbf{v}$. On the other hand, no use is made of conditions OLC, UD and IIA in Theorem 4.1, nor is the social preference relation even assumed to be an ordering. Adding HE to the list of conditions listed above precipitates the leximin (see Hammond (1976)). In this respect, Theorem 4.3 can be seen as providing an alternative characterisation of the leximin, albeit in a more specific framework than that commonly used in the social choice theory. More precisely, we have the following result, where we remind the reader that the elements of $\mathscr{R}$ are reflexive, transitive but not necessarily complete:

Theorem 6.1. *The social preference relation* $R \in \mathscr{R}$ *satisfies conditions A, SP and HE if and only if it is the leximin.*

Proof. That the leximin satisfies conditions A, SP and HE is readily inferred from Theorem 4.3. Indeed, it is said there that, given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, if $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a finite sequence of permutations, Hammond progressive transfers and/or increments, then $\mathbf{u}$ is preferred to $\mathbf{v}$ by the leximin. In other words, the leximin satisfies conditions A, SP and HE, hence $R_{LM} \in \mathscr{R}_+^*$. To establish that conditions A, SP and HE imply the leximin amounts to showing that $\mathscr{R}_+^*$ contains only the leximin. Note that, because the leximin is a complete relation, given two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have *either* $\mathbf{u} R_{LM} \mathbf{v}$, *or* $\mathbf{v} R_{LM} \mathbf{u}$. Suppose that there exists an element $R \in \mathscr{R}_+^*$ distinct from the leximin. Then, there are two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$ such that $\mathbf{u} R_{LM} \mathbf{v}$ and $\neg [\mathbf{u} R \mathbf{v}]$. But this is clearly impossible, for invoking Theorem 4.3 again, we know that, if $\mathbf{u} R_{LM} \mathbf{v}$, then $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a finite sequence of permutations, Hammond progressive transfers and/or increments. Since $R \in \mathscr{R}_+^*$, it must be that $\mathbf{u} R \mathbf{v}$, which contradicts our assumption. □

Making use of Theorem 4.4, one can prove along a similar reasoning that the antileximax is the only relation satisfying conditions A, SAP and HE.

Theorem 6.2. *The social preference relation* $R \in \mathscr{R}$ *satisfies conditions A, SAP and HE if and only if it is the antileximax.*

Contrary to what happens with Theorems 4.3 and 4.4, it must be noted that it is not possible to derive a characterisation of the leximin-antileximax starting with Theorem 4.2. The difficulty originates in the fact that the leximin-antileximax is not the only relation satisfying conditions A and HE. Then, it is possible to find a social preference in $R^* \in \mathscr{R}^*$ distinct from R_{LMX} such that $\mathbf{u} P^* \mathbf{v}$ and $\neg [\mathbf{u} R_{LMX} \mathbf{v}]$, for some $\mathbf{u}, \mathbf{v} \in \mathscr{U}$. For instance, both the leximin and the antileximax are consistent with the leximin-antileximax. Indeed, since $\mathscr{R}_+^* \subset \mathscr{R}^*$ and $\mathscr{R}_-^* \subset \mathscr{R}^*$, it follows that, for all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, we have $\mathbf{u} R_{LM} \mathbf{v}$ and $\mathbf{u} R_{ALX} \mathbf{v}$ whenever $\mathbf{u} R_{LMX} \mathbf{v}$ (see Remark 3.1). Choose $\mathbf{u} = (a, b, b, \ldots, b, d, d)$ and $\mathbf{v} = (a, a, c, \ldots, c, c, d)$, where $a < b < c < d$. Then, we get $\mathbf{u} P_{LM} \mathbf{v}$, $\mathbf{v} P_{ALX} \mathbf{u}$, while $\mathbf{u}$ and $\mathbf{v}$ cannot be ranked by the leximin-antileximax.

7. Concluding Remarks

Suppose that we are interested in the comparisons of distributions of a continuous and ordinal attribute with a particular concern for inequality consideration. In such a context, the principle of transfers, according to which inequality decreases as the result of a progressive transfer, is meaningless because it is not invariant to arbitrary changes in the scale of measurement. The fact that one distribution is obtained from another by means of a progressive transfer does not imply that it will be possible to perform the same operation after these distributions have been subjected to the same increasing transformations of the individuals' utilities. The difficulty originates in that the equality of distances between the utility of the beneficiary and that of the recipient of the transfer does not survive to such transformations. Abandoning the equality restriction, while still insisting fact that individuals must be brought closer without affecting their relative positions, leads us to what we refer to as Hammond progressive transfers. The corresponding Hammond equity principle constitutes therefore in an ordinal setting the natural analogue of the principle of transfers.

We have shown in the paper that (i) imposing that the social preference relation satisfies Hammond equity and anonymity, and (ii) requiring unanimity over all such social preferences inexorably lead to what we have called the leximin-antileximax. In other words, one profile is judged to be better than another if the leximin and the antileximax agree to rank the first profile above the second. Adding strong Pareto to Hammond equity and anonymity precipitates the leximin, while the adjunction of strong antiPareto results in the antileximax. The leximin-antileximax and leximin can be considered the analogues in an ordinal framework of the majorisation and weak submajorisation criteria better known as the Lorenz and generalised Lorenz quasi-orderings, respectively, in the inequality and welfare literature. Similarly, the antileximax criterion is the adaptation of weak supermajorisation when the available information is of an ordinal nature.

An important difference between the cardinal and ordinal frameworks when accounting for inequality is the ability of the resulting criteria to discriminate among the distributions. It is well-known – and this is sometimes considered a weakness – that the Lorenz criterion, as well as its extensions, does not permit one to completely order the distributions under comparison. While a similar critique – though somehow weaker – may be addressed to the leximin-antileximax criterion, it must be emphasised that the leximin and the antileximax are able to rank *all* the distributions under comparison. Thus, the addition of strong Pareto – similarly, of strong antiPareto – to Hammond equity et anonymity allows one to break all the incomparabilities that may result from the application of the leximin-antileximax criterion. This does not occur in the standard framework where the introduction of a concern for efficiency in addition to the principle of transfers is unable to resolve all the cases of incomparabilities that arise from the use of the Lorenz criterion.

