# The Measurement of Mobility: A Class of Distance Indices.

Marcello D'Agostino — Università di Ferrara
Valentino Dardanoni — Università di Palermo

March 15, 2006

**Abstract**

Although there are many good reasons for being interested in the concept of social mobility, its precise definition — and *a fortiori* its measurement — appears to be somewhat elusive. In this paper we propose an axiomatic characterization of this concept which leads to a suitable class of "canonical" mobility indices. Such indices are defined over pairs $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$ is a vector representing the values of a relevant indicator of the social and economic status of individuals, and $\mathbf{y}$ is a vector representing the values of the same indicator for "the next generation". The axioms in our approach take into account the distinction between "structural" mobility (i.e. "how far apart" $\mathbf{x}$ and $\mathbf{y}$ are) and "exchange" mobility (i.e. the result of "swaps" in the relative positions of families). Another interesting aspect is that we distinguish between "absolute" and "relative" mobility, depending on whether the values of the status indicator are based on data concerning each individual taken in isolation or on data concerning both an individual and its "generational context". As far as absolute mobility is concerned, the main result consists in showing that our axioms force a rule for making mobility comparisons which is based on the mean squared distance between vectors, which is equivalent to using the familiar Euclidean distance when comparing societies of equal size. As for relative mobility, the main result consists in showing that our axioms characterize a class of mobility indices that includes Pearson's correlation coefficient whose use in the measurement of (im)mobility is widespread and dates back at least to the late XIX century.

KEYWORDS: Mobility measurement, Absolute and relative mobility, Axiomatic approach, Mean squared distance, Euclidean distance, Correlation coefficient.

JEL numbers: D31, D63.

ACKNOWLEDGMENTS: Previous versions of this paper have circulated before, sometimes with a different title. We have received comments and suggestions from many seminars' audiences, and help from many colleagues. A full list of thanks will appear in the final version of the paper.

> You are, all of you in this community, brothers. But when god fashioned you, he added gold in the composition of those of you who are qualified to be Rulers (which is why their prestige is greatest); he put silver in the Auxiliaries, and iron and bronze in the farmers and other workers. Now since you are all of the same stock, though your children will commonly resemble their parents, occasionally a silver child will be born of golden parents, or a golden child of silver parents, and so on. Therefore the first and most important of god's commandments to the Rulers is that in the exercise of their function as Guardians their principal care must be to watch the mixture of metals in the characters of their children. If one of their own children has traces of bronze or iron in its make-up, they must harden their hearts, assign it its proper value, and degrade it to the ranks of the industrial and agricultural class where it properly belongs: similarly, if a child of this class is born with gold or silver in its nature, they will promote it appropriately to be a Guardian or an Auxiliary. And this they must do because there is a prophecy that the State will be ruined when it has Guardians of silver or bronze.
>
> Plato, *The Republic*, III, 415bc

# 1 Introduction

At least since Plato it has been recognized that there are many reasons to be interested in social mobility: socially mobile societies are both equitable ("if a child from bronze parents is born with gold or silver in its nature, must be promoted to be a Guardian or an Auxiliary") and efficient ("Rulers should be made of gold regardless of their origin"). The study of the determinants and effects of social mobility has recently interested economists working in the political economy tradition. In an ethical perspective, social mobility is often interpreted as a factor determining equality of opportunity. In fact, some scholars even argue that equality of opportunity should be the main, if not exclusive, point of political concern in a society [Stokey, 1998]. On the other hand, in more socially mobile communities, the support for redistribution might be lower [Alesina and La Ferrara, 2005, Benabou and Ok, 2001a]. In addition, social mobility may promote economic efficiency and stimulate economic growth, and an understanding of social mobility issues is key for analyzing public educational policies [Benabou, 2002, Checchi *et al.*, 1999].

It is widely believed that socioeconomic mobility is somewhat an elusive concept, difficult to define, let alone to measure: as remarked by Fields and Ok [1999b] in a recent survey "... the mobility literature does not provide a unified discourse of analysis. ... a considerable rate of confusion confronts a newcomer in the field."[1] This is in stark contrast with the literature on income inequality, where a consensus has emerged on what concepts of inequality mean, on the correct theoretical procedures to measure it, and on how to go from theory to empirical applications.

When discussing mobility issues, a basic distinction is usually made between intergenerational and intragenerational mobility. The first concept concerns the study of how the distribution of some relevant measure of individual status changes between different generations in a given society. Alternatively, intragenerational mobility studies how the distribution of individual status changes among a group of individuals over a given period of their lifetime. In general, the simplest framework to capture either of these aspects is to consider how, in a society of $n$ individuals, a vector $\mathbf{x}$ is transformed into another vector $\mathbf{y}$, where the $i$-th element $x_i$ denotes the value of a relevant indicator of the social and economic status of individual $i$, and $y_i$ denotes its value in the next generation (intergenerational case) or in the next time period (intragenerational case). Typical variables employed in most mobility studies for measuring socio-economic status are income, wage, consumption, education, and occupational prestige. Henceforth, we will focus on intergenerational mobility and follow the usual convention of analyzing father-to-son movements in status as unit of analysis.

When analyzing mobility data, the interplay between the distributions of $\mathbf{x}$ and $\mathbf{y}$ can be described by two quite different concepts, first introduced in the

[1] See also [Maasoumi, 1998] for a survey on mobility measurement.

sociology literature. *Structural mobility* refers to how far apart $\mathbf{x}$ is from $\mathbf{y}$. For example, if a country is experimenting a substantial economic growth, there will be a greater number of high status positions available to the sons than there were to the fathers, and this determines some kind of social change. However, two hypothetical societies could display the same amount of structural mobility because they have the same marginal distributions, but they could differ in how families interchange their relative positions. This second aspect is called *exchange mobility* by sociologists and refers to the positive association between fathers' and sons' status in the society.

Economists who have discussed mobility indices have generally borrowed from the literature on inequality measurement the fundamental distinction between "absolute" and "relative" indices. The absolute concept of income inequality is concerned with the differences between individual incomes, while the relative concept is concerned with their ratios. The two concepts can be formally defined in terms of *invariance properties*: an inequality index $I : \mathbb{R}^n \mapsto \mathbb{R}$ is a relative inequality index if it is *scale invariant*, i.e. $I(\mathbf{x}) = I(\lambda\mathbf{x})$ for all positive $\lambda$'s, while $I(\mathbf{x})$ is an absolute inequality index if it is *translation invariant*, i.e. $I(\mathbf{x}) = I(\mathbf{x} + \theta)$ for all positive $\theta$'s. Most literature on the measurement of mobility has followed the same path, by calling "relative" a mobility index which is scale invariant, and "absolute" index which is translation invariant.

The distinction between absolute and relative inequality is a clear-cut one, as shown by the well-known result that an inequality index cannot be both scale and translation invariant.[2] By contrast, in the measurement of mobility the situation appears to be more problematic; as a matter of fact we can have mobility indices which are at the same time scale *and* translation invariant or even satisfy much more general conditions. For example, one of the more commonly used indices of (im)mobility, namely Pearson's correlation coefficient,

[2] A similar result holds also in the field of poverty measurement [Zheng, 1994]. Foster [1998] however argues that translation and scale invariance is only one of the multiple notions of absolute and relative poverty that arise in choosing poverty lines and in aggregating the data into an overall poverty index.

is invariant under different affine transformations of $\mathbf{x}$ and $\mathbf{y}$. Moreover, it is not clear how the distinction between translation and scale invariant indices is related to the distinction between structural and exchange mobility.

In this paper we propose a different approach. In particular, we interpret the distinction between absolute and relative mobility as a distinction between two different ways of evaluating an individual's status. We argue that the *absolute* status of an individual is a synthetic indicator which can be derived by looking at data regarding the individual *taken in isolation.* For instance, we can evaluate the absolute status of the $i$-th father in a population as a function of appropriate data such as his income, wage, consumption, education, occupational prestige etc. Thus, we shall assume that, given a set of $k$ status indicators $\bar{u}_{i1}, \ldots, \bar{u}_{ik}$ for individual $i$, there is an "aggregating function" $\phi$ such that the absolute status of $i$ is given by: $u_i = \phi(\bar{u}_{i1}, \ldots, \bar{u}_{ik})$.

The value $u_i$ of this function is a measure of absolute status since it can be determined without explicit reference to the other individuals in the same population. It is beyond the scope of this paper to suggest meaningful ways to aggregate various status indicators into a synthetic measure,[3] for example by using panel data to estimate permanent income from repeated observations on annual income; hence, in the context of this paper, our "data" will be the values $u_i$ of the aggregating function $\phi$ applied to the "crude" data $\bar{u}_{i1}, \ldots, \bar{u}_{ik}$.

On the other hand, we argue that the *relative* status of an individual refers to his relative standing *with respect to* his generational context. For example, deriving the *relative* status of the $i$-th father in a population requires not only knowledge of his absolute status $u_i$, but also of the vector $\mathbf{u}$ representing the absolute statuses of all the individuals in his generation. So, we shall assume that there is a *relative status function* $S$ such that the relative status of $i$ is given by $S(u_i, \mathbf{u})$. In the context of this paper, then, absolute (relative) mobility simply refers to how a vector of absolute (relative) status indicators for the fathers' generation is transformed into another vector of absolute (relative)

[3] However we will discuss a (significant) special case in section 2.3.

status indicators for the sons.[4]

In the light of our proposed distinction between absolute and relative mobility, the rest of the paper contains two main sections. In the next section we consider absolute mobility, by postulating four rather natural axioms. The two key axioms relate to structural and exchange mobility, while the other two are subgroup consistency and population invariance properties which are standard in this sort of analysis. The main result of the next section (Theorem 1) consists in showing that these axioms characterize a rule for making absolute mobility comparisons which is based on the *mean squared distance* of fathers' and sons' status vectors: given two societies $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{w}, \mathbf{z})$, respectively composed of $n$ and $m$ families, $(\mathbf{x}, \mathbf{y})$ will display at least as much absolute mobility as $(\mathbf{w}, \mathbf{z})$ if and only if

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \geq \frac{1}{m}\sum_{i=1}^{m}(w_i - z_i)^2.$$

Notice that when comparing vectors of the same size this rule coincides precisely with the ranking induced by the Euclidean distance; since the Euclidean distance is massively applied in economics, econometrics, statistics and in all the sciences in general, our result may be of independent interest. We end the section by introducing and motivating a kind of *canonical index* for representing the absolute mobility ordering characterized in Theorem 1.

The third section deals with *relative* mobility. The main challenge is how to evaluate the relative status of an individual by taking into account the whole distribution of the absolute statuses of all the individuals in his generation. We shall show that five rather natural axioms are sufficient to give a precise solution to our problem and to provide a natural link between absolute and relative mobility. The main result of the final section (Theorem 2) will then consist in showing that these axioms characterize a *class* of relative mobility

[4] A further use of the terminology of absolute and relative mobility is in the recent sociological literature, where absolute mobility refers to the amount and rates of movement between different social classes, while relative mobility refers to the degree of inequality, according to class origins, in a person's chances of acquiring a better social class.

indices of the following form:

$$R(\mathbf{x},\mathbf{y}) = \frac{c}{n} \cdot \sum_{i}^{n} \left( \frac{x_i - \lambda(\mathbf{x})}{\delta(\mathbf{x})} - \frac{y_i - \lambda(\mathbf{y})}{\delta(\mathbf{y})} \right)^2$$

where $\lambda$ and $\delta$ denote, respectively, a location and a dispersion statistic and $c$ is a positive constant. By appropriate choice of $\lambda$ and $\delta$, it is immediately seen that the proposed class contains some rather well-known elements: for example, if we choose $\lambda$ to be equal to the mean and $\delta$ to be equal to the standard deviation it is easily seen that the index is equivalent to Pearson's correlation coefficient, once we adjust for the fact that the correlation coefficient actually measures immobility rather than mobility. Use of the correlation coefficient to measure immobility dates back at least to Galton's model [Galton, 1886] for analyzing the intergenerational transmission of hereditary stature, and it is fair to say that modern versions of Galton's model, based on OLS regression of sons' on fathers' income, are the workhorse of applied mobility analyses by economists, see e.g. Zimmerman [1992] and Solon [2000]. Thus, our class of relative mobility indices gives an axiomatic base to this practice.

