

Measuring Mobility

by

Frank A. Cowell

STICERD
London School of Economics
Houghton Street
London, WC2A 2AE, UK
email: f.cowell@lse.ac.uk

and

Emmanuel Flachaire

GREQAM, Aix-Marseille University
Centre de la Vieille Charité
2, rue de la Charité
13236 Marseille cedex 02, France
email: emmanuel.flachaire@univ-cezanne.fr

July 2012

Abstract

Our new approach to mobility measurement involves separating out the valuation of positions in terms of individual status (using income, social rank, or other criteria) from the issue of movement between positions. The quantification of movement is addressed using a general concept of distance between positions and a parsimonious set of axioms that characterise the distance concept and yield a class of aggregative indices. This class of indices induces a superclass of mobility measures over the different status concepts consistent with the same underlying data. We investigate the statistical inference of mobility indices using two well-known status concepts, related to income mobility and rank mobility.

Keywords: Income mobility, rank mobility, measurement, axiomatic approach

JEL codes: D63

Acknowledgments: We are grateful for helpful comments from Guillermo Cruces, Abigail McKnight, Patrick Moyes, Dirk Van de gaer, Polly Vizard and seminar participants at STICERD and the University of Geneva.

1 Introduction

Mobility is an important concept in several branches of social science and economics. The way it has been conceived has, to some extent, depended on the particular application or even the particular data set under consideration. So, different parts of the literature have focused on income or wealth mobility, wage mobility, educational mobility, mobility in terms of social class. As a consequence of this diversity, the measurement of mobility is an intellectual problem that has been addressed from many different standpoints.¹ Mobility measures are sometimes defined, explicitly or implicitly, in relation to a specific dynamic model,² sometimes as an abstract distributional concept similar to inequality, polarisation, dispersion and so on.

This paper focuses on the second interpretation of mobility measurement - mobility measures in the abstract. It develops an approach that is sufficiently flexible to cover income or wealth mobility on the one hand and, on the other, various types of “rank” mobility where the underlying data are categorical. Our approach separates out the fundamental components of the mobility-measurement problem, proposes an axiomatic framework for the core theoretical issues and examines the statistical properties of several classes of measures that emerge from the implementation of the theory.

In the extensive literature mobility is characterised either in terms of one’s income or in terms of one’s position in the distribution, or sometimes both. In some approaches the distinction between mobility and income volatility - movement of incomes (?) - becomes a little fuzzy. This is perhaps a mistake since mobility is essentially something that characterises society, or the individual’s relationship to the society (?), whereas volatility can be seen as something that relates just to an individual; mobility would be meaningless for Robinson Crusoe, but income volatility might be very important.

In the light of this the essential ingredients for a theory of mobility measurement are as follows:

1. a time frame of two or more periods;
2. a measure of an individual’s status within society;
3. an aggregation of changes in individual status over the time frame.

In this paper we consider a standard two-period problem and focus on the interplay between ingredients 2 and 3, the status measure and the basis for aggregation of movements.

A brief word on the notion of “status” is important here. Status may be defined in a variety of ways, depending on the focus of interest of a particular mobility study. It could be something that is directly observable and measurable for each individual, independent of information about anyone else, a person’s income or wealth, perhaps. Alternatively it could be that a person’s status is only well defined in relation to information about others - one’s location in the income distribution, for example. Our approach is sufficiently flexible to cover either of these interpretations, as we will explain in detail in Section 2.

¹For a survey see ?).

²See, for example, ?), ?).

Because mobility is inherently quite a complex phenomenon it is common to find it broken down into constituent parts, for example into structural and exchange mobility.³ However, this traditional breakdown is not so important here. What is crucial in our approach is the notional separation of the status concept from the aggregation method. Nevertheless, there is a link to the structural/exchange distinction as presented in the literature. Exchange mobility can be characterised as a type of average of individual distances “travelled” in the reranking process (?).⁴ The method of aggregation that we will apply is also based on an elementary distance concept that could have a similar natural interpretation in terms of exchange mobility. As a consequence, different implementations of our classes of mobility measure would allow different ways of breaking down overall mobility into exchange and structural mobility.

The contribution of this paper is, first, to characterise mobility comparisons for an arbitrary definition of status, using a parsimonious set of axioms and, second, to show how to implement our new approach using sample data and applying different concepts of status. The paper is organised as follows. Section 2 sets out in more detail the basic ideas underlying our approach. Section 3 contains the theoretical foundations of the approach and the formal derivation of a “superclass” - a collection of classes - of mobility indices. The properties of the superclass are discussed in Section 4 and statistical inference for key members of the superclass are discussed in section 5. Section 7 concludes.

2 The Approach: Individual status and mobility

As suggested in the introduction we need very few basic concepts to set out our approach. We will assume that there is some quantity, to be called “income,” that is cardinally measurable and interpersonally comparable. However, this is used only as a device to show the range of possibilities with our approach; in fact the informational requirements for our approach are very modest: only ordinal data are required. We need to characterise in a general way a set of classes and a way of representing individual movements between the classes. So the word “income” here is just a convenient shorthand for initiating the discussion; in what follows “income” can be replaced with any other quantity that is considered to be interpersonally comparable.

Let there be an ordered set of K income classes; each class k is associated with income level x_k where $x_k < x_{k+1}$, $k = 1, 2, \dots, K - 1$. Let $p_k \in \mathbb{R}_+$ be the size of class k , $k = 1, 2, \dots, K$ and

$$\sum_{k=1}^K p_k = n,$$

where n is the size of the population. Let $k_0(i)$ be the income class occupied by person i ; then the information about a distribution is completely characterised by the vector $(x_{k_0(1)}, \dots, x_{k_0(n)})$. Clearly this includes the special case where classes are individual

³For an illuminating discussion see (?). On the definition of exchange mobility see (?),

⁴Several ad hoc measures of income mobility pursue the idea of average distance (?). ? (?), ? proposed a mobility index whose distance concept is based on the absolute differences of logarithms of incomes.

incomes if $p_k = 0$ or 1 in each class and the individual is assumed to be at the lower bound of the class.

To represent mobility we need income distributions in two time periods 0 and 1 (“before” and “after”) and the location of any person i in the two distributions. Let $k_0(i)$ and $k_1(i)$ be the income class occupied by person i in periods 0 and 1 respectively. Mobility is completely characterised by $(x_{k_0(1)}, \dots, x_{k_0(n)})$ and $(x_{k_1(1)}, \dots, x_{k_1(n)})$. However this does not necessarily mean that we should use some simple aggregation of the x_k or aggregation of a transformation of the x_k in order to compute a mobility index. We could instead carry out a relabelling of the income classes using information from the income distribution. For example we could do this using the number of persons in, or below, each income class, according to the distribution in period 0:

$$n_0(x_k) := \sum_{h=1}^k p_h, \quad k = 1, \dots, K. \quad (1)$$

We could also do a similar relabelling using information about the 1-distribution. Suppose that the class sizes (p_1, \dots, p_K) in period 0 change to (q_1, \dots, q_K) in period 1, Then the new way of relabelling the income classes is given by

$$n(x_k) := \sum_{h=1}^k q_h, \quad k = 1, \dots, K. \quad (2)$$

All of this gives us a way of thinking about the second of the essential ingredients of the mobility problem mentioned in the introduction: how to measure an individual’s status within society. The combination of the two-period framework and the different types of information about the classes enables us to specify a number of status concepts that can be used to generate different types of mobility measure. Let us use u_i and v_i to denote individual i ’s status in the 0-distribution and the 1-distribution respectively and, for each $i = 1, 2, \dots, n$, define the ordered pair $z_i := (u_i, v_i)$; then consider four examples of status concepts:

- *Distribution-independent, static (1)*. The simplest and, perhaps, most obvious case is where we just use the x values to evaluate individual status, both the before and after the event:

$$z_i = (x_{k_0(i)}, x_{k_1(i)}). \quad (3)$$

The information about distribution (before or after) is irrelevant to the evaluation of individual status. This simple case results in a model of the movements of incomes.

- *Distribution-independent, static (2)*. Clearly case 1 can be extended to include any case that involves a simple transformation of income:

$$z_i = (\varphi(x_{k_0(i)}), \varphi(x_{k_1(i)})), \quad (4)$$

where the monotonic increasing function φ could be chosen for arbitrary convenience, (such as log or exp), economic interpretation (utility of x) or to insure that the transformed variable has appropriate statistical properties. The φ function is

used to “revalue” the income concept and in general one would expect the mobility index to be dependent upon the choice of φ ; this amounts to requiring that mobility be characterised as a cardinal concept. But such an approach is inappropriate for some types of mobility problem: if one is studying social status or educational attainment then one particular cardinalisation may appear to be arbitrary. To require that a mobility index be based on purely ordinal concepts - to be independent of the cardinalisation φ - might seem rather demanding and to imply a somewhat vague approach to the measurement problem. However there is a way forward that, as we will see, leads to sharp conclusions. This way forward uses the distribution itself as a means of valuing the K classes; there are two important further cases that we will consider.