The social preference relations discussed in the paper – the Suppes quasi-ordering, the leximin, the antileximax, and the leximin-antileximax – are unidimensional criteria. This is not a problem in a social choice setting as long as a person's utility incorporates everything is deemed to be relevant for her well-being. Following persuasive arguments from Sen, this welfarist position has been criticised on the grounds that it fails to take into account some important dimensions of a person's well-being (see, e.g., Sen (1985)) and it has been suggested that a multidimensional approach to inequality might be preferable. A much neglected difficulty in the multidimensional approach is the recognition that the components of a person's well-being – one may think of income, health, or cognitive ability, among other things – involve measurement scales of different nature ranging from cardinal to ordinal scales when they are not categorical. Extending the approach followed in the paper to the comparison of distributions of attributes involving different measurement scales is certainly the way to go in the future.

References

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

Amiel, Y. and Cowell, F. A. (1999). *Thinking About Inequality*. Cambridge University Press, Cambridge.

Bossert, W. and Weymark, J. A. (2004). Utility in social choice. In S. Barberà, P. J. Hammond,

and C. Seidl, editors, *Handbook of Utility Theory. Volume 2, Extensions*, pages 1099–1177. Kluwer Academic Publishers, Boston/Dordrecht/London.

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

Donaldson, D. and Weymark, J. A. (1998). A quasiordering is the intersection of orderings. *Journal of Economic Theory*, **78**, 382–387.

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.

Gevers, L. (1979). On interpersonal comparability and social welfare orderings. *Econometrica*, **47**, 75–89.

Gravel, N., Magdalou, B., and Moyes, P. (2018). Ranking distributions of an ordinal variable. Mimeo.

Hammond, P. J. (1976). Equity, Arrow's conditions and Rawls' difference principle. *Econometrica*, **44**, 793–803.

Hammond, P. J. (1979). Equity in two person situations: Some consequences. *Econometrica*, **47**, 1127–1135.

Kolm, S.-C. (1997). *Justice and Equity*. The MIT Press, Cambridge, Massachusetts.

Kolm, S.-C. (1999). The rational foundations of income inequality measurement. In J. Silber, editor, *Handbook of Income Inequality Measurement*, pages 19–100. Kluwer Academic Publishers, Boston/Dordrecht/London.

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.

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

Moyes, P. (1999). Stochastic dominance and the Lorenz curve. In J. Silber, editor, *Handbook of Income Inequality Measurement*, pages 199–222. Kluwer Academic Publishers, Boston/Dordrecht/London.

Moyes, P. (2013). Rearrangements and sequential rank order dominance. *Journal of Mathematical Economics*, **49**, 278–290.

Roberts, K. (1980). Possibility theorems with interpersonally comparable welfare levels. *Review of Economic Studies*, **47**, 409–420.

Sen, A. K. (1977). On weights and measures: Informational constraints in social welfare analysis. *Econometrica*, **45**, 1539–1572.

Sen, A. K. (1985). *Commodities and Capabilities*. North-Holland, Amsterdam.

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

Suppes, P. (1966). Some formal models of grading principles. *Synthese*, **16**, 284–306.

Tungodden, B. (2000). Egalitarianism: Is leximin the only option? *Economics and Philosophy*, **16**, 229–245.

Additional Material

Part of the following notes have been motivated to a large extent by the comments of our referees. While they have not be inserted in the final version of the paper, it appears to us that they complement our paper and that they might be of interest to our potential readers.

1. **Continuous versus categorical variables** Our general impression is that many people in the profession wrongly assimilate *cardinality* with *continuity*, on the one hand, and *ordinality* with a *list of ordered categories*, on the other hand. We distinguish in Table 1 the different cases

Table 1: Inequality for different types of variables

	Cardinal	Ordinal
Continuous	Standard inequality literature	This paper
Categorical	Chakravarty and Zoli (2012) Moyes (2012, Section 4) Yalonetzky (2013)	Allison and Foster (2004) Cowell and Flachaire (2017) Gravel, Magdalou, and Moyes (2018)

that may arise and provide examples of papers that propose criteria for measuring inequality in each case.[1]

Before we discuss the cases indicated in Table 1, we would like to stress that our paper is theoretically inclined: in this respect, it must be mainly seen as providing characterisations of criteria related to the leximin. From an illustrative perspective, it must be acknowledged that the distinction between continuous and categorical variables is not totally unambiguous. Depending on the authors and the circumstances, the same attribute (e.g., health status, PISA scores) can be used to illustrate, *either* the case of an ordinal and continuous variable, *or* the case of an ordinal and categorical variable. Considerable space in the literature has been devoted to the case of a continuous variable measurable on a (at least) cardinal scale and there is nothing to be added there.

The fact that we have a fixed list of *ordered categories* does not mean that making *comparisons of utilities differences* – one utility per category – is a meaningless exercise. For instance, consider a variable like family size and assume that we all agree to order the different family categories on the basis of the number of individuals as indicated in Table 2 where the symbols

Table 2: A cardinal and categorical variable

Family Categories	A	C	$C+1$	$C+2$	$C+3$	$C+4$
Family Size	1	3	3	4	5	6
Categories Ranking	6	5	4	3	2	1

"A" indicate a household comprising a single adult, "C" a household constituted of a (married) couple, and "$C+K$" a household constituted of a couple with K children. Actually, we could go even further and consider that the difference in family neediness is greater between a couple

---

[1] We do not intend to provide here an exhaustive list of contributions falling into any of these categories and we apologise to our colleagues for the omissions.

and a single than between a couple with one child and a couple without children, which is not total nonsense.[2] Pursuing in this direction opens the route to dominance criteria like those examined in Moyes (2012, Section 4), where the cardinal nature of family neediness is fully exploited. Incidently, we note that a number of papers that claim to measure inequality in an ordinal framework actually use a similar approach: a good example is Yalonetzky (2013).

