Measuring Mobility[*]

by

Frank A. Cowell[†]

and

Emmanuel Flachaire[‡]

April 2011

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel: (020) 7955 6674

PEP 09

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

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

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

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: Mobility measures, axiomatic approach, bootstrap
JEL codes: D63

Public Economics Programme

The Public Economics Programme was established in 2009. It is located within the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD) at the London School of Economics and Political Science. The programme is directed by Frank Cowell and Henrik Kleven. The Discussion Paper series is available free of charge. To subscribe to the PEP paper series, or for further information on the work of the Programme, please contact our Research Secretary, Leila Alberici on:

Telephone:	UK+20 7955 6674
Fax:	UK+20 7955 6951
Email:	l.alberici@lse.ac.uk
Web site:	http://sticerd.lse.ac.uk/PEP

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 (Fields and Ok 1999b) - 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 (Dardanoni 1993), 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.

[1]For a survey see Fields and Ok (1999a).

[2]See, for example, Atoda and Tachibanaki (1991), Bénabou and Ok (2001).

Because mobility is inherently quite a complex phenomenon it is common to find that the phenomenon is 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 (see above). 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 (Ayala and Sastre 2008, Van de gaer et al. 2001).[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 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 $x_k < x_{k+1}$, $k = 1, 2, ..., K-1$. Let $p_k \in \mathbb{R}_+$ be the size of class k, $k = 1, 2, ..., K$ 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 $\left(x_{k_0(1)}, ..., x_{k_0(n)}\right)$. Clearly this includes the special case where classes are individual incomes if $p_k = 0$ or 1 in each class and the individual is assumed to be at the lower bound of the class.

[3]For an illuminating discussion see Van Kerm (2004). On the definition of exchange mobility see Tsui (2009),

[4]Several ad hoc measures of income mobility pursue the idea of average distance (Mitra and Ok 1998). Fields and Ok (1996, 1999b) proposed a mobility index whose distance concept is based on the absolute differences of logarithms of incomes.

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 $\left(x_{k_0(1)}, ..., x_{k_0(n)}\right)$ and $\left(x_{k_1(1)}, ..., x_{k_1(n)}\right)$. 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:

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

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

$$n_1(x_k) := \sum_{h=1}^{k} q_h, \; k = 1, ..., K. \tag{2}$$

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 $i = 1, 2, ..., n$, 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:
$$z_i = \left(x_{k_0(i)}, x_{k_1(i)}\right). \tag{3}$$
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:
$$z_i = \left(\varphi\left(x_{k_0(i)}\right), \varphi\left(x_{k_1(i)}\right)\right), \tag{4}$$
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 rather 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 require 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 :
$$z_i = \left(n_0\left(x_{k_0(i)}\right), n_0\left(x_{k_1(i)}\right)\right). \tag{5}$$
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 it 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:
$$z_i = \left(n_0\left(x_{k_0(i)}\right), n_1\left(x_{k_1(i)}\right)\right). \tag{6}$$
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. In particular 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

$$v_i = u_i, i = 1, ..., n. \tag{7}$$

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

$$x_k = \lambda x_{k-1}, k = 2, ..., K, \lambda > 1. \tag{8}$$

[5]If there were an exogenous revaluation of the K classes so that $(x_1, ..., x_K)$ in period 0 changes to $(y_1, ..., y_K)$ in period 1 - perhaps because of inflation or economic growth - then clearly one could also consider a *distribution-independent, dynamic* case where $z_i = \left(x_{k_0(i)}, y_{k_1(i)}\right)$. 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
x_1	A	_
x_2	B	A
x_3	C	B
x_4	_	C
x_5	_	_

Table 1: Zero mobility?

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

$$v_i = \lambda u_i, i = 1, ..., n, \lambda > 0. \tag{9}$$

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.

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 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

$$U := \left\{ \mathbf{u} | \ \mathbf{u} \in \mathbb{R}^n_+, u_1 \leq u_2 \leq ... \leq u_n \right\}. \tag{10}$$

[6] An early treatment of this type of problem for the specific case where status equals income is given in Cowell (1985). 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 Cowell (1985); second, the current paper deals with any arbitrary representation of status (including ordinal status) rather than being specific to income.

Individual mobility is completely characterised by the ordered pairs z_i, $i = 1, 2, ..., n$ 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

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

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

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

where the function $m : Z \to \mathbb{R}$ 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 ith 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 ith 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 $\mathbf{z}, \mathbf{z}' \in Z^n$ be such that, for some i and j, $u_i = v_i$, $u_j = v_j$, $u_i' = u_i + \delta$, $v_i' = v_i + \delta$, $u_j' = u_j - \delta$, $v_j' = v_j - \delta$ and, for all $h \neq i, j$, $u_h' = u_h$, $v_h' = v_h$. 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 seems 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*

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

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

$$\phi_i(u, u) = a_i + b_i u. \tag{12}$$

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

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

where ϕ_i is defined as in (11), (12) and $\phi : \mathbb{R} \to \mathbb{R}$ 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*

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

where H_i is a real-valued function.