- *Distribution-dependent, static.* If we wish to use information from the income distribution to evaluate a person’s status then, for example we might take the number of persons with incomes no greater than that of i :

$$z_i = (n_0(x_{k_0(i)}), n_0(x_{k_1(i)})) . \quad (5)$$

Here we use the cumulative numbers in class to “value” the class. It results in a concept that is consistent with a purely ordinal approach to mobility - one that is independent of arbitrary monotonic, order-preserving transformations of the x_k . As an aside note that this case can be naturally extended to the case where the 1-distribution is used to evaluate the classes: just replace n_0 with n_1 in both parts of expression (5).

- *Distribution-dependent, dynamic.* An extension of the previous case that is arguably more important is where both n_0 and n_1 are used in status evaluation:

$$z_i = (n_0(x_{k_0(i)}), n_1(x_{k_1(i)})) . \quad (6)$$

In (6) we are taking into account the change in “valuation” of each status class that arises from the changing income distribution.⁵

It is clear that status is different from “income”: we could, if we wish, make the two things identical, but that would be an explicit normative assumption. It is also clear that different status concepts could produce quite different interpretations of mobility from the same basic data. For example the meaning of zero mobility depends on the way individuals’ status is defined. For example, in each of the cases (3) to (6) it makes sense say that there is zero mobility if

$$v_i = u_i, i = 1, \dots, n. \quad (7)$$

⁵If there were an exogenous revaluation of the K classes so that (x_1, \dots, x_K) in period 0 changes to (y_1, \dots, y_K) in period 1 - perhaps because of inflation or economic growth - then clearly one could also consider a *distribution-independent, dynamic* case where $z_i = (x_{k_0(i)}, y_{k_1(i)})$. However, this is intrinsically less interesting and cases where the income scale changes are probably better handled as in the next paragraph.

	period 0	period 1	period 1'	period 1''
x_1	A	—	C	—
x_2	B	A	A	C
x_3	C	B	B	A
x_4	—	C	—	B
x_5	—	—	—	—

Table 1: Mobility patterns

Consider the $n = 3$ scenario depicted in Table 1: three individuals A, B, C move up the income classes from period 0 to period 1. If status is defined as (6) then there is zero mobility; if it is defined as (5) it is clear that mobility is positive. Now suppose that

$$x_k = \lambda x_{k-1}, k = 2, \dots, K, \lambda > 1. \quad (8)$$

Then, in the cases (3) and (4), it *may* make sense to consider

$$v_i = \lambda u_i, i = 1, \dots, n, \lambda > 0 \quad (9)$$

as representing zero mobility; this would apply, for example, if one made the judgment that uniform proportional income growth for all members of society is irrelevant for mobility. Each of these answers makes sense in its own way. As a second example, suppose that there is pure exchange mobility as far as income positions are concerned: is the mobility involved in going from the situation in period 0 to that in period 1' the same as the mobility involved in going from period 0 to period 1''? Again we can imagine sensible approaches to mobility that would respond affirmatively to this question and sensible approaches that would respond negatively.

It is also clear that allowing for different definitions of status will induce different types of mobility measure. Moreover the four illustrative examples of status concepts are not exhaustive. What we will see in the theoretical development of Section 3 is that for any given definition of status we can derive an associated *class* of mobility measures. Taking this with the diversity of status concepts that may be derived from a given data set we are, in effect, characterising a *superclass* of classes of mobility measures.

To make progress we may exploit the separability of the concept of status from the concepts of individual and aggregate mobility.

3 Mobility measures - theory

3.1 Individual mobility

Accordingly, let us address the third essential ingredient of the mobility problem mentioned in the introduction: the aggregation of changes in individual status.⁶ For the

⁶An early treatment of this type of problem for the specific case where status equals income is given in (?). However, the present treatment is more general, in two ways. First, the axiomatisation here does not require differentiability or additivity, which were arbitrarily imposed in (?); second, the current paper deals with any arbitrary representation of status (including ordinal status) rather than being specific to income.

analysis that follows the status measure that is imputed can be arbitrary, subject only that it be weakly increasing in the income levels x_k - for example it does not matter whether or not is dependent on the cardinalisation of income. So, we assume that a measure of individual status has been agreed, determined by the information available from the income distribution at any moment; we also assume that we have an observation of the status of each person i in periods 0 and 1; we need a coherent method of quantifying the implicit status changes as “mobility.”

In this approach the set of status distributions is given by

$$U := \{ \mathbf{u} \mid \mathbf{u} \in \mathbb{R}_+^n, u_1 \leq u_2 \leq \dots \leq u_n \}. \quad (10)$$

Individual mobility is completely characterised by the ordered pairs z_i , $i = 1, 2, \dots, n$ as defined in section 2. The set of possible income movements Z is taken to be a connected subset of $\mathbb{R}_+ \times \mathbb{R}_+$ and we define

$$Z^n := Z \times Z \times \dots \times Z.$$

We may refer to any $\mathbf{z} \in Z^n$ as a *movement profile*. Overall mobility for a given profile can be described in terms of the individual mobility of each person as a vector

$$\mathbf{m}(\mathbf{z}) := (m(z_1), \dots, m(z_n)).$$

where the function $m : Z \rightarrow \mathbb{R}$ is such that $m(z_i)$ is strictly increasing in $|u_i - v_i|$. Other than this property, the form of the individual-mobility function m is left open at the moment. It will be determined through an axiomatisation of a mobility ordering that will then induce a particular structural form on individual and overall mobility.

As we have noted, a particular advantage of our approach is that the axiomatisation of the ordering can be completely separated from the specification of the status concepts. Of course it will be the case that some axioms are particularly appropriate in the case of certain types of status measure and we will discuss these on a case-by-case basis.

3.2 A mobility ordering

First we need to characterise a method of comparing movement profiles. Let us consider mobility as a weak ordering \succeq on Z^n ; denote by \succ the strict relation associated with \succeq and denote by \sim the equivalence relation associated with \succeq . We also need one more piece of notation: for any $\mathbf{z} \in Z^n$ denote by $\mathbf{z}(\zeta, i)$ the member of Z^n formed by replacing the i th component of \mathbf{z} by $\zeta \in Z$.

Axiom 1 [*Continuity*] \succeq is continuous on Z^n .

Axiom 2 [*Monotonicity*] If $\mathbf{z}, \mathbf{z}' \in Z^n$ differ only in their i th component then $m(u_i, v_i) > m(u'_i, v'_i) \iff \mathbf{z} \succ \mathbf{z}'$.

In other words, if two movement profiles differ only in respect of person i 's status, then the profile that registers higher individual mobility for i is the profile that exhibits greater mobility. This is a very weak requirement.

Axiom 3 [*Independence*] For $\mathbf{z}, \mathbf{z}' \in Z^n$ such that: $\mathbf{z} \sim \mathbf{z}'$ and $z_i = z'_i$ for some i then $\mathbf{z}(\zeta, i) \sim \mathbf{z}'(\zeta, i)$ for all $\zeta \in [z_{i-1}, z_{i+1}] \cap [z'_{i-1}, z'_{i+1}]$.

Suppose that the profiles \mathbf{z} and \mathbf{z}' are equivalent in terms of overall mobility and that there is some person i such that i 's status pair $z_i = (u_i, v_i)$ is the same in \mathbf{z} and in \mathbf{z}' . Then, the same small change Δz_i in i 's status pair in both profiles \mathbf{z} and \mathbf{z}' still leaves \mathbf{z} and \mathbf{z}' as equivalent in terms of overall mobility.

Axiom 4 [*Local immobility*] Let $\mathbf{z}, \mathbf{z}' \in Z^n$ where for some i , $u_i = v_i$, $u'_i = u_i + \delta$, $v'_i = v_i + \delta$ and, for all $j \neq i$, $u'_j = u_j$, $v'_j = v_j$. Then $\mathbf{z} \sim \mathbf{z}'$.

The principle states that if two profiles are identical except for the status of i and j who are both immobile then a status-preserving spread involving only i and j (a notional improvement in the status of i and worsening of the status of j by the same amount) has no effect on the evaluation of mobility. Person i has the same before-status and after-status in \mathbf{z} and the same is true for j ; by construction the same is true for both i and j in \mathbf{z}' ; for every other person individual mobility is the same in \mathbf{z} and \mathbf{z}' ; so it is reasonable to require that \mathbf{z} and \mathbf{z}' represent the same overall mobility.

Theorem 1 Given Axioms 1 to 4 (a) \succeq is representable by the continuous function given by

$$\sum_{i=1}^n \phi_i(z_i), \forall \mathbf{z} \in Z^n \quad (11)$$

where, for each i , $\phi_i : Z \rightarrow \mathbb{R}$ is a continuous function that is strictly increasing in $|u_i - v_i|$ and (b)

$$\phi_i(u, u) = a_i + b_i u. \quad (12)$$

Corollary 1 Since \succeq is an ordering it is also representable by

$$\phi \left(\sum_{i=1}^n \phi_i(z_i) \right) \quad (13)$$

where ϕ_i is defined as in (11), (12) and $\phi : \mathbb{R} \rightarrow \mathbb{R}$ continuous and strictly monotonic increasing.