A typical example of the case addressed in the paper – that of a *continuous and ordinal variable* – is the comparison of utility profiles in the social choice and welfare theory. As another instance, we can think of (i) comparisons of performance across different disciplines (e.g., philosophy, geography, mathematics) of a given group of students or (ii) comparisons of achievements in a given discipline of different classes. On the one hand, one observes that teachers from different disciplines have specific habits for marking copies: traditionally, philosophers have a tendency to give lower marks than their colleagues from other fields. On the other hand, there is evidence that teachers of the same discipline do not use the full range of available marks: some teachers are cautious and give notes concentrated around the average scale while others are, *either* too demanding, *or* too generous. Consider two classes $A = \{a, b, c, d\}$ and $B = \{\alpha, \beta, \gamma, \delta\}$ and three teachers X, Y and Z who are allowed to give marks from 0 to 10. The marks attributed to the different students in classes A and B by the three teachers are indicated in Table 3. Let us denote by $X(A)$, $Y(A)$ and $Z(A)$ (resp. $X(B)$,

Table 3: The marking of two classes by four teachers

Teacher	Class A				Class B			
	a	b	c	d	α	β	γ	δ
X	1	3	8	9	1	5	6	9
Y	1	3	7	9	1	6	7	8
Z	2	4	6	7	2.5	3.5	4	9

$Y(B)$ and $Z(B)$) the distributions of marks given by the teachers X, Y and Z to the class A (resp. B). Whatever the (relative) inequality index one uses, the marks given by X are more equally distributed in class B than in class A: this follows from the fact that the (relative) Lorenz curve of $X(B)$ is nowhere below and somewhere above that of $X(A)$. Considering next the distributions $Y(A)$ and $Y(B)$, it can be checked that, depending on which particular inequality index is chosen, $Y(A)$ is more equal or less equal than $Y(B)$. Similarly, $Z(A)$ is more equal than $Z(B)$ according to some inequality indices while the opposite verdict arises for other indices.[3] This does not come as a surprise because the relative Lorenz criterion fails to declare which distribution among $Y(A)$ and $Y(B)$ on the one hand, and among $Z(A)$ and $Z(B)$ on the other hand, is the more equal. It is therefore difficult to obtain an indisputable verdict as far as which class is the most unequal when one gives a cardinal meaning to the notes observed. However, inspection of Table 3 suggests that there is something common in the way teachers X and Y assess students. Indeed, both teachers share the opinion that class B is more equal than class A in the sense that distributions of marks given to class B can

---

[2] In this respect, the use of equivalence scales in empirical work makes it possible to compare the differences – when they are not the ratios – of well-being between households of different composition.

[3] This is typically the case when one appeals to the AKS family of inequality indices for comparing the distributions of marks between the classes A and B. Depending on the value of the inequality aversion parameter, the marks are more evenly distributed in class A than in class B, and conversely.

be derived from those given to class A by means of Hammond progressive transfers. On the other hand, it is impossible to declare on the basis of the marks given by teacher Z which of the classes A and B is the least unequal: indeed, distribution $Z(A)$ cannot be derived from distribution $Z(B)$ by means of Hammond progressive transfers, and *vice versa*.

The fact that teachers do not mark copies in the same way has long been recognised and this observation has led to the application of standardisation procedures in order to force the distributions of marks to look the same and prevent systematic biases in the overall evaluation. These different procedures all amount to re-scaling the distributions of marks in particular ways and this re-scaling has an impact on both the mean and dispersion of the marks obtained. To give an example, the distribution of the PISA scores obtained in the different countries are adjusted in such a way that the worldwide distribution of scores has mean and standard deviation equal to 500 and 100, respectively. While the impact on the ranking of countries of the standardisation of the PISA scores is not fully comprehended, it is nevertheless clear that the comparison of the rankings across time is questionable. Indeed, other things being the same, the adjunction of a country into the database can potentially affect the original ranking of the countries obtained on the basis of the countries' mean scores. By the same token, it is possible to distort the overall picture of a class's achievements in various disciplines or of different classes in a given discipline by re-scaling the exams' marks. By relying on the sole ordinal information provided by the distributions of scores or marks, one avoids the difficulties pointed out above.

As reminded above, the case of *ordered categorical data* – individuals are allocated into a finite number of ordered categories – is the subject of Gravel, Magdalou, and Moyes (2018). Subscribing to an ordinal scale amounts to accepting that the only relevant information, when deciding which distribution is more equal than another, is the rank of the category that is attributed to each individual. We claim that, even within such a poorly informative setting, the notion of a more equal distribution still makes sense. Now, the precise meaning it must be given is clearly open to discussion. The notion of inequality reduction that we have retained – what we have called a Hammond progressive transfer – is imported from the social choice and welfare literature. A transformation that makes two individuals closer – each individual makes a move in direction of the other one – without modifying their respective positions is considered an inequality reducing operation. Going back to Table 3, one observes that distributions $X(B)$ and $Y(B)$ are obtained from distributions $X(A)$ and $Y(A)$, respectively, by means of a Hammond progressive transfer.

The fact that the variable of interest is continuous or categorical makes a huge difference when measuring inequality. Indeed, depending on the nature of the variable, the same normative conditions result in criteria that appear to be very different. Actually, this must be tempered as the leximin, antileximin and leximin-antileximax can be seen respectively as the limits of the H^+-, H^-- and H-criteria introduced in Gravel, Magdalou, and Moyes (2018).

2. **Quantiles versus cumulative distributions** We could re-write all the paper by adopting the dual approach that consists in working with the distributions of densities and the CDFs. Then, we would associate to the quantile distribution $X(p)$ ($p \in [0,1]$) the density distribution $f_X(s)$ and the cumulative distribution $F_X(s)$ ($s \in \mathbb{R}$). Consider the following distributions: $\mathbf{x} = (2,2,5,8,8,9)$, $\mathbf{y} = (1,4,5,7,7,7)$, $\mathbf{z} = (1,3,3,7,9,9)$ and $\mathbf{u} = (1,4,4,5,5,8)$. Let $D(\mathbf{x}) =$

$\{2,5,8,9\}$, $D(\mathbf{y}) = \{1,4,5,7\}$, $D(\mathbf{z}) = \{1,3,7,9\}$, $D(\mathbf{u}) = \{1,4,5,8\}$ and $\mathscr{D} = D(\mathbf{x}) \cup D(\mathbf{y}) \cup D(\mathbf{z}) \cup D(\mathbf{u}) = \{1,2,3,4,5,7,8,9\}$. The corresponding relative frequencies – or densities – are indicated in Table 4. We next define the corresponding CDFs $F_X(s)$, $F_Y(s)$, $F_Z(s)$ and $F_U(s)$