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 $m(\mathbf{z}') = tm(\mathbf{z}_0')$: $\mathbf{z} \sim \mathbf{z}'$.*

The principle states this. Suppose we have two profiles $\mathbf{z}_0$ and $\mathbf{z}_0'$ that are regarded as equivalent under $\succeq$. Then scale up (or down) 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]

Theorem 3 *Given Axioms 1 to 6 $\succeq$ is representable by*

$$\Phi(\mathbf{z}) = \phi\left(\sum_{i=1}^{n} u_i^{\alpha} v_i^{1-\alpha}\right) \tag{15}$$

where $\alpha \neq 1$ is a constant.

3.4 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 :

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

Theorem 3 implies that, for all $\mathbf{z} \in Z(\bar{u}, \ \bar{v})$, the mobility index must take the form

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

where $\bar{u}, \ \bar{v}$ are parameters of the function $\bar{\phi}$ that is the counterpart of ϕ in (15). 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

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

[7] Also 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.

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 $v_i = \lambda u_i$, $i = 1,..,n$ (where $\lambda = \bar{v}/\bar{u}$) then, from (16), we have

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

which implies

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

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

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

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

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

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

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

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

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).

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 (Chakravarty 1984, Van Kerm 2009). 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.[8] 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 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 Demuynck and Van de gaer 2010) 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) unambiguously[9] 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

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

where the weights $w^{\mathrm{P}}, w^{\mathrm{R}}$ and the between-group mobility component $M_\alpha^{\mathrm{between}}$ 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,

[8] See, for example, Atkinson (1981, 1983), Bibby (1975), D'Agostino and Dardanoni (2009), Kearl and Pope (1984), Shorrocks (1978).

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

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

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

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

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

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 of the mobility index (borrowing the terminology of Bossert and Pfingsten 1990, Eichhorn 1988). However, by writing

$$\alpha(c):=\gamma+\beta c \tag{26}$$

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

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

Using the standard expansion

$$\log\left(1+t\right)=t-\frac{t^2}{2}+\frac{t^3}{3}-... \tag{28}$$

and (26) we find that (27) becomes

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

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

$$\beta\left[u-\mu_u-v+\mu_v\right] \tag{30}$$

and

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

From (30) and (31) we can see that in the limit (24) becomes[10]

$$\frac{1}{n\beta^2}\sum_{i=1}^{n}\left[\mathrm{e}^{\beta[u_i-\mu_u-v_i+\mu_v]}-1\right], \tag{32}$$

[10] See also equation (56) of Cowell (1985).

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

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

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

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

As $\beta \to 0$ 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

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

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.

Finally, should mobility indices be "ethical" indices?[11] Some recent contributions to the literature have built this in - notably the "extended Atkinson" approach of Gottschalk and Spolaore (2002), which is based on Atkinson and Bourguignon (1982). 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.[12] It enables us to capture *directional sensitivity* in the mobility context:[13] 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.

[11] The classic paper by King (1983) is explicitly based on a social-welfare function..Other early contributions using a welfare-based approach see Markandya (1982, 1984) and Chakravarty et al. (1985) later reinterpreted by Ruiz-Castillo (2004).

[12] In 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.

[13] See also Demuynck and Van de gaer (2010) and Schluter and Van de gaer (2011).

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,

$$u_i = x_{0i} \qquad \text{and} \qquad v_i = x_{1i}, \tag{36}$$

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 $g(.)$ 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

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

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

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

Under standard regularity conditions, the Central Limit Theorem can be applied and thus the M_α index will follows 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 $\mu_{u^\alpha v^{1-\alpha}}$, the variance estimator for M_α is :

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

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

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

The covariance matrix $\hat{\Sigma}$ is defined as follows:[14]

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

We can use this variance estimator of M_α to compute a test statistic or a confidence interval.[15]

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

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

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

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

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

The variance estimator of this index is defined as follows:

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

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:

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

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

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

The variance estimator of this index is defined as follows:

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

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,

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

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

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

where $k = 1, 2$ and $I(.)$ 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 $\{\frac{1}{n}, \frac{2}{n}, \ldots, 1\}$. 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.