This additive structure means that we can proceed to evaluate aggregate mobility by taking one person at a time. The following axiom imposes a fairly weak structural requirement, namely that the ordering remains unchanged by some uniform scale change to status in both periods simultaneously. As Theorem 2 shows it is enough to induce a rather specific structure on the function representing \succeq .

Axiom 5 [*Status scale irrelevance*] For any $\mathbf{z}, \mathbf{z}' \in Z^n$ such that $\mathbf{z} \sim \mathbf{z}'$, $t\mathbf{z} \sim t\mathbf{z}'$ for all $t > 0$.

Axiom 5 is completely natural in the case of distribution-dependent measures of status such as (5) or (6) since it enables one to characterise mobility in terms of population proportions rather than absolute numbers. In the case where status is given by x one is clearly making a judgment about the mobility implications of across-the-board changes in real income.

Theorem 2 *Given Axioms 1 to 5 \succeq is representable by*

$$\phi \left(\sum_{i=1}^n u_i^c H_i \left(\frac{u_i}{v_i} \right) \right) \quad (14)$$

where $c \in \mathbb{R}$, H_i is a real-valued function and either (a) $c = 1$, or (b) $c = 0$, or (c) $H_i(\cdot) = \log(\cdot)$.

This result is important but limited since the function H_i is essentially arbitrary: we need to impose more structure.

3.3 (u, v) vectors and mobility

We now focus on the way in which one compares the (u, v) vectors in different parts of the distribution. The form of (14) suggests that movement should be characterised terms of proportional differences:

$$m(z_i) = \max \left(\frac{u_i}{v_i}, \frac{v_i}{u_i} \right).$$

This is the form for m that we will assume from this point onwards. We also introduce:

Axiom 6 [*Mobility scale irrelevance*] *Suppose there are $\mathbf{z}_0, \mathbf{z}'_0 \in Z^n$ such that $\mathbf{z}_0 \sim \mathbf{z}'_0$. Then for all $t > 0$ and \mathbf{z}, \mathbf{z}' such that $m(\mathbf{z}) = tm(\mathbf{z}_0)$ and $m(\mathbf{z}') = tm(\mathbf{z}'_0)$: $\mathbf{z} \sim \mathbf{z}'$.*

Axiom 7 [*Anonymity*] *For any $\mathbf{z} \in Z^n$ and any $\mathbf{z}' \in Z^n$ that can be obtained as a permutation of the components of $\mathbf{z}' \in Z^n$, $\mathbf{z} \sim \mathbf{z}'$.*

Mobility scale irrelevance means this. Suppose we have two profiles \mathbf{z}_0 and \mathbf{z}'_0 that are regarded as equivalent under \succeq ; then rescale all the individual mobility (the status movements) in \mathbf{z}_0 and \mathbf{z}'_0 by the same factor t ; the resulting pair of profiles \mathbf{z} and \mathbf{z}' will also be equivalent.⁷ This might at first seem to be restrictive, but in fact this is not so - see Section 4. Anonymity simply means that switching the personal labels from one status-pair to another within a movement profile has no effect on mobility rankings.

Theorem 3 *Given Axioms 1 to 6 \succeq is representable by (14) where, for all i , either $H_i(r) = r^b$, $b \neq 0$ and $c \in \{0, 1\}$, or $H_i(r) = \log r$.*

3.4 Interpretation

Theorem 3 means that the implied mobility ordering can be represented by some transformation of the expression

$$\sum_{i=1}^n u_i^c [u_i/v_i]^b, \quad (15)$$

⁷Also note that Axiom 6 can be stated equivalently by requiring that, for a given $\mathbf{z}_0, \mathbf{z}'_0 \in Z^n$ such that $\mathbf{z}_0 \sim \mathbf{z}'_0$, either (a) any \mathbf{z} and \mathbf{z}' found by rescaling the u -components will be equivalent or (b) any \mathbf{z} and \mathbf{z}' found by rescaling the v -components will be equivalent.

(with the log function replacing the case where $b = 0$). This can be interpreted as an evaluation function over two-period status, a weighted function of the ratio of status in the two periods; the weights could be population weights ($c = 0$)⁸ or status weights ($c = 1$). It is difficult to see a good economic intuition for cases where $c \neq 0, 1$ (case (c) in Theorem 2). However the case $c = 1$, where the evaluation function is homogeneous of degree one in status, makes good sense: if there is no mobility (so that $u_i = v_i$ for all i) then (16) just becomes aggregate status.

3.5 An aggregate mobility index

We can now use the function representing mobility rankings to generate an aggregate mobility index. Consider the behaviour of the index over the following subset of Z :

$$Z(\bar{u}, \bar{v}) := \left\{ \mathbf{z} \in Z \mid \sum_{i=1}^n z_i = (\bar{u}, \bar{v}) \right\}.$$

Taking the case $c = 1$ then, for all $\mathbf{z} \in Z(\bar{u}, \bar{v})$, the mobility index must take the form

$$\Phi(\mathbf{z}) = \bar{\phi} \left(\sum_{i=1}^n u_i^\alpha v_i^{1-\alpha}; \bar{u}, \bar{v} \right), \quad (16)$$

where \bar{u}, \bar{v} are parameters of the function $\bar{\phi}$ that is the counterpart of ϕ in (14). It is reasonable to require that $\Phi(\mathbf{z})$ should take the value zero when \mathbf{z} there is no mobility between the 0-distribution and the 1-distribution. If we take the standard interpretation of zero mobility as given in (7) in which case the form (16) requires that

$$\bar{\phi} \left(\sum_{i=1}^n u_i; \bar{u}, \bar{u} \right) = 0, \quad (17)$$

in other words we have the restriction $\bar{\phi}(\bar{u}; \bar{u}, \bar{u}) = 0$; but this restriction does not impose much additional structure on the problem. By contrast, suppose we take a broader interpretation of zero mobility given in (9), namely that if the 1-distribution is obtained rescaling each component in the 0-distribution by a factor $\lambda > 0$ then there is no mobility; in other words suppose we say that the aggregate of status is not relevant in the evaluation of mobility. This interpretation requires that, if $v_i = \lambda u_i$, $i = 1, \dots, n$ (where $\lambda = \bar{v}/\bar{u}$) then, from (16), we have

$$\bar{\phi} \left(\lambda^{1-\alpha} \sum_{i=1}^n u_i; \bar{u}, \bar{v} \right) = 0 \quad (18)$$

which implies

$$\bar{\phi}(\bar{u}^\alpha \bar{v}^{1-\alpha}; \bar{u}, \bar{v}) = 0. \quad (19)$$

This can only be true for all α if ϕ in (15) and $\bar{\phi}$ in (16) can be written in the form

$$\psi \left(\sum_{i=1}^n \left[\frac{u_i}{\mu_u} \right]^\alpha \left[\frac{v_i}{\mu_v} \right]^{1-\alpha} \right), \quad (20)$$

⁸This is the approach implied in ?).

where $\mu_u := \frac{1}{n} \sum_{i=1}^n u_i$ and $\mu_v := \frac{1}{n} \sum_{i=1}^n v_i$.

A suitable cardinalisation of (20) gives the aggregate mobility measure

$$M_\alpha := \frac{1}{\alpha [\alpha - 1] n} \sum_{i=1}^n \left[\left[\frac{u_i}{\mu_u} \right]^\alpha \left[\frac{v_i}{\mu_v} \right]^{1-\alpha} - 1 \right], \alpha \in \mathbb{R}, \alpha \neq 0, 1 \quad (21)$$

where we have the following limiting forms for the cases $\alpha = 0$ and $\alpha = 1$, respectively

$$M_0 = -\frac{1}{n} \sum_{i=1}^n \frac{v_i}{\mu_v} \log \left(\frac{u_i}{\mu_u} / \frac{v_i}{\mu_v} \right), \quad (22)$$

$$M_1 = \frac{1}{n} \sum_{i=1}^n \frac{u_i}{\mu_u} \log \left(\frac{u_i}{\mu_u} / \frac{v_i}{\mu_v} \right). \quad (23)$$

Expressions (21)-(23) constitute a *class* of aggregate mobility measures where an individual family member is characterised by choice of α : a high positive α produces an index that is particularly sensitive to downward movements (where u exceeds v) and a negative α yields an index that is sensitive to upward movements (where v exceeds u). If we refer back to the illustrative sketch in Table 1 it is clear that M_α has the property that the change from period 0 to period 1 in Table 1 will register zero mobility; but this is purely a matter of normalisation - an alternative normalisation of (15) could have been chosen if we want uniform growth in status to register as positive mobility. However the change from period 0 to period 1' - pure exchange mobility - will produce $M_\alpha > 0$.

4 Discussion

The nature of the superclass referred to in the introduction is now clear: expressions (21)-(23) characterise a class of indices, *for a given definition of the status variables u and v* ; the superclass is the collection of all such classes for the different status concepts that are supported by the data. We can generate a different class of mobility indices just by replacing the status concept, for example by choosing a different specification from section 2. Let us briefly review the issues raised by the structure of our superclass in the light of the mobility-measurement literature.

