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

Aix-Marseille University
AMSE, CNRS and EHESS
Centre de la Vieille Charité
2, rue de la Charité
13236 Marseille cedex 02, France
email: emmanuel.flachaire@univ-amu.fr

Revised August 2017

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. We also show how our superclass provides a more consistent and intuitive approach to mobility, in contrast to other measures in the literature, and illustrate its performance using recent data from China.

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

JEL codes: D63

Acknowledgments: Flachaire acknowledges the support of the Institut Universitaire de France (IUF) and of the projects ANR-11-IDEX-0001-02 and ANR-16-CE41-0005, managed by the French National Research Agency (ANR). 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 depended on the particular application or even the particular data set under consideration. 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 including cases where the underlying data are categorical. There are two reasons why such an approach is needed.

First, simple pragmatic approaches to mobility can be seriously misleading. As an example, consider the commonly used measure of mobility $1-\beta$, where β is an elasticity coefficient, computed as the ordinary least-squares estimation of the slope coefficient from a linear regression of log-income in period 1 on log-income on period 0. It has been used in almost every empirical study of intergenerational income mobility (Jäntti and Jenkins 2015). However, this index has a major drawback. Even if a large value of $1-\beta$ may provide evidence of significant mobility, a low value does not necessarily imply low mobility. To see why, take three persons with log-incomes equal to, respectively, $(1, 1.5, 2)$ in period 0 and $(1, 3, 2)$ in period 1. In this case, the index $1-\beta$ is equal to zero, suggesting there is no mobility, while there is clear evidence of income mobility (the second person's log-income doubles, while the others remain unchanged). As a further example consider another widespread mobility measure, $1-\rho$, where ρ is the Pearson correlation coefficient. This has the same drawback: if log-incomes are $(1, 1.5, 2)$ in period 0 and $(1, 2, 3)$ in period 1, then this would give a mobility measurement equal to zero. The problem comes from the fact that elasticity and correlation coefficients are designed to capture a linear relationship between two variables: non-linear relationships may remain undetected. The two examples show that the elasticity and correlation-based indices are inadequate to measure income mobility; they illustrate the need to develop mobility measures with appropriate properties.

Second, some of the extensive literature[3] on mobility confounds issues in the analysis that should be kept distinct. For example, in some contributions mobility is tied specifically to income (Fields and Ok 1999b), in others mobility is exclusively in terms of position in the distribution. However, this way of approaching mobility analysis mixes up the definition of mobility along with the definition of an individual's status. As a further example, in some approaches the distinction between mobility and income volatility becomes fuzzy. This is unfortunate since mobility is essentially something that characterises society, or the individual's relationship to the society (Dardanoni 1993),

[1] For a survey see Fields and Ok (1999a) or Jäntti and Jenkins (2015).

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

[3] We examine the performance of some of the principal mobility measures in Section 7.

whereas volatility can be seen as something that could relate to a single individual; mobility would be meaningless for Robinson Crusoe, but income volatility might be very important.

In setting out our approach let us make a brief list of the essential ingredients of a theory of mobility measurement:

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[4] and focus on the interplay between ingredients 2 and 3, the status measure and the basis for aggregation of movements.

The contribution of the paper is, first, to characterise mobility comparisons for an arbitrary definition of status, using an axiomatic framework and, second, to show how to implement our new approach using sample data and applying different concepts of status. Our approach separates out the fundamental components of the mobility-measurement problem, proposes a parsimonious set of axioms for the core theoretical issues and examines the statistical properties of several classes of measures that emerge from the implementation of the theory. The paper is organised as follows. Section 2 sets out in 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 Sections 5 and 6. In Section 7 we examine the performance of other mobility measures suggested in the literature and consider a real-world application of our approach. Section 8 concludes.

2 Individual status and mobility

The concept of "status" is important in an analysis of mobility: it may be defined in a variety of ways, depending on the focus of interest of a particular study. Status could be something that is directly observable and measurable for each individual, independent of information about anyone else, such as a person's income or wealth. Alternatively it could be that an individual'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.

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.[5] However, this traditional breakdown is not so important here. What is crucial in our

[4]The two-period case is taken as standard in almost every empirical study of intergenerational income mobility and in discussion of the concepts such as the Great Gatsby Curve (Corak 2006, 2013; Jäntti and Jenkins 2015) .

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

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 an average of individual distances "travelled" in the reranking process (Ayala and Sastre 2008, Van de gaer et al. 2001).[6] 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.

We will introduce a simple framework that allows for a variety of definitions of status that may be useful in different contexts of mobility analysis. 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 $k(i)$ be the income class occupied by person i; then the information about a distribution is completely characterised by the vector $\left(x_{k(1)}, x_{k(2)}, ..., x_{k(n)}\right)$ where n is the size of the population. Clearly this includes the special case where classes are individual incomes if the number in each class is 0 or 1 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 classes occupied by person i at periods 0 and 1 respectively. Mobility is completely characterised by $\left(x_{k_0(1)}, x_{k_0(2)}, ..., x_{k_0(n)}\right)$ and $\left(x_{k_1(1)}, x_{k_1(2)}, ..., x_{k_1(n)}\right)$. 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:

$$N_0(x_k) := \sum_{h=1}^{k} n_{0h}, \; k = 1, ..., K, \tag{1}$$

where $n_{0k} \in \mathbb{R}_+$ denotes the number of persons in period 0 who are in class k, $k = 1, 2, ..., K$ and $\sum_{k=1}^{K} n_{0k} = n$. We could also do a similar relabelling using information about the 1-distribution. Suppose that the class sizes $(n_{01}, n_{02}, ..., n_{0K})$ in period 0 change to $(n_{11}, n_{12}, ..., n_{1K})$ in period 1, Then the new way of relabelling the income

[6]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.

classes is given by

$$N_1(x_k) := \sum_{h=1}^{k} n_{1h},\ k = 1, ..., K, \tag{2}$$

where $n_{1k} \in \mathbb{R}_+$ denotes the number of persons in period 1 who are in class k, $k = 1, 2, ..., K$ and $\sum_{k=1}^{K} n_{1k} = n$.