Table 4: Relative frequencies of distributions $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ and $\mathbf{u}$

$f_X(2) = 2/6$; $f_X(5) = 1/6$; $f_X(8) = 2/6$; $f_X(9) = 1/6$; $f_X(s) = 0$ otherwise;
$f_Y(1) = 1/6$; $f_Y(4) = 1/6$; $f_Y(5) = 1/6$; $f_Y(7) = 3/6$; $f_Y(s) = 0$ otherwise;
$f_Z(1) = 1/6$; $f_Z(3) = 2/6$; $f_Z(7) = 1/6$; $f_Z(9) = 2/6$; $f_Z(s) = 0$ otherwise;
$f_U(1) = 1/6$; $f_U(4) = 2/6$; $f_U(5) = 2/6$; $f_U(8) = 1/6$; $f_U(s) = 0$ otherwise.

in the usual way (see, e.g., Moyes (1999, Section 2)). Then, we proceed as in classic stochastic dominance and we have to rewrite all our definitions, results and proofs in order to fit this new framework. While we believe it is possible to do this, we also think it would make things more complicated and we do not really see what we could gain from this exercise except to contrast in a more direct way our results with those of the classic stochastic dominance approach.

To come back to Gravel, Magdalou, and Moyes (2018), suppose that there are 6 such categories and let $\mathscr{C} := \{a, \square, \triangledown, \alpha, \boxtimes, \diamondsuit\}$ indicate the set of categories. Assume furthermore that the following ordering of categories is agreed upon: $a \triangleleft \alpha \triangleleft \square \triangleleft \triangledown \triangleleft \boxtimes \triangleleft \diamondsuit$, where $\triangleleft$ is a complete, transitive and antisymmetric binary relation on $\mathscr{C}$. Let us denote by $r(h, \triangleleft) := \#\{f \in \mathscr{C} \mid f \triangleleft h\} + 1$ the rank of category $h \in \mathscr{C}$. Let $n = 8$ and consider the distributions $\mathbf{x} := (a, \square, \square, \square, \triangledown, \boxtimes, \diamondsuit, \diamondsuit)$ and $\mathbf{y} := (\alpha, \alpha, \square, \triangledown, \triangledown, \boxtimes, \boxtimes, \boxtimes)$. The categories' ranks and the densities of distributions $\mathbf{x}$ and $\mathbf{y}$ are indicated in Table 5. We can alterna-

Table 5: Example of distributions for ordered categories

Category h	a	α	$\square$	$\triangledown$	$\boxtimes$	$\diamondsuit$
Category's ranks $r(h, \triangleleft)$	1	2	3	4	5	6
$f_X(r(h, \triangleleft))$	1/8	0/6	3/6	1/6	1/6	2/6
$f_Y(r(h, \triangleleft))$	0/6	2/6	1/6	2/6	3/6	0/6

tively work with, *either* the distributions of the categories' ranks $\tilde{\mathbf{x}} := (1,3,3,3,4,5,6,6)$ and $\tilde{\mathbf{y}} := (2,2,3,4,4,5,5,5)$, *or* the lotteries $(1,2,3,4,5,6 \mid f_X(1), f_X(2), f_X(3), f_X(4), f_X(5), f_X(6))$ and $(1,2,3,4,5,6 \mid f_Y(1), f_Y(2), f_Y(3), f_Y(4), f_Y(5), f_Y(6))$.

3. **Comparison with related papers in the literature** To the extent that Gravel, Magdalou, and Moyes (2018) devote apparently too little space to the discussion of related results in the literature, we take this opportunity to comment on a selection of contributions related to the afore mentioned paper.

The work by Zheng (2011) is in line with several studies in health economics where health inequalities are not considered in isolation but rather in association with a selection of socioeconomic variables. The question is not to know if health is unequally distributed among the population but rather to see if – and to which extent – the individuals' circumstances (income, education, area of residence) have an impact on their health statuses. This approach is closely

related to the area of research initiated by Atkinson and Bourguignon (1987) and concerned with the comparisons of bidimensional distributions where one attribute (typically, income) is cardinal and the other (for instance, health achievement) is ordinal. More generally, one is interested in the aversion to bidimensional inequality which, in the utilitarian framework, has typically been captured by the submodularity of the utility function. Other things being the same, transformations of the joint distribution of income and health status that decrease the correlation between the two variables result in a welfare improvement according to the utilitarian rule if and only if the utility function is submodular. Invoking the *other things being the same* clause, this welfare improvement can be assimilated with a decrease in (bidimensional) inequality. As is well-known, unanimity among these utilitarian rules coincide with bidimensional first order stochastic dominance. The latter criterion is a purely ordinal one: arbitrary changes in the scales of measurement of the two variables preserves the ranking of the distributions under comparison. This is no longer true when one considers higher degrees of dominance for now the cardinal nature of the second variable is clearly involved (see, e.g., Atkinson and Bourguignon (1982), Bourguignon (1989), Gravel and Moyes (2012)).

The approach of Allison and Foster (2004) is the most closely related to ours even though they do not consider explicitly transformations of distributions that capture the idea of inequality reduction. Clearly, increments below the median category and decrements above the median category decrease inequality as measured by these authors. This implies in particular that Hammond progressive transfers around the median – the giver stays above the median and the receiver below it – reduce inequality. It is a weaker notion of inequality reduction than ours because of the constraint that the median is preserved. This constraint limits the possibilities of obtaining a conclusive verdict: it is therefore not surprising that our criteria are more discriminating than that of Allison and Foster (2004). The authors do not show whether a sequence of such transformations – increments below the median and/or decrements above the median – allows one to derive the dominating distribution from the dominating one. In addition, their dominance criterion mixes (in)efficiency and equality considerations: increments and decrements reduce inequality as soon as they take place below and above the median, respectively.

Of a completely different nature is the contribution of Cowell and Flachaire (2017). Here, each individual is given a status assimilated with her relative position in the society. More precisely, the status of an individual who falls in category h, let's say, is equal to the percentage of those individuals who fall in categories ranked (stricly) below h. In other words, an individual's status is nothing but the value of the decumulative distribution function – the *survival function* – evaluated at the category ranked h. This position clearly depends on the distribution and an individual who occupies the same category in two distinct distributions is likely to end up with two different statuses. Contrary to Allison-Foster's approach and ours, there is no need here that the list of categories be fixed. This might be considered an advantage of the approach of Cowell and Flachaire (2017) since it makes it possible to use databases where the list of categories differ. On the other hand, it is not very clear what one can infer about inequality from the comparisons of distributions defined over different sets of categories.