First, is there a good argument for taking an ordinal-status class of indices from the superclass? In so far as mobility is concerned with ranks rather than income levels then making status an ordinal concept is exactly the thing to do (? , ?). However, there is a variety of ways of attempting to define status ordinally. For example a large section of the mobility adopts a ‘‘mobility table’’ or ‘‘transition matrix’’ approach to mobility.⁹ This focuses attention on the size p_k of each class k and the number of the p_k that move to other classes. However, this approach could be sensitive to the merging or splitting of classes or the adjustment of class boundaries. Consider the case where in the original set of classes $p_k = 0$ and $p_{k+1} > 0$; if the mobility index is sensitive to small values of p and the income boundary between classes k and $k + 1$ is adjusted some of the population

⁹See, for example, ? (? , ?), (?), (?), (?), (?).

mass that was formerly in class $k + 1$ is now tipped into class k and there could be a big jump in the mobility index. This will not happen if the index is defined in terms of u_i and v_i as in (5) or (6).

Our axioms induce an additive structure for the mobility index, which might be thought to be restrictive. Individual mobility depends only on the individual's status in the before- and after-distributions. Should mobility perhaps also depend on the person's rank relative to others? (see for example ?) However, as we have just explained, i 's status may well depend on i 's relative position in the distribution according to some formulations of u and v . So, rank *can* enter into the formulation of the mobility index, but through the definition of status, not the formulation of individual mobility. In fact the additive structure makes it particularly straightforward to interpret the underlying composition of mobility; the reason for this is that the expressions in (21)-(23) are clearly decomposable by arbitrary population subgroups. This means, for example, that we may choose a number i^* and partition U in (10) unambiguously¹⁰ into a poor group P (for $i \leq i^*$) and a rich group R (for $i > i^*$) and, using an obvious notation, express overall mobility as

$$M_\alpha = w^P M_\alpha^P + w^R M_\alpha^R + M_\alpha^{\text{between}},$$

where the weights w^P, w^R and the between-group mobility component $M_\alpha^{\text{between}}$ are functions of the status-means μ_u, μ_v for each of the two groups and overall; comparing M_α^P and M_α^R enables one to say precisely where in the distribution mobility has taken place.

Our axioms also induce a homothetic structure, which once again might be thought to be rather restrictive for some interpretations of u and v . We have effectively introduced scale independence which could be considered unobjectionable when u and v are evaluated in terms of numbers of persons, but might be questioned if u and v are to be interpreted in terms of income or wealth, say: why not have a translation-independent mobility index? However, the fact that our approach defines a superclass, not just a single class, of mobility measures can be used to handle this issue. As we have discussed, the methodology is valid for arbitrary methods of valuing the K classes. So, for example, we may replace the u and v by $u + c$ and $v + c$ where c is a non-negative constant. In which case (21) will be replaced by

$$\frac{\theta(c)}{n} \sum_{i=1}^n \left[\left[\frac{u_i + c}{\mu_u + c} \right]^{\alpha(c)} \left[\frac{v_i + c}{\mu_v + c} \right]^{1-\alpha(c)} - 1 \right], \alpha(c) \in \mathbb{R}, \alpha(c) \neq 0, 1 \quad (24)$$

where $\gamma \in \mathbb{R}, \beta \in \mathbb{R}_+$, the term $\alpha(c)$ indicates that the sensitivity parameter may depend upon the location parameter c and $\theta(c)$ is a normalisation term given by

$$\theta(c) := \frac{1 + c^2}{\alpha(c)^2 - \alpha(c)}; \quad (25)$$

for $\alpha(c) = 0$ and $\alpha(c) = 1$ there are obvious special cases of (24) corresponding to (22) and (23). If we take a given value of c then we have generated an "intermediate" version

¹⁰Clearly "poor" and "rich" refer to status in the before-distribution and we could have used a finer partition into more than two groups.

of the mobility index (borrowing the terminology of ?, ?). However, by writing

$$\alpha(c) := \gamma + \beta c \quad (26)$$

and analysing the behaviour as $c \rightarrow \infty$ we may say more. Consider the main expression inside the summation in (24); taking logs we may write this as

$$\log \left(\frac{1 + \frac{v}{c}}{1 + \frac{\mu_v}{c}} \right) + \alpha(c) \left[\log \left(1 + \frac{u}{c} \right) + \log \left(1 + \frac{\mu_v}{c} \right) - \log \left(1 + \frac{v}{c} \right) - \log \left(1 + \frac{\mu_u}{c} \right) \right]. \quad (27)$$

Using the standard expansion

$$\log(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - \dots \quad (28)$$

and (26) we find that (27) becomes

$$\log \left(\frac{1 + \frac{v}{c}}{1 + \frac{\mu_v}{c}} \right) + \left[\beta + \frac{\gamma}{c} \right] \left[u + \mu_v - v - \mu_u - \frac{u^2}{2c} - \frac{\mu_v^2}{2c} + \frac{v^2}{2c} + \frac{\mu_u^2}{2c} \dots \right]. \quad (29)$$

For finite $\gamma, \beta, u, v, \mu_u, \mu_v$ we find that (29) becomes

$$\beta [u - \mu_u - v + \mu_v] \quad (30)$$

and

$$\lim_{c \rightarrow \infty} \theta(c) = \lim_{c \rightarrow \infty} \frac{1 + \frac{1}{c^2}}{\left[\beta + \frac{\gamma}{c} \right]^2 - \frac{1}{c} \left[\beta + \frac{\gamma}{c} \right]} = \frac{1}{\beta^2}. \quad (31)$$

From (30) and (31) we can see that in the limit (24) becomes¹¹

$$\frac{1}{n\beta^2} \sum_{i=1}^n [e^{\beta[u_i - \mu_u - v_i + \mu_v]} - 1], \quad (32)$$

for any $\beta \neq 0$. Let $q_i := u_i - \mu_u - v_i + \mu_v$ so that (32) can be written

$$\frac{1}{n\beta^2} \sum_{i=1}^n [e^{\beta q_i} - 1] = \frac{1}{n\beta^2} \sum_{i=1}^n \left[1 + \beta q_i + \frac{1}{2!} \beta^2 q_i^2 + \frac{1}{3!} \beta^3 q_i^3 + \frac{1}{4!} \beta^4 q_i^4 + \dots - 1 \right], \quad (33)$$

using a standard expansion. Noting that $\frac{1}{n} \sum_{i=1}^n q_i = 0$, the right-hand side of (33) becomes

$$\frac{1}{n} \sum_{i=1}^n \left[\frac{1}{2!} q_i^2 + \frac{1}{3!} \beta q_i^3 + \frac{1}{4!} \beta^2 q_i^4 + \dots \right]. \quad (34)$$

As $\beta \rightarrow 0$ it is clear that (34) tends to $\frac{1}{2n} \sum_{i=1}^n q_i^2$. So the limiting form of (32) for $\beta = 0$ is

$$\frac{1}{2} \text{var}(u_i - v_i). \quad (35)$$

¹¹See also equation (56) of ?).

So expressions (32) and (35) show that a class of translation-independent mobility measures - where mobility is independent of uniform absolute additions to/subtractions from everyone's income - is also contained within our superclass. The importance of this is that, because we may redefine status arbitrarily, Axiom 6 in effect only requires a weak form of homogeneity where all the contours are homothetic to the point (c, c, \dots, c) , a property that is satisfied by almost all mobility indices.

Finally, should mobility indices be “ethical” indices?¹² Some recent contributions to the literature have built this in - notably the “extended Atkinson” approach of (?), which is based on (?). Although our approach has not started from a basis in welfare economics, ethical considerations can be incorporated through two channels, (1) the definition of status and (2) the parameter α . (1) The use income itself or a rank-dependent construct as the status variable is essentially a normative choice. (2) Depending on the interpretation of u and v we can endow the sensitivity parameter α with welfare interpretation, using an analogy with the well-known relation between welfare-based inequality measures and the generalised-entropy approach.¹³ It enables us to capture *directional sensitivity* in the mobility context:¹⁴ high positive values result in a mobility index that is sensitive to downward movements from period 0 to period 1; negative α is sensitive to upward movements. Picking a value for this parameter is again a normative choice.

5 Statistical Inference

In this section we establish the asymptotic distribution of our mobility measures, taking the situation where there are as many classes as there are observations. For two well-known status concepts, associated with movements of incomes and with rank mobility, we show that M_α is asymptotically Normal.

5.1 Income mobility

Let us consider the *distribution-independent, static* status, as defined in (3). The income values at period 0 and 1 are used to evaluate individual status,

$$u_i = x_{0i} \quad \text{and} \quad v_i = x_{1i}, \quad (36)$$

it corresponds to a model of movement of incomes. Let us define the following moment: $\mu_{g(u,v)} = n^{-1} \sum_{i=1}^n g(u_i, v_i)$, where $g(\cdot)$ is a specific function. We proceed by taking the cases (21)-(23) separately.

