UNIVERSITY OF ABERDEEN

Business School

Department of Economics

Social Welfare Analaysis for Ordered Response Data

R. H. Abul Naga and T. Yalcin

Discussion Paper in Economics No 16-2

February 2016

ISSN 0143-4543

Social Welfare Analysis for Ordered Response Data

Ramses H. Abul Naga * and Tarik Yalcin †, ‡

February 3, 2016

Abstract

A social welfare ordering suited for ordered response data is introduced. Paretian probability transfers (PPTs) and equalizing probability transfers (EPTs) respectively capture society's preference for efficiency and equity. The resulting relation is shown to generate a complete lattice over the set of cumulative distributions. A Gamma curve (akin to the generalized Lorenz curve–enabling the researcher to explore welfare orderings) and a class of social welfare functions are also introduced. The methodology is illustrated using Egyptian data on nutritional status.

Keywords: Social welfare measurement, ordered response health data, Paretian probability transfers, equalizing probability transfers, Gamma curves, complete lattices.

JEL codes: I3, I1.

---

*Business School, Health Economics Research Unit and Centre for European Labour Market Research (University of Aberdeen).

†Institute of Health Economics and Management (Université de Lausanne).

‡We are grateful to Francesco Andreoli, Nadia Ghanmi, Nicolas Gravel, Christophe Muller, Ernesto Savaglio, Joe Swierzbinski and Claudio Zoli for helpful comments. We are also grateful to participants of the 12th Louis-André Gérard-Varet Journées in Public Economics, at the University of Aix-Marseille for discussions.

1. Introduction

Ordered response data abound in socio-economic surveys. They are often the subject of inequality and social welfare comparisons. Examples of such data include self-reported health status, happiness, educational attainment, satisfaction with local government and local public goods and socio-economic status.

Allison and Foster (2004) and Zheng (2011) among others have noted that the classical tools of welfare analysis tailored for income data are not suitable for this specific context of ordered response data because scales used to measure inequality in relation to such data are entirely arbitrary. Thus, following Allison and Foster (2004), a new literature has emerged in relation to the measurement of inequality and polarization specific to the context of ordered response data, and utilizing functions of the cumulative distribution as the argument of the dispersion index (see among others Apouey, 2007; Abul Naga and Yalcin, 2008; Cowell and Flachaire, 2012 and Kobus and Miłoś 2012). The present paper contributes to this literature pertaining to the measurement of *pure inequalities in health* by addressing the question of social welfare measurement in relation to ordered response data. We mention however that there is also a large body of literature pertaining to the measurement of *socio-economic inequalities in health* (see for instance Bommier and Stecklov, 2002; Erreygers, 2009, and Zheng, 2011) for which the concepts developed in this paper may also be of relevance.

Specifically, the central question of the present paper is the following: how do we order two distributions X and Y in terms of social welfare when the underlying data are ordered response data? A social welfare function is generally taken to be Paretian and egalitarian. Thus in order to elaborate a successful methodology for social welfare analysis, we need to formulate appealing analogues of the Paretian and egalitarian properties of the social welfare function in the context of ordered response data. Furthermore, for such a methodology to be attractive to the data analyst, we need to develop an analogue for the generalized Lorenz curve as well as a family of social welfare indices in order to investigate welfare orderings in the data and possibly to quantify differences in social welfare attainment in alternative distributions.

There are several transformations of the cumulative distribution function that we take to increase welfare. We only consider transformations of the cumulative distribution as the latter is invariant to all cardinalizations of the scale that preserve the initial order of the n states.

Consider firstly two states k and l such that $k + 1 \leq l - 1$. We may express several transformations of the cumulative distribution that we take to be welfare improving:

(i) an upward displacement of one person from state k to state $k+1$, while holding other individuals in their initial state,

(ii) a downward move of one person from state l to $l-1$ and a simultaneous move from state k to $k+1$ of another person,

(iii) a downward move of one person from state l to $l-1$ and a simultaneous move of one person or more from state k to $k+1$.

We call the first type of transformation a Paretian probability transfer (PPT), and the second type an equalizing probability transfer (EPT)[1]. The third type of transformation is clearly the sum of a PPT and an EPT, and as a consequence will also entail a welfare improvement. In this paper, it is called a welfare improving probability transfer (WPT).

The resulting welfare ordering allows us to order different hypothetical social scenarios of a given society, where each scenario is defined with respect to a specific cumulative distribution. Because our welfare ordering is an incomplete relation, some scenarios may be ordered and others not. However the social cost (in money terms) of moving from one cumulative distribution Y to a preferred distribution X is not quantifiable in our proposed framework, as the scale used to construct the cumulative distributions is only useful to order the n discrete states, but is otherwise arbitrary.

While our proposed welfare ordering is incomplete, it is shown in the paper that it possesses a fundamental lattice property. That is, when two distributions X and Y cannot be ordered, we can always calculate two hypothetical distributions, $X \vee Y$ and $X \wedge Y$ which can be very informative in the context of applied welfare analysis. Specifically, $X \vee Y$ (read "X join Y") is the distribution with least welfare level that is ordered superior to both X and Y; that is $X \vee Y$ is the least upper bound to the set $\{X, Y\}$ in terms of our welfare ordering. Likewise, $X \wedge Y$ (read "X meet Y") is the greatest lower bound to the set $\{X, Y\}$ in terms of our proposed welfare relation.

This paper also introduces related tools of empirical welfare analysis to be utilized in relation to our proposed social welfare ordering. Specifically,

- We introduce the concept of the Gamma curve, a graphical device analogous to the Generalized Lorenz curves for investigating social welfare orderings

[1] This concept is sometimes also referred to as an *exchange* in related literatures. See Section 2 below for further discussion.

in applied work.

- We characterize the class of order-preserving functions in the context of our proposed social welfare ordering for ordered response data.
- We introduce a two-parameter family of social welfare functions with the purpose of expanding the data-analyst's tool kit in the analysis of ordered response data.
- We illustrate with the help of three simple steps an application of our new methodology in the context of analyzing body mass related health outcomes in a sample of women from the Egyptian Integrated Household Survey.

In Section 2 of the paper we introduce our social welfare ordering and the underlying transformations of the cumulative distribution that are welfare improving. In Section 3 we introduce the Gamma curve and we dwell on the lattice property of our social welfare ordering and its significance for empirical work. In Section 4 we characterize the family of order-preserving functions for our proposed welfare ordering. The results derived there are used to introduce a two-parameter family of social welfare functions that may be utilized in work pertaining to applied welfare analysis. Section 6 contains our illustrative application. There, we study the regional variation of social welfare in the context of Egyptian data on anthropometric health, where the data are grouped into five body mass index (BMI) categories arranged in order of increasing health (type III obese, type II obese, type I obese, overweight and not overweight). Section 7 concludes the paper. An appendix contains proofs of our main results.

2. Social welfare ordering

Our starting point is to consider a situation whereby the economic status of a person is measured according to an ordered scale $c \equiv (c_1, ..., c_n)$. We denote $C \equiv \{c : 0 < c_1 < c_2 < ... < c_n < \infty\}$ the set of ordered increasing scales. Because the scale is entirely arbitrary, calculations of summary statistics (mean and other moments) will not form the basis of our measurement of social welfare. Instead, following earlier work in this literature, started with Allison and Foster (2004), the proportions underlying each outcome will be the key inputs to our measurement of welfare indices.

Let $\mathbb{D}$ denote the set of cumulative distributions defined over n ordered states, and let $X = [X_1, ..., X_{n-1}, 1]$ be a cumulative distribution. We define the real-valued function $\sigma : \mathbb{D} \to [1, n]$

$$\sigma(X) = \sum_{i=1}^{n} X_i \tag{2.1}$$

We also define $int(\sigma)$ and $frac(\sigma)$ respectively as the integer and decimal parts of σ.

Example 1 Suppose $n = 5$, and there results a frequency distribution $x = [0.20, 0.04, 0.02, 0.73, 0.01]$. Then, $X = [0.20, 0.24, 0.26, 0.99, 1.00]$ and $\sigma(X) = 2.69$. Furthermore, $int(\sigma) = 2$ and $frac(\sigma) = 0.69$.

Social welfare will rise if other things equal, one person's socio-economic status improves. This is the Paretian property of the social welfare function. The extent of social welfare improvement is characterized by the magnitude of *Paretian probability transfers*, defined below:

Definition 2.1 (Paretian probability transfers).

Let $X, Y \in \mathbb{D}$. We say that Y is obtained from X via a single Paretian probability transfer, written as $X \lhd_{PPT} Y$, if and only if for some state i , and for $0 \leq \varepsilon \leq X_i - X_{i-1}$ and for all states $j \neq i$ we have

$$\begin{aligned} Y_i &= X_i - \varepsilon \\ Y_j &= X_j \quad j \neq i \end{aligned}$$

We will say that Y has higher attainment than X, if and only if Y is obtained from X via a finite sequence of Paretian probability transfers. We will write this as $X \prec_{PPT} Y$.

Observe that $\sigma(.)$ finds an intuitive interpretation in the light of the concept of PPT:

- $PPTs$ reduce σ, and therefore we can consider $\sigma(.)$ to be a summary statistic on the level of economic attainment.
- $\sigma(X) - 1$ is the average number of transitions (to higher states) required in order to reach the maximum welfare level in the population, a distribution $\top = [0, ..., 0, 1]$.

Furthermore,

- Define a family of transformations on the set of scales $G = \{g : C \longrightarrow C\}$. Then $Y(g(c)) = Y(c)$ and as a consequence $\sigma(Y; c) = \sigma(Y; g(c))$ for all function $g \in G$ and for all $Y \in \mathbb{D}$.