Variable population size The criteria discussed in the paper can be easily adapted to the general case where the population size is allowed to vary through the introduction of the

principle of population (PP) of Dalton (1920). Let us say that distribution $\mathbf{u} := (u_1, \ldots, u_n)$ is a *replication* of distribution $\mathbf{v} := (v_1, \ldots, v_m)$ if there exists $q \geqslant 2$ such that $\mathbf{u} = (\mathbf{v}; \ldots; \mathbf{v}) \in \mathbb{R}^{qm}$. The binary relation R defined over $\bigcup_{n=2}^{\infty} \mathbb{R}^n$ verifies the principle of population if $\mathbf{u} \, \mathrm{I} \, \mathbf{v}$ whenever $\mathbf{u}$ is a replication of $\mathbf{v}$ and I is the symmetric component of R. Then, given two distributions $\mathbf{u}, \mathbf{v} \in \bigcup_{n=2}^{\infty} \mathbb{R}^n$, we say that $\mathbf{u}$ (weakly) leximin-antileximax dominates $\mathbf{v}$, which we write $\mathbf{u} \, \mathrm{R}_{\mathrm{LMX}}^* \, \mathbf{v}$, if and only if, *either* $U(p) = V(p)$, for all $p \in [0,1]$, *or*

(1a) $$\exists \, 0 < s < t < 1 \mid U(s) > V(s); \ U(t) < V(t); \text{ and}$$

(1b) $$U(p) = V(p), \forall p \in [0,s) \cup (t,1],$$

where $U(p)$ and $V(p)$ are the *quantile* functions corresponding to $\mathbf{u}$ and $\mathbf{v}$ defined in the usual way (see, e.g., Moyes (1999, Section 2.2). Denote respectively by $\mathrm{I}_{\mathrm{LMX}}^*$ and $\mathrm{P}_{\mathrm{LMX}}^*$ the symmetric and asymmetric components of $\mathrm{R}_{\mathrm{LMX}}^*$. To illustrate things, consider the distributions $\mathbf{u} = (1,2,3,4)$ and $\mathbf{v} = (1,b,4)$. Let $\tilde{\mathbf{u}} = (1,1,1,2,2,2,3,3,3,4,4,4)$ and $\tilde{\mathbf{v}} = (1,1,1,1,b,b,b,b,4,4,4,4)$ be the 3-fold and 4-fold replicates of $\mathbf{u}$ and $\mathbf{v}$, respectively. Then, we have

(2) $$\tilde{u}_4 > \tilde{v}_4; \ \tilde{u}_9 < \tilde{v}_9 \text{ and } \tilde{u}_i = \tilde{v}_i, \forall i \in \{1,2,3\} \cup \{10,11,12\}.$$

Equivalently

(3) $$U\left(\frac{1}{4}\right) > V\left(\frac{1}{4}\right); \ U\left(\frac{3}{4}\right) < V\left(\frac{3}{4}\right) \text{ and } U(p) = V(p), \forall p \in \left[0, \frac{1}{4}\right) \cup \left(\frac{3}{4}, 1\right],$$

and therefore $\mathbf{u} \, \mathrm{I}_{\mathrm{LMX}}^* \, \tilde{\mathbf{u}} \, \mathrm{P}_{\mathrm{LMX}}^* \, \tilde{\mathbf{v}} \, \mathrm{I}_{\mathrm{LMX}}^* \, \mathbf{v}$ as long as $1 < b < 4$. One similarly adapts the definitions of the Suppes, leximin and antileximax criteria to accommodate the possibility that the populations are of different sizes.

Let $\mathscr{R}^*$ be the set of social preference relations that verify conditions A, HE and PP. The following result is a possible generalisation of our Theorem 4.2 in the case where population sizes may differ.

Theorem 1. *For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$, the following three statements are equivalent:*

(a) *There exist two replications $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ of the profiles $\mathbf{u}$ and $\mathbf{v}$, respectively, such that $\tilde{\mathbf{u}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a finite sequence of permutations and/or Hammond progressive transfers.*

(b) $\mathbf{u} \, \mathrm{R} \, \mathbf{v}$, *for all $\mathrm{R} \in \mathscr{R}^*$.*

(c) $\mathbf{u} \, \mathrm{R}_{\mathrm{LMX}}^* \, \mathbf{v}$.

HINT. We indicate by $n(\mathbf{u})$ and $n(\mathbf{v})$ the dimensions of profiles $\mathbf{u}$ and $\mathbf{v}$, respectively. Let $\tilde{\mathbf{u}}$ be the $n(\mathbf{v})$-replicate of $\mathbf{u}$ and $\tilde{\mathbf{v}}$ be the $n(\mathbf{u})$-replicate of $\mathbf{v}$. The profiles $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{v}}$ have the same dimension equal to $n(\mathbf{u}) \times n(\mathbf{v})$ and Theorem 4.2 applies.

(a) $\Longrightarrow$ (b). This follows from the definition of $\mathscr{R}^*$.

(a) $\Longrightarrow$ (c). According to statement (a), $\tilde{\mathbf{u}}$ is obtained from $\tilde{\mathbf{v}}$ by means of a finite sequence of permutations and/or Hammond progressive transfers. We know from Theorem 4.2 that this equivalent to $\tilde{\mathbf{u}} \, \mathrm{R}_{\mathrm{LMX}}^* \, \tilde{\mathbf{v}}$, which implies that $\mathbf{u} \, \mathrm{R}_{\mathrm{LMX}}^* \, \mathbf{v}$.

(c) $\Longrightarrow$ (a). Suppose that $\mathbf{u} \, \mathrm{R}_{\mathrm{LMX}}^* \, \mathbf{v}$. This implies that $\tilde{\mathbf{u}} \, \mathrm{R}_{\mathrm{LMX}}^* \, \tilde{\mathbf{v}}$ and statement (a) follows from Theorem 4.2.