Case M_α ($\alpha \neq 0, 1$). We can rewrite the index (21) as

$$M_\alpha = \frac{1}{\alpha(1-\alpha)} \left[\frac{n^{-1} \sum u_i^\alpha v_i^{1-\alpha}}{\mu_u^\alpha \mu_v^{1-\alpha}} - 1 \right]$$

¹²The classic paper by (?) is explicitly based on a social-welfare function. Other early contributions using a welfare-based approach see ? (?), (?) and (?) later reinterpreted by ?).

¹³In particular notice that in the case where $u_i = x_i$ and $\forall v_i = \mu_v$, M_α in (21)-(23) becomes the class of generalised-entropy inequality indices.

¹⁴See also (?) and ?).

from which we obtain M_α as a function of three moments:

$$M_\alpha = \frac{1}{\alpha(\alpha-1)} \left[\frac{\mu_u^{\alpha} \mu_v^{1-\alpha}}{\mu_u^\alpha \mu_v^{1-\alpha}} - 1 \right]. \quad (37)$$

Under standard regularity conditions, the Central Limit Theorem can be applied and thus the M_α index will follow asymptotically a Normal distribution. Under these circumstances the asymptotic variance can be calculated by the delta method. Specifically, if $\hat{\Sigma}$ is the estimator of the covariance matrix of μ_u , μ_v and $\mu_u^{\alpha} \mu_v^{1-\alpha}$, the variance estimator for M_α is :

$$\widehat{\text{Var}}(M_\alpha) = D \hat{\Sigma} D^\top \quad \text{with} \quad D = \left[\frac{\partial M_\alpha}{\partial \mu_u} ; \frac{\partial M_\alpha}{\partial \mu_v} ; \frac{\partial M_\alpha}{\partial \mu_u^{\alpha} \mu_v^{1-\alpha}} \right] \quad (38)$$

where the matrix D can be written as functions of sample moments. We have

$$D = \left[\frac{-\mu_u^{\alpha} \mu_v^{1-\alpha} \mu_u^{-\alpha-1} \mu_v^{\alpha-1}}{(\alpha-1)} ; \frac{\mu_u^{\alpha} \mu_v^{1-\alpha} \mu_u^{-\alpha} \mu_v^{\alpha-2}}{\alpha} ; \frac{\mu_u^{-\alpha} \mu_v^{\alpha-1}}{\alpha(\alpha-1)} \right].$$

The covariance matrix $\hat{\Sigma}$ is defined as follows:¹⁵

$$\hat{\Sigma} = \frac{1}{n} \begin{bmatrix} \mu_u^2 - (\mu_u)^2 & \mu_{uv} - \mu_u \mu_v & \mu_u^{1+\alpha} \mu_v^{1-\alpha} - \mu_u \mu_u^{\alpha} \mu_v^{1-\alpha} \\ \mu_{uv} - \mu_u \mu_v & \mu_v^2 - (\mu_v)^2 & \mu_u^{\alpha} \mu_v^{2-\alpha} - \mu_v \mu_u^{\alpha} \mu_v^{1-\alpha} \\ \mu_u^{1+\alpha} \mu_v^{1-\alpha} - \mu_u \mu_u^{\alpha} \mu_v^{1-\alpha} & \mu_u^{\alpha} \mu_v^{2-\alpha} - \mu_v \mu_u^{\alpha} \mu_v^{1-\alpha} & \mu_u^{2\alpha} \mu_v^{2-2\alpha} - (\mu_u^{\alpha} \mu_v^{1-\alpha})^2 \end{bmatrix} \quad (39)$$

We can use this variance estimator of M_α to compute a test statistic or a confidence interval.¹⁶

Similar developments permit us to derive the variance estimators of the limiting forms of the mobility index.

Case M_0 . We can rewrite M_0 as a function of four moments:

$$M_0 = \frac{\mu_v \log v - \mu_v \log u}{\mu_v} + \log \left(\frac{\mu_u}{\mu_v} \right) \quad (40)$$

The variance estimator of this index is defined as follows:

$$\widehat{\text{Var}}(M_0) = D_0 \hat{\Sigma}_0 D_0^\top \quad \text{with} \quad D_0 = \left[\frac{\partial M_0}{\partial \mu_u} ; \frac{\partial M_0}{\partial \mu_v} ; \frac{\partial M_0}{\partial \mu_v \log v} ; \frac{\partial M_0}{\partial \mu_v \log u} \right] \quad (41)$$

We have

$$D_0 = \left[\frac{1}{\mu_u} ; \frac{-\mu_v \log v + \mu_v \log u - \mu_v}{\mu_v^2} ; \frac{1}{\mu_v} ; -\frac{1}{\mu_v} \right],$$

and the estimator of the covariance matrix of the four moments $\hat{\Sigma}_0$ is equal to:

$$\frac{1}{n} \begin{bmatrix} \mu_u^2 - (\mu_u)^2 & \mu_{uv} - \mu_u \mu_v & \mu_{uv} \log v - \mu_u \mu_v \log v & \mu_{uv} \log u - \mu_u \mu_v \log u \\ \mu_{uv} - \mu_u \mu_v & \mu_v^2 - (\mu_v)^2 & \mu_v^2 \log v - \mu_v \mu_v \log v & \mu_v^2 \log u - \mu_v \mu_v \log u \\ \mu_{uv} \log v - \mu_u \mu_v \log v & \mu_v^2 \log v - \mu_v \mu_v \log v & \mu_v^2 \log v - (\mu_v \log v)^2 & \mu_v^2 \log u \log v - \mu_v \log u \mu_v \log v \\ \mu_{uv} \log u - \mu_u \mu_v \log u & \mu_v^2 \log u - \mu_v \mu_v \log u & \mu_v^2 \log u \log v - \mu_v \log u \mu_v \log v & \mu_v^2 \log u - (\mu_v \log u)^2 \end{bmatrix}$$

¹⁵If the observations are assumed independent, we have $\widehat{\text{Cov}}(\mu_u, \mu_v) = \frac{1}{n} \widehat{\text{Cov}}(u_i, v_i)$. In addition, we use the fact that, by definition $\text{Cov}(U, V) = E(UV) - E(U)E(V)$.

¹⁶Note that we assume that the observations are independent, in the sense that $\text{Cov}(u_i, u_j) = 0$ and $\text{Cov}(v_i, v_j) = 0$ for all $i \neq j$, but this independence assumption is not between the two samples: $\text{Cov}(u_i, v_i)$ can be different from 0.

Case M_1 . We can rewrite M_1 as a function of four moments:

$$M_1 = \frac{\mu_u \log u - \mu_u \log v}{\mu_u} + \log \left(\frac{\mu_v}{\mu_u} \right) \quad (42)$$

The variance estimator of this index is defined as follows:

$$\widehat{\text{Var}}(M_1) = D_1 \hat{\Sigma}_1 D_1^\top \quad \text{with} \quad D_1 = \left[\frac{\partial M_1}{\partial \mu_u} ; \frac{\partial M_1}{\partial \mu_v} ; \frac{\partial M_1}{\partial \mu_u \log u} ; \frac{\partial M_1}{\partial \mu_u \log v} \right] \quad (43)$$

We have

$$D_1 = \left[\frac{-\mu_u \log u + \mu_u \log v - \mu_u}{\mu_u^2} ; \frac{1}{\mu_v} ; \frac{1}{\mu_u} ; -\frac{1}{\mu_u} \right],$$

and the estimator of the covariance matrix of the four moments $\hat{\Sigma}_1$ is equal to:

$$\frac{1}{n} \begin{bmatrix} \mu_u^2 - (\mu_u)^2 & \mu_{uv} - \mu_u \mu_v & \mu_u^2 \log u - \mu_u \mu_u \log u & \mu_u^2 \log v - \mu_u \mu_u \log v \\ \mu_{uv} - \mu_u \mu_v & \mu_v^2 - (\mu_v)^2 & \mu_{uv} \log u - \mu_v \mu_u \log u & \mu_{uv} \log v - \mu_v \mu_u \log v \\ \mu_u^2 \log u - \mu_u \mu_u \log u & \mu_{uv} \log u - \mu_v \mu_u \log u & \mu(u \log u)^2 - (\mu_u \log u)^2 & \mu_u^2 \log u \log v - \mu_u \log u \mu_u \log v \\ \mu_u^2 \log v - \mu_u \mu_u \log v & \mu_{uv} \log v - \mu_v \mu_u \log v & \mu_u^2 \log u \log v - \mu_u \log u \mu_u \log v & \mu(u \log v)^2 - (\mu_u \log v)^2 \end{bmatrix}$$

5.2 Rank mobility

Let us consider the *distribution-dependent, dynamic* status, as defined in (6), that is, u_i (resp. v_i) is the number of individuals with incomes less or equal to the income of i at period one (resp. at period two). In other words, ranks are used to evaluate individual status. Because of the scale independence property of M_α , we may use proportions rather than numbers to define status,

$$u_i = \hat{F}_0(x_{0i}) \quad \text{and} \quad v_i = \hat{F}_1(x_{1i}) \quad (44)$$

where $\hat{F}_0(\cdot)$ and $\hat{F}_1(\cdot)$ are the empirical distribution functions of individual incomes in period 0 and 1,

$$\hat{F}_k(x) = \frac{1}{n} \sum_{j=1}^n I(x_{kj} \leq x) \quad (45)$$

where $k = 1, 2$ and $I(\cdot)$ is an indicator function, equals to 1 if its argument is true and to 0 otherwise. Then, u_i (resp v_i) is the rank of i 's income in the set of incomes at period 0 (resp. 1), divided by the total number of incomes n . Let us consider that we have no ties in the sample, u and v are thus defined by two differently ordered sets of the same values $\{\frac{1}{n}, \frac{2}{n}, \dots, 1\}$. The values in u and v are non i.i.d., and thus, the method of moments used previously in the case of income mobility does not apply.