This gives a method of dealing with 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. Use u_i and v_i to denote individual i's status in the 0-distribution and the 1-distribution respectively. In this approach the set of *status distributions* is given by

$$U := \left\{\mathbf{u} |\ \mathbf{u} \in \mathbb{R}_+^n, u_1 \leq u_2 \leq ... \leq u_n\right\} \tag{3}$$

and person i's *history* is given 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:

  $$z_i = \left(x_{k_0(i)}, x_{k_1(i)}\right). \tag{4}$$

  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:

  $$z_i = \left(\varphi\left(x_{k_0(i)}\right), \varphi\left(x_{k_1(i)}\right)\right), \tag{5}$$

  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 any 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 leads to sharp conclusions: this 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 we might take the number of persons with incomes no greater than that of i :

$$z_i = \left(N_0\left(x_{k_0(i)}\right), N_0\left(x_{k_1(i)}\right)\right). \tag{6}$$

  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 the right-hand side of (6).

- *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:

$$z_i = \left(N_0\left(x_{k_0(i)}\right), N_1\left(x_{k_1(i)}\right)\right). \tag{7}$$

  In (7) we are taking into account the change in "valuation" of each status class that arises from the changing income distribution.[7]

Status is, in principle, distinct from "income": we could, if we wish, define status as equal to income, but that would be an explicit normative assumption. It is also clear that different status concepts could produce 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 (4) to (7) it makes sense say that there is zero mobility if

$$v_i = u_i, i = 1, ..., n. \tag{8}$$

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 (7) then there is zero mobility; if it is defined as (6) it is clear that mobility is positive. Now suppose that

$$x_k = \lambda x_{k-1}, k = 2, ..., K, \lambda > 1. \tag{9}$$

Then, in the cases (4) and (5), it *may* make sense to consider

$$v_i = \lambda u_i, i = 1, ..., n, \lambda > 0 \tag{10}$$

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

[7] 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	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

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. We will also see that, in a practical implementation (section 7.2 below), for a given data set very different conclusions can be drawn about mobility trends, just by changing the status concept while keeping the structure of the mobility measure the same.

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

3 Mobility measures: theory

3.1 Aggregation of histories

Let us address the third essential ingredient of the mobility problem mentioned in the introduction: the aggregation of the changes in status encapsulated in individual histories.[8] 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 it is dependent on the cardinalisation of income. Assume that a measure of individual status has been agreed, determined by the information available from the income distribution at any moment; also assume that there is an observation of

[8] 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; second, the current paper deals with any arbitrary representation of status (including ordinal status) rather than being specific to income; this requires a treatment of the case where the mobility measure is defined on categorical data.

the status of each person i in periods 0 and 1; we need a coherent method of quantifying the implicit status changes as "mobility."

Individual movements or changes in status are completely characterised by the histories z_i, $i = 1, 2, ..., n$ as defined in section 2. The set of possible status movements Z depends on the nature of status: if status is income, consumption or wealth then Z would be a connected subset of $\mathbb{R} \times \mathbb{R}$; but if status is determined by the person's position in the distribution (in terms of absolute numbers or proportions) then Z would equal $\mathbb{Q}_+ \times \mathbb{Q}_+$ where $\mathbb{Q}_+$ is the set of non-negative rationals; our analysis will take care of both these cases. Furthermore define

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

we may refer to any $\mathbf{z} \in Z^n$ as a *movement profile*. Clearly overall mobility for a given profile could be described in terms of the status changes of each individual i: the information encapsulated in i's history z_i. In the introduction we argued that mobility is essentially a social rather than individual concept. So, rather than trying to give meaning to an "individual mobility function", applying this to each person's history and then applying an aggregation function to all n observations of "individual mobility", a more direct approach will be used. We just need to specify a set of principles for comparing the elements of Z^n: stating these principles as formal axioms we may characterise an aggregate mobility ordering and the ordering may yield a family of mobility measures. To be useful the principles underlying the ordering should respect the mobility implicit in an individual history, but the exact form in which individual status changes are to be quantified will emerge from the aggregate mobility ordering.

A particular advantage of our approach is that the formal axiomatisation of the mobility ordering (presented in sections 3.2 and 3.3) 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 Mobility ordering: basic structure

In this section and section 3.3 we characterise an ordering that enables us to compare movement profiles. Use $\succeq$ to denote a weak ordering on Z^n; denote by $\succ$ the strict relation associated with $\succeq$ and denote by $\sim$ the equivalence relation associated with $\succeq$. We first consider the interpretation of five axioms that underpin the approach; we then state a basic result that follows from them.[9]

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 and $u_i' = u_i$ then, if $v_i > v_i' \geq u_i$, or if $v_i < v_i' \leq u_i$, $\mathbf{z} \succ \mathbf{z}'$*

Suppose we know that, with the sole exception of person i, each person's history in profile $\mathbf{z}$ is the same as it is in profile $\mathbf{z}'$. Person i's history can be described as follows:

[9] The proof is in the Appendix.

i starts with the same period-0 status in $\mathbf{z}$ and in $\mathbf{z}'$ and then moves up to a higher status in period 1; but i's period-1 status in profile $\mathbf{z}$ is even higher than it is in $\mathbf{z}'$. Then Axiom 2 implies that mobility would be higher in $\mathbf{z}$ than in $\mathbf{z}'$ (A corresponding story can be told for downward movement). Note that, in particular Axiom 2 will ensure that the ordering $\succeq$ satisfies a "minimal mobility" requirement. If one starts with a profile representing complete stasis ($\forall i : \ u_i = v_i$) then a change in any person's status must increase mobility.[10]

Axiom 3 [Independence] *Let $\mathbf{z}(\zeta, i)$ denote the profile formed by replacing the ith component of $\mathbf{z}$ by the history $\zeta \in Z$. For $\mathbf{z}, \mathbf{z}' \in Z^n$ suppose 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 $\zeta \in Z$*

Suppose that the profiles $\mathbf{z}$ and $\mathbf{z}'$ are equivalent in terms of overall mobility and that there is some person i with the same history z_i in both $\mathbf{z}$ and $\mathbf{z}'$. Then the same change Δz_i in i's history in both $\mathbf{z}$ and $\mathbf{z}'$ leaves the two modified profiles as equivalent in terms of overall mobility.