(b) $\implies$ (c). That (a) implies (c) guarantees that R^*_{LMX} verifies condition PP, A and HE, hence $R^*_{LMX} \in \mathscr{R}^*$. Therefore, if statement (b) holds, then so does statement (c). $\qquad\square$

Inclusion quasi-ordering and unanimity The notion of a Hammond progressive transfer is closely related to that of *pairwise inclusion* introduced by Kolm (1999, Section 3.2). Before we discuss at more length the latter, it is useful to note that all the criteria considered in the paper verify a weak form of separability. Given any profile $\mathbf{u} \in \mathscr{U}$ and any individual $i \in N$, let $\mathbf{u}_{-i} := (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n)$ indicate the profile where individual i has been removed. Consider the following condition that is a slight variant of Ebert (1988, Property 4):

COMONOTONE SEPARABILITY (CS). For all $\mathbf{u}^\circ, \mathbf{v}^\circ, \mathbf{u}^*, \mathbf{v}^* \in \mathscr{U}$, all $a, b \in \mathbb{R}$ and all $i, j, r \in N$ such that $\mathbf{u}^\circ_{-i} = \mathbf{u}^*_{-i}$, $\mathbf{v}^\circ_{-j} = \mathbf{v}^*_{-j}$, $u^\circ_i = \tilde{u}^\circ_r = a = \tilde{v}^\circ_r = v^\circ_j$ and $u^*_i = \tilde{u}^*_r = b = \tilde{v}^*_r = v^*_j$, we have $\mathbf{u}^\circ \, R \, \mathbf{v}^\circ$ if and only if $\mathbf{u}^* \, R \, \mathbf{v}^*$.

According to condition CS, if there exist two individuals who have the same utilities in profiles $\mathbf{u}^\circ$ and $\mathbf{v}^\circ$ and also in profiles $\mathbf{u}^*$ and $\mathbf{v}^*$ – these utilities being not necessarily the same – and if in addition they occupy the same positions on the utility scales, then their utilities must not count. In other words, those individuals who are unconcerned by the choice between two profiles must not have any impact on the ranking of these profiles. It must be emphasised that each of the individuals who is unconcerned must be so in two respects: not only she has the same utility in the two profiles, but also she occupies the same position on the utility scale in both profiles.[4]

Let $\mathbf{x} = (1, 2, 5, 8, 9)$, $\mathbf{y} = (1, 4, 5, 7, 9)$, $\mathbf{z} = (1, 2, 5, 7, 9)$ and $\mathbf{w} = (1, 4, 5, 8, 9)$ and assume further that the relation R used for comparing the distributions verifies condition CS. There are good reasons to consider that both $\tilde{\mathbf{z}} := (2, 7)$ and $\tilde{\mathbf{w}} := (4, 8)$ are more equal than $\tilde{\mathbf{x}} := (2, 8)$, for individuals 2 and 4 are closer in distributions $\tilde{\mathbf{z}}$ and $\tilde{\mathbf{w}}$ than they are in distribution $\tilde{\mathbf{x}}$. More precisely, we observe that distribution $\tilde{\mathbf{w}}$ is obtained from distribution $\tilde{\mathbf{x}}$ by giving two units of the attribute to individual 2, while taking one unit from individual 4 allows us to obtain $\tilde{\mathbf{z}}$ from $\tilde{\mathbf{x}}$. What is important to note here is that the increment and the decrement are constrained in such a way that the endowments of individuals 2 and 4 in the distributions $\tilde{\mathbf{w}}$ and $\tilde{\mathbf{z}}$ belong to the interval $[2, 8] = [x_2, x_4]$. Then, invoking CS, we conclude that $\mathbf{z} \, R \, \mathbf{x}$ and $\mathbf{w} \, R \, \mathbf{x}$. By the same token, $\tilde{\mathbf{y}} := (4, 7)$ is more equal than $\tilde{\mathbf{z}} := (2, 7)$ and $\tilde{\mathbf{w}} := (4, 8)$, which implies that $\mathbf{y} \, R \, \mathbf{z}$ and $\mathbf{y} \, R \, \mathbf{w}$. On the other hand, it is impossible to decide which of $\tilde{\mathbf{z}}$ and $\tilde{\mathbf{w}}$ is the most equal, and therefore $\mathbf{z}$ and $\mathbf{w}$ cannot be compared.

Consider two profiles $\mathbf{u}, \mathbf{v} \in \mathscr{U}$ and, from now on, assume without loss of generality that there are non-decreasingly arranged. Then, we say that $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a *weak pairwise inclusion* if there exist two individuals i and j such that

$$(4) \qquad\qquad [u_i, u_j] \subseteq [v_i, v_j] \text{ and } u_h = v_h, \text{ for all } h \neq i, j.$$

For later use, we write $(u_i, u_j) \succsim (v_i, v_j)$ when $[u_i, u_j] \subseteq [v_i, v_j]$ and we denote respectively by $\succ$ and $\sim$ the asymmetric and symmetric components of $\succsim$. The inclusion is *semi-strict* if (4) holds and, either $u_i \neq v_i$, or $u_j \neq v_j$. If $u_i \neq v_i$ and $u_j \neq v_j$, then the pairwise inclusion is *strict* and we are back to Hammond progressive transfers. Note that in both cases, we have $(u_i, u_j) \succ (v_i, v_j)$. These definitions seem to be in accordance with what Kolm writes in

---

[4] This condition is far from being innocuous and it is well-known that some social welfare orderings – the best-known being the Borda rule – violate this condition.

the second paragraph of Kolm (1999, Page 43). We have found no mention of the notion of pairwise inclusion in Kolm (1969). On the other hand, the notion of inclusion in relation with ordinal inequalities is discussed at some length in Sections 8 and 9 of Kolm (1997, Chapter III-C) and a number of results are given there.

A recurrent limitation of a progressive transfer – be it a Hammond transfer or a standard one – is the fact that, while it reduces inequality between the individuals involved, the transfer may generate an increase in the inequalities between each of these two individuals and the rest of the population.[5] Consider for instance the profiles $\mathbf{u} = (1,2,5,6,7,8,10,11)$ and $\mathbf{v} = (1,2,3,6,7,9,10,11)$, where $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a Hammond progressive transfer involving individuals 3 and 6. For each pair of indices $(h,k) \in \{1,2,\ldots,7\} \times \{2,3,\ldots,8\}$, Table 6 indicates the ranking of the couples (u_h, u_k) and (v_h, v_k) by the relation $\succsim$ defined above. It must be observed that, while $(u_3, u_6) \succ (v_3, v_6)$, there exist a number of pairs of indices (h,k) for which the opposite ranking obtains.