?) have established the asymptotic normality in the non i.i.d. case of the following multivariate rank statistic,

$$T_n = \frac{1}{n} \sum_{i=1}^n c_{in} \phi_1(u_i) \phi_2(v_i), \quad (46)$$

where c_{in} are given real constants, ϕ_1 and ϕ_2 are (scores) functions defined on $(0,1)$, which are allowed to tend to infinity near 0 and 1 but not too quickly. Indeed, the following assumption is required: there exists positive numbers K_1 , a_1 and a_2 , such that

$$\phi_1(t) \leq \frac{K_1}{[t(1-t)]^{a_1}} \quad \text{and} \quad \phi_2(t) \leq \frac{K_1}{[t(1-t)]^{a_2}} \quad \text{with} \quad a_1 + a_2 < \frac{1}{2} \quad (47)$$

for $t \in (0,1)$. This condition implies that $\phi_1(t)$ and $\phi_2(t)$ should tend to infinity near 0 at a rate slower than the functions t^{-a_1} and t^{-a_2} . Moreover, they have shown that the variance of T_n is finite, even if not analytically tractable.

In the following, we show that M_α can be written as a function of T_n and we check when the condition defined in (47) is respected. Let us first notice that,

$$\mu_u = \mu_v = \frac{1}{n} \sum_{i=1}^n \frac{i}{n} = \frac{n+1}{2n}. \quad (48)$$

Case M_α ($\alpha \neq 0, 1$). From (37) and (48), we obtain M_α as a function of one moment:

$$M_\alpha = \frac{1}{\alpha(\alpha-1)} \left[\frac{2n}{n+1} \mu_{u^\alpha v^{1-\alpha}} - 1 \right]. \quad (49)$$

From (49) and (46), it is clear that

$$M_\alpha = \frac{1}{\alpha(\alpha-1)} [T_n - 1], \quad (50)$$

with $c_{in} = \frac{2n}{n+1}$, $\phi_1(u_i) = u_i^\alpha$ and $\phi_2(v_i) = v_i^{1-\alpha}$. The condition defined in (47) is respected for $\alpha \in]-0.5, 1.5[$. Indeed, for $\alpha > 0$, we have $0 < \phi_1(u_i) \leq 1$ and we can use $a_1 = 0$. Then, the condition requires $a_2 < 1/2$, that is, $-(1-\alpha) < 1/2$. For $\alpha < 0$, we have $0 < \phi_2(v_i) \leq 1$ and we can use $a_2 = 0$, the condition requires $a_1 < 1/2$, that is, $-\alpha < 1/2$. Note that, when $0 < \alpha < 1$, the two functions ϕ_1 and ϕ_2 are bounded, they both provide values in $(0, 1)$.

Case M_0 . From (40), (48) and (46), we have

$$M_0 = \frac{2n}{n+1} (k - \mu_{v \log u}) = l - T_n, \quad (51)$$

where k and l are real constants¹⁷ and $c_{in} = \frac{2n}{n+1}$, $\phi_1(u_i) = \log u_i$ and $\phi_2(v_i) = v_i$. The condition (47) is respected because $\phi_2(v_i) \leq 1$ and $\phi_1(u_i)$ tends to infinity near 0 at a slower rate than $-1/\sqrt{u_i}$, which implies $a_1 < 1/2$.

Case M_1 . From (42), (48) and (46), we have

$$M_1 = \frac{2n}{n+1} (k - \mu_{u \log v}) = l - T_n, \quad (52)$$

¹⁷ $k = \mu_{v \log v} = \mu_{u \log u} = n^{-1} \sum_{i=1}^n \frac{i}{n} \log \frac{i}{n}$ and $l = \frac{2nk}{n+1}$

where $c_{in} = \frac{2n}{n+1}$, $\phi_1(u_i) = u_i$ and $\phi_2(v_i) = \log v_i$. The condition (47) is respected because $\phi_1(u_i) \leq 1$ and $\phi_2(v_i)$ tends to infinity near 0 at a slower rate than $-1/\sqrt{v_i}$, which implies $a_2 < 1/2$.

Our rank mobility indices M_α can be rewritten as linear functions of T_n and the condition (47), required to establish the asymptotic normality of T_n , is respected for $-0.5 < \alpha < 1.5$. It follows that M_α is asymptotically normal, for $-0.5 < \alpha < 1.5$. Even if the asymptotic variance is not analytically tractable, the existence of the asymptotic distribution provides an asymptotic justification for using the bootstrap to perform statistical inference.

6 Finite sample performance

We now turn to the way mobility indices within the superclass perform in practice. We study the finite sample properties of M_α for the two families of measures within the superclass: a family of income-mobility measures and a family of rank-mobility measures. We do this for the case where there are as many classes as observations.

The coverage error rate of a confidence interval is the probability that the random interval does not include, or cover, the true value of the parameter. A method of constructing confidence intervals with good finite sample properties should provide a coverage error rate close to the nominal rate. For a confidence interval at 95%, the nominal coverage error rate is equal to 5%. In this section, we use Monte-Carlo simulation to approximate the coverage error rate of asymptotic and bootstrap confidence intervals in several experimental designs.

Three methods are considered to calculate confidence intervals: asymptotic, percentile bootstrap and studentized bootstrap methods. The asymptotic confidence interval is equal to

$$CI_{asym} = [M_\alpha - c_{0.975} \widehat{\text{Var}}(M_\alpha)^{1/2}; M_\alpha + c_{0.975} \widehat{\text{Var}}(M_\alpha)^{1/2}] \quad (53)$$

where $c_{0.975}$ is a critical value obtained from the Student distribution $T(n-1)$. Asymptotic confidence intervals do not always perform well in finite samples. When asymptotic confidence intervals give poor coverage, bootstrap confidence intervals can be expected to perform better. A variety of bootstrap intervals can be used - for a comprehensive discussion, see ?). A first method, called the *percentile bootstrap* method, does not require the computation and the use of the (asymptotic) standard error of the mobility measure estimated. We generate B bootstrap samples, by resampling in the original data, and then, for each resample, we compute the mobility index. We obtain B bootstrap statistics, M_α^b , $b = 1, \dots, B$. The percentile bootstrap confidence interval is equal to

$$CI_{perc} = [c_{0.025}^b; c_{0.975}^b] \quad (54)$$

where $c_{0.025}^b$ and $c_{0.975}^b$ are the 2.5 and 97.5 percentiles of the EDF of the bootstrap statistics. A second method, called the *studentized bootstrap* method, makes use of the asymptotic standard error of the mobility measure estimated. We generate B bootstrap samples, by resampling in the original data, and then, for each resample, we compute a t -statistic. We obtain B bootstrap t -statistics $t_\alpha^b = (M_\alpha^b - M_\alpha) / \widehat{\text{Var}}(M_\alpha^b)^{1/2}$, $b = 1, \dots, B$,

where M_α is the mobility index computed with the original data. The studentized bootstrap confidence interval is equal to

$$CI_{stud} = [M_\alpha - c_{0.975}^* \widehat{\text{Var}}(M_\alpha)^{1/2}; M_\alpha - c_{0.025}^* \widehat{\text{Var}}(M_\alpha)^{1/2}] \quad (55)$$

where $c_{0.025}^*$ and $c_{0.975}^*$ are the 2.5 and 97.5 percentiles of the EDF of the bootstrap t -statistics. It is also called a bootstrap- t or a percentile- t confidence interval. The main difference between the two bootstrap methods is that the studentized bootstrap confidence interval is based on an asymptotically pivotal statistic, not the percentile bootstrap confidence interval. Indeed, the t -statistics follow asymptotically a known distribution, which does not depend on unknown parameters. This property is known to provide superior statistical performance of the bootstrap over asymptotic confidence intervals (?). Note that both bootstrap confidence intervals are asymmetric. Then, they should provide more accurate confidence intervals than the asymptotic confidence interval when the exact distribution of the statistic is not symmetric. For well-known reasons - see ?) or ?) - the number of bootstrap resamples B should be chosen so that $(B + 1)/100$ is an integer. In what follows, we set $B = 199$.