# 2 Absolute mobility

Let $\mathcal{M}_n \subseteq \mathbb{R}^n \times \mathbb{R}^n$ be the set of all mobility data[5] for societies of $n$ families, and let $\mathcal{M} := \cup_n \mathcal{M}_n$. The elements of $\mathcal{M}$ will be pairs $(\mathbf{x},\mathbf{y})$ where $\mathbf{x}$ and $\mathbf{y}$ are vectors of the same size called, respectively, the *marginal distribution* of the fathers' and of the sons' absolute status. The pairs $(\mathbf{x},\mathbf{y})$ will be called *societies*. We shall write $x_i$ and $y_i$ to denote, respectively, the absolute status of the father and of the son in the $i$-th family of the society $(\mathbf{x},\mathbf{y})$. We shall also use the notation $|\mathbf{x}|$ to denote the size of the vector $\mathbf{x}$.

In what follows we shall use lightface letters $a, b, c, d$, etc. without subscripts, to denote 1-element vectors, so that societies in $\mathcal{M}_1$ (families) will be

[5]Recall that when we consider a society $(\mathbf{x},\mathbf{y})$ we shall assume that the elements of $\mathbf{x}$ and $\mathbf{y}$ are real numbers representing the absolute status of the corresponding individuals, i.e. that we have already applied the aggregating function to the crude data.

denoted by $(a, b), (c, d)$, etc. Of course the same family can be seen both as a society in $\mathcal{M}_1$ and as an element of a society in $\mathcal{M}$. We shall write $(x_i, y_i)$ to denote the $i$-th family of the society $(\mathbf{x}, \mathbf{y})$.

We shall denote by $\mathbf{xy}$ the concatenation of two vectors $\mathbf{x}$ and $\mathbf{y}$. We shall also find it convenient to write $(\mathbf{x}, \mathbf{y}) \circ (\mathbf{w}, \mathbf{z})$ to denote the society obtained by concatenating the corresponding vectors $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{w}, \mathbf{z})$, that is: $(\mathbf{x}, \mathbf{y}) \circ (\mathbf{w}, \mathbf{z}) := (\mathbf{xw}, \mathbf{yz})$. Therefore:

$$(\mathbf{x}, \mathbf{y}) = (x_1, y_1) \circ \cdots \circ (x_n, y_n).$$

A *mobility ordering* $\preceq$ is a weak ordering (a transitive and complete binary relation) on $\mathcal{M}$.[6] For any $(\mathbf{x}, \mathbf{y}), (\mathbf{w}, \mathbf{z}) \in \mathcal{M}$, $(\mathbf{x}, \mathbf{y}) \preceq (\mathbf{w}, \mathbf{z})$ means that $(\mathbf{w}, \mathbf{z})$ has at least the same mobility as $(\mathbf{x}, \mathbf{y})$. As usual we denote by $\sim$ and $\prec$, respectively, the symmetric and asymmetric part of $\preceq$. We assume that $\preceq$ is *continuous* and *anonymous*. Recall that $\preceq$ is continuous if for all $(\mathbf{w}, \mathbf{z}) \in \mathcal{M}$, the sets $\{(\mathbf{x}, \mathbf{y}) \in \mathcal{M} : (\mathbf{x}, \mathbf{y}) \preceq (\mathbf{w}, \mathbf{z})\}$ and $\{(\mathbf{x}, \mathbf{y}) \in \mathcal{M} : (\mathbf{x}, \mathbf{y}) \succeq (\mathbf{w}, \mathbf{z})\}$ are both closed, and anonymous if $(x_1, y_1) \circ \cdots \circ (x_n, y_n) \sim (x_{\pi(1)}, y_{\pi(1)}) \circ \cdots \circ (x_{\pi(n)}, y_{\pi(n)})$ for every permutation $\pi$ and every $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}_n$, $n \in \mathbb{N}$. Note that the anonymity of the mobility ordering implies that the concatenation operator $\circ$ is symmetric, i.e. $(\mathbf{x}, \mathbf{y}) \circ (\mathbf{x}', \mathbf{y}') \sim (\mathbf{x}', \mathbf{y}') \circ (\mathbf{x}, \mathbf{y})$.

The assumption of continuity and anonymity of the ordering are invariably made in the literature on inequality, poverty and mobility measurement; continuity is justified by the need for robustness to small data perturbations, while anonymity is justified on the basis of the "equal" treatment of all families in

[6] An important branch of the literature on mobility measurement drops the completeness assumption to considers partial orderings [Shorrocks, 1978, Atkinson, 1983, Conlisk, 1990, Dardanoni, 1993, Benabou and Ok, 2001b, Formby *et al.*, 2003]. Formby *et al.* [2004] consider issues of statistical inference for some of these mobility orderings. [Chakravarty *et al.*, 1985, Cowell, 1985, Fields and Ok, 1996, Fields and Ok, 1999a, Gottschalk and Spolaore, 2002, King, 1983, Maasoumi and Zandvakili, 1986, Mitra and Ok, 1998, Ruiz-Castillo, 2004] consider mobility indices in much the same spirit as the present paper. We omit to quote the huge empirical literature on mobility comparisons.

the society. We take these properties as basic in the derivation of our results, and in the sequel we assume they hold without further reference:

**Axiom 0** $\preceq$ *is continuous and anonymous.*

In the sequel we shall also write $\preceq_n$ to denote the restriction of $\preceq$ to $\mathcal{M}_n$, i.e. to societies of size $n$.

## 2.1 Axioms

In this secton we list and comment four basic axioms that a reasonable mobility ordering should satisfy. Our first axiom captures the essence of structural mobility. It simply says that, in a society consisting of a single family, mobility grows with the distance between father's and son's absolute status:

**Axiom 1** *For all $a, b, c, d \in \mathbb{R}$,*

$$(a, b) \preceq (c, d) \text{ if and only if } |a - b| \leq |c - d|.$$

Notice that this axiom implies the symmetry of upward and downward mobility; as remarked by Fields and Ok [1999a], this symmetry is unexceptionable if one does not distinguish between 'good' and 'bad' movement of absolute status. In section 2.3 below, we will discuss the interpretation of this axiom in the special case where absolute status is a function of a given indicator such as income or wage.

Our second axiom is well-known in the inequality [Shorrocks, 1988], poverty [Foster and Shorrocks, 1991] and mobility [Fields and Ok, 1999a] measurement literature where it is known as "subgroup consistency". (Though the axiom is widely accepted, for a critical discussion see Foster and Sen [1997].)

**Axiom 2** *For all $k, j$ and all $(\mathbf{x}, \mathbf{y}), (\mathbf{w}, \mathbf{z}) \in \mathcal{M}_k, (\mathbf{x}', \mathbf{y}'), (\mathbf{w}', \mathbf{z}') \in \mathcal{M}_j$,*

$$(\mathbf{x}, \mathbf{y}) \prec (\mathbf{w}, \mathbf{z}) \text{ and } (\mathbf{x}', \mathbf{y}') \sim (\mathbf{w}', \mathbf{z}') \text{ imply } (\mathbf{x}, \mathbf{y}) \circ (\mathbf{x}', \mathbf{y}') \prec (\mathbf{w}, \mathbf{z}) \circ (\mathbf{w}', \mathbf{z}'). \quad (1)$$

Observe that, by continuity of the mobility ordering, this axiom implies that strict ordering $\prec$ on both sides of (1) can be replaced by $\preceq$.

The next axiom expresses the usual *replication invariance* property which is known to be the key for comparing societies of different size. Letting $\overbrace{\mathbf{u}\cdots\mathbf{u}}^{n}$ denote the result of concatenating the vector $\mathbf{u}$ with itself $n$ times, we have:

**Axiom 3** *For every* $(\mathbf{x},\mathbf{y}) \in \mathcal{M}$ *and every* $n \in \mathbb{N}$*:*

$$(\mathbf{x},\mathbf{y}) \sim (\overbrace{\mathbf{x}\cdots\mathbf{x}}^{n}, \overbrace{\mathbf{y}\cdots\mathbf{y}}^{n})$$

Our final axiom deals with exchange mobility; it concentrates on what happens to mobility when, given fixed marginal distributions $(\mathbf{x},\mathbf{y})$, there is an interchange in status in the transition from $\mathbf{x}$ to $\mathbf{y}$. Consider two societies $(\mathbf{x},\mathbf{y}),(\mathbf{x},\mathbf{y}')$ such that for some $i,j$, with $i \neq j$,

- $x_i < x_j$
- $y_i < y_j$
- $y'_i = y_j$
- $y'_j = y_i$
- for all $k \neq i,j$, $(x_k, y_k) = (x_k, y'_k)$.

Given that in the two families $(x_i, y_i),(x_j, y_j)$ the poorer father has the poorer son, exchanging the sons' status will intuitively yield an increase in mobility. Let $\Delta_f = x_j - x_i > 0$ and $\Delta_s = y_j - y_i = y'_i - y'_j > 0$; under these circumstances we say that $(\mathbf{x},\mathbf{y}')$ has been obtained from $(\mathbf{x},\mathbf{y})$ by means of a *swap of type* $(\Delta_f, \Delta_s)$, and write $(\mathbf{x},\mathbf{y}) \xrightarrow{(\Delta_f,\Delta_s)} (\mathbf{x},\mathbf{y}')$. Such swaps are well-known in the mathematical statistics [Tchen, 1980] and economics [Epstein and Tanny, 1980, Atkinson, 1983, Dardanoni, 1993] literature and it is often assumed that they are always mobility-increasing. An independent question is whether or not swaps of the same type have the same effect on the mobility ordering. Our axiom assumes that swaps of the same type have the same effect on mobility (see Fig. 1 below).

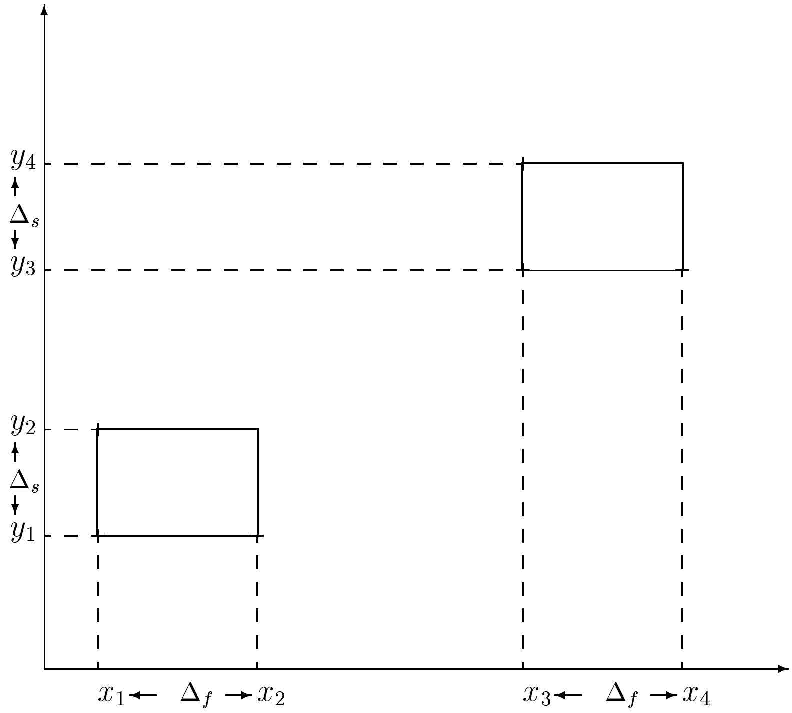


Figure 1: In a society made of four families: $(x_1, y_1) \circ (x_2, y_2) \circ (x_3, y_3) \circ (x_4, y_4)$ consider two alternative swaps of type $(\Delta_f, \Delta_s)$. Axiom 4 implies that $(x_1, y_2) \circ (x_2, y_1) \circ (x_3, y_3) \circ (x_4, y_4) \sim (x_1, y_1) \circ (x_2, y_2) \circ (x_3, y_4) \circ (x_4, y_3)$.