Ruymgaart and van Zuijlen (1978) have established the asymptotic normality in the non i.i.d. case of the following multivariate rank statistic,

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

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

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

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,

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

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

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

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

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

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

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

where k and l are real constants[16] 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

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

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

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

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

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

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 Davison and Hinkley (1997). 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, $b = 1, \ldots, B$. The percentile bootstrap confidence interval is equal to

$$IC_{perc} = [c_{0.025}^b \,;\; c_{0.975}^b] \tag{54}$$

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 $t_\alpha^b = (M_\alpha^b - M_\alpha)/\widehat{\mathrm{Var}}(M_\alpha^b)^{1/2}$, $b = 1, \ldots, B$, where M_α is the mobility index computed with the original data. The studentized bootstrap confidence interval is equal to

$$IC_{stud} = [M_\alpha - c_{0.975}^* \widehat{\mathrm{Var}}(M_\alpha)^{1/2} \,;\; M_\alpha - c_{0.025}^* \widehat{\mathrm{Var}}(M_\alpha)^{1/2}] \tag{55}$$

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 (Beran 1987). 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 Davison and Hinkley (1997) or Davidson and MacKinnon (2000) - 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

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

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 increases 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.

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

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, 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 $\alpha \in [0, 1]$.

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

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 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.

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 ($\rho = 0.99$).

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.

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.

- the coverage error rate is close to 0.05 for $\rho = 0.8, 0.9$ and $\alpha = 0, 0.5, 1$.

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

Aczél, J. (1966). *Lectures on Functional Equations and their Applications*. Number 9 in Mathematics in Science and Engineering. New York: Academic Press.

Aczél, J. and J. G. Dhombres (1989). *Functional Equations in Several Variables*. Cambridge: Cambridge University Press.

Atkinson, A. B. (1981). On intergenerational income mobility in Britain. *Journal of Post Keynesian Economics 3*, 194–218.

Atkinson, A. B. (1983). The measurement of economic mobility. In A. B. Atkinson (Ed.), *Social Justice and Public Policy*, Chapter 3, pp. 61–75. Hemel Hempstead, UK.: Harvester Wheatsheaf.

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

Atoda, N. and T. Tachibanaki (1991). Earnings distribution and inequality over time: Education versus relative position and cohort. *International Economic Review 32*, 475–489.

Ayala, L. and M. Sastre (2008). The structure of income mobility: empirical evidence from five UE countries. *Empirical Economics 35*, 451–473.

Bénabou, R. and E. A. Ok (2001). Mobility as progressivity: Ranking income processes according to equality of opportunity. NBER Working Paper W8431, National Bureau of Economic Research.

Beran, R. (1987). Prepivoting to reduce level error of confidence sets. *Biometrika 74*, 457–468.

Bibby, J. (1975). Methods of measuring mobility. *Quality and Quantity 9*, 107–136.

Bossert, W. and A. Pfingsten (1990). Intermediate inequality: concepts, indices and welfare implications. *Mathematical Social Science 19*, 117–134.

Chakravarty, S. R. (1984). Normative indices for measuring social mobility. *Economics Letters 15*, 175–180.

Chakravarty, S. R., B. Dutta, and J. Weymark (1985). Ethical indices of income mobility. *Social Choice and Welfare 2*, 1–21.

Cowell, F. A. (1985). Measures of distributional change: An axiomatic approach. *Review of Economic Studies 52*, 135–151.

D'Agostino, M. and V. Dardanoni (2009). The measurement of rank mobility. *Journal of Economic Theory 144*, 1783–1803.

Dardanoni, V. (1993). Measuring social mobility. *Journal of Economic Theory 61*, 372–394.

Davidson, R. and J. G. MacKinnon (2000). Bootstrap tests: How many bootstraps? *Econometric Reviews 19*, 55–68.

Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods*. Cambridge: Cambridge University Press.

Demuynck, T. and D. Van de gaer (2010, January). Rank dependent relative mobility measures. Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium 10/628, Ghent University, Faculty of Economics and Business Administration.

Eichhorn, W. (1978). *Functional Equations in Economics*. Reading Massachusetts: Addison Wesley.

Eichhorn, W. (1988). On a class of inequality measures. *Social Choice and Welfare 5*, 171–177.

Fields, G. S. and E. A. Ok (1996). The meaning and measurement of income mobility. *Journal of Economic Theory 71*(2), 349–377.

Fields, G. S. and E. A. Ok (1999a). The measurement of income mobility: an introduction to the literature. In J. Silber (Ed.), *Handbook on Income Inequality Measurement*. Dewenter: Kluwer.

Fields, G. S. and E. A. Ok (1999b). Measuring movement of incomes. *Economica 66*, 455–472.

Fishburn, P. C. (1970). *Utility Theory for Decision Making*. New York: John Wiley.