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

Consider a profile $\mathbf{z}$ in which person i is immobile: change i's status by the same amount in both the 0-distribution and the 1-distribution (so that i is still immobile after the change in status). Then the new profile $\mathbf{z}'$ should exhibit the same mobility as the original $\mathbf{z}$.

Theorem 1 *Given Axioms 1 to 4 then $\forall \mathbf{z} \in Z^n$ the mobility ordering $\succeq$ is representable by an increasing monotonic transform of*

$$\sum_{i=1}^{n} \phi_i(z_i), \tag{11}$$

where the ϕ_i are continuous functions $Z \to \mathbb{R}$, defined up to an affine transformation, each of which is increasing (decreasing) in v_i if $v_i > (<)\ u_i$ and that has the property $\phi_i(u, u) = b_i u$, where $b_i \in \mathbb{R}$.

3.3 Mobility ordering: scale

Theorem 1, the first part of the characterisation of the mobility ordering, shows that it can be represented as the sum of the evaluation of individual histories where ϕ_i is the history-evaluation function for i; but theorem 1 leaves the specification of ϕ_i open. The second part of our characterisation of the mobility ordering involves the comparison of profiles at different levels of status. To do this let us use the notation $\mathbf{z} \times (\lambda_0, \lambda_1)$ for

[10] It was shown by example in the introduction that the regression coefficient violates this minimal-mobility property.

the movement profile that is derived from $\mathbf{z}$ if all the 0-components (u_i) are multiplied by λ_0 and all the 1-components (v_i) are multiplied by λ_1. Then we can introduce one more axiom that encapsulates the idea that the ordering of profiles remains unchanged by some scale change to status in both periods.

We then have a theorem (proof in the Appendix) showing that the evaluation function ϕ_i in (11) takes a particularly convenient form.

Axiom 5 [Status scale irrelevance] *For any* $\mathbf{z}, \mathbf{z}' \in Z^n$ *such that* $\mathbf{z} \sim \mathbf{z}'$, $\mathbf{z} \times (\lambda_0, \lambda_1) \sim \mathbf{z}' \times (\lambda_0, \lambda_1)$, *for all* $\lambda_0, \lambda_1 > 0$.

Axiom 5 is completely natural in the case of distribution-dependent measures of status such as (6) or (7) since it enables one to characterise mobility rankings 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 but it is a judgment that is often seen as perfectly unremarkable: suppose we consider that China and the US show the same mobility over a given time interval; if we rescale the incomes in both countries at the beginning of the interval then China and the US still show the same mobility; the same thing is true of the incomes at the end of the interval. However, scale irrelevance of the mobility ordering clearly does not mean that the resulting mobility index is independent of scale changes: more is required for this, as discussed in section 3.4.

Theorem 2 *Given Axioms 1 to 5* $\succeq$ *is representable by (11), where* ϕ_i *is given by*

$$\phi_i(u, v) = c_i \left[u^{\alpha} v^{1-\alpha} - \alpha u - [1 - \alpha] v \right]. \tag{12}$$

where $\alpha, c_i \in \mathbb{R}$.

3.4 Aggregate mobility index

Theorem 2 means that the mobility ordering $\succeq$ implied by the five axioms in sections 3.2 and 3.3 can be represented by the expression $\sum_{i=1}^{n} \phi_i(u, v)$, with the ϕ_i given by (12). Since $\succeq$ is an ordering it is also representable by some continuous increasing transformation of this expression. We now examine what normalisation is appropriate in order to construct an aggregate inequality index. There are three steps.

First, it is arguable that mobility should be blind as to individual identity. If the definition of status incorporates all relevant information about an individual, the labelling $i = 1, ..., n$ is irrelevant and anonymity is an innocuous assumption. It simply means that mobility depends only on individual status histories; switching the personal labels from one history to another within a movement profile has no effect on mobility rankings: if a profile $\mathbf{z}'$ can be obtained as a permutation of the components of another

profile $\mathbf{z}$, then they should be treated as equally mobile. If so, then all the c_i should be equal and mobility can be represented as a transform of

$$c\sum_{i=1}^{n}\left[u_i^{\alpha}v_i^{1-\alpha}-\alpha u_i-[1-\alpha]\,v_i\right] \tag{13}$$

Second, consider the effect of population size. A simple replication of profiles $\mathbf{z}$ does not change the essential facts of mobility. Clearly α cannot depend on the size of the population, but the constant c may depend on n. If any profile is replicated r times and the index remains unchanged under replication we have

$$c\,(n)\sum_{i=1}^{n}\left[u_i^{\alpha}v_i^{1-\alpha}-\alpha u_i-[1-\alpha]\,v_i\right]=c\,(nr)\,r\sum_{i=1}^{n}\left[u_i^{\alpha}v_i^{1-\alpha}-\alpha u_i-[1-\alpha]\,v_i\right].$$

So, to ensure that the representation of $\succeq$ is in a form that is constant under replication, we need to have c proportional to $1/n$. Choosing for convenience the constant of proportionality as $\frac{1}{\alpha[\alpha-1]}$ we may write the index as some transform of this "basic-form" mobility index:

$$\frac{1}{\alpha\,[\alpha-1]}\left[\frac{1}{n}\sum_{i=1}^{n}u_i^{\alpha}v_i^{1-\alpha}-\alpha\mu_u-[1-\alpha]\,\mu_v\right] \tag{14}$$

where

$$\mu_u \;:=\; \frac{1}{n}\sum_{i=1}^{n}u_i, \tag{15}$$

$$\mu_v \;:=\; \frac{1}{n}\sum_{i=1}^{n}v_i. \tag{16}$$

Notice that (14) is strictly increasing (decreasing) in u_i if $u_i > v_i$ ($u_i < v_i$) and (14) is strictly decreasing (increasing) in v_i if $u_i < v_i$ ($u_i > v_i$); this behaviour is natural in view of monotonicity (Axiom 2). Furthermore it is clear that the basic form (14) has the property that mobility is zero if $v_i = u_i$ for all i.