Table 6: Non unanimous pairwise inequalities reductions

	2	3	4	5	6	7	8
1	$(1,2)\sim(1,2)$	$(1,5)\prec(1,3)$	$(1,6)\sim(1,6)$	$(1,7)\sim(1,7)$	$(1,8)\succ(1,9)$	$(1,10)\sim(1,10)$	$(1,11)\sim(1,11)$
2		$(2,5)\prec(2,3)$	$(2,6)\sim(2,6)$	$(2,7)\sim(2,7)$	$(2,8)\succ(2,9)$	$(2,10)\sim(2,10)$	$(2,11)\sim(2,11)$
3			$(5,6)\succ(3,6)$	$(5,7)\succ(3,7)$	$(5,8)\succ(3,9)$	$(5,10)\succ(3,10)$	$(5,11)\succ(3,11)$
4				$(6,7)\sim(6,7)$	$(6,8)\succ(6,9)$	$(6,10)\sim(6,10)$	$(6,11)\sim(6,11)$
5					$(7,8)\succ(7,9)$	$(7,10)\sim(7,10)$	$(7,11)\sim(7,11)$
6						$(8,10)\prec(9,10)$	$(8,11)\prec(9,11)$
7							$(10,11)\sim(10,11)$

More generally, if $\mathbf{u}$ is obtained from $\mathbf{v}$ by means of a Hammond progressive transfer, i.e., there exist i,j such that $v_i < u_i \leqslant u_j < v_j$ and $u_h = v_h$, for all $h \neq i,j$, then it can checked that

(5a) $\qquad\qquad (u_i, u_j) \succ (v_i, v_j);$

(5b) $\qquad\qquad (u_h, u_k) \sim (v_h, v_k), \forall h,k \in \{1,2,\ldots,i-1\}\ (h < k);$

(5c) $\qquad\qquad (u_h, u_i) \prec (v_h, v_i), \forall h \in \{1,2,\ldots,i-1\};$

(5d) $\qquad\qquad (u_h, u_j) \prec (v_h, v_j), \forall h \in \{1,2,\ldots,j-1\}\ (h \neq i);$

(5e) $\qquad\qquad (u_h, u_k) \sim (v_h, v_k), \forall h,k \in \{i+1,\ldots,j-1\}\ (h < k);$

(5f) $\qquad\qquad (u_i, u_k) \succ (v_i, v_k), \forall k \in \{i+1,i+2,\ldots,n\}\ (k \neq j);$

(5g) $\qquad\qquad (u_j, u_k) \succ (v_j, v_k), \forall k \in \{j+1,\ldots,n-1,n\};$

(5h) $\qquad\qquad (u_h, u_k) \sim (v_h, v_k), \forall h,k \in \{j+1,\ldots,n-1,n\}\ (h < k).$

An obvious solution in order to avoid the kind of situation described above would be to declare that inequality decreases if and only if all pairwise inequalities are simultaneously reduced. The two following quasi-orderings, that are due to Kolm (1997, Chapter III-C), precisely aim at capturing this basic idea. Let us first declare that $\mathbf{u}$ *unanimously pairwise*

---

[5] The notion of a uniform on the right progressive transfer (see, e.g., Magdalou and Moyes (2009)) provides a means of avoiding this difficulty in the standard cardinal framework.

inclusion dominates $\mathbf{v}$, which we write $\mathbf{u} \geq_{\text{ICL}} \mathbf{v}$, if $[u_i, u_j] \subseteq [v_i, v_j]$, for all i, j. Next, say that $\mathbf{u}$ *bi-truncation dominates* $\mathbf{v}$, which we write $\mathbf{u} \geq_{\text{BTR}} \mathbf{v}$, if there exist two individuals i, j and $v_i \leqslant a < b \leqslant v_j$ such that:

(6a) $\qquad\qquad\qquad\qquad u_h = a, \quad \text{for } h = 1, 2, \ldots, i;$

(6b) $\qquad\qquad\qquad\qquad u_h = v_h, \text{ for } h = i+1, \ldots, j-1; \text{ and}$

(6c) $\qquad\qquad\qquad\qquad u_h = b, \quad \text{for } h = j, j+1, \ldots, n.$

This gives rise to the kind of situation described in Table 7 below.

Table 7: Illustration of the bi-truncation quasi-ordering

v_1	$\leqslant$	$\cdots$	$\leqslant$	v_i	$<$	a	$\leqslant$	v_{i+1}	$\leqslant$	$\cdots$	$\leqslant$	v_{j-1}	$\leqslant$	b	$<$	v_j	$\leqslant$	$\cdots$	$\leqslant$	v_n
$\wedge$				$\wedge$				$\parallel$				$\parallel$				$\vee$				$\vee$
u_1	$=$	$\cdots$	$=$	u_i	$=$	a	$\leqslant$	u_{i+1}	$\leqslant$	$\cdots$	$\leqslant$	u_{j-1}	$\leqslant$	b	$=$	u_j	$=$	$\cdots$	$=$	u_n

Kolm (1997, page 250) claims that the following equivalence holds:

(7) $\qquad\qquad\qquad$ For all $\mathbf{u}, \mathbf{v} \in \mathscr{U}$; $\mathbf{u} \geq_{\text{ICL}} \mathbf{v}$ if and only if $\mathbf{u} \geq_{\text{BTR}} \mathbf{v}$.