- $\sigma(X)$ has a useful invariance property to some transformations of X that we shall coin *equalizing probability transfers*.

Now consider the egalitarian property of the social welfare function. The concept of median preserving spreads introduced by Allison and Foster (2004) features prominently in the literature on inequality measurement for ordered response data. We may define the inequality relation AF on $\mathbb{D}$ as follows: X is more egalitarian than Y, written $(X, Y) \in AF$ where

$$AF = \left\{ \begin{array}{c} (X, Y) \in \mathbb{D}^2 : med(X) = med(Y) = m, \\ X_i \leq Y_i \quad i = 1, \cdots, m-1 \\ X_i \geq Y_i \quad i = m, \cdots, n \end{array} \right\}$$

It follows that a social welfare concept based on $PPTs$ and median preserving spreads, entails a relation

$$AF \cap PPT = \left\{ \begin{array}{c} (X, Y) \in \mathbb{D}^2 : X_i \leq Y_i \quad i = 1, \cdots, m-1 \\ X_i = Y_i \quad i = m, \cdots, n \end{array} \right\}$$

such that X has higher welfare than Y if and only if X is obtained from Y via a sequence of $PPTs$ below the median state m, but X must otherwise be identical to Y. This approach would appear to be unnecessarily restrictive for our purposes. Therefore, below we choose to conceptualize changes in inequality using a concept of *equalizing probability transfers*.

Consider two distributions X and Y that have the same level of economic attainment, i.e. such that $\sigma(X) = \sigma(Y)$. Clearly, these two distributions can attain different levels of social welfare, depending on their underlying levels of dispersion. For the purpose of increasing social welfare while maintaining σ at a constant level, we now define the following concept of equity increasing transfers:

Definition 2.2 (Equalizing probability transfers).

Let $X, Y \in \mathbb{D}$. We say that Y is obtained from X via an equalizing probability transfer, written as $Y \lhd_{EPT} X$, if and only if for $k < l$, for $0 \leq \delta \leq \min\{X_k - X_{k-1}, X_{l+1} - X_l\}$ and for all $j \neq k, l$ we have

$$\begin{aligned} Y_k &= X_k - \delta \\ Y_l &= X_l + \delta \\ Y_j &= X_j \end{aligned}$$

We will say that Y is more egalitarian than X if and only if Y is obtained from X via a finite sequence of equalizing probability transfers, and we will write $Y \prec_{EPT} X$.

This concept of transfers features prominently in the literature on the measurement of health polarization (Apouey, 2007). In terms of proportions, one single EPT will involve transferring probability mass δ from state k to adjacent state $k+1$ and simultaneously transferring δ from state $l+1$ to adjacent state l. This concept of transfers is also clearly related to transfer concepts found in the theory of integer majorization (Folkman and Fulkerson 1969, Chakravarty and Zoli, 2012). We shall return to this point further below after we define our welfare ordering.

Our first result identifies bounds on the extent of redistribution that is feasible given a particular level $t = \sigma(X) = \sigma(Y)$ of economic attainment.

Proposition 2.3

Let $t \in [1, n]$ and let $X \in \mathbb{D}$, *with* $\sigma(X) = t$. *Then there are two distributions* $\hat{\Pi}(t)$ *and* $\tilde{\Pi}(t)$ *such that*

$$\hat{\Pi}(t) \prec_{EPT} X \prec_{EPT} \tilde{\Pi}(t)$$

where

$$\hat{\Pi}(t) = \begin{cases} 0 & i < n - \mathrm{int}(t) \\ \mathrm{frac}(t) & i = n - \mathrm{int}(t) \\ 1 & i > n - \mathrm{int}(t) \end{cases}$$

and

$$\tilde{\Pi}(t) = \left(\frac{t-1}{n-1}, \cdots, \frac{t-1}{n-1}, 1 \right)$$

Furthermore, if $\sigma(U) = t$ *and* $U \prec_{EPT} \hat{\Pi}(t)$, *then* $U = \hat{\Pi}(t)$. *If* $\sigma(V) = t$ *and* $\tilde{\Pi}(t) \prec_{EPT} V$, *then* $\tilde{\Pi}(t) = V$.

In Example 1 for instance, the most and least egalitarian distributions that prevail given $\sigma(X) = 2.69$ are respectively $\hat{\Pi}(2.69) = [0, 0, 0.69, 1, 1]$ and $\tilde{\Pi}(2.69) = [0.4225, 0.4225, 0.4225, 0.4225, 1]$. Clearly, for a given level of economic attainment $t = 2.69$, the set of frequency distributions cannot be more compressed than $[0, 0, 0.69, 0.31, 0]$ and cannot be more spread out than $[0.4225, 0, 0, 0, 0.5775]$.

Thus $\hat{\Pi}(t)$ and $\tilde{\Pi}(t)$ will respectively yield the highest and lowest levels of social welfare for all distributions X such that $\sigma(X) = t$. Further improvements of welfare must result from Paretian probability transfers, and hence from different (i.e lower) values of σ. The essence of welfare improving probability transfers is to capture both types of probability shifts:

Definition 2.4 (welfare improving probability transfers).

Let $X, Y \in \mathbb{D}$. We say that Y is obtained from X via a single welfare improving probability transfer, written as $X \lhd_W Y$, if and only if for $k < l$, for $0 \leq \delta \leq \min\{X_k - X_{k-1}, X_{l+1} - X_l\}$, for $\varepsilon \geq 0$ and for $0 \leq \delta + \varepsilon \leq X_k - X_{k-1}$ we have

$$\begin{aligned} Y_k &= X_k - \delta - \varepsilon \\ Y_l &= X_l + \delta \\ Y_j &= X_j \quad j \neq k, l \end{aligned}$$

We will say that Y welfare dominates X, written $X \prec_W Y$ if and only if Y is obtained from X via a finite sequence of welfare improving probability transfers.

To clarify the relation between our inequality and welfare orderings and related orderings on integers, it useful to turn to a simple example. Consider then the case $n = 3$, and some scale $c = (2, 4, 6)$. We have a hypothetical two person economy, where we compare two distributions of a given cake of fixed size. As the cumulative distribution is a function of c, we may write a data vector as $d(Y(c))$, or more simply as $d(Y; c)$, where it is understood that d depends on c via the dependence of Y on c.

Let then $d(Y, c) = [c_3,\ c_1] = [6, 2]$ and let $d(X; c) = [c_2,\ c_2] = [4, 4]$. The underlying cumulative distributions are $Y(c) = [0.5, 0.5, 1]$ and $X(c) = [0, 1, 1]$. We let $s[d(Y; c)]$ denote the size of the cake and we observe that $s[d(Y; c)] = s[d(X; c)] = 8$. Likewise, we have $\sigma(X; c) = 2 = \sigma(Y; c)$. In the theory of integer majorization (Marshall et al. 2011, Chapter 5, Section D) it is possible to define an order over integers, a relation $\prec_I$ whereby $d_1 \prec_I d_2$ if $s(d_1) = s(d_2)$, and d_1 is obtained from d_2 via several progressive transfers of a fixed integer number. It is clear here that $d(X; c) \prec_I d(Y; c)$ and likewise that $X(c) \prec_{EPT} Y(c)$. It is also true that any linear transformation of c into another integer scale $\chi = (ac_1 + b,\ ac_2 + b,\ ac_3 + b)$ will entail that $s[d(Y; \chi)] = s[d(X; \chi)]$ and also will preserve integer majorization: $d(X; \chi) \prec_I d(Y; \chi)$.

Now consider a transformation $g \in G$ of c into say $\zeta = g(c) = [2, 3, 6]$. Then this cardinalization entails $d(X; \zeta)$ and $d(Y; \zeta)$ are cakes of different sizes: $s[d(Y; \zeta)] = 8$ while $s[d(X; \zeta)] = 6$ and thus the two vectors are no longer comparable in terms of integer majorization. On the other hand, as discussed above, the cumulative distribution is invariant to all scale transformations $g : C \longrightarrow C$, so that $X(c) = X(\zeta)$, and accordingly $\sigma(X; c) = \sigma(X; \zeta) = \sigma(Y; \zeta) = \sigma(Y; c)$ and the $\prec_{EPT}$ relation is preserved for all scale transformations $g \in G$. To sum up, our example illustrates that while integer majorization entails an ordering of the cumulative distributions in terms of $EPTs$, the latter $\prec_{EPT}$ relation is a more general concept in that the ordering of two cumulative distributions in terms of $EPTs$ can occur while the ordering of the underlying data in terms of integer majorization is not possible.

Likewise, the same arguments can be made in relation to social welfare orderings (Shorrocks, 1983) and their extension to integers (Chakravarty and Zoli, 2012). We cumulate the incomes vectors $d(Y;c)$ and $d(X,c)$ from the bottom up to obtain $4 > 2$ and $4+4 \geq 2+6$, to conclude that the generalized Lorenz curve of $d(X,c)$ lies above that of $d(Y,c)$ and we write $d(Y,c) \prec_u d(X,c)$. However, the generalized Lorenz curves of the two distributions $d(X;\zeta)$, $d(Y;\zeta)$ intersect under the scale $\zeta = [2,3,6]$, and the distributions are no longer comparable by the relation $\prec_u$. Nonetheless, it remains the case that $Y(c) \prec_W X(c)$ and $Y(\zeta) \prec_W X(\zeta)$.