Third, we may wish to normalise so that the index remains unchanged under a scale change $\lambda_0 > 0$ in the 0-distribution and under a scale change $\lambda_1 > 0$ in the 1-distribution. Whether one takes this third step or not depends on the way in which the mobility concept is characterised, as we highlighted in the discussion of Table 1: simple income growth may or may not be counted as mobility. If we accept that simple income growth counts as mobility then expression (14) gives us a complete *class* of measures that are normalised to ensure anonymity and independence of population size. But if we do not, then scale normalisation is appropriate: we strengthen the scale-irrelevance property (Axiom 5) already imposed on mobility orderings to *scale-independence* of the resulting mobility index.

A transformation of (14) that involves the means μ_u, μ_v and that preserves the "zero" property must take the form

$$\psi\left(\frac{1}{n}\sum_{i=1}^{n}u_i^{\alpha}v_i^{1-\alpha}-\theta\,(\mu_u,\mu_v)\,,\mu_u,\mu_v\right) \tag{17}$$

where ψ is monotonic in its first argument and has the property that $\psi(0, \mu_u, \mu_v) = 0$, and where θ is a function that is homogeneous of degree 1 with the property that $\theta(\mu, \mu) = \mu$. Setting $\lambda_0 = 1/\mu_u$ and $\lambda_1 = 1/\mu_v$ it is clear that (17) becomes

$$\psi\left(\frac{1}{n}\sum_{i=1}^{n}\left[\frac{u_i}{\mu_u}\right]^{\alpha}\left[\frac{v_i}{\mu_v}\right]^{1-\alpha} - \theta(1,1), 1, 1\right) = \overline{\psi}\left(\frac{1}{n}\sum_{i=1}^{n}\left[\left[\frac{u_i}{\mu_u}\right]^{\alpha}\left[\frac{v_i}{\mu_v}\right]^{1-\alpha} - 1\right]\right), \quad (18)$$

where $\overline{\psi}(t) := \psi(t, 1, 1)$. A suitable cardinalisation[11] of ψ in (18) gives the aggregate mobility measure

$$M_\alpha := \frac{1}{\alpha[\alpha-1]n}\sum_{i=1}^{n}\left[\left[\frac{u_i}{\mu_u}\right]^{\alpha}\left[\frac{v_i}{\mu_v}\right]^{1-\alpha} - 1\right], \ \alpha \in \mathbb{R}, \alpha \neq 0, 1, \quad (19)$$

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

$$M_0 = -\frac{1}{n}\sum_{i=1}^{n}\frac{v_i}{\mu_v}\log\left(\frac{u_i}{\mu_u}\Big/\frac{v_i}{\mu_v}\right), \quad (20)$$

$$M_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{u_i}{\mu_u}\log\left(\frac{u_i}{\mu_u}\Big/\frac{v_i}{\mu_v}\right). \quad (21)$$

Expressions (19)-(21) constitute the class of aggregate mobility measures that are independent of population size and independent of the scale of status. An individual member of the class is characterised by the choice of α: a high positive α produces an index that is particularly sensitive to downward movements and a negative α yields an index that is sensitive to upward movements– see Section 4. 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 if (9) holds; but this is purely a matter of normalisation – as we have just seen, an alternative normalisation could be 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 $M_\alpha > 0$.

As we have just noted, this third step should not be treated as simply a minor technicality. Normalisation by mean status in order to ensure scale independence introduces a complication: the type of status variation considered in the statement of Axiom 2 will change the mean of u and/or the mean of v. This would mean, for example, that such a status variation will affect not only the numerator of each fraction in (19)-(21), but also the denominator. Such normalised indices still have minimal-mobility property: starting from a situation of zero mobility, changing any one person's status will result in positive mobility. But to cover cases where mobility is not close to zero, for the normalised index we need to reconsider the monotonicity principle. This can be done by replacing Axiom 2 with the following Axiom 6:

Axiom 6 [Monotonicity-2] *If $\mathbf{z}, \mathbf{z}' \in Z^n$ differ only in their ith and jth components and $u_i' = u_i$, $u_j' = u_j$, $v_i' - v_i = v_j - v_j'$ then, if $v_i > v_i' \geq u_i$ and if $v_j < v_j' \leq u_j$, $\mathbf{z} \succ \mathbf{z}'$.*

[11] This cardinalisation ensures that M_α is well-defined and non-negative for all values of α and that, for any profile $\mathbf{z}$, M_α is continuous in α.

Note that the modified version of monotonicity in Axiom 6 will again ensure that minimal-mobility property is satisfied. Also Axiom 6 is clearly satisfied by the normalised index. Consider, for example, the case where an infinitesimal change in period-1 status is given by $dv_i = v_i' - v_i < 0$, $dv_j = -dv_i > 0$ then the result of this infinitesimal change on mobility is given by

$$n\mu_v \left[\frac{\partial M_\alpha}{\partial v_i} - \frac{\partial M_\alpha}{\partial v_j} \right] = \frac{1}{\alpha} \left[\frac{\mu_v}{\mu_u} \right]^\alpha \left[\left[\frac{u_j}{v_j} \right]^\alpha - \left[\frac{u_i}{v_i} \right]^\alpha \right] \tag{22}$$

If all the conditions on the components of the profiles in Axiom 6 are satisfied then the change in mobility given in (22) is clearly negative, as required.[12]

4 Discussion

The nature of the superclass referred to in the introduction is now clear: expressions (19)-(21) 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.

4.1 Ordinal status

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 introducing an ordinal concept of status. For example a large section of the mobility adopts a "mobility table" or "transition matrix" approach to mobility.[13] This focuses attention on the size n_k of each class k and the number of the n_k that move to other classes.