**Axiom 4** *For all* $n \geq 4$*, all* $(\mathbf{x}, \mathbf{y}), (\mathbf{x}, \mathbf{y}'), (\mathbf{x}, \mathbf{y}'') \in \mathcal{M}_n$*, and all* $\Delta_f, \Delta_s \in \mathbb{R}_+$*, if*

- $(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}')$
- $(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}'')$

*then:*

$$(\mathbf{x}, \mathbf{y}') \sim (\mathbf{x}, \mathbf{y}'').$$

This axiom seems unexceptionable if we recall that absolute status is a transformation of some underlying individual indicators (see again section 2.3 below). Taken in isolation, this axiom says only that swaps of the same type have the same effect and does not specify whether the effect of a given swap is positive. However, in the next section (Corollary 1) we show that, under Axioms 1– 4, *every* swap has a *positive* effect on mobility.

## 2.2 Results

The main result of this section is that the only ordering $\preceq$ satisfying Axioms 1–4 is the one defined as follows: for any $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{w}, \mathbf{z}) \in \mathcal{M}$

$$(\mathbf{x}, \mathbf{y}) \preceq (\mathbf{w}, \mathbf{z}) \Longleftrightarrow d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{w}, \mathbf{z})$$

where, for $(\mathbf{u}, \mathbf{v}) \in \mathcal{M}_k$, $d(\mathbf{u}, \mathbf{v})$ is the *mean squared distance* of $\mathbf{u}$ and $\mathbf{v}$, namely

$$d(\mathbf{u}, \mathbf{v}) = \frac{1}{k}\sum_{i=1}^{k}(u_i - v_i)^2.$$

**Theorem 1** *A mobility ordering $\preceq$ satisfies Axioms 1–4 if and only if, for every $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{w}, \mathbf{z}) \in \mathcal{M}$ the following equivalence holds true:*

$$(\mathbf{x}, \mathbf{y}) \preceq (\mathbf{w}, \mathbf{z}) \Longleftrightarrow d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{w}, \mathbf{z}).$$

**Proof** The proof that the class of orderings induced by the mean squared distance satisfies the axioms is left to the reader. As a hint, for Axiom 4, just observe that the incremental mobility effect of a generic swap $(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}')$ is given by:

$$\begin{aligned} & d(\mathbf{x}, \mathbf{y}') - d(\mathbf{x}, \mathbf{y}) = \\ & = \frac{1}{k}[(x_i - y_i - \Delta_s)^2 + (x_i + \Delta_f - y_i)^2] - [(x_i - y_i)^2 + (x_i + \Delta_f - y_i - \Delta_s)^2] = \\ & = \frac{2}{k}\Delta_f \Delta_s, \end{aligned} \quad (2)$$

where $k$ is the size of $\mathbf{x}$ and $\mathbf{y}$. Therefore, the effect of the swap depends only on $\Delta_f$ and $\Delta_s$.

We now prove that the class of orderings induced by the mean squared distance is the only one satisfying the axioms. First, observe that Axiom 1 implies

$$(a, b) \sim (|a - b|, 0) \quad (3)$$

and Axiom 2 implies that for all $(\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}') \in \mathcal{M}_n$ and all $n \in \mathbb{N}$,

$$(x_i, y_i) \sim (x_i', y_i') \text{ for all } i = 1, \cdots, n, \Longrightarrow (\mathbf{x}, \mathbf{y}) \sim (\mathbf{x}', \mathbf{y}').$$

Therefore, given (3) above, we have that for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}_n$ and all $n \in \mathbb{N}$,

$$(\mathbf{x}, \mathbf{y}) \sim (|x_1 - y_1|, 0) \circ \cdots \circ (|x_n - y_n|, 0). \tag{4}$$

Hence, using Debreu's utility representation theorem [Debreu, 1954], continuity of $\preceq$ implies that there exists a continuous function $M_n : \mathbb{R}^n \mapsto \mathbb{R}$ such that, for all $(\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}') \in \mathcal{M}_n$,

$$(\mathbf{x}, \mathbf{y}) \preceq (\mathbf{x}', \mathbf{y}') \Longleftrightarrow M_n(|x_1 - y_1|, \cdots, |x_n - y_n|) \leq M_n(|x'_1 - y'_1|, \cdots, |x'_n - y'_n|).$$

Thus, the function $M_n : \mathbb{R}^n \mapsto \mathbb{R}$ is a numerical representation of $\preceq_n$ (i.e. the restriction of $\preceq$ to $\mathcal{M}_n$). Moreover, given Axioms 1 and 2, $M_n$ must be increasing in each argument. Then we can follow Foster and Shorrocks [1991], and in particular the proof of their Proposition 1, to show that under Axioms 1–3, the mobility ordering admits of an additive representation. We reproduce here the main steps of their argument for the sake of completeness (see also the proof of Proposition 3 in Fields and Ok [1999a] for a similar derivation).

First, Axiom 2 implies that, for all $\mathbf{u}, \mathbf{u}' \in \mathbb{R}^h$ and $\mathbf{v}, \mathbf{v}' \in \mathbb{R}^k$ such that $h + k = n$,

$$M_h(\mathbf{u}) < M_h(\mathbf{u}') \text{ and } M_k(\mathbf{v}) = M_k(\mathbf{v}') \Rightarrow M_n(\mathbf{uv}) < M_n(\mathbf{u}'\mathbf{v}').$$

Following Foster and Shorrocks' derivation of their equation (12), this implies that $M_n$ must satisfy

$$M_n(\mathbf{uv}) \leq M_n(\mathbf{u}'\mathbf{v}) \Rightarrow M_n(\mathbf{uv}') \leq M_n(\mathbf{u}'\mathbf{v}').$$

Since this condition must hold for all $\mathbf{u}, \mathbf{u}' \in \mathbb{R}^h$ and $\mathbf{v}, \mathbf{v}' \in \mathbb{R}^k$ with $h + k = n$, one can apply Gorman's separability Theorem [Gorman, 1968], to show that the function $M_n$ must be additively separable, that is, for all $n \geq 3$ there must exist continuous and increasing functions $F_n : \mathbb{R} \mapsto \mathbb{R}$ and $g_n : \mathbb{R}_+ \mapsto \mathbb{R}$ such that for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}_n$,

$$M_n(|x_1 - y_1|, \cdots, |x_n - y_n|) = F_n[\sum_{i=1}^{n} g_n(|x_i - y_i|)].$$

Let $f_n := ng_n$, so that

$$M_n(|x_1 - y_1|, \cdots, |x_n - y_n|) = F_n[\sum_{i=1}^{n} \frac{1}{n} f_n(|x_i - y_i|)]. \tag{5}$$

Then, following the derivation of equation (21) from equation (14) in Foster and Shorrocks [1991], we have that the replication invariance axiom (Axiom 3) allows us to choose the functions $F_n$ and $f_n$ to be independent of $n$ (i.e. such that for all $m, n \in \mathbb{N}$, $F_m = F_n = F$ and $f_m = f_n = f$, for some fixed $F$ and $f$) and to extend (5) to the case of $n < 3$. Therefore, the mobility ordering $\preceq$ satisfies Axioms 1, 2 and 3 if and only if there is an increasing and continuous function $f : \mathbb{R}_+ \mapsto \mathbb{R}$ such that for all $(\mathbf{x}, \mathbf{y})$ in $\mathcal{M}_m$ and $(\mathbf{w}, \mathbf{z}) \in \mathcal{M}_n$,

$$(\mathbf{x}, \mathbf{y}) \preceq (\mathbf{w}, \mathbf{z}) \Longleftrightarrow \frac{1}{m} \sum_{i=1}^{m} f(|x_i - y_i|) \leq \frac{1}{n} \sum_{i=1}^{n} f(|w_i - z_i|). \tag{6}$$

To end the proof of the theorem, it suffices then to prove that the function $f$ in equation (6) must be quadratic. Let $a, b, c \in \mathbb{R}_+$ be arbitrary non negative real numbers with $a \geq c$. Suppose $(\mathbf{x}, \mathbf{y})$, $(\mathbf{x}, \mathbf{y}')$, $(\mathbf{x}, \mathbf{y}'') \in \mathcal{M}_n$, with $n \geq 4$, are such that for some $(\mathbf{w}, \mathbf{z}) \in \mathcal{M}_{n-4}$,

1. $(\mathbf{x}, \mathbf{y}) = (\mathbf{w}, \mathbf{z}) \circ (a, 0) \circ (a + c, c) \circ (b, b) \circ (b + c, b + c)$;

2. $(\mathbf{x}, \mathbf{y}') = (\mathbf{w}, \mathbf{z}) \circ (a, c) \circ (a + c, 0) \circ (b, b) \circ (b + c, b + c)$;

3. $(\mathbf{x}, \mathbf{y}'') = (\mathbf{w}, \mathbf{z}) \circ (a, 0) \circ (a + c, c) \circ (b, b + c) \circ (b + c, b)$

so that $(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}')$ and $(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}'')$ with $\Delta_f = \Delta_s = c$. Hence, we have:

$$\sum_{i=1}^{n} f(|x_i - y_i'|) - \sum_{i=1}^{n} f(|x_i - y_i|) = f(a + c) + f(a - c) - 2f(a)$$

and

$$\sum_{i=1}^{n} f(|x_i - y_i''|) - \sum_{i}^{n} f(|x_i - y_i|) = 2f(c) - 2f(0)$$

so that, for all $a \geq c \geq 0$, in order to satisfy Axiom 4 we must have

$$f(a + c) - f(a) = f(a) - f(a - c) + 2f(c) - 2f(0).$$

We shall show that this functional equation has a unique solution $f(t) = \alpha t^2 + \beta$ for some constants $\alpha > 0$ and $\beta$.