We note at this stage that all three relations we have introduced over the set of probability distributions, namely $\prec_{PPT}$, $\prec_{EPT}$ and $\prec_W$ do not necessarily allow us to order any given pair of distributions X and Y. We state this result more formally as follows:

Proposition 2.5

Each of the relations $\prec_{PPT}$, $\prec_{EPT}$ and $\prec_W$ is a partial ordering over the set $\mathbb{D}$ of cumulative distributions.

Our immediate task therefore will be to provide a simple numerical representation of our concept of welfare improving probability transfers. This is the purpose of the next section of the paper.

3. The Gamma curve

The Generalized Lorenz curve is used as a graphical device to investigate welfare orderings between income distributions. By analogy, the Gamma curve is introduced here to investigate welfare orderings for ordered response data (Gamma for Generalized Lorenz).

Let $X, Y \in \mathbb{D}$, with $X \lhd_W Y$. Define the set $H \subseteq \mathbb{R}^n_+$ given by

$$H = \left\{ \begin{array}{cc} \eta_1, ..., \eta_n \ : & \\ 0 \leq \eta_1 \leq 1, & \\ \eta_j - \eta_{j-1} \leq \eta_{j+1} - \eta_j & j = 2, ..., n \end{array} \right\} \tag{3.1}$$

and the n-dimensional vector function $\Gamma : \mathbb{D} \to H$ defined by

$$\Gamma(X) = \left(\begin{array}{ccc} \Gamma_1(X) & ... & \Gamma_n(X) \end{array} \right) \tag{3.2}$$

where $\Gamma_j(X) : \mathbb{D} \to \eta_j$ is defined by

$$\Gamma_j(X) = \sum_{i=1}^{j} X_i \tag{3.3}$$

Since $X \lhd_W Y$ (Y differs from X by a single welfare improving probability transfer), for $i = 1, ..., k-1$ we have $\Gamma_i(Y) = \Gamma_i(X)$. However, as $Y_k = X_k - \delta - \varepsilon$ and $Y_l = X_l + \delta$ we also have for $k \leq j \leq l-1$

$$\Gamma_j(Y) = \Gamma_j(X) - \delta - \varepsilon$$

and for $i = l, ..., n$ we have $\Gamma_i(Y) = \Gamma_i(X) - \varepsilon$. If $X \prec_W Y$, it follows that a finite number of welfare improving probability transfers have been undertaken on X to obtain Y. This is our first result:

Proposition 3.1
Let $X, Y \in \mathbb{D}$. Then $\Gamma(Y) \leq \Gamma(X)$ if and only if $X \prec_W Y$.

If X differs from Y only by a finite number of equalizing probability transfers (δ transfers), then welfare is higher under Y only because Y is more egalitarian than X. Then we obtain the following corollary to Proposition 3.1:

Corollary 3.2
Let $X, Y \in \mathbb{D}$. Then $\Gamma(Y) \leq \Gamma(X)$ with $\Gamma_n(Y) = \Gamma_n(X)$ if and only if $Y \prec_{EPT} X$.

Note of course that not every pair of distributions X and Y may be ordered by the relation $\prec_W$. In this case, we write $X \ ||_W Y$. Specifically, since $\prec_W$ is a partial ordering, $X \ ||_W Y$ will result in intersecting Gamma curves (other than as prescribed by Corollary 3.2). In practice, this means that two social welfare functions that are increasing in welfare improving probability transfers may order X and Y differently (more on this in our applications section 6).

Nonetheless, when $X \ ||_W Y$ it is possible to construct a distribution L such that the Gamma curve of L lies above both that of $\Gamma(X)$ and $\Gamma(Y)$, and conversely there is a distribution U that has a Gamma curve that lies below the Gamma curves of X and Y. These distributions bound the distributions X and Y in a meaningful sense that we shall clarify.

To clarify the status of L and U, we introduce for any $X, Y \in \mathbb{D}$ the concept of a set of lower bounds $\{X, Y\}^{lo}$ and a set of upper bounds $\{X, Y\}^{up}$. Specifically, $\{X, Y\}^{lo}$ is the set of distributions that X and Y dominate in terms of social welfare, and conversely $\{X, Y\}^{up}$ is the set of distributions that are welfare improving over X and Y. Formally, we have:

$$\{X, Y\}^{up} = \{Z \in \mathbb{D} : X \prec_W Z \text{ and } Y \prec_W Z\} \tag{3.4}$$
$$\{X, Y\}^{lo} = \{Z \in \mathbb{D} : Z \prec_W X \text{ and } Z \prec_W Y\} \tag{3.5}$$

L and U bound X and Y in the sense that L is an element of $\{X, Y\}^{lo}$ and U is an element of $\{X, Y\}^{up}$.

Lemma 3.3
Let $X, Y \in \mathbb{D}$. Then there exist two distributions U and L such that

$$L \prec_W X \prec_W U$$
$$L \prec_W Y \prec_W U$$

where $L = \Gamma^{-1}(\max\{\Gamma(X), \Gamma(Y)\})$ and $U = \Gamma^{-1}(\max\{\Gamma(Z) : Z \in \{X, Y\}^{up}\})$.

Furthermore, define $X \wedge Y$ as the greatest element of the set $\{X, Y\}^{lo}$ and $X \vee Y$ as the smallest element of the set $\{X, Y\}^{up}$, where "smallest" and "greatest" are taken in the sense of the ordering relation $\prec_W$. In other terms for any $Z \in \{X, Y\}^{lo}$ we have $Z \prec_W (X \wedge Y)$ and likewise for any $Z \in \{X, Y\}^{up}$ we have $(X \vee Y) \prec_W Z$. The next result clarifies that the existence of $L = (X \wedge Y)$ and $U = (X \vee Y)$ for any X, Y insures that $\prec_W$ endows $\mathbb{D}$ with a particular structure known as a lattice ordering.

Proposition 3.4
The ordered set $(\mathbb{D}, \prec_W)$ is a lattice, with operations $\vee$ and $\wedge$ defined by

$$X \wedge Y = \Gamma^{-1}(\max\{\Gamma(X), \Gamma(Y)\})$$
$$X \vee Y = \Gamma^{-1}(\max\{\Gamma(Z) : Z \in \{X, Y\}^{u}\})$$

for all $X, Y \in \mathbb{D}$.

Thus, in applied work, when $X \,||_W Y$, that is when the Gamma curves of the two distributions intersect, we can construct artificial distributions $L = X \wedge Y$ and $U = X \vee Y$ which provide the tightest bounds in terms of the welfare ordering relation. To clarify these concepts we now turn to an example.

Example 2 Suppose $n = 5$, $X = [0.20\ 0.24\ 0.26\ 0.99\ 1.00]$ and $Y = [0.25\ 0.25\ 0.26\ 0.84\ 1.00]$ .Then $X \,||_W Y$. Furthermore

$$\Gamma(X) = \begin{pmatrix} 0.20 & 0.44 & 0.70 & 1.69 & 2.69 \end{pmatrix}$$
$$\Gamma(Y) = \begin{pmatrix} 0.25 & 0.50 & 0.76 & 1.60 & 2.60 \end{pmatrix}$$

so that the two Gamma curves intersect. The Gamma curves of the distributions $L = X \wedge Y$ and $U = X \vee Y$ are easily derived:

$$\begin{aligned}\Gamma(L) &= \begin{pmatrix} 0.25 & 0.50 & 0.76 & 1.69 & 2.69 \end{pmatrix} \\ \Gamma(U) &= \begin{pmatrix} 0.20 & 0.44 & 0.70 & 1.60 & 2.60 \end{pmatrix}\end{aligned}$$

so that $L = [0.25\ 0.25\ 0.26\ 0.93\ 1.00]$ and $U = [0.20\ 0.24\ 0.26\ 0.90\ 1.00]$.

As it turns out, the ordering relation $(\mathbb{D},\prec_W)$ allows us to define the analogues of L and U above, not only for a given pair of distributions X and Y, but more generally we can bound any subset S of $\mathbb{D}$. A lattice with the property that any subset $S \subseteq \mathbb{D}$ has a least upper bound and a greatest lower bound is known as a complete lattice.

Proposition 3.5

The ordered set $(\mathbb{D},\prec_W)$ *is a complete lattice with bottom element* $\perp = (1, ..., 1)$ *and top element* $\top = (0, ..., 0, 1)$

In particular, it is possible to normalize any social welfare function $W : \mathbb{D} \to \mathbb{R}$ from below and above by respectively $W(\perp)$ and $W(\top)$. The axiomatization of the social welfare function is the topic of our next section.

4. Order-preserving social welfare functions

In empirical work, after inspecting Gamma curves, the researcher may want to summarize the level of welfare underlying each distribution using some function $\omega : \mathbb{D} \longrightarrow \mathbb{R}$. In what follows, it is useful to treat $\mathbb{D}$ as a subset of $\mathbb{R}^n$ and to define the interior of $\mathbb{D}$ as the following set:

$$int\mathbb{D} = \{X \in \mathbb{R}^n : 0 < X_1 < \cdots < X_{n-1} < 1\}. \tag{4.1}$$

There are several key properties the social welfare function ω may be required to satisfy:

- CON : $\omega(X)$ is continuously differentiable on the interior of $\mathbb{D}$ (continuous differentiability).
- $PAR : Y \prec_{PPT} X \Rightarrow \omega(X) \geq \omega(Y)$ for all $X, Y \in \mathbb{D}$ (Paretian property).
- $EQUAL : X \prec_{EPT} Y \Rightarrow \omega(X) \geq \omega(Y)$ for all $X, Y \in \mathbb{D}$ (preference for equality).