To see how our approach can be reconciled with the standard transition-matrix approach consider the following. Let there be income classes $X_k := [x_k, x_{k+1})$, $k = 1, ..., K$. The transition matrix is $\Pi := \|\pi_{k\ell}\|$, where each element $\pi_{k\ell}$ is the conditional probability that an individual moves to class ℓ of the 1-distribution given that he was initially in class k of the 0-distribution:

$$\pi_{k\ell} = \frac{\Pr(y_0 \in X_k, y_1 \in X_\ell)}{\Pr(y_0 \in X_k)}$$

[12]There is a parallel with welfare analysis and inequality measurement. In that context let the v-distribution denote a distribution individual utilities, where each person's utility is independent of anyone else's utility; the u-distribution is irrelevant. Then the counterpart of Axiom 2 for a social-welfare function is the Pareto principle; if utility is cardinal (so that the mean is well-defined) then the counterpart of Axiom 6 is the principle of progressive transfers.

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

So $\sum_{\ell=1}^{K} \pi_{k\ell} = 1$, $\Pr(y_0 \in X_k) = \pi_k$ and the initial income distribution (in grouped form) is given by $(X_k, \pi_k), k = 1, ..., K$. The estimate of $\pi_{k\ell}$ is given by $m_{k\ell}/n_{0k}$, where $m_{k\ell}$ is the number of persons who are in class k in period 0 and in class ℓ in period 1, and n_{0k} is (as in section 2) the total number of people in class k in period 0; so $\sum_{j=1}^{K} m_{kj} = n_{0k}$. The estimate of π_k is given by n_{0k}/n. Commonly used summary statistics to capture the mobility implied by Π are:

$$S_0(\Pi) = \frac{K - \sum_{k=1}^{K} m_{kk}/n_{0k}}{K-1} \tag{23}$$

$$S_1(\Pi) = \frac{K - K\sum_{k=1}^{K} m_{kk}/n}{K-1} \tag{24}$$

$$S_2(\Pi) = \frac{\frac{1}{n}\sum_{k=1}^{K}\sum_{\ell=1}^{K} m_{k\ell}|k-\ell|}{K-1} \tag{25}$$

– see Formby et al. (2004). We can also use our mobility index. Let there be income classes X_k as defined above and the status of any individual in class X_k in period 0 being denoted by u_k, and by v_k in period 1, then (19) becomes,

$$M_\alpha = \frac{1}{\alpha[\alpha-1]}\frac{1}{n}\sum_{k=1}^{K}\sum_{\ell=1}^{K} m_{k\ell}\left[\frac{u_k}{\mu_u}\right]^{\alpha}\left[\frac{v_\ell}{\mu_v}\right]^{1-\alpha} - \frac{1}{\alpha[\alpha-1]}, \tag{26}$$

The two mobility indices $S_2(\Pi)$ and M_α are closely related: the weights $|k-\ell|$ in (25) are replaced by $[u_k/\mu_u]^{\alpha}[v_\ell/\mu_v]^{1-\alpha}$ in (26) and the normalisation is different.

However, the transition-matrix approach could be sensitive to the merging or splitting of classes or the adjustment of class boundaries. If there is a spike in the distribution at $x_k - \epsilon$ and the interval boundaries are changed so that x_k becomes $x_k - \delta$ where $\delta \geq \epsilon$ then we can get a big change in estimated mobility.

4.2 Decomposability

Our axioms induce an additive structure for the mobility index, which might be thought to be restrictive. 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 2012) As explained above, i's status may 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 only through the definition of status. 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 (19)-(21) are clearly decomposable by arbitrary population subgroups.

Let there be K groups and let the proportion of population falling in group k be p_k, the class of scale-independent mobility measures (19) can be expressed as:

$$M_\alpha = \sum_{k=1}^{K} p_k\left[\frac{\mu_{u,k}}{\mu_u}\right]^{\alpha}\left[\frac{\mu_{v,k}}{\mu_v}\right]^{1-\alpha} M_{\alpha,k} + \frac{1}{\alpha^2-\alpha}\left(\sum_{k=1}^{K} p_k\left[\frac{\mu_{u,k}}{\mu_u}\right]^{\alpha}\left[\frac{\mu_{v,k}}{\mu_v}\right]^{1-\alpha} - 1\right) \tag{27}$$

for $\alpha \neq 0, 1$, where $\mu_{u,k}$ ($\mu_{v,k}$) is the mean status in period-0 (period-1) in group k, and μ_u, μ_v are the corresponding population means defined in (15), (16) (so that $\mu_u = K^{-1}\sum_{k=1}^{K} p_k \mu_{u,k}$, $\mu_v = K^{-1}\sum_{k=1}^{K} p_k \mu_{v,k}$). In particular, notice that in the case where $u = x$ and $v = \mu_x$, we obtain the standard formula of decomposability for the class of GE inequality indices (Cowell 2011). We have the following limiting forms for the cases $\alpha = 0$ and $\alpha = 1$, respectively

$$M_0 = \sum_{k=1}^{K} p_k \left[\frac{\mu_{v,k}}{\mu_v}\right] M_{0,k} - \sum_{k=1}^{K} p_k \left[\frac{\mu_{v,k}}{\mu_v}\right] \log\left(\frac{\mu_{u,k}}{\mu_u} \Big/ \frac{\mu_{v,k}}{\mu_v}\right) \tag{28}$$

$$M_1 = \sum_{k=1}^{K} p_k \left[\frac{\mu_{u,k}}{\mu_u}\right] M_{1,k} + \sum_{k=1}^{K} p_k \left[\frac{\mu_{u,k}}{\mu_u}\right] \log\left(\frac{\mu_{u,k}}{\mu_u} \Big/ \frac{\mu_{v,k}}{\mu_v}\right) \tag{29}$$

This means, for example, that we may partition U in (3) unambiguously into an upward status group U (for $u_i \leq v_i$) and a downward status group D (for $u_i > v_i$) and, using an obvious notation, express overall mobility as

$$M_\alpha = w^{\mathsf{U}} M_\alpha^{\mathsf{U}} + w^{\mathsf{D}} M_\alpha^{\mathsf{D}} + M_\alpha^{\mathsf{btw}}, \tag{30}$$

where the weights w^{U}, w^{D} and the between-group mobility component M_α^{btw} are functions of the status-means for each of the two groups and overall; comparing M_α^{U} and M_α^{D} enables one to say precisely where mobility has taken place.