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

Kearl, J. R. and C. L. Pope (1984). Mobility and distribution. *Review of Economics and Statistics 66*, 192–199.

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

Markandya, A. (1982). Intergenerational exchange mobility and economic welfare. *European Economic Review 17*, 301–324.

Markandya, A. (1984). The welfare measurement of changes in economic mobility. *Economica 51*, 457–471.

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

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

Ruymgaart, F. H. and M. C. A. van Zuijlen (1978). Asymptotic normality of multivariate linear rank statistics in the non-i.i.d. case. *Annals of Statistics 6*, 588–602.

Schluter, C. and D. Van de gaer (2011). Structural mobility, exchange mobility and subgroup consistent mobility measurement, US - German mobility rankings revisited. *Review of Income and Wealth forthcoming*.

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

Tsui, K. (2009). Measurement of income mobility: A re-examination. *Social Choice and Welfare 33*, 629–645.

Van de gaer, D., E. Schokkaert, and M. Martinez (2001). Three meanings of intergenerational mobility. *Economica 68*, 519–537.

Van Kerm, P. (2004). What lies behind income mobility? Reranking and distributional change in Belgium, Western Germany and the USA. *Economica 71*, 223–239.

Van Kerm, P. (2009). Income mobility profiles. *Economics Letters 102*, 93–95.

A Proofs

Proof. [Proof of Theorem 1] . Axioms 1 to 4 imply that $\succeq$ can be represented by a continuous function $\Phi : Z^n \to \mathbb{R}$ that is increasing in $|u_i - v_i|$, $i = 1, ..., n$. Using Axiom 3 part (a) of the result follows from Theorem 5.3 of Fishburn (1970). Now take $\mathbf{z}'$ and $\mathbf{z}$ in as specified in Axiom 4. Using (11) it is clear that $\mathbf{z} \sim \mathbf{z}'$ if and only if

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

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. [Proof of 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 $\mathbb{R} \to \mathbb{R}$. Hence, using the ϕ_i given in (11), we have for all $\mathbf{z}$:

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

or, equivalently

$$\phi_i(tu_i, tv_i) = \psi(t)\,\phi_i(u_i, v_i), i = 1, ..., n. \tag{57}$$

So, in view of Aczél and Dhombres (1989), page 346 there must exist $c \in \mathbb{R}$ and a function $H_i : \mathbb{R}_+ \to \mathbb{R}$ such that

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

From (12) and (58) it is clear that

$$\phi_i(u_i, u_i) = u_i^c H_i(1) = a_i + b_i u_i, \tag{59}$$

which implies $c = 1$. Putting (58) with $c = 1$ into (13) gives the result. ■

Proof. [Proof of Theorem 3] Take the special case where, in distribution $\mathbf{z}_0'$ the individual movement takes the same value r for all n. If (u_i, v_i) represents a typical component in $\mathbf{z}_0$ then $\mathbf{z}_0 \sim \mathbf{z}_0'$ implies

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

where ψ is the solution in r to

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

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

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

Using (61) we have

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

Introduce the following change of variables

$$u_i := u_i H_i\left(\frac{u_i}{v_i}\right), i = 1, ..., n \tag{64}$$

and write the inverse of this relationship as

$$\frac{u_i}{v_i} = \psi_i\left(u_i\right), i = 1, ..., n \tag{65}$$

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

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

Also define the following functions

$$\theta_0\left(u, t\right) := \sum_{i=1}^{n} u_i H_i\left(t\psi\left(u\right)\right) \tag{67}$$

$$\theta_i\left(u, t\right) := u_i H_i\left(t\psi_i\left(u\right)\right), i = 1, ..., n. \tag{68}$$

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

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

which has as a solution

$$\theta_i\left(u, t\right) = b_i\left(t\right) + B\left(t\right) u, i = 0, 1, ..., n$$

where

$$b_0\left(t\right) = \sum_{i=1}^{n} b_i\left(t\right)$$

— see Aczél (1966), page 142. Therefore we have

$$H_i\left(t\frac{u_i}{v_i}\right) = \frac{b_i\left(t\right)}{u_i} + B\left(t\right)H_i\left(\frac{u_i}{v_i}\right), i = 1, ..., n \tag{69}$$

From Eichhorn (1978), Theorem 2.7.3 the solution to (69) is of the form

$$H_i\left(v\right) = \begin{array}{ll} \beta_i v^{\alpha-1} + \gamma_i, & \alpha \neq 1 \\ \beta_i \log v + \gamma_i & \alpha = 1 \end{array} \tag{70}$$

where $\beta_i > 0$ is an arbitrary positive number. Substituting for $H_i\left(\cdot\right)$ from (70) into (14) for the case where β_i is the same for all i gives the result. ■