- $NORM$: $\omega(X) \geq 0$ and $\omega(X) = 0$ if and only if $X = \perp$ for all $X \in \mathbb{D}$ (normalization).

An important class of order preserving functions in this area of research is the class of Schur-convex functions. In the context of the set of cumulative distributions $\mathbb{D}$, this class may be defined in the following manner:

Definition 4.1 *Let $\omega : \mathbb{D} \longrightarrow \mathbb{R}$ be a real valued function, continuously differentiable on the interior of $\mathbb{D}$. Then ω is Schur-convex if and only if its partial derivatives satisfy the condition*

$$\omega_1(Z) \leq \omega_2(Z) \leq ... \leq \omega_{n-1}(Z)$$

for all $Z \in int\mathbb{D}$. Furthermore, ω is decreasing and Schur-convex if and only if its partial derivatives satisfy the condition

$$\omega_1(Z) \leq \omega_2(Z) \leq ... \leq \omega_{n-1}(Z) \leq 0$$

for all $Z \in int\mathbb{D}$.

Our first result states conditions on the social welfare function such that it satisfies the axioms CON, PAR *and* $EQUAL$.

Proposition 4.2 *Let $\omega : \mathbb{D} \longrightarrow \mathbb{R}$ be some function used to measure social welfare. Then ω satisfies the axioms CON, PAR and $EQUAL$ if and only if ω is continuously differentiable on the interior of $\mathbb{D}$, decreasing and Schur-convex.*

It follows from our discussion in Section 3 above that for any $X \in \mathbb{D}$, $\perp \prec_W X$ and $X \prec_W \top$. Accordingly, the inequality $\omega(\perp) \leq \omega(X) \leq \omega(\top)$ holds for any X and we may normalize a social welfare function to take a zero value at the worst case distribution. Thus, we have,

Corollary 4.3 *Let $W : \mathbb{D} \longrightarrow \mathbb{R}$ be some function used to measure social welfare. Then W satisfies the axioms CON, PAR, $EQUAL$ and $NORM$ if and only if*

$$W(X) = \omega(X) - \omega(\perp)$$

where ω is continuously differentiable on the interior of $\mathbb{D}$, decreasing and Schur-convex.

We conclude this section with the following theorem that summarizes the various results of Section 3 and the present section.

Theorem 4.4 *Let $X, Y \in \mathbb{D}$. The following statements are equivalent:*
(i) $\Gamma(Y) \leq \Gamma(X)$
(ii) $X \prec_W Y$,
(iii) $W(X) \leq W(Y)$ *for all decreasing and schur-convex functions* $W : \mathbb{D} \to \mathbb{R}$.

The result is also of independent interest, as it shows that in theory, the function W need not be continuously differentiable. However, in practice it is easiest to verify Schur convexity when the function is differentiable. The family of social welfare functions we introduce in the section below is precisely constructed to verify the properties listed in Corollary 4.3, differentiability being one such important property.

5. A Family of social welfare functions

We may summarize the discussion around Section 2 as follows: (1) For any $X \in \mathbb{D}$, the function $\sigma(X)$ is decreasing in $PPTs$, and (2) for any two distributions X and Y such that $\sigma(X) = \sigma(Y)$, X generates a higher social welfare level than Y if it is less dispersed (i.e. if X is obtained from Y via a sequence of $EPTs$). In the light of the results of Proposition 4.2 and its corollary 4.3, we may now define the following class of social welfare functions to be used in applied work:

Proposition 5.1
The family of social welfare functions $W_{\alpha,\beta} : \mathbb{D} \longrightarrow \mathbb{R}$ such that

$$W_{\alpha,\beta}(X) = \frac{n - \sum_{i=1}^{n} {X_i}^{\alpha}}{(\sigma(X))^{\beta}} \qquad 0 < \alpha \leq 1, \beta \geq 0, \tag{5.1}$$

satisfies the axioms CON, PAR, $EQUAL$ and $NORM$, where n is the number of ordered states.

In simple terms, choosing larger values of α makes the index less sensitive to the presence of probability mass in the bottom of the distribution. Higher values of β make the index more sensitive to Paretian Probability transfers. To ground our understanding, it is instructive to examine some specific members of this family.

Consider first the index that results from setting $\alpha = 1$:

$$W_{1,\beta}(Z) = \frac{n - \sigma(Z)}{(\sigma(Z))^{\beta}} \qquad \beta \geq 0 \tag{5.2}$$

The above index is entirely summarized by the function σ. Since σ is a sufficient statistic for the level of welfare in this particular case, $W_{1,\beta}$ only weakly satisfies $EQUAL$ as the index is neutral with respect to equalizing probability transfers. Thus, when $\alpha = 1$ our judgement on the level of welfare is entirely characterized by the way we judge Paretian probability transfers. In particular, when we set $\alpha = \beta = 1$, we have a simple index taking the form $W_{1,1}(Z) = (n/\sigma(z)) - 1$.

Next consider setting β at its smallest admissible value, 0. The resulting index is of the form

$$W_{\alpha,0}(Z) = n - \sum_{i=1}^{n} X_i^{\alpha} \qquad 0 < \alpha \leq 1 \tag{5.3}$$

Interestingly, the above welfare index satisfies strict versions of PAR and $EQUAL$ as long as we set $0 < \alpha < 1$. This may be deduced from Proposition 4.2 by observing that for $0 < \alpha < 1$ and $\beta = 0$, the function $W_{\alpha,\beta}(Z)$ is decreasing and Schur-convex over $\mathbb{D}$. As stated above, for values of α close to zero, the index is more sensitive to the bottom of the distribution. Conversely, as α increases the index is less sensitive to the presence of probability mass at the bottom of the distribution, and at the limit, we find $W_{1,0}(Z) = n - \sigma(Z)$.

As we allow β to approach infinity the index becomes more sensitive to $PPTs$. At the limit the index takes on one of two values: the value 0 for all $Z \neq \top$, and the value $n - 1$ when $Z = \top$.

We summarize this discussion with the following corollary to Proposition 5.1:

Corollary 5.2

Consider the class of social welfare functions of Proposition 5.1. The following results hold:

(i) *For any* $Z \in \mathbb{D}$, *we have that* $0 \leq W_{\alpha,\beta}(Z) \leq n - 1$, with $W_{\alpha,\beta}(\perp) = 0$ and $W_{\alpha,\beta}(\top) = n - 1$. *Furthermore,*

(ii) *For any* $Z \in \mathbb{D}$ *such that* $Z \neq \top$, *and for all* $0 \leq \alpha \leq 1$, $\quad W_{\alpha,\infty}(Z) = 0$.

(iii) *For any* $Z \in \mathbb{D}$ *such that* $Z \neq \perp$, *and for all* $\beta \geq 0$, $\quad W_{1,\beta}(Z) = \dfrac{n - \sigma(Z)}{\sigma(Z)^{\beta}}$.

6. An illustrative application

The implementation of the methodology set out in the previous section can be broken down into three simple steps. Firstly, It involves cumulating the frequencies

$x_1, ..., x_n$ to construct the cumulative responses $X_1, ..., X_n$. Secondly, the data analyst computes the Gamma curves pertaining to each of the distributions that form the basis of the empirical investigations. The Gamma curve jth ordinate is obtained by cumulating the cumulative responses up to state j; i.e. $\Gamma_j(X) = X_1 + ... + X_j$. This exercise will reveal which distributions can be ordered by the relation $\prec_W$. Thirdly, the researcher is interested in computing social welfare indices incorporating different social value judgements; this can be done here by varying the parameters α and β in the class of social welfare functions $W_{\alpha,\beta}(X_1, ..., X_n)$ introduced in Section 5.

To illustrate the proposed methodology we use data from the 1997 wave of the Egyptian Integrated Household Survey [EIHS] [2]. We have a total of 1066 observations pertaining to the body mass [3] of adult women from four geographic areas of Egypt. The larger share of Egypt's population is located in the thin strip of land bounding the Nile valley south of Cairo (Upper Egypt), in Cairo Suez, Alexandria and the main towns in the Delta valley north of Cairo (Lower Egypt.) For upper Egypt we have two distributions pertaining to individuals located in rural areas (Rural Upper Egypt, $RUPE$) and urban centres (urban Upper Egypt $UUPE$). The data pertaining to Lower Egypt on the other hand are grouped into data pertaining to metropolitan centres, namely Cairo, Alexandria and Suez, (metropolitan Lower Egypt, MLE) while the remaining observations are gathered in a distribution pertaining to non-metropolitan Lower Egypt, $NMLE$.

Step 1: Calculating cumulative distributions

The data were then grouped into five BMI categories arranged in order of increasing health (Type III obese, Type II obese, type I obese, overweight and not overweight). This classification is standard (see WHO, 2010) where the move from one BMI group to the next (lower health here) is associated with increased risk of diabetes, cardiovascular disease and mortality relative to being in the top health category (not overweight.) The cumulative frequencies for each of the four distributions are reported in Table 1.

[2] The 1997 Egypt Integrated Household Survey (EIHS) was undertaken by the International Food Policy Research Institute in collaboration with United States Agency for International Development (USAID), the Ministry of Agriculture and Land Reclamation of the Government of Egypt and the Ministry of Trade and Supply of the Government of Egypt. The EIHS survey was funded under USAID Grant No. 263-G-00-96-00030-00. See International Food Policy Research Institute (1997) for a description of the survey.

[3] Body mass is defined here as weight in kilograms normalized by squared height, measured in squared meters.