4.3 Choice of α

Let us consider a sample where every individual's upward mobility is matched by a symmetric downward mobility of someone else ($\forall i, \exists j$ such that $u_j = v_i, v_j = u_i$).

In this particular case of (perfect) symmetry between downward and upward status movements, we have $\mu_u = \mu_v$. Then, it is clear from (19) that 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).[14]

To go further, let us consider the upward status group U (for $u_i \leq v_i$) and the downward status group D (for $u_i > v_i$), as defined in (30). From (19) we have[15]

$$M_\alpha^{\mathsf{U}} = M_{1-\alpha}^{\mathsf{D}} \tag{31}$$

It suggests that mobility measurement of upward movements and of symmetric downward movements would be identical with $\alpha = 0.5$ ($M_{0.5}^{\mathsf{U}} = M_{0.5}^{\mathsf{D}}$). Furthermore, mobility measurement of upward movements with $\alpha = 1$ would be identical to mobility measurement of symmetric downward movements with $\alpha = 0$ ($M_1^{\mathsf{U}} = M_0^{\mathsf{D}}$).

[14] With symmetric downward/upward mobility, $M_\alpha = \frac{1}{\alpha[\alpha-1]n}\sum_{i=1}^{n}\left(\frac{v_i}{\mu_v}\left[\frac{u_i}{v_i}\right]^\alpha - 1\right)$.

[15] More generally, if we generate a "reverse profile" $\mathbf{z}'(\mathbf{z}) := \{z_i' = (v_i, u_i) \mid z_i = (u_i, v_i), i = 1, ..., n\}$ by reversing each person's history – swapping the us and vs in (19) – we have $M_\alpha(\mathbf{z}'(\mathbf{z})) = M_{1-\alpha}(\mathbf{z})$.

In the mobility index M_α, the weights given to upward mobility and to downward mobility can be studied through its decomposability property. With symmetric upward/downward status movements, from (27) and (30), we can see that[16]

1. for $\alpha = 0.5$, we have $w^{\mathsf{U}} = w^{\mathsf{D}}$,
2. for $\alpha < 0.5$, we have $w^{\mathsf{U}} > w^{\mathsf{D}}$,
3. for $\alpha > 0.5$, we have $w^{\mathsf{U}} < w^{\mathsf{D}}$.

In other words, $\alpha = 0.5$ puts the same weight on both upward and downward mobility components in (27), while $\alpha < 0.5$ ($\alpha > 0.5$) puts more weights on upward (downward) mobility component. The sensitivity parameter α enables us to capture *directional sensitivity* in the mobility context:[17] high positive values result in a mobility index that is more sensitive to downward movements from period 0 to period 1; negative α is more sensitive to upward movements. Picking a value for this parameter is a normative choice.

4.4 Homotheticity

The axioms also induce a homothetic structure, which once again might be thought to be rather restrictive for some interpretations of u and v. Furthermore, the normalisations introduced in section 3.4 impose 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: 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 (19) will be replaced by

$$\frac{\theta(c)}{n}\sum_{i=1}^{n}\left[\left[\frac{u_i + c}{\mu_u + c}\right]^{\alpha(c)}\left[\frac{v_i + c}{\mu_v + c}\right]^{1-\alpha(c)} - 1\right], \alpha(c) \in \mathbb{R}, \alpha(c) \neq 0, 1 \tag{32}$$

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

$$\theta(c) := \frac{1 + c^2}{\alpha(c)^2 - \alpha(c)}; \tag{33}$$

for $\alpha(c) = 0$ and $\alpha(c) = 1$ there are obvious special cases of (32) corresponding to (20) and (21). 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

$$\alpha(c) := \gamma + \beta c \tag{34}$$

[16]From (27) and (30), we have $w^{\mathsf{U}} = p_1(\mu_{u,1}/\mu_u)^\alpha(\mu_{v,1}/\mu_v)^{1-\alpha}$ and $w^{\mathsf{D}} = p_2(\mu_{u,2}/\mu_u)^\alpha(\mu_{v,2}/\mu_v)^{1-\alpha}$. With symmetric downward/upward mobility, we also have $p_1 = p_2$, $\mu_{u,1} = \mu_{v,2} < \mu_{v,1} = \mu_{u,2}$ and $\mu_u = \mu_v$. Then, $w^{\mathsf{U}}/w^{\mathsf{D}} = (\mu_{u,2}/\mu_{v,2})^{1-2\alpha}$, which is greater (less) than one if $1 - 2\alpha > (<) 0$.

[17]See also : Bhattacharya and Mazumder (2011), Corak et al. (2014), Demuynck and Van de gaer (2012) and Schluter and Van de gaer (2011).

and analysing the behaviour as $c \to \infty$ we may say more. Consider the main expression inside the summation in (32); taking logs we may write this as

$$\log\left(\frac{1+\frac{v}{c}}{1+\frac{\mu_v}{c}}\right) + \alpha(c)\left[\log\left(1+\frac{u}{c}\right) + \log\left(1+\frac{\mu_v}{c}\right) - \log\left(1+\frac{v}{c}\right) - \log\left(1+\frac{\mu_u}{c}\right)\right]. \tag{35}$$

Using the standard expansion

$$\log(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} - ... \tag{36}$$

and (34) we find that (35) becomes

$$\log\left(\frac{1+\frac{v}{c}}{1+\frac{\mu_v}{c}}\right) + \left[\beta + \frac{\gamma}{c}\right]\left[u + \mu_v - v - \mu_u - \frac{u^2}{2c} - \frac{\mu_v^2}{2c} + \frac{v^2}{2c} + \frac{\mu_u^2}{2c}...\right]. \tag{37}$$