In our experiments, samples are drawn from a Bivariate Lognormal distribution with parameters

$$(x_0, x_1) \sim LN(\mu, \Sigma) \quad \text{with} \quad \mu = (0, 0) \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad (56)$$

where μ and Σ are the mean and the square root of the covariance matrix of the variable's natural logarithm. The case $\rho = 1$ corresponds to zero mobility and the case $\rho = 0$ corresponds to incomes in periods 0 and 1 (resp. x_0 and x_1) being independently generated. Then mobility should increase as ρ decreases. The asymptotic distribution is undefined for the case of zero mobility ($\rho = 1$); it is thus interesting to study the statistical properties in case of "nearly" zero mobility ($\rho = 0.99$). In the experiments, we consider different mobility indices ($\alpha = -1, -0.5, 0, 0.5, 1, 1.5, 2$), different sample sizes ($n = 100, 200, 500, 1\,000, 5\,000, 10\,000$) and different mobility levels ($\rho = 0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.99$).

For fixed values of α , n and ρ , we draw 10 000 samples from the bivariate lognormal distribution. For each sample we compute M_α and its confidence interval at 95%. The coverage error rate is computed as the proportion of times the true value of the mobility index is not included in the confidence intervals. The true value of the mobility index is approximated from a sample of a million observations. Confidence intervals perform well in finite sample if the coverage error rate is close to the nominal value, that is, close to the value 0.05.

6.1 Income mobility

Let us consider the *distribution-independent, static* status, as defined in (36). Here the income values are used to evaluate individual status: this corresponds to a model of movement of incomes.

Table 2 shows coverage error rates of asymptotic confidence intervals at 95%. If the asymptotic distribution is a good approximation of the exact distribution of the statistic,

α		-1	0	0.5	1	2
$n = 100,$	$\rho = 0$	0.3686	0.1329	0.1092	0.1357	0.3730
$n = 100,$	$\rho = 0.2$	0.3160	0.1334	0.1136	0.1325	0.3194
$n = 100,$	$\rho = 0.4$	0.2664	0.1353	0.1221	0.1351	0.2889
$n = 100,$	$\rho = 0.6$	0.2175	0.1346	0.1275	0.1361	0.2263
$n = 100,$	$\rho = 0.8$	0.1718	0.1349	0.1304	0.1345	0.1753
$n = 100,$	$\rho = 0.9$	0.1528	0.1321	0.1308	0.1329	0.1531
$n = 100,$	$\rho = 0.99$	0.1355	0.1340	0.1331	0.1324	0.1333
$n = 200,$	$\rho = 0$	0.3351	0.1077	0.0923	0.1107	0.3153
$n = 500,$	$\rho = 0$	0.2594	0.0830	0.0696	0.0818	0.2631
$n = 1000,$	$\rho = 0$	0.2164	0.0703	0.0609	0.0726	0.2181
$n = 5000,$	$\rho = 0$	0.1713	0.0554	0.0469	0.0522	0.2066
$n = 10000,$	$\rho = 0$	0.1115	0.0532	0.0527	0.0534	0.1151

Table 2: Coverage error rate of asymptotic confidence intervals at 95% of income mobility measures. The nominal error rate is 0.05, 10.000 replications

the coverage error rate should be close to the nominal error rate, 0.05. From Table 2, we can see that:

- asymptotic confidence intervals always perform poorly for $\alpha = -1, 2$,
- the coverage error rate is stable as ρ varies (for $\alpha = 0, 0.5, 1$ and $n = 100$),
- the coverage error rate decreases as n increases,
- the coverage error rate is close to 0.05 for $n \geq 5.000$ and $\alpha = 0, 0.5, 1$.

These results suggest that asymptotic confidence intervals perform well in very large sample, with $\alpha \in [0, 1]$.

The dismal performance of asymptotic confidence intervals for small and moderate samples is sufficient to motivate the use of bootstrap methods. Table 3 shows coverage error rates of asymptotic and bootstrap confidence intervals at 95%. We select the value $\rho = 0.8$, because it gives the poorest results for asymptotic confidence intervals with $\alpha \in [0, 1]$ in Table 2. It is clear from Table 3 that:

- percentile bootstrap and asymptotic confidence intervals perform similarly,
- studentized bootstrap confidence intervals outperform other methods,

These results show that studentized bootstrap confidence intervals provide significant improvements over asymptotic confidence intervals.

6.2 Rank mobility

Let us consider the *distribution-dependent, dynamic* status, as defined in (44). Here ranks (the income positions) are used to evaluate individual status; it corresponds to a

α	-1	0	0.5	1	2
$n = 100, \rho = 0.8$					
Asymptotic	0.1718	0.1349	0.1304	0.1345	0.1753
Boot-perc	0.1591	0.1294	0.1215	0.1266	0.1552
Boot-stud	0.0931	0.0751	0.0732	0.076	0.0952
$n = 200, \rho = 0.8$					
Asymptotic	0.1315	0.0973	0.0927	0.0973	0.1276
Boot-perc	0.1222	0.0943	0.0900	0.0950	0.1176
Boot-stud	0.0794	0.0666	0.0660	0.0688	0.0791
$n = 500, \rho = 0.8$					
Asymptotic	0.1127	0.0847	0.0828	0.0857	0.1124
Boot-perc	0.1054	0.0814	0.0813	0.0843	0.1036
Boot-stud	0.0765	0.0641	0.0629	0.0630	0.0779
$n = 1.000, \rho = 0.8$					
Asymptotic	0.0880	0.0678	0.0659	0.0672	0.0864
Boot-perc	0.0862	0.0672	0.0661	0.0689	0.0851
Boot-stud	0.0680	0.0585	0.0589	0.0596	0.0693

Table 3: Coverage error rate of asymptotic and bootstrap confidence intervals at 95% of income mobility measures. 10 000 replications, 199 bootstraps

model of rank mobility. Since the variance of M_α is not analytically tractable, we cannot use asymptotic and studentized bootstrap confidence intervals. We use the percentile bootstrap method.

α	-0.5	0	0.5	1	1.5
$\rho = 0$	0.5592	0.1575	0.1088	0.1583	0.5282
$\rho = 0.2$	0.3176	0.1122	0.0884	0.1135	0.3231
$\rho = 0.4$	0.1883	0.0931	0.0755	0.0913	0.1876
$\rho = 0.6$	0.1122	0.0767	0.0651	0.0741	0.1118
$\rho = 0.8$	0.0671	0.0593	0.0555	0.0590	0.0652
$\rho = 0.9$	0.0432	0.0430	0.0431	0.0441	0.0446
$\rho = 0.99$	0.0983	0.0985	0.0981	0.0984	0.0992

Table 4: Coverage error rate of percentile bootstrap confidence intervals at 95% of rank-mobility measures. 10 000 replications, 199 bootstraps and 100 observations.

Table 4 shows coverage error rates of percentile bootstrap confidence intervals at 95% with $n = 100$ observations. We can see that:

- the coverage error rate can be very different for different values of ρ and α ,
- it decreases as ρ increases, except for the case of “nearly” zero mobility ($\rho = 0.99$).
- the coverage error rate is close to 0.05 for $\rho = 0.8, 0.9$ and $\alpha = 0, 0.5, 1$.

α	-0.5	0	0.5	1	1.5
$n = 100, \rho = 0$	0.5592	0.1575	0.1088	0.1583	0.5282
$n = 200$	0.4613	0.1143	0.0833	0.1180	0.4723
$n = 500$	0.3548	0.0868	0.0645	0.0814	0.3644
$n = 1000$	0.3135	0.0672	0.0556	0.0735	0.3170
$n = 100, \rho = 0.9$	0.0432	0.0430	0.0431	0.0441	0.0446
$n = 200$	0.0454	0.0441	0.0456	0.0454	0.0459
$n = 500$	0.0500	0.0499	0.0485	0.0480	0.0483
$n = 1000$	0.0511	0.0509	0.0539	0.0538	0.0538
$n = 100, \rho = 0.99$	0.0983	0.0985	0.0981	0.0984	0.0992
$n = 200$	0.0981	0.0971	0.0970	0.0974	0.0977
$n = 500$	0.0855	0.0838	0.0833	0.0822	0.083
$n = 1000$	0.0788	0.0777	0.0762	0.0767	0.0771

Table 5: Coverage error rate of percentile bootstrap confidence intervals at 95% of rank-mobility measures. 10 000 replications, 199 bootstraps.

These results suggest that percentile bootstrap confidence intervals perform well in small sample in the presence of low but significant mobility levels ($\rho = 0.8, 0.9$) and for $\alpha \in [0, 1]$.

Table 5 shows coverage error rates of percentile bootstrap confidence intervals at 95% as the sample size increases. We can see that:

- the coverage error rate gets closer to 0.05 as the sample size increases,
- the coverage error rate is smaller when $\alpha = 0, 0.5, 1$.

These results show that percentile bootstrap confidence intervals have better statistical properties as the sample size increases.

7 Conclusion

What makes our approach to mobility measurement novel is not the introduction of a new specific index but rather a way of rethinking the representation of the problem and then the theoretical and statistical treatment of this representation of mobility. The key step involves a logical separation of fundamental concepts, (1) the measure of individual status and (2) the aggregation of changes in status.