An examination of Table 1 clearly confirms in the context of Egypt that obesity-related ill-health is mostly an urban phenomenon. Type III obesity stands at 7.5% in the metropolitan Lower distribution MLE, but at 1.6% for the rural Upper distribution $RUPE$. Specifically, It is important to note that a first order stochastic dominance ordering reveals that the distribution pertaining to rural Upper Egypt, $RUPE$, has the highest social welfare level: each X_i component is smaller than the corresponding Y_i component, where Y is any of the remaining three distributions. The same conclusion, in reverse, applies to the distribution MLE revealing that the metropolitan Lower Egypt has the worst possible state of social welfare. Table 1 also reports cumulative distributions GLB and LUB; these distributions will be defined once we have inspected Gamma curves.

Step 2: Inspecting Gamma curves

An inspection of Figure 1 plotting the Gamma curves pertaining to these four distributions reveals that, in this illustrative example, social welfare is unambiguously lowest in metropolitan Lower Egypt, and highest in rural Upper Egypt. On the other hand, the Gamma curves pertaining to the remaining two distributions intersect: $NMLE||_W UUPE$. Since the ordering relation $\prec_W$ has the lattice property, (Proposition 3.4), it is possible to construct the least upper bounds (the $\vee$ operation) and greatest lower bounds (the $\wedge$ operation) in terms of social welfare to these two unordered distributions. We thus define two additional distributions, namely $GLB = NMLE \wedge UUPE$ and $LUB = NMLE \vee UUPE$. Together with the worst possible state of affairs, the distribution $\perp = [1\ 1\ 1\ 1\ 1]$, and the top distribution $\top = [0\ 0\ 0\ 0\ 1]$ respectively, the lattice structure pertaining to the set $\mathcal{L} = \{\perp, MLE, GLB, NMLE, UUPE, LUB, RUPE, \top\}$ can thus be exhibited using a Hasse diagram (Figure 2). Note that while the distributions $\perp$, and $\top$ bound the set $\{NMLE, UUPE\}$ from below and above and likewise $\{MLE, RUPE\}$ bound this same set, neither of these bounds is tight in terms of our proposed welfare ordering. The lattice property of the $(\mathbb{D}, \prec_W)$ however ensures that the tightest possible bounds on the pair $\{NMLE, UUPE\}$ is provided by GLB and LUB.

Step 3: Social welfare calculations

To obtain an order of magnitude on the underlying level of social welfare these greatest lower bound and least upper bound on the pair $\{NMLE, UUPE\}$ entail, we calculate in Table 2 the underlying level of welfare for each of the distributions in the set $\mathcal{L}$ using the class of social welfare functions $W_{\alpha,\beta}(.)$ introduced in Section 5. Because orderings are only available for a subset of the pair-wise comparisons, it is clearly necessary to compute a range of indices exhibiting sensitivity to different areas of the health distribution. Thus, in Table 2 we compute $W_{0.05,1}(.)$, $W_{0.05,0}(.)$, $W_{0.50,1}(.)$, and $W_{0.50,0}(.)$. There are several points which these calculations serve to illustrate.

- For all values of α and β considered in Proposition 5.1 The $W_{\alpha,\beta}(.)$ family bounds social welfare between 0 for the bottom distribution $\perp$ and $n-1=4$ for the top distribution $\top$.

Though in a sense trivial, the above normalization allows us to clarify how each of the social welfare indices cardinalize well-being into a number ranging between these theoretical bounds of 0 and 4.

- All four social welfare indices mirror the pair-wise ordering of distributions summarized in the Hasse diagram of Figure 2.

For instance, the indices for metropolitan Lower Egypt (MLE) all indicate less social welfare than the corresponding functions for the other three distributions. Conversely the indices for rural upper Egypt ($RUPE$) always indicate a higher level of welfare. Taking for instance the $W_{0.50,0}(.)$ measure (last column of Table 2), we calculate $W_{0.50,0}(\perp) = 0$, $W_{0.50,0}(\top) = 4$, and the remaining numbers are bounded between the values $W_{0.50,0}(MLE) = 1.7368$ and $W_{0.50,0}(RUPE) = 2.6196$. This example highlights a further point.

- Because in practice some distributions may not be ordered, the computation of welfare using a single index can be rather misleading.

To clarify this point, consider a comparison of non-metropolitan Lower Egypt and urban Upper Egypt: social welfare stands at 0.1405 in $NMLE$ and 0.1432 in the $UUPE$ according to the $W_{0.05,1}(.)$ measure, but social welfare is marginally higher in $NMLE$ (0.3108 versus 0.3099) according to the $I_{0.05,0}(.)$ index.

Our last point exploits the lattice property of the social welfare ordering $\prec_W$ to elaborate further on the problem of ordering distributions.

- In case two distributions X and Y cannot be ordered, the computation of social welfare at $W_{\alpha,\beta}(X \wedge Y)$ and $W_{\alpha,\beta}(X \vee Y)$ provides the tightest lower and upper bounds on the underlying level of welfare of X and Y.

Clearly $NMLE$ and $UUPE$ are not comparable. However, they both unambiguously dominate a hypothetical public policy scenario that generates the distribution $GLB = NMLE \wedge UUPE$. Likewise, $NMLE$ and $UUPE$ are both inferior to a hypothetical public policy that generates a distribution $LUB = NMLE \vee UUPE$. Furthermore, GLB and LUB are the tightest possible brackets on the level of welfare a given social welfare function will attribute to the states of affairs summarized by the distributions X and Y. Thus, while $W_{0.05,1}(.)$ and $W_{0.05,0}(.)$ rank $NMLE$ and $UUPE$ differently, we clearly have $W_{0.05,1}(GLB) = 0.1367$, $W_{0.05,1}(LUB) = 0.1470$, and these two figures provide the tightest possible bounds on the level of welfare underlying $NMLE$ and $UUPE$ when social preferences are captured by the function $W_{0.05,1}$. Likewise, $W_{0.05,0}(GLB) = 0.3024$, $W_{0.05,0}(LUB) = 0.3181$ are the tightest possible bounds on the level of welfare underlying $NMLE$ and $UUPE$ in relation to the social welfare function $W_{0.05,0}(.)$.

7. Conclusion

Following Allison and Foster (2004), a new literature has emerged in relation to the measurement of inequality and polarization specific to the context of ordered response data, and utilizing the cumulative distribution function as the argument of the dispersion index. The present paper advances this literature by addressing the question of social welfare measurement in relation to ordered response data.

Two types of transformations of the data were considered as improving society's welfare: Paretian probability transfers were introduced to capture the Paretian property of the social welfare function; equalizing probability transfers capture society's preference for equality. Our resulting social welfare ordering was shown to generate a complete lattice over the set of cumulative distributions functions, enabling the researcher to construct the tightest possible bounds (i.e least upper and greatest lower bounds) on any number of distributions that cannot be ordered.

We have also introduced the Gamma curve as a means of enabling the researcher to investigate social welfare orderings of distributions in the context of applied work. Our paper has also characterized the set of order-preserving functions for our social welfare relation. Specifically, this consists of the family of decreasing Schur convex functions defined over the set of cumulative distribution functions. This result was further exploited by introducing a two-parameter family of social welfare functions with the purpose of expanding the data-analyst's tool kit in the analysis of ordered response data. Finally, we have proposed an illustrative application of our new methodology in the context of analyzing body mass related health outcomes in a sample of women from the Egyptian Integrated Household Survey.

We note here that we have obtained our family of social welfare functions from a partial axiomatization of our welfare measures. Thus, there may be additional axioms researchers may want to consider from the related inequality literature (typically sub-group decomposability; cf. Kobus and Miłoś, 2012) in order to narrow this class further. Furthermore, social welfare indices may be derived in relation to order statistics of the data other than via the cumulative distribution function (Cowell and Flachaire, 2012).

There are also other contexts than developing country data, and other areas of application where our proposed methodology may be equally applied. The measurement of welfare related happiness (Alesina et al. 2004, Stevenson B. and J. Wolfers, 2008), and satisfaction with the delivery of public services (Jones et al. 2012), are clearly research areas where the methodology we have developed here may be of relevance.

8. Appendix

This appendix gathers proofs of our various results. Throughout, we assume i, j, k, l, m and n all denote positive integers, and we begin by stating and proving two preliminary lemmas.

Lemma A1 *Let $\mathbb{D}$ denote the set of cumulative distribution functions pertaining to l observations and n ordered probability states, where l,$n \in \mathbb{N}$ are finite. Then the resulting set of cumulative distributions $\mathbb{D}$ is finite.*

Proof It suffices to note that there are exactly l^n configurations of the responses of the l observations. Clearly, l^n is finite, and is an upper bound to the number of elements in the set $\mathbb{D}$.

It is convenient for the proof of Lemma $A2$ below to treat $\Gamma(X)$ and $\Gamma(Y)$ as two separate functions from $\mathbb{N}_n \equiv \{1, ..., n\}$ to $[1, n]$, and to write these respectively as ΓX and ΓY, where for $V \in \mathbb{D}$, we define $\Gamma V(1) \equiv V_1$ and for $i > 1$, $\Gamma V(i) \equiv V_1 + \cdots + V_i$. With this notation, we define $\Theta(i) = \max\{\Gamma X(i), \Gamma Y(i)\}$ for all $i \in \mathbb{N}_n$. Note furthermore that for any distribution V, ΓV is an increasing function. Also, as $\Gamma V(i+1) - \Gamma V(i) = V_{i+1}$, the function ΓV is increasing at an increasing rate, that is, it is increasing and convex over $\mathbb{N}_n$.

Lemma A2 *Let $X, Y \in \mathbb{D}$. For all $i \in \mathbb{N}_n$ define* $\Theta(i) = \max\{\Gamma X(i), \Gamma Y(i)\}$.