For finite $\gamma, \beta, u, v, \mu_u, \mu_v$ we find that (37) becomes

$$\beta\left[u - \mu_u - v + \mu_v\right] \tag{38}$$

and

$$\lim_{c\to\infty}\theta(c) = \lim_{c\to\infty}\frac{1+\frac{1}{c^2}}{\left[\beta+\frac{\gamma}{c}\right]^2 - \frac{1}{c}\left[\beta+\frac{\gamma}{c}\right]} = \frac{1}{\beta^2}. \tag{39}$$

From (38) and (39) we can see that in the limit (32) becomes[18]

$$M'_\beta := \frac{1}{n\beta^2}\sum_{i=1}^{n}\left[\mathrm{e}^{\beta[u_i-\mu_u-v_i+\mu_v]} - 1\right], \tag{40}$$

for any $\beta \neq 0$. Let $q_i := u_i - \mu_u - v_i + \mu_v$ so that (40) can be written

$$\frac{1}{n\beta^2}\sum_{i=1}^{n}\left[\mathrm{e}^{\beta q_i} - 1\right] = \frac{1}{n\beta^2}\sum_{i=1}^{n}\left[1 + \beta q_i + \frac{1}{2!}\beta^2 q_i^2 + \frac{1}{3!}\beta^3 q_i^3 + \frac{1}{4!}\beta^4 q_i^4 + ... - 1\right], \tag{41}$$

using a standard expansion. Noting that $\frac{1}{n}\sum_{i=1}^{n} q_i = 0$, the right-hand side of (41) becomes

$$\frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{2!}q_i^2 + \frac{1}{3!}\beta q_i^3 + \frac{1}{4!}\beta^2 q_i^4 + ...\right]. \tag{42}$$

As $\beta \to 0$ it is clear that (42) tends to $\frac{1}{2n}\sum_{i=1}^{n} q_i^2$. So the limiting form of (40) for $\beta = 0$ is

$$M'_0 := \frac{1}{2}\mathrm{var}\left(u_i - v_i\right). \tag{43}$$

So expressions (40) and (43) 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 5 only requires a weak structural assumption, where all the contours are homothetic to the point $(c, c, ..., c)$, a property that is satisfied by almost all mobility indices.

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

5 Statistical Inference

If we carry out a simple computation of the values of a mobility measure, computed from two different samples, we are almost always going to find greater mobility in one sample, even if the two samples come from the same population. Clearly simple computation alone is not enough in order to draw useful conclusions from the raw data: statistical methods are required to test the hypothesis that the two values are not statistically different. 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 (4). The income values at period 0 and 1 are used to evaluate individual status,

$$u_i = x_{0i} \qquad \text{and} \qquad v_i = x_{1i}, \tag{44}$$

it corresponds to a model of movement of incomes in society as a whole. 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 (19)-(21) separately.

Case M_α ($\alpha \neq 0, 1$). We can rewrite the index (19) 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:

$$M_\alpha = \frac{1}{\alpha(\alpha-1)}\left[\frac{\mu_{u^\alpha v^{1-\alpha}}}{\mu_u^\alpha \mu_v^{1-\alpha}} - 1\right]. \tag{45}$$

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 :

$$\widehat{\mathrm{Var}}(M_\alpha) = \frac{1}{n} D\hat{\Sigma}D^\top \qquad \text{with} \qquad D = \left[\frac{\partial M_\alpha}{\partial \mu_u}\ ;\ \frac{\partial M_\alpha}{\partial \mu_v}\ ;\ \frac{\partial M_\alpha}{\partial \mu_{u^\alpha v^{1-\alpha}}}\right] \tag{46}$$

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:[19]

$$\hat{\Sigma} = \begin{bmatrix} \mu_{u^2} - (\mu_u)^2 & \mu_{uv} - \mu_u\mu_v & \mu_{u^{1+\alpha}v^{1-\alpha}} - \mu_u\mu_{u^\alpha v^{1-\alpha}} \\ \mu_{uv} - \mu_u\mu_v & \mu_{v^2} - (\mu_v)^2 & \mu_{u^\alpha v^{2-\alpha}} - \mu_v\mu_{u^\alpha v^{1-\alpha}} \\ \mu_{u^{1+\alpha}v^{1-\alpha}} - \mu_u\mu_{u^\alpha v^{1-\alpha}} & \mu_{u^\alpha v^{2-\alpha}} - \mu_v\mu_{u^\alpha v^{1-\alpha}} & \mu_{u^{2\alpha}v^{2-2\alpha}} - (\mu_{u^\alpha v^{1-\alpha}})^2 \end{bmatrix} \tag{47}$$

We can use this variance estimator of M_α to compute a test statistic or a confidence interval.[20]

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:

$$M_0 = \frac{\mu_{v\log v} - \mu_{v\log u}}{\mu_v} + \log\left(\frac{\mu_u}{\mu_v}\right) \tag{48}$$

The variance estimator of this index is defined as follows:

$$\widehat{\mathrm{Var}}(M_0) = \frac{1}{n} D_0 \hat{\Sigma}_0 D_0^\top \qquad \text{with} \qquad D_0 = \left[\frac{\partial M_0}{\partial \mu_u}\ ;\ \frac{\partial M_0}{\partial \mu_v}\ ;\ \frac{\partial M_0}{\partial \mu_{v\log v}}\ ;\ \frac{\partial M_0}{\partial \mu_{v\log u}}\right] \tag{49}$$

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:

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

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

$$M_1 = \frac{\mu_{u\log u} - \mu_{u\log v}}{\mu_u} + \log\left(\frac{\mu_v}{\mu_u}\right) \tag{50}$$

The variance estimator of this index is defined as follows:

$$\widehat{\mathrm{Var}}(M_1) = \frac{1}{n} D_1 \hat{\Sigma}_1 D_1^\top \qquad \text{with} \qquad D_1 = \left[\frac{\partial M_1}{\partial \mu_u}\ ;\ \frac{\partial M_1}{\partial \mu_v}\ ;\ \frac{\partial M_1}{\partial \mu_{u\log u}}\ ;\ \frac{\partial M_1}{\partial \mu_{u\log v}}\right] \tag{51}$$

We have

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

[19] 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 $\mathrm{Cov}(U, V) = E(UV) - E(U)E(V)$.