Consider the sequence $x_m = mc$ with $m \in \mathbb{N}_+$. We first evaluate the difference between two adjacent terms of the sequence $f(x_m)$:

$$\begin{aligned}
f(mc) - f((m-1)c) &= f((m-1)c) - f((m-2)c) + 2(f(c) - f(0)) \\
&= f((m-2)c) - f((m-3)c) + 2(f(c) - f(0)) + 2(f(c) - f(0)) \\
&\vdots \\
&= f(c) - f(0) + 2(m-1)(f(c) - f(0)) \\
&= (2m-1)(f(c) - f(0))
\end{aligned}$$

Now, notice that:

$$\begin{aligned}
f(mc) &= \sum_{i=1}^{m} \big(f(ic) - f((i-1)c)\big) + f(0) \\
&= \sum_{i=1}^{m} (2i-1)(f(c) - f(0)) + f(0) \\
&= 2\frac{m^2+m}{2}(f(c) - f(0)) - m(f(c) - f(0)) + f(0)
\end{aligned}$$

so that we obtain the following functional equation:

$$f(mc) = m^2(f(c) - f(0)) + f(0) \tag{7}$$

Let $c = \frac{k}{m}$ for some $k \in \mathbb{N}_+$. Since $mc = k$, then

$$m^2(f(c) - f(0)) = f(k) - f(0). \tag{8}$$

Then, using equation (7) with $c = 1$ and $m = k$ we get

$$m^2(f(c) - f(0)) = k^2(f(1) - f(0)). \tag{9}$$

Hence:

$$f(c) = c^2(f(1) - f(0)) + f(0) \tag{10}$$

for all rational $c$. Since $f(1) - f(0)$ is a strictly positive constant (given that $f$ is increasing), and $f$ is continuous, it follows that, for all $t \in \mathbb{R}_+$:

$$f(t) = \alpha t^2 + \beta \tag{11}$$

with $\alpha, \beta$ constants and $\alpha > 0$. $\square$

**Corollary 1** *Axioms 1–4 imply that for all* $n \geq 2$, $(\mathbf{x}, \mathbf{y}), (\mathbf{x}, \mathbf{y}') \in \mathcal{M}_n$, *and any* $\Delta_f, \Delta_s > 0$,

$$(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}') \Rightarrow (\mathbf{x}, \mathbf{y}) \prec (\mathbf{x}, \mathbf{y}').$$

**Proof** When $(\mathbf{x}, \mathbf{y}) \xrightarrow{(\Delta_f, \Delta_s)} (\mathbf{x}, \mathbf{y}')$, recall from equation (2) that

$$d(\mathbf{x}, \mathbf{y}') - d(\mathbf{x}, \mathbf{y}) = \frac{2}{k} \Delta_f \Delta_s.$$

$\square$

Theorem 1 characterizes a criterion for making absolute mobility comparisons which is based on the mean squared distance between the fathers' and sons' absolute status vectors. Notice that, when comparing vectors of the same size, this rule coincides precisely with the ranking implied by the Euclidean distance; since the Euclidean distance is massively applied in economics, econometrics, statistics and in all the sciences in general, our result may be of independent interest.

It is interesting to note that the mean squared distance ordering incorporates elements of both structural (Axiom 1) and exchange (Axiom 4) mobility. This can be appreciated by noting that the mean squared distance $d(\mathbf{x}, \mathbf{y})$ can be decomposed as:

$$\frac{1}{n} \sum_{i}^{n} (x_i - y_i)^2 = (\mu(\mathbf{x}) - \mu(\mathbf{y}))^2 + (\sigma(\mathbf{x}) - \sigma(\mathbf{y}))^2 + 2\sigma(\mathbf{x})\sigma(\mathbf{y})(1 - \rho(\mathbf{x}, \mathbf{y})) \tag{12}$$

where $\mu$, $\sigma$ and $\rho$ denote respectively the mean, standard deviation and Pearson's correlation coefficient. Thus, absolute mobility is higher the greater is the difference between the marginal distributions of $\mathbf{x}$ and $\mathbf{y}$ (high structural mobility), and the weaker is their association (high exchange mobility).

In practical applications, it is often convenient to use a *mobility index*, that is, a real valued function $A : \mathcal{M} \mapsto \mathbb{R}$, rather than mobility orderings as we have done so far. Theorem 1 then tells us that any mobility index representing a mobility ordering satisfying Axioms 1–4 must rank mobility data in precisely the same way as $d(\mathbf{x}, \mathbf{y})$; in other words, if a mobility index $A$ represents $\preceq$, then there must exist a continuous and increasing function $F : \mathbb{R}_+ \mapsto \mathbb{R}$ such that

$$A(\mathbf{x}, \mathbf{y}) = F\big(d(\mathbf{x}, \mathbf{y})\big). \tag{13}$$

The practical application of the mobility index $A$ requires the choice of a suitable function $F$. Now, it is apparent that $d(\mathbf{x}, \mathbf{y})$ enjoys a very useful practical property, namely *subgroup decomposability*.[7] Then, the reasoning of Foster and Shorrocks [1991], page 696, suggests the appropriate choice of a functional form for $F$: the mobility index $A$ in equation (13) is subgroup decomposable if and only if there are constants $c > 0$ and $r$ such that

$$A(\mathbf{x}, \mathbf{y}) = c \cdot d(\mathbf{x}, \mathbf{y}) + r. \tag{14}$$

Note that if the natural normalization $A(\mathbf{u}, \mathbf{u}) = 0$ is employed, in equation (14) above $r$ must be equal to zero.

Let us call *canonical* a mobility index which is subgroup decomposable and normalized. Hence, an absolute mobility index $A$ is canonical if and only if there is a constant $c > 0$ such that

$$A(\mathbf{x}, \mathbf{y}) = c \cdot d(\mathbf{x}, \mathbf{y}). \tag{15}$$

Since a canonical mobility index $A$ takes into account both exchange and structural mobility, its properties are different from those of the well-known and much employed index of absolute mobility of Fields and Ok [1996]; in particular, as shown in Corollary 1, contrary to Fields and Ok's index, our index has the property that swaps have positive effects on mobility. Finally,

[7] An absolute mobility index $A : \mathcal{M} \mapsto \mathbb{R}$ is *subgroup decomposable* if for all $n, m \in \mathbb{N}$, $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}_n$ and $(\mathbf{x}', \mathbf{y}') \in \mathcal{M}_m$, $A\big((\mathbf{x}, \mathbf{y}) \circ (\mathbf{x}', \mathbf{y}')\big) = \frac{n}{n+m} A(\mathbf{x}, \mathbf{y}) + \frac{m}{n+m} A(\mathbf{x}', \mathbf{y}')$.

notice that our absolute mobility index $A$ is a member of the general class of mobility indices considered by Mitra and Ok [1998], though the structure of the axioms and their motivation is quite different.

### 2.3 A special case

Recall that, as argued in the introduction, the absolute status of the $i$th individual in a generation is given by the value $u_i$ of some aggregating function $\phi$ of a set of indicators $\bar{u}_{i1}, \ldots, \bar{u}_{ik}$ regarding the individual himself taken in isolation, i.e. $u_i = \phi(\bar{u}_{i1}, \ldots, \bar{u}_{ik})$. In many practical applications mobility analysis is conducted in terms of a single indicator, most typically income. In such a context, one possibility would be to take the function $\phi$ as the identity function, that is, to identify absolute status with income. But this choice may be objectionable in the light of the two key axioms which characterize our concept of absolute mobility. For example, some people may not agree that a movement from \$1000 to \$1100 has the same mobility effect as a movement from \$1 to \$101. However, this is what Axiom 1 would say if one identified absolute status with income. Similarly, Axiom 4 would imply that a swap in a society $(\$1, \$3) \circ (\$2, \$4)$ has the same effect of a swap in $(\$1001, \$1003) \circ (\$1002, \$1004)$.

This kind of objection would lead to reject the identification of absolute status with income. Someone may argue that a more appropriate choice would be, for example, to take $\phi$ as the log function so that a movement from $x_i = 1000 = \log(\bar{x}_i)$ to $y_i = 1100 = \log(\bar{y}_i)$ would have the same mobility effect as the movement from $w_i = 1 = \log(\bar{w}_i)$ to $z_i = 101 = \log(\bar{z}_i)$ (and similarly for the swap example). In general, the most appropriate choice of $\phi$ depends on the application context and must be justified independently. Axioms 1 and 4 can therefore be seen as general constraints on this choice (i.e., the resulting notion of absolute status must satisfy these axioms).[8]

In the remaining of this section we give two examples of different criteria

[8] Linking the concept of absolute status to individual welfare, by interpreting $\phi$ as a utility function, could help this process.

which lead to justify different choices for the $\phi$ function when income is the only status indicator. For instance, one commonly shared view is that in assessing the movement of income between a father and a son one should look at *ratios* rather than *differences*: going from \$100 to \$200 has the same mobility effect as a movement from \$1000 to \$2000. However, since the concept of mobility we are endorsing in this paper is symmetric with respect to upward and downward mobility, one would also like to say that a movement from \$100 to \$200 has the same mobility effect as a movement from \$200 to \$100. The next Proposition shows that this concept of mobility forces to identify $\phi$ with the log function:

**Proposition 1** *Let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be strictly increasing. Given Axiom 1, the following two conditions are equivalent:*

1. *For all $\bar{a}, \bar{b}, \bar{c}, \bar{d} > 0$,*

$$\big(\phi(\bar{a}), \phi(\bar{b})\big) \preceq \big(\phi(\bar{c}), \phi(\bar{d})\big) \Leftrightarrow Max(\bar{a}/\bar{b}, \bar{b}/\bar{a}) \leq Max(\bar{c}/\bar{d}, \bar{d}/\bar{c});$$

2. *$\phi = \log$.*

**Proof** Given Axiom 1, the first condition can equivalently be written as:

$$|\phi(\bar{a}) - \phi(\bar{b})| \leq |\phi(\bar{c}) - \phi(\bar{d})| \Leftrightarrow Max(\bar{a}/\bar{b}, \bar{b}/\bar{a}) \leq Max(\bar{c}/\bar{d}, \bar{d}/\bar{c}) \qquad (16)$$

If $\phi = \log$, it is easily checked that the above equivalence is satisfied. To prove the converse, let $x, y > 0$ be two arbitrary positive real numbers and, in (16), let $\bar{a} = x/y$, $\bar{b} = 1$, $\bar{c} = x$ and $\bar{d} = y$. There are two cases:

*Case $x \geq y$*: It follows from (16) that

$$\phi(x/y) - \phi(1) = \phi(x) - \phi(y) \Leftrightarrow (x/y)/1 = x/y;$$

*Case $x < y$*: It follows from (16) that

$$\phi(1) - \phi(x/y) = \phi(y) - \phi(x) \Leftrightarrow 1/(x/y) = y/x.$$

Now let, w.l.o.g., $\phi(1) = 0$. Then in both cases we get

$$\phi(x/y) = \phi(x) - \phi(y). \qquad (17)$$

Let now $z = 1/y$ and use (17) to obtain

$$\phi(xy) = \phi(x/z) = \phi(x) - \phi(z) = \phi(x) - \phi(1/y) = \phi(x) - \phi(1) + \phi(y),$$

that is

$$\phi(xy) = \phi(x) + \phi(y).$$

This is a Cauchy equation whose only solution is as claimed [Aczel, 1966]. □

Proposition 1 tells us that if the absolute status function is taken to be the log of income, then absolute income mobility can be measured by a canonical index such as

$$A(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^{n} \big( \log(\bar{x}_i) - \log(\bar{y}_i) \big)^2$$

which is reminiscent of Fields and Ok's [1999a] scale invariant ("relative" in their terminology) index. On the other hand, if one insists that the absolute distance in incomes is the key for mobility comparisons, that is $\big(\phi(\bar{a}), \phi(\bar{b})\big) \preceq \big(\phi(\bar{c}), \phi(\bar{d})\big) \Leftrightarrow |\bar{a} - \bar{b}| \leq |\bar{c} - \bar{d}|$, then one is forced to conclude that $\phi$ must be an affine function and absolute income mobility is measured by a canonical index such as

$$A(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^{n} (\bar{x}_i - \bar{y}_i)^2.$$

# 3 Relative mobility

In the previous section we have considered the problem of ranking mobility data using the information given by the *absolute* status of the individuals in each society, and have we proposed a mobility measure based on the distance between fathers' and sons' absolute status.

An alternative view of the concept of individual status maintains that, in order to measure the status of a given individual in a population, we should have information not only on the individual taken in isolation, but also on his relative position compared to the other individuals in his generation. Accordingly, *relative mobility* measures the movements of the *relative status* of individuals from one generation to the next.