The status concept is derived directly from information available in the marginal distributions. It could involve the simplest derivation - the assumption that status equals income. Or it could involve something more sophisticated, incorporating the person's location in the income distribution. This is a matter for normative judgment.

The aggregation of changes in status involves the application of standard principles to status pairs. From this one derives a superclass of mobility measures - a class of classes of measures. As we have seen this is generally applicable to a wide variety of status concepts and, for any given status concept, the members of the class are indexed by a

parameter α that determines the type of mobility measure. Each measure in each class of the superclass involves a kind of averaging of individual mobilities and the evaluation of individual mobility depends on status in the two periods, but no more (in our approach rank may be important for status but not for quantifying movement). Every measure in the superclass has attractive scale properties that imply structural regularity, but no more than that; once again this is because status can be separated from - if not divorced from - income and wealth.

We have also shown that the principal status types that are likely to be adopted in practice will result in statistically tractable mobility indices. Bootstrap confidence intervals perform well in moderate sample sizes for α in the interval $[0, 1]$, in the cases of both income mobility and rank mobility.

References

A Proofs

Proof. [Theorem 1] . Axioms 1 to 4 imply that \succeq can be represented by a continuous function $\Phi : Z^n \rightarrow \mathbb{R}$ that is increasing in $|u_i - v_i|$, $i = 1, \dots, n$. Using Axiom 3 part (a) of the result follows from Theorem 5.3 of ?). Now take \mathbf{z}' and \mathbf{z} as specified in Axiom 4. Using (11) it is clear that $\mathbf{z} \sim \mathbf{z}'$ if and only if

$$\phi_i(u_i + \delta, u_i + \delta) - \phi_i(u_i, u_i) = 0$$

which can only be true if

$$\phi_i(u_i + \delta, u_i + \delta) - \phi_i(u_i, u_i) = f(\delta)$$

for arbitrary u_i and δ . This is a standard Pexider equation and its solution implies (12). ■

Proof. [Theorem 2] Using the function Φ introduced in the proof of Theorem 1, Axiom 5 implies

$$\begin{aligned}\Phi(\mathbf{z}) &= \Phi(\mathbf{z}') \\ \Phi(t\mathbf{z}) &= \Phi(t\mathbf{z}')\end{aligned}$$

and so, since this has to be true for arbitrary \mathbf{z}, \mathbf{z}' we have

$$\frac{\Phi(t\mathbf{z})}{\Phi(\mathbf{z})} = \frac{\Phi(t\mathbf{z}')}{\Phi(\mathbf{z}')} = \psi(t)$$

where ψ is a continuous function $\mathbb{R} \rightarrow \mathbb{R}$. Hence, using the ϕ_i given in (11), we have for all \mathbf{z} :

$$\phi_i(tz_i) = \psi(t) \phi_i(z_i), \quad i = 1, \dots, n.$$

or, equivalently

$$\phi_i(tu_i, tv_i) = \psi(t) \phi_i(u_i, v_i), \quad i = 1, \dots, n. \quad (57)$$

So, in view of ?), page 346 there must exist $c \in \mathbb{R}$ and a function $H_i : \mathbb{R}_+ \rightarrow \mathbb{R}$ such that

$$\phi_i(u_i, v_i) = u_i^c H_i\left(\frac{u_i}{v_i}\right). \quad (58)$$

Using (58) in (12) gives

$$u_i^c H_i(1) = a_i + b_i u_i, \quad (59)$$

which implies that ϕ_i must take the form implied in one of these three cases:

Case (a) $b_i \neq 0$.

If b_i is nonzero then c must equal 1 and $H_i(1)$ must be nonzero. Therefore $\phi_i(u_i, v_i) = u_i H_i\left(\frac{u_i}{v_i}\right)$.

Case (b) $b_i = 0$ and $H_i(1) \neq 0$.

If $b_i = 0$ then $c = 0$. Therefore $a_i = H_i(1)$ and so $\phi_i(u_i, v_i) = H_i\left(\frac{u_i}{v_i}\right)$.

Case (c) $b_i = 0$ and $H_i(1) = 0$.

In this case a_i must be zero so that $\phi_i(u_i, v_i) = u_i^c \log(u_i/v_i)$. ■

Proof. [Theorem 3] Clearly the result is true in case (c) of Theorem 2. Now consider the other two cases.

Case (a) $\phi_i(u_i, v_i) = u_i H_i\left(\frac{u_i}{v_i}\right)$.

Take the special movement profile where, in distribution \mathbf{z}'_0 the individual movement takes the same value r for all n persons. If (u_i, v_i) represents a typical component in \mathbf{z}_0 then $\mathbf{z}_0 \sim \mathbf{z}'_0$ implies

$$r = \psi\left(\sum_{i=1}^n u_i H_i\left(\frac{u_i}{v_i}\right)\right) \quad (60)$$

where ψ is the solution in r to

$$\sum_{i=1}^n u_i H_i\left(\frac{u_i}{v_i}\right) = \sum_{i=1}^n u_i H_i(r) \quad (61)$$

In (61) we can take the u_i as fixed weights. Using Axiom 6 in (60) requires

$$tr = \psi\left(\sum_{i=1}^n u_i H_i\left(t\frac{u_i}{v_i}\right)\right), \text{ for all } t > 0. \quad (62)$$

Using (61) we have

$$\sum_{i=1}^n u_i H_i\left(t\psi\left(\sum_{i=1}^n u_i H_i\left(\frac{u_i}{v_i}\right)\right)\right) = \sum_{i=1}^n u_i H_i\left(t\frac{u_i}{v_i}\right) \quad (63)$$

Introduce the following change of variables

$$\tau_i := u_i H_i\left(\frac{u_i}{v_i}\right), i = 1, \dots, n \quad (64)$$

and write the inverse of this relationship as

$$\frac{u_i}{v_i} = \psi_i(\tau_i), i = 1, \dots, n \quad (65)$$

Substituting (64) and (65) into (63) we get

$$\sum_{i=1}^n u_i H_i\left(t\psi\left(\sum_{i=1}^n \tau_i\right)\right) = \sum_{i=1}^n u_i H_i(t\psi_i(\tau_i)). \quad (66)$$

Also define the following functions

$$\theta_0(\tau, t) := \sum_{i=1}^n u_i H_i(t\psi(\tau)) \quad (67)$$

$$\theta_i(\tau, t) := u_i H_i(t\psi_i(\tau)), i = 1, \dots, n. \quad (68)$$

Substituting (67),(68) into (66) we get the Pexider functional equation

$$\theta_0\left(\sum_{i=1}^n \tau_i, t\right) = \sum_{i=1}^n \theta_i(\tau_i, t) \quad (69)$$

which has as a solution

$$\theta_i(\tau, t) = b_i(t) + B(t)\tau, i = 0, 1, \dots, n$$

where

$$B(t) = \sum_{i=1}^n b_i(t)$$

– see ?), page 142. Therefore we have

$$H_i\left(t\frac{u_i}{v_i}\right) = \frac{b_i(t)}{u_i} + B(t)H_i\left(\frac{u_i}{v_i}\right), i = 1, \dots, n \quad (70)$$

From ?), Theorem 2.7.3 the solution to (70) is of the form

$$H_i(r) = \begin{cases} \beta_i r^b + \gamma_i, & b \neq 0 \\ \beta_i \log r + \gamma_i & b = 0 \end{cases} \quad (71)$$

where $\beta_i > 0$. Applying Axiom 7 β_i must be the same for all i and the result follows.

Case (b) $\phi_i(u_i, v_i) = H_i\left(\frac{u_i}{v_i}\right)$.

Instead of (63) we now have

$$\sum_{i=1}^n H_i\left(t\psi\left(\sum_{i=1}^n H\left(\frac{u_i}{v_i}\right)\right)\right) = \sum_{i=1}^n H_i\left(t\frac{u_i}{v_i}\right) \quad (72)$$

Introduce the following change of variables

$$\tau_i := H_i\left(\frac{u_i}{v_i}\right), i = 1, \dots, n \quad (73)$$

and write the inverse of this relationship as

$$\frac{u_i}{v_i} = \psi_i(\tau_i), i = 1, \dots, n \quad (74)$$

Substituting (73) and (74) into (72) we get

$$\sum_{i=1}^n H_i\left(t\psi\left(\sum_{i=1}^n \tau_i\right)\right) = \sum_{i=1}^n H_i(t\psi_i(\tau_i)). \quad (75)$$

Also define the following functions

$$\tilde{\theta}_0(\tau, t) := \sum_{i=1}^n H(t\psi(\tau)) \quad (76)$$

$$\tilde{\theta}_i(\tau, t) := H_i(t\psi_i(\tau)), i = 1, \dots, n. \quad (77)$$

Once again we obtain the functional equation (69) but with the functions θ_i replaced by $\tilde{\theta}_i$, $i = 0, 1, \dots, n$; the remainder of the proof follows case (a).

■