Because it is pretty obvious that $\mathbf{u} \geq_{\text{BTR}} \mathbf{v}$ implies that $\mathbf{u} \geq_{\text{ICL}} \mathbf{v}$, we confine ourselves with the proof of the converse implication. When $n = 2$, $\mathbf{u} \geq_{\text{ICL}} \mathbf{v}$ reduces to $[u_1, u_2] \subseteq [v_1, v_2]$. Therefore, suppose first that $n = 3$ and that $\mathbf{u} \geq_{\text{ICL}} \mathbf{v}$. This amounts to requiring that:

(8a) $\qquad\qquad\qquad\qquad \{1,2\}: v_1 \leqslant u_1 \leqslant u_2 \leqslant v_2;$

(8b) $\qquad\qquad\qquad\qquad \{1,3\}: v_1 \leqslant u_1 \leqslant u_3 \leqslant v_3; \text{ and}$

(8c) $\qquad\qquad\qquad\qquad \{2,3\}: v_2 \leqslant u_2 \leqslant u_3 \leqslant v_3;$

from which we deduce that

(9) $\qquad\qquad\qquad\qquad v_1 \leqslant u_1 \leqslant v_2 = u_2 \leqslant u_3 \leqslant v_3.$

Assume next that $n = 4$ and that $\mathbf{u} \geq_{\text{ICL}} \mathbf{v}$, in which case we must have:

(10a) $\qquad\qquad\qquad\qquad \{1,2\}: v_1 \leqslant u_1 \leqslant u_2 \leqslant v_2;$

(10b) $\qquad\qquad\qquad\qquad \{1,3\}: v_1 \leqslant u_1 \leqslant u_3 \leqslant v_3;$

(10c) $\qquad\qquad\qquad\qquad \{1,4\}: v_1 \leqslant u_1 \leqslant u_4 \leqslant v_4;$

(10d) $\qquad\qquad\qquad\qquad \{2,3\}: v_2 \leqslant u_2 \leqslant u_3 \leqslant v_3;$

(10e) $\qquad\qquad\qquad\qquad \{2,4\}: v_2 \leqslant u_2 \leqslant u_4 \leqslant v_4; \text{ and}$

(10f) $\qquad\qquad\qquad\qquad \{3,4\}: v_3 \leqslant u_3 \leqslant u_4 \leqslant v_4;$

and we conclude that

(11) $\qquad\qquad\qquad\qquad v_1 \leqslant u_1 \leqslant u_2 = v_2 \leqslant v_3 = u_3 \leqslant u_4 \leqslant v_4.$

Repeating the same argument, for arbitrary $n \geqslant 2$, we finally arrive at

(12) $\qquad\qquad\qquad v_1 \leqslant u_1 = a \leqslant u_2 = v_2 \leqslant \cdots \leqslant v_{n-1} = u_{n-1} \leqslant b = u_n \leqslant v_n,$

which implies (6) with $i = 1$ and $j = n$. Admittedly, in what precedes we have considered the weak version of the unanimous inclusion quasi-ordering. Actually, semi-strict – ad fortiori strict – pairwise inclusion raises difficulties. For instance, a look at (8) should convince the reader that the only possibilities for strict inequalities are between v_1 and u_1 on the one hand, and u_3 and v_3, on the other hand. One might be tempted to apply the same technique to Hammond progressive transfers and say that distribution **u** is *unanimously Hammond more equal than distribution* **v** if $v_i < u_i \leqslant u_j < v_j$, for all $i \neq j$. However, it follows from the discussion above that such a criterion cannot exist.

While it is sensible to consider that it reduces inequality, it must be emphasised that pairwise inclusion comes into conflict with the notion of efficiency, something that, as far as we are concerned, makes this avenue a *cul-de-sac*: we indeed draw a distinction between increments (decrements) and Hammond progressive transfers. According to pairwise inclusion, all increments are considered inequality reducing transformations as long as there exists some individual whose initial endowment is not smaller than that of the individual who benefited from the increment. The only exception is when the increment affects the best-off individual or when the increment is so large that the beneficiary is made better-off than the initially best-off individual. A similar remark applies *mutatis mutandis* in the case of decrements. Note that, in the approach of Allison and Foster (2004), increments and decrements reduce inequality subject to the proviso that (i) increments take place below the median, (ii) decrements occur above the median, and (iii) neither of them jump over the median. Therefore, the implicit notion of *inequality reducing transformation* considered by these authors is a particular case of the pairwise inclusion of Kolm.

Contrary to the notion of a Hammond progressive transfer, the increment and the decrement need not be matched in a pairwise inclusion. A possibility in order to disentangle inequality reduction and efficiency improvements in the pairwise inclusion approach would be to impose that, for it to be considered an increment, the increase in person i's utility must be such that there exists no person j whose utility is no greater than person i's new utility. Formally, we would say that **u** is obtained from **v** by means of such an increment if $u_i = v_i + \Delta > v_i$ for some i and there exists no j such that $u_i \leqslant v_j$. According to this definition, the only admissible increment is such that $u_i = v_i + \Delta > v_n$, hence $\Delta > v_n - v_i$: all other increases in person i' utility would result in pairwise inclusions.

References

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

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.

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

Chakravarty, S. R. and Zoli, C. (2012). Stochastic dominance relations for integer variables. *Journal of Economic Theory*, **147**, 1331–1341.

Cowell, F. A. and Flachaire, E. (2017). Inequality with ordinal data. *Economica*, **334**, 290–321.

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

Ebert, U. (1988). Measurement of inequality: An attempt at unification and generalization. *Social Choice and Welfare*, **5**, 59–81.

Gravel, N. and Moyes, P. (2012). Ethically robust comparisons of bidimensional distributions with an ordinal attribute. *Journal of Economic Theory*, **147**, 1384–1426.

Gravel, N., Magdalou, B., and Moyes, P. (2018). Ranking distributions of an ordinal variable. Mimeo.

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. (1997). *Justice and Equity*. The MIT Press, Cambridge, Massachusetts.

Kolm, S.-C. (1999). The rational foundations of income inequality measurement. In J. Silber, editor, *Handbook of Income Inequality Measurement*, pages 19–100. Kluwer Academic Publishers, Boston/Dordrecht/London.

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

Moyes, P. (1999). Stochastic dominance and the Lorenz curve. In J. Silber, editor, *Handbook of Income Inequality Measurement*, pages 199–222. Kluwer Academic Publishers, Boston/Dordrecht/London.

Moyes, P. (2012). Comparisons of heterogeneous distributions and dominance criteria. *Journal of Economic Theory*, **147**, 1351–1383.

Yalonetzky, G. (2013). Stochastic dominance with ordinal variables: Conditions and a test. *Econometric Reviews*, **32**, 126–163.

Zheng, B. (2011). A new approach to measure socioeconomic inequality in health. *Journal of Economic Inequality*, **9**, 555–577.