We have assumed, so far, that the relevant information concerning each single individual in a generation is synthesized by the marginal distribution $\mathbf{u}$ of the absolute status values. With a slight abuse of language, we shall often speak of the "individual" $u_i$ and of the "generation" $\mathbf{u}$. In this context, the relative status of an individual $u_i$ in a generation $\mathbf{u}$ should then be a function $S$ *both* of $u_i$ *and* of his context $\mathbf{u}$.[9]

How then should we evaluate the relative status of $u_i$ in $\mathbf{u}$? How should we compare the relative status of $u_i$ in $\mathbf{u}$ with that of $v_j$ in $\mathbf{v}$? How much information do we need about $\mathbf{u}$ and $\mathbf{v}$? On one extreme, we may insist that we must know *all* the elements of $\mathbf{u}$ and $\mathbf{v}$. On the other, we may think that we need *no* information about the marginal distributions, in which case the very notion of relative status is meaningless and we fall back to the absolute case. In most situations, however, we may be happy with partial information about the marginal distributions in terms of *summarizing statistics*.

A summarizing statistic is formally defined as a permutation invariant function $\tau : \mathbb{R}^* \mapsto \mathbb{R}$ (where $\mathbb{R}^* = \bigcup_{i=1}^{\infty} \mathbb{R}^i$). We shall assume that there is always a fixed *finite amount of statistical information* about marginal distributions (no matter how large they are), which is *sufficient* to make all the necessary discriminations required for the evaluation of relative status. Such fixed set of summarizing statistics, call it $\mathcal{T} = \{\tau_1, \dots, \tau_k\}$, is used as a "sieve" in order to filter out the irrelevant information about marginal distributions. Clearly we want $\mathcal{T}$ to be *non-redundant*:

**Definition 1** *A finite set $\mathcal{T} = \{\tau_1, \dots, \tau_k\}$ of statistics is non-redundant if for no $\tau_i \in \mathcal{T}$ there is a function $h$ such that for all $\mathbf{u}$, $\tau_i(\mathbf{u}) = h(\tau_{i.1}(\mathbf{u}), \dots, \tau_{i.(k-1)}(\mathbf{u}))$, where $i.j = j$ if $j \neq i$, and $i + 1$ if $j = i$.*

The evaluation of relative status always depends on the partition of $\mathbb{R}^*$ induced by the statistical information conveyed by the chosen set $\mathcal{T}$: the larger the amount of information, the finer the partition. Suppose our set of statistics

[9]Recall that each element $u_i$ of $\mathbf{u}$ is itself a transformation of raw data concerning the $i$-th individual, as explained in the introduction.

consists of mean and standard deviations and we know that $\mathbf{u}$ and $\mathbf{v}$ have the same mean and the same standard deviation. Then, we may be willing to conclude that if an individual $u_i$ in $\mathbf{u}$ has the same absolute status as another individual $v_j$ in $\mathbf{v}$, the relative status of $u_i$ is equal to the relative status of $v_j$.

Given some suitable non-redundant set of statistics $\mathcal{T} = \{\tau_1, \ldots, \tau_k\}$, let $\mathbf{u} \equiv_{\mathcal{T}} \mathbf{v}$ mean that the vectors $\mathbf{u}$ and $\mathbf{v}$ are $\mathcal{T}$-equivalent, that is, $\tau_i(\mathbf{u}) = \tau_i(\mathbf{v})$ for all $i = 1, \ldots, k$, and, for any vector $\mathbf{s}$, let $[\mathbf{s}]_{\mathcal{T}}$ denote the class of all vectors which are $\mathcal{T}$-equivalent to $\mathbf{s}$. Thus, each equivalence class $[\mathbf{s}]_{\mathcal{T}}$ is a domain whose elements are indistinguishable from the point of view of the summarizing statistics, and $\mathcal{T}$-equivalent distributions will be treated exactly in the same way by the relative status function. In other words, we maintain that the "context" relative to which the status of an individual is evaluated is not a vector, but an *equivalence class* of vectors identified by the values of the chosen statistics.

**Definition 2** *Let $\mathcal{T}$ be a non-empty and non-redundant finite set of statistics. A* relative status function based on $\mathcal{T}$ *is a function $S : \mathbb{R} \times \mathbb{R}^* \mapsto \mathbb{R}$ such that for all $\mathbf{u} \in \mathbb{R}^*$, all $i \in \{1, \ldots, |\mathbf{u}|\}$ and all $j \in \{1, \ldots, |\mathbf{v}|\}$:*

$$S(u_i, \mathbf{u}) = S(v_j, \mathbf{v})$$

*whenever $u_i = v_j$ and $\mathbf{u} \equiv_{\mathcal{T}} \mathbf{v}$.*

(Observe that if $\mathcal{T}$ were empty, all vectors in $\mathbb{R}^*$ would be $\mathcal{T}$-equivalent, and so the function $S$ would no longer be a *relative* status function, bringing us back to absolute mobility.)

The above definition implies that for every relative status function $S : \mathbb{R} \times \mathbb{R}^* \mapsto \mathbb{R}$ based on $\mathcal{T} = \{\tau_1, \ldots, \tau_k\}$, there exists a function $\Phi : \mathbb{R}^{k+1} \mapsto \mathbb{R}$, such that:

$$S(u_i, \mathbf{u}) = \Phi(u_i, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})). \tag{18}$$

That $\Phi$ defined above is a function follows from the observation that the value of $S(u_i, \mathbf{u})$, in accordance with the definition of a status function S, is uniquely determined by the value of $u_i$ and by the values of the relevant statistics for the vector $\mathbf{u}$. From now on, we shall use the two formulations $S(u_i, \mathbf{u})$ and $\Phi(u_i, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u}))$ interchangeably.

In the context of this paper, we find it convenient to restrict our attention to summarizing statistics which enjoy the two basic properties illustrated in the following definition:

**Definition 3** *A statistic* $\tau$ *is*

1. homogeneous *(of the first degree) if* $\tau(a\mathbf{u}) = a\tau(\mathbf{u})$ *for every vector* $\mathbf{u}$ *and every scalar* $a > 0$*;*

2. weakly additive *if* $\tau(\mathbf{u} + \mathbf{a}) = \tau(\mathbf{u}) + \tau(\mathbf{a})$*, for every vector* $\mathbf{u}$ *and every vector* $\mathbf{a} = [a, a, \ldots, a]$.

This is not a severe restriction since most statistics one may want to use in this context either enjoy the two properties above (such as mean, mode, quantiles, maximum and minimum, standard deviation, interquantile dispersion) or may be expressed as functions of statistics which enjoy them (such as variance, skewness, kurtosis, moments, positive power means).[10] Hence we assume that the set of statistics $\tau_1, \ldots, \tau_k$ in Equation (18) are all homogeneous and weakly additive.

We may look at the relationship between status functions and $\mathcal{T}$ from "the user's end". It is the user who decides what statistical information about the marginal distributions is relevant for the purpose of status evaluation and specifies a relevant set $\mathcal{T}$ which should be used in defining a suitable status function.

As argued above, relative mobility measures the movements of relative status, and therefore the relative mobility of a society $(\mathbf{x}, \mathbf{y})$ should be expressed

[10] To mention just one example, the $m$-order moment $\frac{1}{n}\sum_{i=1}^{n} x_i^m$ can be expressed as a function of the mean $\mu(\mathbf{x})$ and the set of statistics $(\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu(\mathbf{x}))^j)^{1/j}$ for $j = 2, \cdots, m$, which enjoy our two properties.

by some suitable function of $S(\mathbf{x})$ and $S(\mathbf{y})$, where we abbreviate with $S(\mathbf{u})$ the result of applying the relative status function to all the elements of $\mathbf{u}$:

$$\begin{aligned} S(\mathbf{u}) &= \big(S(u_1, \mathbf{u}), \ldots, S(u_n, \mathbf{u})\big) \\ &= \big(\Phi(u_1, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})), \ldots, \Phi(u_n, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u}))\big). \end{aligned} \tag{19}$$

**Definition 4** *Let $\mathcal{T}$ be a non-empty and non-redundant finite set of statistics for $\mathbb{R}^*$ and $S$ a relative status function based on $\mathcal{T}$. A* relative mobility index based on $S$ *is a function $R : \mathcal{M} \mapsto R$ such that*

$$R(\mathbf{x}, \mathbf{y}) = \Psi(S(\mathbf{x}), S(\mathbf{y}))$$

*for some function $\Psi : \mathcal{M} \mapsto R$.*

Thus, in our approach, measuring relative mobility amounts to specifying:

1. a suitable non-empty and non-redundant finite set $\mathcal{T}$ of homogeneous and weakly additive statistics, which we call the *relevant set*;

2. a suitable relative status function $S$ based on $\mathcal{T}$;

3. a suitable relative mobility index $R$ based on $S$.

We shall refer to such a triple $(\mathcal{T}, S, R)$ as to a *measurement system for relative mobility* (*R-system* for short). There are several questions concerning R-systems, whose answer will ultimately depend on the axioms we shall impose on them:

- Can any relevant set $\mathcal{T}$ be part of an R-system? How free are we in determining what kind of information is relevant to describe marginal distributions?

- The extension of the relevant set $\mathcal{T}$ of an R-system — that is the consideration of some additional statistic as relevant for the purpose of status evaluation — leads to a finer partition of the space of all marginal distributions, and therefore requires the specification of a new status function. The new R-system that results from this revision will typically overthrow the assessments made on the basis of the old R-system, so making the measurement of relative mobility potentially very unstable. Consideration of this problem leads to the following question: are there *maximal* relevant sets? Is there a *maximal* amount of statistical information we can consistently use in an R-system?

- Given a relevant set $\mathcal{T}$, how free are we in determining an R-system that includes $\mathcal{T}$? In other words, how free are we in determining the form of the relative status function $S$ and of the relevant mobility index $R$?

Under a few natural axioms about relative mobility systems, all these questions receive a precise answer. First, *there are* (very small) relevant sets of statistics which provide us with *maximal information* about the marginal distributions, as rich as it can be consistently used within an R-system. Moreover, such maximal relevant sets dictate the form of the relative status function; the latter, in turn, dictates the form of the relative mobility index $R$. Hence, the choice of a maximal relevant set of statistics *forces* the adoption of an R-system whose relative mobility index is uniquely determined.

In the analysis of relative mobility developed in the next sections we restrict ourselves to the case where marginal distributions are *non-degenerate*, i.e. such that not all of their elements are equal. The technical reason of this restriction will be apparent after the proof of the main result of this section. At this stage, we just notice that when marginal distributions are degenerate, the very notion of "relative status" becomes elusive. Hence, in the sequel, we shall consider only societies $(\mathbf{x}, \mathbf{y})$ such that their marginal distributions $\mathbf{x}$ and $\mathbf{y}$ are non-degenerate. Observe that, in accordance with this restriction, we shall deal with marginal distributions containing at least two elements (since one-element marginal distributions are, by definition, degenerate).

## 3.1 Axioms

Our axioms on R-systems are divided into two groups. The first group of axioms focuses on the relative status function, while the second group are "consistency axioms": they specify the domains in which the relative mobility index of an R-system must be consistent with the absolute mobility index axiomatized in the first part of the paper.

### 3.1.1 Axioms on the relative status function

The first axiom states the uncontroversial assumption that *ceteris paribus*, i.e. within $\mathcal{T}$-equivalent generational contexts, the relative status of an individual grows with his absolute status:

**Axiom 5** *For every R-system $(\mathcal{T}, S, R)$, every non-degenerate $\mathbf{u}, \mathbf{v} \in \mathbb{R}^*$ and every $i, j$ $(1 \leq i \leq |\mathbf{u}|,\ 1 \leq j \leq |\mathbf{v}|)$,*

$$u_i > v_j \text{ and } \mathbf{u} \equiv_{\mathcal{T}} \mathbf{v} \Rightarrow S(u_i, \mathbf{u}) > S(v_j, \mathbf{v}).$$

Our second axiom states that the relative status function $S$ of an R-system should be "sensitive" to the value of each single statistic. Suppose, for instance, that the relevant set of the R-system contains only mean and standard deviation and that $\mathbf{u}$ and $\mathbf{v}$ are two vectors with the same mean but different standard deviation. Suppose also that for some $i, j$, $u_i = v_j$. It is natural to expect that $S$ is sensitive enough to deliver different relative status values for these arguments: *ceteris paribus* the difference in standard deviation should yield a difference in the respective relative status. The next axiom is just a generalization of this assumption.