Furthermore, construct the vector Z such that $Z(1) \equiv \Theta(1)$ and $Z(i) \equiv \Theta(i) - \Theta(i-1)$ for all $i = 2, ..., n$. Then $Z \in \mathbb{D}$ for all $X, Y \in \mathbb{D}$.

Proof To show that Z is a distribution, we need to verify three properties, namely: (i) $Z_n = 1$, (ii) $Z_1 \geq 0$, and (iii) $Z_{i+1} \geq Z_i$ for all $i = 1, ..., n-1$.

(i) Let $\Theta(i) = \max\{\Gamma X(i), \Gamma Y(i)\}$ for all $i \in \mathbb{N}_n$. Since $V_n = 1$ for all $V \in \mathbb{D}$, it follows that for all $X, Y \in \mathbb{D}$ there holds $\Theta(n) - \Theta(n-1) = 1$, and thus $Z_n = 1$.

(ii) Since $\Theta(1) = \max\{X_1, Y_1\} \geq 0$, it follows that $Z_1 = \Theta(1) \geq 0$ for all $X, Y \in \mathbb{D}$.

(iii) We have the following equivalences: $Z_{i+1} \geq Z_i \iff \Theta(i+1) - \Theta(i) \geq \Theta(i) - \Theta(i-1) \iff \frac{1}{2}\Theta(i+1) + \frac{1}{2}\Theta(i-1) \geq \Theta(i)$. But the latter inequality is a version of Jensen's inequality: it holds true for all $X, Y \in D$ since ΓX and ΓY are increasing and convex over $\mathbb{N}_n$, and accordingly $\Theta(i) = \max\{\Gamma X(i), \Gamma Y(i)\}$ is convex over $\mathbb{N}_n$. $\square$

Proof of Proposition 2.3 For $X \in \mathbb{D}$, with $\sigma(X) = t$ and $X \neq \tilde{\Pi}$, we must have $X_1 < \frac{t-1}{n-1}$ and $X_{n-1} > \frac{t-1}{n-1}$. In full generality, for all X, there is an index i with $1 \leqslant j < i < k < n-1$ such that $X_j < \frac{t-1}{n-1}$ and $X_k \geqslant \frac{t-1}{n-1}$. Let $\delta_j \equiv \frac{t-1}{n-1} - X_j$ and $\tilde{\Pi}^1 \equiv (\frac{t-1}{n-1} - \delta_1, \frac{t-1}{n-1}, ..., \frac{t-1}{n-1} + \delta_1, 1)$, then there holds $\tilde{\Pi}^1 \lhd_{EPT} \tilde{\Pi}$. With at most $(n-1)$ $EPTs$, $\delta = \displaystyle\sum_{i=1}^{i} \delta_j$ can be easily distributed over the components $\tilde{\Pi}_k$, $k = 1, .., n-1$ of $\tilde{\Pi}$ to get the X_k and X with $\sigma(X) = t$. Then $X \prec_{EPT} \tilde{\Pi}$.

We next show that $\hat{\Pi} \prec_{EPT} X$. For any $X \in \mathbb{D}$, let $\delta_i = X_i$ for $i = 1, .., (n - \text{int}(t) - 1)$. These δ_i can be distributed over the last $\text{int}(t)$ components of X to get the 1, and the $\text{frac}(t)$ goes on $X_{n-int(t)}$. Then, we obtain $\hat{\Pi} \prec_{EPT} X$.

Finally, it is easy to see that from $\hat{\Pi} = (0, 0, .., \text{frac}(t), 1, 1, .., 1)$ it is impossible to make any EPT to obtain any distribution in $\mathbb{D}$. So if $U \prec_{EPT} \hat{\Pi}(t)$, *then* $U = \hat{\Pi}(t)$. Similarly there is no $V \in \mathbb{D}$ with $\sigma(V) = t$ and $\tilde{\Pi} \lhd_{EPT} V$ because $V_1 < \frac{t-1}{n-1}$. So if $\tilde{\Pi}(t) \prec_{EPT} V$ then $\tilde{\Pi}(t) = V$. $\square$

Proof of Proposition 2.5 Each of these binary relations are based on component-wise arithmetic inequalities. They are easily verified to be reflexive, anti-symmetric and transitive ordering relations. $\square$

Proof of Proposition 3.1 $(\Leftarrow)$ If $X \lhd_W Y$ we have for some k, j and l such that $k \leq j \leq l-1$:

$$\begin{aligned}\Gamma_1(Y) &= \Gamma_1(X)\\ &\vdots\\ \Gamma_{k-1}(Y) &= \Gamma_{k-1}(X)\end{aligned}$$

However, as $Y_k = X_k - \delta - \varepsilon$ and $Y_l = X_l + \delta$ we also have for $k \leq j \leq l-1$

$$\Gamma_j(Y) = \Gamma_j(X) - \delta - \varepsilon$$

and

$$\begin{aligned}\Gamma_l(Y) &= \Gamma_l(X) - \varepsilon\\ &\vdots\\ \Gamma_n(Y) &= \Gamma_n(X) - \varepsilon\end{aligned}$$

Using successive $\lhd_W$ relations, we then obtain that $\Gamma(Y) \leq \Gamma(X)$.

($\Rightarrow$) If $\Gamma(Y) \leq \Gamma(X)$, in full generality for all i, there are $\varepsilon_i, \delta \geq 0$ such that $\Gamma_i(Y) = \Gamma_i(X) - \delta - \varepsilon_i$. Using successive $\lhd_W$ relations, we obtain $X \prec_W Y$.
$\square$

Proof of Lemma 3.3 If X and Y are ordered, it is clear that L and U exist: if $X \prec_W Y$ then $L = X$ and $U = Y$, and vice versa if $Y \prec_W X$.

If $X \|_W Y$, then from Lemma A2 it follows that there is a distribution $L \in \mathbb{D}$ defined as $L = \Gamma^{-1}(\max\{\Gamma(X), \Gamma(Y)\})$ such that $L \prec_W X$ and $L \prec_W Y$. It is clear that $L \in \{X, Y\}^{lo}$.

Now consider the set of upper bounds for $\{X, Y\}$. Because $(\mathbb{D}, \prec_W)$ has a top element $\top = (0, ..., 0, 1)$ it follows that every subset S of $\mathbb{D}$ has an upper bound in $\mathbb{D}$, and thus that $\{X, Y\}^{up}$ is non-empty. In particular, since from Lemma A1 $\mathbb{D}$ is a finite set, $\{X, Y\}^{up}$ is also finite, and repeated application of Lemma A2 entails that there is a distribution $U \in \mathbb{D}$, such that $\Gamma(U) = \max(\{\Gamma(Z) : Z \in \{X, Y\}^{up}\})$. It then follows that $X \prec_W U$ and $Y \prec_W U$ as required.
$\square$

Proof of Proposition 3.4 Lemma A2 ensures that for all $X, Y \in \mathbb{D}$ there is a distribution $Z \in \mathbb{D}$ such that $\Gamma(Z) = \max\{\Gamma(X), \Gamma(Y)\}$. Thus it follows that $X \wedge Y = \Gamma^{-1}(\max\{\Gamma(X), \Gamma(Y)\})$.

Since Lemma A1 entails that $\mathbb{D}$ is a finite set, it also follows that for all $S \subseteq \mathbb{D}$, $\wedge S$ exists in $\mathbb{D}$. Furthermore, since $(\mathbb{D}, \prec_W)$ has for top element $\top = [0, ..., 0, 1]$, we have that every subset $S \subseteq \mathbb{D}$ has an upper bound in $\mathbb{D}$ and S^{up} is non-empty. In particular, $\{X,Y\}^{up}$ is non-empty and there exists some $Z \in \mathbb{D}$ such that $Z = \wedge\{X,Y\}^{up} = \Gamma^{-1}(\max\ \{\Gamma(Z) : Z \in \{X,Y\}^{up}\})$. But $\wedge\{X,Y\}^{up} = X \vee Y$ and we have that the partial ordering $(\mathbb{D}, \prec_W)$ is a lattice.

Proof of Proposition 3.5 We observe that Lemma A1 and Proposition 3.4 entail together that $(\mathbb{D}, \prec_W)$ is a lattice defined over a finite set. Hence, it follows that $(\mathbb{D}, \prec_W)$ is a complete lattice. $\square$

Proof of Proposition 4.2 $(\Rightarrow)$ By definition, CON entails differentiability of the function ω on the interior of $\mathbb{D}$. Let $X, Y \in \mathbb{D}$. Since $X \prec_{PPT} Y \iff Y \leq X \Rightarrow \omega(Y) \geq \omega(X)$, CON and PAR jointly entail that the function ω is differentiable on $int\mathbb{D}$ and decreasing.

Consider next the axiom $EQUAL$. By definition, $Y \prec_{EPT} X \iff X \prec_{HLP} Y$, where $\prec_{HLP}$ is the majorization ordering of Hardy Littlewood and Polya (Marshall et al., 2011; Chapter 1). Preference for equity entails $Y \prec_{EPT} X \Longrightarrow \omega(X) \leq \omega(Y)$. Therefore $X \prec_{HLP} Y \Longrightarrow \omega(X) \leq \omega(Y)$. Then ω is a majorization order preserving function; equivalently, ω is Schur-convex. Together the three axioms then entail ω is differentiable on the interior of $\mathbb{D}$, decreasing and Schur-convex.