**Axiom 6** *For every R-system $(\mathcal{T}, S, R)$, every non-degenerate $\mathbf{u}, \mathbf{v} \in \mathbb{R}^*$ and every $i, j$ $(0 \leq i \leq |\mathbf{u}|, 0 \leq j \leq |\mathbf{v}|)$, if*

- $u_i = v_j$,
- *for some $\tau \in \mathcal{T}$, $\mathbf{u} \equiv_{\mathcal{T}\setminus\{\tau\}} \mathbf{v}$ and $\mathbf{u} \not\equiv_{\mathcal{T}} \mathbf{v}$,*

*then* $S(u_i, \mathbf{u}) \neq S(v_j, \mathbf{v})$.

The last axiom of the first group states that the relative status function of an R-system should be *stable*, i.e. that for all $\mathbf{u}$, $S(\mathbf{u})$ is a fixpoint for $S$. This simply means that we should expect the application of $S$ to all the elements of a vector $\mathbf{u}$ to yield a standardized context for the measurement of mobility, so that re-applying the relative status function to such standardized context is of no effect.

**Axiom 7** *For every R-system* $(\mathcal{T}, S, R)$ *and every non-degenerate* $\mathbf{u} \in \mathbb{R}^*$, $S(S(\mathbf{u})) = S(\mathbf{u})$.

### 3.1.2 Consistency axioms

To introduce our first consistency axiom, consider the domain consisting of all the pairs $(\mathbf{x}, \mathbf{y})$ such that $S(\mathbf{x}) = \mathbf{x}$ and $S(\mathbf{y}) = \mathbf{y}$, i.e. such that the relative status of all individuals coincide with their absolute status. Since the whole dichotomy relative *vs.* absolute mobility is motivated by the distinction between relative and absolute status, it is natural to require that within the domain in which this distinction is immaterial, relative mobility should coincide with absolute mobility. Our axiom therefore requires that, within this domain, the metric defined by any relative mobility index $R$ should be equal to the metric defined by some (canonical) absolute mobility index $A$.

**Axiom 8** *For every R-system* $(\mathcal{T}, S, R)$ *and for all* $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}$ *such that* $\mathbf{x}, \mathbf{y}$ *are non-degenerate marginal distributions with* $S(\mathbf{x}) = \mathbf{x}$ *and* $S(\mathbf{y}) = \mathbf{y}$,

$$R(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y})$$

*for some (canonical) absolute index* $A$.

As for our second consistency axiom, consider an R-system $(\mathcal{T}, S, R)$ and suppose that we are comparing two societies $(\mathbf{x}, \mathbf{y}), (\mathbf{w}, \mathbf{z}) \in \mathcal{M}$, where $\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{z}$ are non-degenerate marginal distributions that are all $\mathcal{T}$-equivalent, i.e. indistinguishable from the point of view of the summarizing statistics. Under these circumstances a (canonical) *absolute index* $A$ should be perfectly suitable, since relative mobility really comes into play only when dealing with societies with different summarizing statistics. On the other hand, the *relative index* $R$ of the R-system, should also be perfectly suitable, since we expect $R$ to measure correctly the mobility of *all* societies, obviously including those that share the same statistics. In other words, it is reasonable to require that, within any of the equivalence classes induced by the equivalence relation $\equiv_{\mathcal{T}}$, the metric defined by the relative index be equal to the metric defined by some (canonical) absolute index. Let $\mathbb{R}^*/\equiv_{\mathcal{T}}$ be the partition of $\mathbb{R}^*$ induced by the equivalence relation $\equiv_{\mathcal{T}}$

**Axiom 9** *For every R-system* $(\mathcal{T}, S, R)$*, every equivalence class* $E \in \mathbb{R}^*/\equiv_{\mathcal{T}}$ *and all* $(\mathbf{x}, \mathbf{y}) \in \mathcal{M}$ *such that* $\mathbf{x}, \mathbf{y}$ *are non-degenerate marginal distributions in* $E$*,*

$$R(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y})$$

*for some canonical absolute index* $A$*.*

## 3.2 Preliminary Results

In order to prove the main theorem of this section we need a definition and three preliminary lemmas.

**Definition 5** *A statistic* $\tau$ *is*

- *a* location statistic *if* $\tau(\mathbf{u} + \mathbf{a}) = \tau(\mathbf{u}) + a$ *for every vector* $\mathbf{u}$ *and every vector* $\mathbf{a} = [a, a, \ldots, a]$*;*
- *a* dispersion statistic *if* $\tau(\mathbf{a}) = 0$ *for every vector* $\mathbf{a} = [a, a, \ldots, a]$*.*

**Lemma 1** *Let* $\mathcal{T}$ *be a relevant set of statistics. We can assume without loss of generality that every statistic* $\tau \in \mathcal{T}$ *is either a dispersion or a location statistic.*

**Proof** Let $\tau$ be any statistic in $\mathcal{T}$. There are two cases: if $\tau(\mathbf{1}) = 0$ then $\tau$ is a dispersion statistic, since we assume that all statistics are homogeneous of first degree and, therefore, $\tau(\mathbf{a}) = a\tau(\mathbf{1}) = 0$; on the other hand, if $\tau(\mathbf{1}) = c \neq 0$, then $\tau(\mathbf{a}) = a\tau(\mathbf{1}) = ca$ and so $\frac{1}{c}\tau(\mathbf{a}) = a$. Therefore, by weak additivity, $\frac{1}{c}\tau(\mathbf{x}+\mathbf{a}) = \frac{1}{c}\tau(\mathbf{x}) + \frac{1}{c}\tau(\mathbf{a}) = \frac{1}{c}\tau(\mathbf{x}) + a$, and so $\frac{1}{c}\tau$ is a location statistic. Now, since $\frac{1}{c}\tau(\mathbf{x})$ is a function of $\tau(\mathbf{x})$, the set $\mathcal{T}'$, obtained from $\mathcal{T}$ by replacing the statistic $\tau$ with $\frac{1}{c}\tau$, is still a relevant set and the partition induced by $\mathcal{T}'$ on $\mathbb{R}^*$ is the same as that induced by $\mathcal{T}$. $\square$

**Lemma 2** *Axioms 7 and 8 imply that, for all non-degenerate* $\mathbf{x}$ *and* $\mathbf{y}$*:*

$$R(\mathbf{x},\mathbf{y}) = c \cdot d\big(S(\mathbf{x}), S(\mathbf{y})\big)$$

*for some constant* $c > 0$*.*

**Proof** By Axiom 8 and equation 15, for all non-degenerate $\mathbf{x}, \mathbf{y}$

$$R\big(S(\mathbf{x}), S(\mathbf{y})\big) = c \cdot d\big(S(\mathbf{x}), S(\mathbf{y})\big)$$

for some $c > 0$. By Definition 4 and Axiom 7, for some function $\Psi$:

$$\begin{aligned} R\big(S(\mathbf{x}), S(\mathbf{y})\big) &= \Psi\big(S(S(\mathbf{x})), S(S(\mathbf{y}))\big) \\ &= \Psi\big(S(\mathbf{x}), S(\mathbf{y})\big) \\ &= R(\mathbf{x},\mathbf{y}). \end{aligned}$$

$\square$

**Lemma 3** *Let* $(\mathcal{T}, S, R)$ *be an R-system, with* $\mathcal{T} = \{\tau_1, \ldots, \tau_k\}$*. Axioms 5–9 imply that, for all non-degenerate* $\mathbf{u}$ *and all* $i = 1, \cdots, |\mathbf{u}|$*,*

$$S(u_i, \mathbf{u}) = \alpha[\tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})]u_i + \beta[\tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})]$$

*for some function* $\alpha : \mathbb{R}^k \mapsto \mathbb{R}_+$ *and some function* $\beta : \mathbb{R}^k \mapsto \mathbb{R}$*.*

**Proof** Recall (see Equation 18 above) that

$$S(u_i, \mathbf{u}) = \Phi(u_i, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u}))$$

for some function $\Phi$. The lemma says that if Axioms 5–9 hold true, then $\Phi$ is an affine function of $u_i$ whenever $\tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})$ are given. Suppose that $\Phi$ is not affine. Then there must exist a vector $\mathbf{u}$ which contains three elements $u_1 < u_2 < u_3$ such that:

$$[\Phi(u_2, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_1, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u}))]/(u_2 - u_1) \neq \\ \neq [\Phi(u_3, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_2, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u}))]/(u_3 - u_2). \quad (20)$$

Suppose, without loss of generality, that $\mathbf{u} = (u_1, u_2, u_3)\mathbf{v}$, let $\mathbf{u}'$ and $\mathbf{u}''$ be permutations of $\mathbf{u}$ such that $\mathbf{u}' = (u_2, u_1, u_3)\mathbf{v}$ and $\mathbf{u}'' = (u_1, u_3, u_2)\mathbf{v}$. Then since $\mathbf{u} \equiv_{\mathcal{T}} \mathbf{u}' \equiv_{\mathcal{T}} \mathbf{u}''$ we have

$$\begin{aligned}\Phi(u_i, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) &= \Phi(u_i, \tau_1(\mathbf{u}'), \ldots, \tau_k(\mathbf{u}')) \\ &= \Phi(u_i, \tau_1(\mathbf{u}''), \ldots, \tau_k(\mathbf{u}'')).\end{aligned} \quad (21)$$

Then, (20) above can be written as:

$$(\Phi(u_2, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_1, \tau_1(\mathbf{u}'), \ldots, \tau_k(\mathbf{u}')))/(u_2 - u_1) \neq \\ \neq (\Phi(u_3, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_2, \tau_1(\mathbf{u}''), \ldots, \tau_k(\mathbf{u}'')))/(u_3 - u_2) \quad (22)$$

so that squaring both sides (recall that by Axiom 5, both sides of (22) are positive) we obtain:

$$(\Phi(u_2, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_1, \tau_1(\mathbf{u}'), \ldots, \tau_k(\mathbf{u}')))^2/(u_2 - u_1)^2 \neq \\ \neq (\Phi(u_3, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_2, \tau_1(\mathbf{u}''), \ldots, \tau_k(\mathbf{u}'')))^2/(u_3 - u_2)^2. \quad (23)$$

Now, observe that (by Lemma 2):

$$R(\mathbf{u}, \mathbf{u}') = 2\frac{c}{k}(\Phi(u_2, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_1, \tau_1(\mathbf{u}'), \ldots, \tau_k(\mathbf{u}')))^2$$

and

$$R(\mathbf{u}, \mathbf{u}'') = 2\frac{c}{k}(\Phi(u_3, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) - \Phi(u_2, \tau_1(\mathbf{u}''), \ldots, \tau_k(\mathbf{u}'')))^2$$

for some constant $c > 0$, where $k = |\mathbf{u}| = |\mathbf{u}'| = |\mathbf{u}''|$. Hence, (23) can be rewritten as:

$$R(\mathbf{u}, \mathbf{u}')/(u_2 - u_1)^2 \neq R(\mathbf{u}, \mathbf{u}'')/(u_3 - u_2)^2. \quad (24)$$

Now, since $\mathbf{u} \equiv_{\mathcal{T}} \mathbf{u}' \equiv_{\mathcal{T}} \mathbf{u}''$, it follows from Axiom 9 that for some constant $c' > 0$ we have

$$R(\mathbf{u}, \mathbf{u}') = c' d(\mathbf{u}, \mathbf{u}'),$$

and

$$R(\mathbf{u}, \mathbf{u}'') = c' d(\mathbf{u}, \mathbf{u}'').$$

But, since $d(\mathbf{u}, \mathbf{u}') = \frac{2}{k}(u_2 - u_1)^2$ and $d(\mathbf{u}, \mathbf{u}'') = \frac{2}{k}(u_3 - u_2)^2$, we have from (24)

$$\frac{(u_2 - u_1)^2}{(u_2 - u_1)^2} \neq \frac{(u_3 - u_2)^2}{(u_3 - u_2)^2}$$

which is a contradiction. □

**Lemma 4** *Let $(\mathcal{T}, S, R)$ be an R-system. Axioms 5–9 imply that for every statistic $\tau \in \mathcal{T}$, there is a constant $c$ such that for all non-degenerate $\mathbf{u}$, $\tau(S(\mathbf{u})) = c$.*

**Proof** Let $(\mathcal{T}, S, R)$ be an arbitrary R-system with $\mathcal{T} = \{\tau_1, \ldots, \tau_k\}$. By Lemma 3 and Definition 4, there are functions $\Phi$, $\alpha_0$ and $\beta_0$ such that, for all non-degenerate $\mathbf{u}$ and all $i = 1, \cdots, |\mathbf{u}|$,

$$S(u_i, \mathbf{u}) = \Phi(u_i, \tau_1(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) = \alpha_0\big[\mathcal{T}(\mathbf{u})\big]u_i + \beta_0\big[\mathcal{T}(\mathbf{u})\big]. \quad (25)$$

Since, by Axiom 6, $\Phi$ is one-to-one in each argument $\tau_j(\mathbf{u})$, we have

$$\tau_1(\mathbf{u}) = g_1(S(u_i, \mathbf{u}), \tau_2(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) \quad (26)$$

for some function $g_1$. Let

$$\xi_0(\mathbf{u}) = \alpha_0(\mathcal{T}(\mathbf{u}))$$

and

$$\chi_0(\mathbf{u}) = \beta_0(\mathcal{T}(\mathbf{u})).$$

From (25) and (26), considering that $\tau_1(\mathbf{u})$ in (26) is independent of any particular element $u_i$ of $\mathbf{u}$, we obtain:

$$\tau_1(\mathbf{u}) = h_1(\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \tau_2(\mathbf{u}), \ldots, \tau_k(\mathbf{u})) \quad (27)$$

for some function $h_1$. Hence, from (25), there are functions $\alpha_1$ and $\beta_1$ such that:

$$S(u_i, \mathbf{u}) = \Phi(u_i, \tau_1(\mathbf{u}), \dots, \tau_k(\mathbf{u})) = \\ = \alpha_1[\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \tau_2(\mathbf{u}), \dots, \tau_k(\mathbf{u})]u_i + \beta_1[\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \tau_2(\mathbf{u}), \dots, \tau_k(\mathbf{u})]. \quad (28)$$

Letting

$$\xi_1(\mathbf{u}) = \alpha_1[\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \tau_2(\mathbf{u}), \dots, \tau_k(\mathbf{u})]$$

and

$$\chi_1(\mathbf{u}) = \beta_1[\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \tau_2(\mathbf{u}), \dots, \tau_k(\mathbf{u})],$$

by the same argument that led us from (25) to (27), and one obtains:

$$\tau_2(\mathbf{u}) = h_2(\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \xi_1(\mathbf{u}), \chi_1(\mathbf{u}), \tau_3(\mathbf{u}), \dots, \tau_k(\mathbf{u})), \quad (29)$$

and so on for every $\tau_j$ until:

$$\tau_k(\mathbf{u}) = h_k(\xi_0(\mathbf{u}), \chi_0(\mathbf{u}), \xi_1(\mathbf{u}), \chi_1(\mathbf{u}), \xi_{k-1}(\mathbf{u}), \chi_{k-1}(\mathbf{u})). \quad (30)$$

Observe that, by definition of $\xi_j$ and $\chi_j$,

$$S(u_i, \mathbf{u}) = \Phi(u_i, \tau_1(\mathbf{u}), \dots, \tau_k(\mathbf{u})) = \xi_j(\mathbf{u})u_i + \chi_j(\mathbf{u})$$

for every $j = 1, \dots, k-1$.

Now, let us consider $\tau_k(S(\mathbf{u}))$. Axiom 7 implies that

$$\xi_j(S(\mathbf{u}))S(u_i, \mathbf{u}) + \chi_j(S(\mathbf{u})) = S(u_i, \mathbf{u}).$$

This is possible only if

$$\xi_j(S(\mathbf{u})) = 1 \text{ and } \chi_j(S(\mathbf{u})) = 0,$$

for every $j = 1, \dots, k-1$. Then, (30) implies that $\tau_k(S(\mathbf{u}))$ must always be equal to a constant, say $c_k$. By replacing $\tau_k(S(\mathbf{u}))$ with its constant value $c_k$ in the equation for $\tau_{k-1}(S(\mathbf{u}))$, we obtain that also $\tau_{k-1}(S(\mathbf{u}))$ must always be equal to a constant $c_{k-1}$, and so on for all the others $\tau_i \in \mathcal{T}$. $\square$

# 4 Main result

We have presented relative status functions as dependent on the choice of what we have called a relevant set of statistics, in such a way that different choices may yield very different relative status functions. As we discussed above, this means that considering some additional information about the population (i.e. expanding the set of statistics employed to describe it) may defeat previous assessments about the relative status of individuals, so making our relative mobility measures potentially very unstable. As already explained, this problem would be partly solved if we were able to show that some sets of statistics are *maximal* relevant sets.

**Definition 6** *We say that a set $\mathcal{T}$ of statistics is a* maximal relevant set *if $\mathcal{T}$ is a relevant set and, under Axioms 5–9, there exists no R-system $(\mathcal{T}', S, R)$ such that $\mathcal{T} \subset \mathcal{T}'$. We say that an R-system $(\mathcal{T}, S, R)$ is* robust *if $\mathcal{T}$ is a maximal relevant set.*

Maximal relevant sets represent upper bounds on the statistical information that can be processed, consistently with our axioms, by an R-system. So, if we were able to specify a maximal relevant set, this would restrict considerably our freedom (as well as our disagreement) about what information should be used in determining relative status. Given a maximal relevant set we might still be left with some freedom (and so with a potential source of disagreement) about the mathematical form of the status function. As suggested above, the strongest possible outcome would be that maximal relevant sets *dictate* the form of the status function and of the relative mobility index.

The main result of this section is stated in the following:

**Theorem 2** *The triple $(\mathcal{T}, S, R)$ is a robust R-system satisfying Axioms 5–9 if and only if*

1. *$\mathcal{T} = \{\delta, \lambda\}$ for some dispersion statistic $\delta$ and some location statistic $\lambda$.*

*2. For every non-degenerate* $\mathbf{u} \in \mathbb{R}^*$*:*

$$S(u_i, \mathbf{u}) = \frac{c(u_i - \lambda(\mathbf{u}))}{\delta(\mathbf{u})} + d$$

*for some constants* $c > 0$ *and* $d$*.*

*3. For every* $(\mathbf{x}, \mathbf{y})$ *in* $\mathcal{M}_n$*, such that* $\mathbf{x}$ *and* $\mathbf{y}$ *are non-degenerate:*

$$R(\mathbf{x}, \mathbf{y}) = \frac{c}{n} \cdot \sum_i^n \left( \frac{x_i - \lambda(\mathbf{x})}{\delta(\mathbf{x})} - \frac{y_i - \lambda(\mathbf{y})}{\delta(\mathbf{y})} \right)^2 \tag{31}$$

*for some constant* $c > 0$*.*

**Proof** Let us assume that $(\mathcal{T}, S, R)$ is a robust R-system satisfying Axioms 5–9. Recall that, by Lemma 2,

$$R(\mathbf{x}, \mathbf{y}) = c \cdot d(S(\mathbf{x}), S(\mathbf{y})).$$

Assume that the statistics in the maximal relevant set $\mathcal{T}$ have been ordered in some arbitrary way, that is $\mathcal{T} = \{\tau_1, \dots, \tau_k\}$. Given Lemma 1, we may distinguish two cases.

*Case 1.* The first statistic in $\mathcal{T}$ is a dispersion statistic, say $\delta$. Since $\delta$ is homogeneous of first degree and weakly additive then, using Lemma 3 and Lemma 4:

$$\begin{aligned} \delta\big(S(\mathbf{u})\big) = \delta(\alpha_{\mathcal{T}}(\mathbf{u})\mathbf{u} + \beta_{\mathcal{T}}(\mathbf{u})) &= \delta(\alpha_{\mathcal{T}}(\mathbf{u})\mathbf{u}) + \delta(\beta_{\mathcal{T}}(\mathbf{u})) \\ &= \alpha_{\mathcal{T}}(\mathbf{u})\delta(\mathbf{u}) \\ &= c \end{aligned} \tag{32}$$

for some constant $c$, where $\alpha_{\mathcal{T}}(\mathbf{u})$ denotes $\alpha[\tau_1(\mathbf{u}), \dots, \tau_k(\mathbf{u})]$ and $\beta_{\mathcal{T}}(\mathbf{u})$ denotes $\beta[\tau_1(\mathbf{u}), \dots, \tau_k(\mathbf{u})]$. Hence $\alpha_{\mathcal{T}}(\mathbf{u}) = c/\delta(\mathbf{u})$.

Observe that, if $\delta'$ is another dispersion statistic in $\mathcal{T}$, then $\alpha_{\mathcal{T}}(\mathbf{u}) = c'/\delta'(\mathbf{u})$ for some constant $c'$ and so $\delta'(\mathbf{u}) = \frac{c'}{c}\delta(\mathbf{u})$. But $\delta'$ is a function of $\delta$. So since $\mathcal{T}$ is non-redundant and $\delta \in \mathcal{T}$, then $\delta' \notin \mathcal{T}$. Therefore, $\mathcal{T}$ cannot contain any other dispersion statistic.

Therefore, by Lemma 1, the next statistic in $\mathcal{T}$, must be a location statistic. Let us call it $\lambda$. Again, since $\lambda$ is homogeneous of first degree and weakly additive, we have that:

$$\begin{aligned} \lambda\big(S(\mathbf{u})\big) = \lambda(\alpha_{\mathcal{T}}(\mathbf{u})\mathbf{u} + \beta_{\mathcal{T}}(\mathbf{u})) &= \lambda(\alpha_{\mathcal{T}}(\mathbf{u})\mathbf{u}) + \lambda(\beta_{\mathcal{T}}(\mathbf{u})) \\ &= \alpha_{\mathcal{T}}(\mathbf{u})\lambda(\mathbf{u}) + \beta_{\mathcal{T}}(\mathbf{u}) \\ &= d \end{aligned} \tag{33}$$

for some constant $d$.

Hence, $\beta_{\mathcal{T}}(\mathbf{u}) = d - \alpha_{\mathcal{T}}(\mathbf{u})\lambda(\mathbf{u})$ and, replacing $\alpha_{\mathcal{T}}(\mathbf{u})$ with its value obtained above, $\beta_{\mathcal{T}}(\mathbf{u}) = d - \frac{c}{\delta(\mathbf{u})}\lambda(\mathbf{u})$. So:

$$S(u_i, \mathbf{u}) = \alpha_{\mathcal{T}}(\mathbf{u})u_i + \beta_{\mathcal{T}}(\mathbf{u}) = \frac{c(u_i - \lambda(\mathbf{u}))}{\delta(\mathbf{u})} + d.$$

Since at this stage the relative status function is completely determined, the maximal relevant set $\mathcal{T}$ cannot contain any other statistic.