$(\Leftarrow)$ The converse statement is easily shown to hold true. $\square$

Proof of Corollary 4.3 Let ω satisfies the axioms CON, PAR, and $EQUAL$. Then, because for all $X \in \mathbb{D}$ we have $\perp \prec_W X$, it follows that $W(\perp) = \omega(X) - \omega(\perp) \geq 0$ and $W(\perp) = 0$. Clearly then, W also satisfies $NORM$. $\square$

Proof of Theorem 4.4 That $(i) \Leftrightarrow (ii)$ follows from Proposition 3.1.

$(ii \Rightarrow iii)$ Since from Lemma A1 $\mathbb{D}$ is a finite set and by assumption $X \prec_W Y$, we have $X = U^o \prec U^1 \prec ... \prec U^m = Y$ where $m \in \mathbb{N}$ is finite and where U^{i+1} differs from U^i by one single WPT, and U^{i+1} is the immediate successor of U^i.

Let $X = (X_1, ..., X_{n-1}, 1)$ and $U^1 = (a_1, ..., a_{n-1}, 1)$. Since X and U^1 differ by one single WPT, there exist states k and l with $k+1 \leq l-1$ such that $a_k = X_k - \delta - \varepsilon$, $a_l = X_l + \delta$, where δ and ε satisfy the restrictions of Definition 2.4, and where finally $a_i = X_i$ for all $i \neq k, l$. It then follows that $\sum\limits_{i=1}^{j} X_i \geq \sum\limits_{i=1}^{j} a_i$ for all $j = 1, ..., n$.

These n inequalities in turn entail that X is weakly super-majorized by U^1 in the sense of Definition 1.A.2.a of Marshall et al.(2011). Furthermore, from Theorem 3A8 of Marshall et al.(2011), it follows that $\phi(U^1) \geq \phi(X)$ for all decreasing and Schur-convex functions $\phi : \mathbb{D} \to \mathbb{R}$. We next proceed to establish the same relation between U^i and U^{i+1} and we arrive at the result that $X \prec_W Y \Rightarrow W(X) \leq W(Y)$ for all $W : \mathbb{D} \to \mathbb{R}$ decreasing and Schur-convex.

($iii \Rightarrow ii$) Assume that $\phi(X) \leq \phi(Y)$ for all decreasing and Schur-convex functions $\phi : \mathbb{D} \to \mathbb{R}$. Then we may consider the following n functions

$$\phi^j(U) = -\sum_{i=1}^{j} U_i, \qquad j = 1, ..., n$$

for which by assumption there holds $\phi^j(X) \leq \phi^j(Y)$. But in turn these n inequalities entail $\Gamma(Y) \leq \Gamma(X)$, or equivalently that $X \prec_W Y$. $\square$

Proof of Proposition 5.1 Consider first some $X \in int\mathbb{D}$. Then,

$$\frac{\partial W_{\alpha,\beta}}{\partial X_i} = -\frac{\alpha}{\sigma(X)^\beta} X_i^{\alpha-1} - \frac{\beta}{\sigma(X)^{\beta+1}} \left(n - \sum_{i=1}^{n} X_i^\alpha \right) \tag{8.1}$$

$$\frac{\partial W_{\alpha,\beta}}{\partial X_i} - \frac{\partial W_{\alpha,\beta}}{\partial X_{i+1}} = \frac{\alpha}{\sigma(X)^\beta} \left(X_{i+1}^{\alpha-1} - X_i^{\alpha-1} \right) \tag{8.2}$$

It is clear from (8.1) that CON holds as $W_{\alpha,\beta}$ is continuously differentiable on $int\mathbb{D}$ for all $\alpha \in (0,1]$ and for all $\beta \geq 0$. Furthermore, for any $X \in int\mathbb{D}$, there holds $1 < \sum_{i=1}^{n} X_i^\alpha < n$ and $1 < \sigma(X) < n$. Thus we deduce that PAR is also satisfied for all $\alpha \in (0,1]$ and for all $\beta \geq 0$.

Next consider $EQUAL$. Since the function $g(t) = t^{\alpha-1}$ is decreasing for all $t \in (0,1)$ and for all $\alpha \in (0,1]$, it follows that the right hand side of (8.2) is non-positive and thus that $W_{\alpha,\beta}$ satisfies $EQUAL$ for any $X \in int\mathbb{D}$.

Now consider boundary points of $\mathbb{D}$. It is useful here to look at three separate cases. First consider some point $Z \in \mathbb{D}$, such that $0 < Z_1 \leq \cdots \leq Z_{n-1} \leq 1$. Since $W_{\alpha,\beta}$ remains differentiable at Z, PAR and $EQUAL$ are also satisfied at this point for all $\alpha \in (0,1]$ and for all $\beta \geq 0$.

Next consider some $l \in \{2, ..., n-1\}$, $\alpha \in (0,1]$, $\beta \geq 0$ and a boundary point $V = (0, \cdots, 0, V_l, \cdots, V_{n-1}, 1) \in \mathbb{D}$ such that $V_l > 0$. It is clear that $W_{\alpha,\beta}$ is not differentiable at V. Take some ε such that $0 < \varepsilon < V_l/2$ and construct the following two distributions:

$$\begin{aligned} U &= (0, \cdots, 0, V_l - \varepsilon, V_{l+1}, \cdots, V_{n-1}, 1) \\ Y &= (0, \cdots, 0, \varepsilon, V_l - \varepsilon, V_{l+1}, \cdots, V_{n-1}, 1) \end{aligned}$$

Clearly, $\sigma(U) < \sigma(V)$, and $\sum_{i=1}^{n} U_i^\alpha < \sum_{i=1}^{n} V_i^\alpha$. Thus $W_{\alpha,\beta}$ satisfies PAR at V. Also, it is the case that $\sigma(V) = \sigma(Y)$ while $\sum_{i=1}^{n} Y_i^\alpha > \sum_{i=1}^{n} V_i^\alpha$ as the function $\sum_{i=1}^{n} X_i^\alpha$ is Schur-concave for all $\alpha \leq 1$. Thus $W_{\alpha,\beta}$ also satisfies $EQUAL$ at V.

Now consider the final case, that of the two boundary points $\top$ and $\perp$ of $\mathbb{D}$. Then, for $0 < \varepsilon < 1$ consider $A = (0, \cdots, 0, \varepsilon, 1)$ and $B = (1 - \varepsilon, 1, \cdots, 1)$. It is clear that $\top$ is a PPT of A while B is a PPT of $\perp$. Also, for all $\alpha \in (0, 1]$ and for all $\beta \geq 0$ we have that $W_{\alpha,\beta}(\top) \geq W_{\alpha,\beta}(A)$ while $W_{\alpha,\beta}(\perp) \leq W_{\alpha,\beta}(B)$. Thus $W_{\alpha,\beta}(.)$ satisfies PAR at $\top$ and $\perp$.

Finally, note in the light of Proposition 2.3 that $\sigma(\perp) = 5$, $\sigma(\top) = 1$, and furthermore that $\hat{\Pi}(5) = \perp$ and $\hat{\Pi}(1) = \top$. In fact, $\top$ is the unique distribution X such that $\sigma(X) = 1$, and likewise $\perp$ is the unique distribution X such that $\sigma(X) = 5$. Thus for all $\alpha \in (0, 1]$ and for all $\beta \geq 0$, $EQUAL$ holds trivially at $\top$ and $\perp$. $\square$

9. References

Abul Naga R. and T. Yalcin (2008): "Inequality Measurement for Ordered Response Health Data" *Journal of Health Economics* 27, 1614-1625.

Allison R. A. and Foster J. (2004): "Measuring Health Inequalities Using Qualitative Data", *Journal of Health Economics* 23, 505-524.

Alesina A., R. Di Tella and R. MacCulloch (2004): " Inequality and Happiness: are Europeans and Americans Different?" *Journal of Public Economics* 88, 2009-2042.

Apouey B. (2007): "Measuring Health Polarization with Self-Assessed Health Data", *Health Economics* 16, 875-894.

Atkinson A. (1970): "On the Measurement of Inequality", *Journal of Economic Theory* 2, 244–263.

Bommier A. and G. Stecklov (2002): "Defining Health Inequality: Why Rawls Succeeds where Social Welfare Theory Fails", *Journal of Health Economics* 21, 497-513.

Chakravarty S. and C. Zoli (2012): "Stochastic Dominance Relations for Integer Variables", *Journal of Economic Theory* 147, 1331–1341.

Cowell F. and E. Flachaire (2012): "Inequality with Ordinal Data", Manuscript.

Davey B. and H. Priestley (2002): *Introduction to Lattices and Order*, Cambridge University Press.

Erreygers G. (2009): "Correcting the Concentration Index", *Journal of Health Economics* 28, 504–515.

Folkman J. and D. Fulkerson (1969): " Edge Colorings in Bipartite Graphs", in Bose R. and T. Dowling *eds.* : *Combinatorial Mathematics and its Applications*, Chapel Hill, University of North Carolina Press.

International Food Policy Research Institute (1997): "Egypt Integrated Household Survey (EIHS): 1997 Overview and Description of Version 3 Data Files", Manuscript.

Jones A., N. Rice, S. Robone and P. Rosa Dias (2011): " Inequality and Polarization in Health Systems Responsiveness: a Cross-Country Analysis", *Journal of Health Economics* 30, 616-625.

Kobus M. and P. Miłoś (2012): "Inequality Decomposition for Population Subgroups" *Journal of Health Economics* 31, 15-21.

Marshall A., I. Olkin and B. Arnold (2011): *Inequalities: Theory of Majorization and its Applications*, New York, Elsevier.

Shorrocks A. (1983): "Ranking Income Distributions", *Economica* 50, 3-17.

Stevenson B. and J. Wolfers (2008): "Happiness Inequality in the United States", IZA discussion paper 3624.

World Health Organization (2010): *Global Database on Body Mass Index*, http://apps.who.int/bmi/index.jsp?introPage=intro_3.html.

Zheng B. (2011): "A New Approach to Measure Socioeconomic Inequality in Health", *Journal of Economic Inequality* 9, 555-577.

Table 1: Cumulative distribution of anthropometric health in various statistical areas of Egypt

Area	Type III obese	Type II obese	Type I obese	Overweight	Not overweight
MLE	0.0748	0.1869	0.4299	0.8131	1
NMLE	0.0398	0.1438	0.3628	0.6659	1
UUPE	0.0492	0.1475	0.2951	0.6721	1
RUPE	0.0157	0.0418	0.154	0.4334	1
GLB	0.0492	0.1475	0.3498	0.6659	1
LUB	0.0398	0.1438	0.3082	0.6721	1

Figure 1: Gamma curves in four statistical areas of Egypt

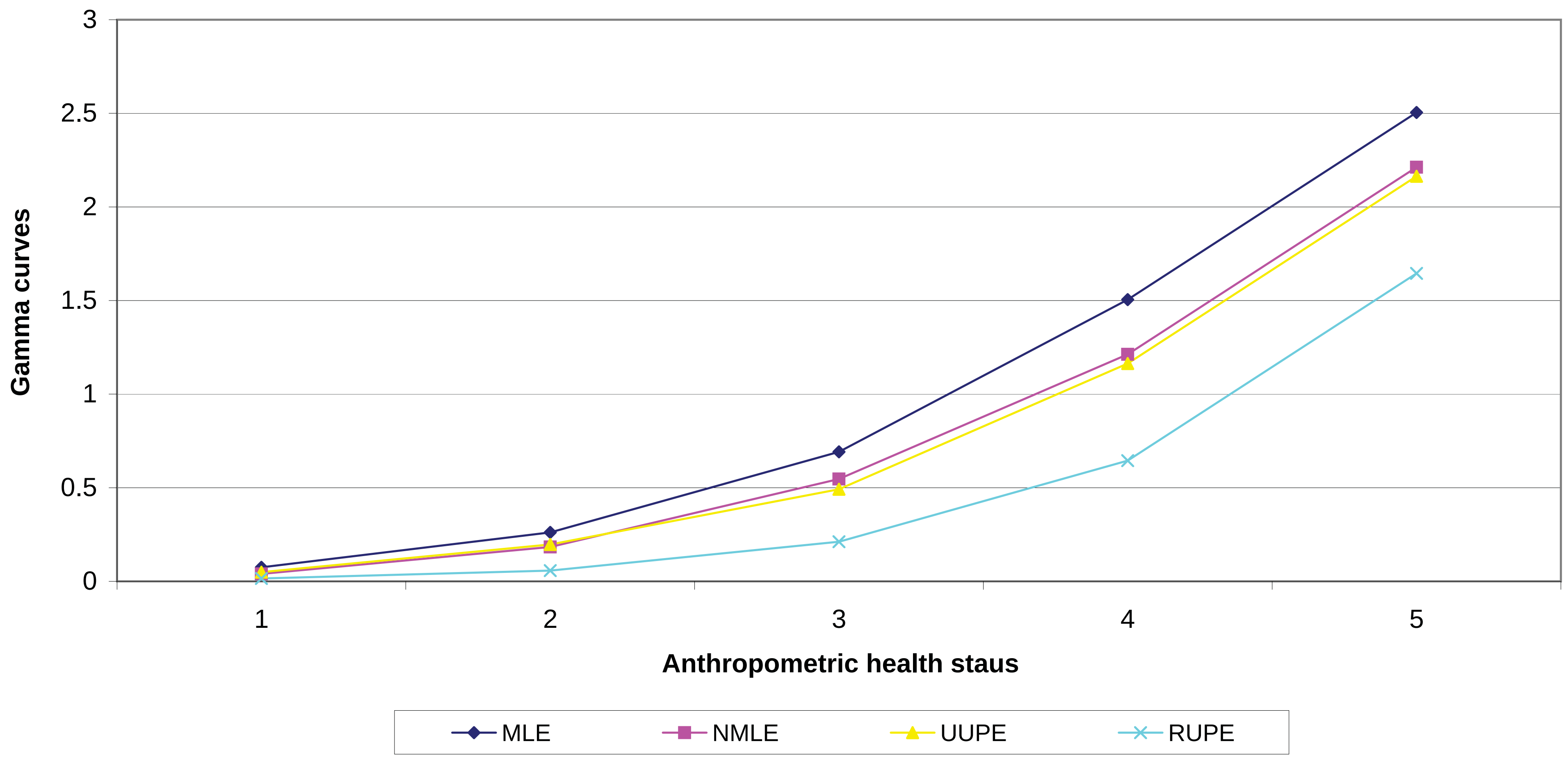

Figure 2: Social welfare lattice of health distributions in Egypt

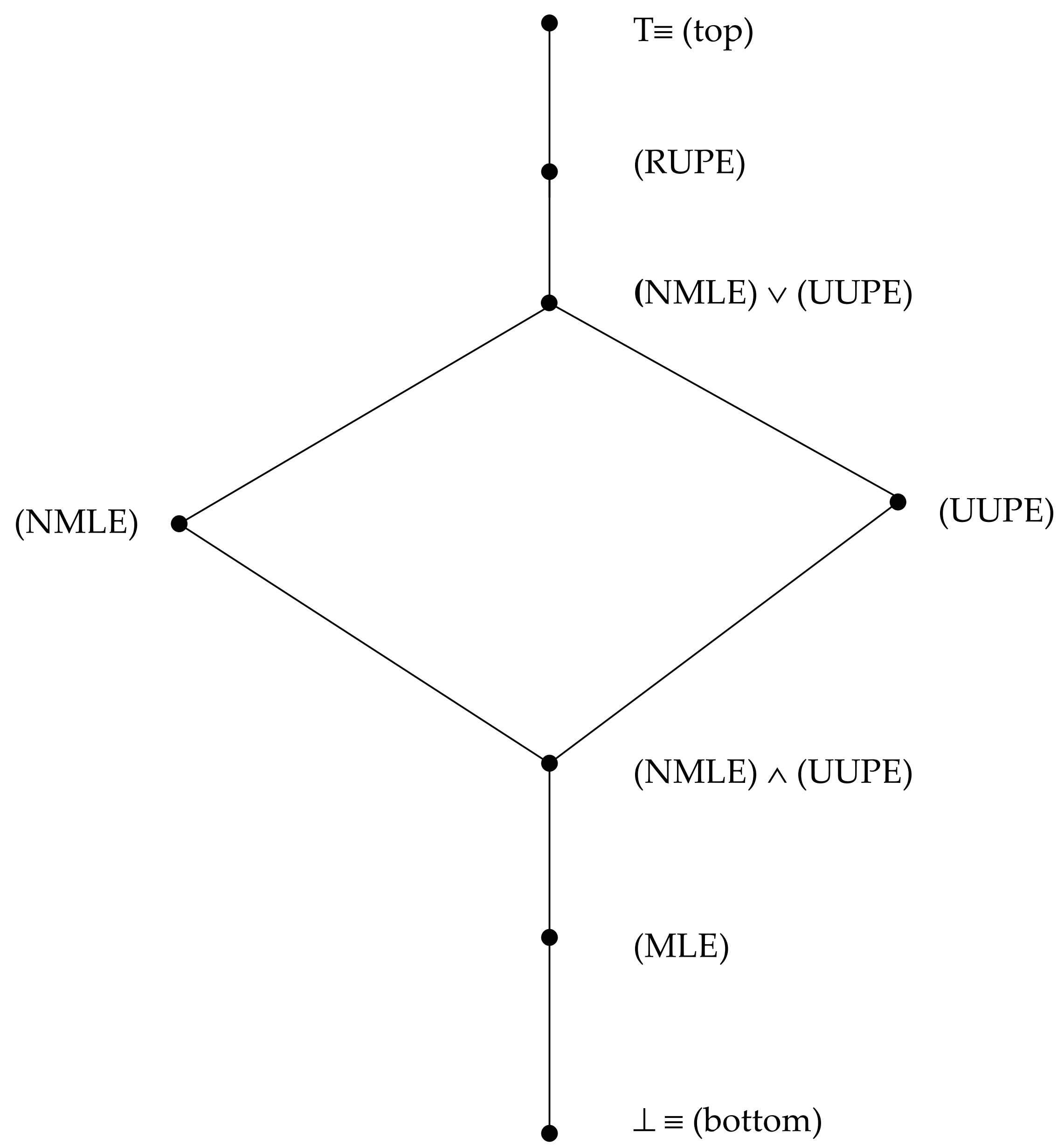

Table 2: Social Welfare calculations for various distributions

Distribution	Social welfare measure			
	$W_{0.05,1}$	$W_{0.05,0}$	$W_{0.50,1}$	$W_{0.50,0}$
$\perp \equiv$ (bottom)	0	0	0	0
MLE	0.1013	0.2537	0.6934	1.7368
(NMLE) $\wedge$ (UUPE)	0.1367	0.3024	0.8980	1.9867
NMLE	0.1405	0.3108	0.9053	2.0028
UUPE	0.1432	0.3099	0.9386	2.0311
(NMLE) $\vee$ (UUPE)	0.1470	0.3181	0.9457	2.0463
RUPE	0.2825	0.4047	1.5926	2.6196
T $\equiv$ (top)	4.0000	4.0000	4.0000	4.0000