*Case 2*: The first statistic in $\mathcal{T}$ is a location statistic. Let us call it $\lambda$. Then, since $\lambda$ is homogeneous of first degree and weakly additive, we have, as in (33) above, that $\beta_{\mathcal{T}}(\mathbf{u}) = d - \alpha_{\mathcal{T}}(\mathbf{u})\lambda(\mathbf{u})$. Let us, then, try to determine $\alpha_{\mathcal{T}}(\mathbf{u})$. Now, if the next statistic in $\mathcal{T}$ (which must be independent of $\lambda$) is a dispersion statistic, then by (32) above, we fall back to case 1. Suppose, then, that the next statistic in $\mathcal{T}$ is a location statistic, call it $\lambda'$. Then, using again the equations (33) above, with $\lambda'$ instead of $\lambda$ and a new constant $e$ instead of $d$, we obtain $\beta_{\mathcal{T}}(\mathbf{u}) = e - \alpha_{\mathcal{T}}(\mathbf{u})\lambda'(\mathbf{u})$. So

$$d - \alpha_{\mathcal{T}}(\mathbf{u})\lambda(\mathbf{u}) = e - \alpha_{\mathcal{T}}(\mathbf{u})\lambda'(\mathbf{u})$$

and

$$\alpha_{\mathcal{T}}(\mathbf{u}) = \frac{e - d}{\lambda'(\mathbf{u}) - \lambda(\mathbf{u})}.$$

Now, observe that, since $\lambda'$ and $\lambda$ are location statistics, their difference is a dispersion statistic, for

$$\begin{aligned} \lambda(\mathbf{x} + \mathbf{a}) - \lambda'(\mathbf{x} + \mathbf{a}) &= \lambda(\mathbf{x}) + \lambda(\mathbf{a}) - \lambda'(\mathbf{x}) - \lambda'(\mathbf{a}) \\ &= \lambda(\mathbf{x}) + a - \lambda'(\mathbf{x}) - a \\ &= \lambda(\mathbf{x}) - \lambda'(\mathbf{x}) \end{aligned} \tag{34}$$

Let $\delta(\mathbf{u}) = \lambda'(\mathbf{u}) - \lambda(\mathbf{u})$ and let $c = e - d$, then we have again that:

$$S(u_i, \mathbf{u}) = \alpha_{\mathcal{T}}(\mathbf{u})u_i + \beta_{\mathcal{T}}(\mathbf{u}) = \frac{c(u_i - \lambda(\mathbf{u}))}{\delta(\mathbf{u})} + d.$$

Since $S$ is again completely determined, the maximal relevant set $\mathcal{T}$ cannot contain any other statistic.

We leave it to the reader to verify that the proposed R-system satisfies all our axioms and also, on the basis of the above argument, that it is a robust R-system. This concludes the proof of Theorem 2. □

Theorem 2 shows that any robust R-system that satisfies our axioms, must have a rather precise form. This theorem characterizes a *class* of relative mobility indices, depending on the choice of the statistics in the maximal relevant set. For example, if we choose $\lambda(\mathbf{u})$ to be equal to the mean of $\mathbf{u}$ and $\delta(\mathbf{u})$ to be equal to its standard deviation, the resulting relative mobility index will be:

$$R(\mathbf{x}, \mathbf{y}) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i - \mu(\mathbf{x})}{\sigma(\mathbf{x})} - \frac{y_i - \mu(\mathbf{y})}{\sigma(\mathbf{y})}\right]^2$$

and simple manipulation shows that using such $R$ is equivalent to using Pearson's correlation coefficient $\rho$, once we adjust for the fact that the correlation coefficient actually measures immobility rather than mobility.[11]

Use of the correlation coefficient to measure immobility dates back at least to Galton's model [Galton, 1886] for analyzing the intergenerational transmission of hereditary stature, and it is fair to say that modern versions of Galton's model, based on OLS regression of sons' on fathers' income, are the workhorse of applied mobility analyses by economists, see e.g. Zimmerman [1992] and Solon [2000]. Thus, our class of relative mobility indices gives an axiomatic base to this practice.

Of course, the choice of the mean and standard deviation as elements of a relevant set of statistics yields only one member of our class; for instance, if we choose $\lambda(\mathbf{u})$ to be equal to $\min(\mathbf{u})$ and $\delta(\mathbf{u})$ to be equal to $\max(\mathbf{u}) - \min(\mathbf{u})$,

[11]Use equation (12) and recall that $\rho(\mathbf{x}, \mathbf{y})$ is invariant to affine transformations of $\mathbf{x}$ and $\mathbf{y}$ to get $R(\mathbf{x}, \mathbf{y}) = 2(1 - \rho(\mathbf{x}, \mathbf{y}))$, so that $\rho(\mathbf{x}, \mathbf{y}) = 1 - \frac{R(\mathbf{x}, \mathbf{y})}{2}$.

then we get a relative mobility index

$$R(\mathbf{x},\mathbf{y}) = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} - \frac{y_i - \min(\mathbf{y})}{\max(\mathbf{y}) - \min(\mathbf{y})}\right]^2$$

which measures the distance between fathers and sons' status after standardizing the relative status $S$ to lie in the $[0, 1]$ interval.

To conclude, the actual choice of a mobility index which satisfies the properties considered in this paper can be seen as a process which involves:

1. In the preliminary phase, the choice of an aggregating function $\phi$ to turn a set of status indicators into a suitable real number representing the "absolute status" of a given individual. If the researcher feels it appropriate to use a single indicator such as income, then the discussion in Section 2.3 and in particular Proposition 1 may help in linking the chosen indicator to absolute status.

2. Having found appropriate vectors for describing the absolute status of fathers and sons in a society, if one is interested in measuring absolute mobility, and finds the set of axioms discussed in Section 2 compelling, then Theorem 1 forces the use of the mean squared distance as a canonical absolute mobility index.

3. If the researcher is interested in capturing relative mobility, in the sense endorsed in this paper, and finds the set of axioms discussed in Section 3 compelling, then Theorem 2 states that two statistics (a location one and a dispersion one) are sufficient to summarize relative status in each generation, and forces the use of the mean squared distance in relative status as a canonical relative mobility index.

As an example, the choice of using an OLS regression model with log-incomes to measure mobility means that absolute status is represented by the log of income, and that mean and standard deviation of log-incomes are used as summarizing statistics.

# References

[Aczel, 1966] Aczel, J. (1966). *Lectures on functional equations and their applications*. Accademic Press, New York.

[Alesina and La Ferrara, 2005] Alesina, A. and La Ferrara, E. (2005). Preference for redistribution in the land of opportunities. *Journal of Public Economics* 89:897–932.

[Atkinson, 1983] Atkinson, A.B. (1983). The measurement of economic mobility. In Atkinson, A.B., editor, *Social Justice and Public Policy*. Wheatsheaf Books Ltd., London.

[Benabou and Ok, 2001a] Benabou, R. and Ok, E. (2001a). Social mobility and the demand for redistribution. *Quarterly Journal of Economics* 116:447–87.

[Benabou and Ok, 2001b] Benabou, R. and Ok, E. (2001b). Mobility as progressivity: Ranking income processes according to equality of opportunity. *NBER working paper* 4381.

[Benabou, 2002] Benabou, R (2002). Tax and education policy in a heterogeneous agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica* 70:481–517.

[Chakravarty *et al.*, 1985] Chakravarty, S.R.; B., Dutta; and Weymark, J. A. (1985). Ethical indices of income mobility. *Social Choice and Welfare* 2:1–21.

[Checchi *et al.*, 1999] Checchi, D.; A., Ichino; and A.Rustichini, (1999). More equal but less mobile ? intergenerational mobility and inequality in italy and in the us. *Journal of Public Economics* 74:351–393.

[Conlisk, 1990] Conlisk, J. (1990). Monotone mobility matrices. *Journal of Mathematical Sociology* 15:173–191.

[Cowell, 1985] Cowell, F. (1985). Measures of distributional change: An axiomatic approach. *Review of Economic Studies* 52:35–51.

[Dardanoni, 1993] Dardanoni, V. (1993). Measuring social mobility. *Journal of Economic Theory* 61:372–94.

[Debreu, 1954] Debreu, G. (1954). Representation of a preference ordering by a numerical function. In R. M. Thrall, C.H.Coombs and Davis, R.L., editors, *Decision Processes.* Wiley, New York.

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

[Fields and Ok, 1996] Fields, G.S. and Ok, E. (1996). The meaning and measurement of income mobility. *Journal of Economic Theory* 71:349–77.

[Fields and Ok, 1999a] Fields, G.S. and Ok, E. (1999a). Measuring movements of incomes. *Economica* 66:455–71.

[Fields and Ok, 1999b] Fields, G.S. and Ok, E. (1999b). The measurement of income mobility. In Silber, J., editor, *Handbook of Income Inequality Measurement.* Kluwer Academic Publishers, Dordrecht.

[Formby *et al.*, 2003] Formby, J.P.; W.J., Smith; and Zheng, B (2003). Economic growth, welfare and the measurement of social mobility. In Amiel, Y. and Bishop, J., editors, *Research in Economic Inequality.* JAI Press, Amsterdam.

[Formby *et al.*, 2004] Formby, J.P.; W.J., Smith; and Zheng, B (2004). Mobility measurement, transition matrices and statistical inference. *Journal of Econometrics* 120:181–205.

[Foster and Sen, 1997] Foster, J.E. and Sen, A. (1997). On economic inequality after a quarter century. In Sen, A., editor, *On Economic Inequality.* Clarendon Press, Oxford.

[Foster and Shorrocks, 1991] Foster, J.E. and Shorrocks, A.F. (1991). Subgroup consistent poverty indices. *Econometrica* 59:687–709.

[Foster, 1998] Foster, J.E. (1998). Absolute versus relative poverty. *American Economic Review* 88:335–341.

[Galton, 1886] Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland* 15:246–63.

[Gorman, 1968] Gorman, W.M. (1968). The structure of utility functions. *Review of Economic Studies* 35:367–390.

[Gottschalk and Spolaore, 2002] Gottschalk, P and Spolaore, E. (2002). On the evaluation of economic mobility. *Review of Economic Studies* 69:191–208.

[King, 1983] King, M.A. (1983). An index of ineqaulity: with applications to horizontal equity and social mobility. *Econometrica* 51:99–116.

[Maasoumi and Zandvakili, 1986] Maasoumi, E. and Zandvakili, S. (1986). A class of generalized measures of mobility with applications. *Economics Letters* 22:97–102.

[Maasoumi, 1998] Maasoumi, E. (1998). On mobility. In Giles, D. and Ullah, A., editors, *The Handbook of Economic Statistics*. Marcel Dekker, New York.

[Mitra and Ok, 1998] Mitra, T. and Ok, E. (1998). The measurement of income mobility: A partial ordering approach. *Economic Theory* 12:77–102.

[Ruiz-Castillo, 2004] Ruiz-Castillo, J. (2004). The measurement of structural and exchange mobility. *Journal of Economic Inequality* 2:219–228.

[Shorrocks, 1978] Shorrocks, A.F. (1978). The measurement of mobility. *Econometrica* 46:1013–1024.

[Shorrocks, 1988] Shorrocks, A.F. (1988). Aggregation issues in inequality measurement. In Eichhorn, W., editor, *Measurement in Economics.* Springer Verlag, New York.

[Solon, 2000] Solon, G. (2000). Intergenerational mobility in the labour market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3c. North Holland, Amsterdam.

[Stokey, 1998] Stokey, N. (1998). Shirtsleeves to shirtsleeves: The economic of social mobility. In N.L. Schwartz, D. Jacobs and Kalai, E., editors, *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures 1983-1997.* Econometric Society Monographs, Cambridge.

[Tchen, 1980] Tchen, A.H.T. (1980). Inequalities for distributions with given marginals. *Annals of Probability* 8:814–27.

[Zheng, 1994] Zheng, B. (1994). Can a poverty index be both relative and absolute? *Econometrica* 62:1453–58.

[Zimmerman, 1992] Zimmerman, D. (1992). Regression towards mediocrity in economic stature. *American Economic Review* 82:409–429.