[20] 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.

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

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

5.2 Rank mobility

Let us consider the *distribution-dependent, dynamic* status, as defined in (7), 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,

$$u_i = \hat{F}_0(x_{0i}) \qquad \text{and} \qquad v_i = \hat{F}_1(x_{1i}) \tag{52}$$

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

$$\hat{F}_k(x) = \frac{1}{n}\sum_{j=1}^{n} I(x_{kj} \leq x) \tag{53}$$

where $k = 1, 2$ and $I(.)$ is an indicator function, equal to 1 if its argument is true and to 0 otherwise. So 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,

$$T_n = \frac{1}{n}\sum_{i=1}^{n} c_{in}\phi_1(u_i)\phi_2(v_i), \tag{54}$$

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

$$\phi_1(t) \leq \frac{K_1}{[t(1-t)]^{a_1}} \quad \text{and} \quad \phi_2(t) \leq \frac{K_1}{[t(1-t)]^{a_2}} \quad \text{with} \quad a_1 + a_2 < \frac{1}{2} \tag{55}$$

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 (55) is respected. Let us first notice that,

$$\mu_u = \mu_v = \frac{1}{n}\sum_{i=1}^{n}\frac{i}{n} = \frac{n+1}{2n}. \tag{56}$$

Case M_α ($\alpha \neq 0, 1$). From (45) and (56) , we obtain M_α as a function of one moment:

$$M_\alpha = \frac{1}{\alpha(\alpha-1)} \left[\frac{2n}{n+1} \mu_{u^\alpha v^{1-\alpha}} - 1 \right]. \tag{57}$$

From (54) and (57) it is clear that

$$M_\alpha = \frac{1}{\alpha(\alpha-1)} \left[T_n - 1 \right], \tag{58}$$

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 (55) 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 (48), (56) and (54), we have

$$M_0 = \frac{2n}{n+1}(k - \mu_{v \log u}) = l - T_n, \tag{59}$$

where k and l are real constants[21] and $c_{in} = \frac{2n}{n+1}$, $\phi_1(u_i) = \log u_i$ and $\phi_2(v_i) = v_i$. The condition (55) 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 (50), (56) and (54), we have

$$M_1 = \frac{2n}{n+1}(k - \mu_{u \log v}) = l - T_n, \tag{60}$$

where $c_{in} = \frac{2n}{n+1}$, $\phi_1(u_i) = u_i$ and $\phi_2(v_i) = \log v_i$. The condition (55) 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 (55), 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

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

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

$$CI_{asym} = [M_\alpha - c_{0.975}\widehat{\mathrm{Var}}(M_\alpha)^{1/2}\,;\; M_\alpha + c_{0.975}\widehat{\mathrm{Var}}(M_\alpha)^{1/2}] \tag{61}$$

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 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

$$CI_{perc} = [c_{0.025}^b\,;\; c_{0.975}^b] \tag{62}$$

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

$$CI_{stud} = [M_\alpha - c_{0.975}^*\widehat{\mathrm{Var}}(M_\alpha)^{1/2}\,;\; M_\alpha - c_{0.025}^*\widehat{\mathrm{Var}}(M_\alpha)^{1/2}] \tag{63}$$

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. So 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

$$(x_0, x_1) \sim LN(\mu, \Sigma) \quad \text{with} \quad \mu = (0,0) \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{64}$$

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 ($\rho = 0, 0.2, 0.4, 0.6, 0.8, 0.9, 0.99$).[22]

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 (44). Here the income values are used to evaluate individual status.

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.

[22] Finite sample performance of estimators of inequality measures based on Lognormal distributions, with a variance equals to one or greater, are as problematic as with Singh-Maddala distributions – see Cowell and Flachaire (2015), Table 6.6, p.414.

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

These results suggest that asymptotic confidence intervals perform well in very large samples, 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 (52). 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 ($\rho = 0.99$).
- the coverage error rate is close to 0.05 for $\rho = 0.8, 0.9$ and $\alpha = 0, 0.5, 1$.

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

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.

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.

7 Mobility measures: comparison and assessment

Why use the approach discussed in sections 3-5? In this section, we address this question in two ways. First, in subsection 7.1, we examine the behaviour of other mobility measures in the light of the foregoing analysis. Then, in section 7.2 we examine the performance of our family of mobility indices in a real-world application.

7.1 Properties of mobility indices

Let us now compare our family of measures to other approaches in the literature. We show that the M_α family has appropriate properties and performs better than other widely used indices. In particular, we show that our scale-independent income mobility index, defined in (19)-(21), shares the same desirable properties as the elasticity and correlation based measures, without having the major drawback noted in the introduction.

In addition to the well-known mobility measures based on the elasticity and correlation coefficients discussed in the introduction, we consider the following mobility measures:

- Fields and Ok (1996) provided a measure of mobility based on income differences:[23]

$$FO_1 = \frac{1}{n} \sum_{i=1} \mid x_{0i} - x_{1i} \mid .$$

- Fields and Ok (1999b) provided a measure of mobility based on differences in log

[23]No mobility is defined when incomes at both periods are shifted by the same value.

incomes:[24]

$$FO_2 = \frac{1}{n} \sum_{i=1} \mid \log x_{1i} - \log x_{0i} \mid .$$

- Shorrocks (1978a) provided mobility measures related to inequality:

$$S_I = 1 - \frac{I(x_0 + x_1)}{\frac{\mu_{x_0}}{\mu_{x_0+x_1}} I(x_0) + \frac{\mu_{x_1}}{\mu_{x_0+x_1}} I(x_1)},$$

  where $I(.)$ is a predefined inequality measure.

- We consider our scale-independent M_0 and our translation-independent M_0' measures, defined, respectively, in (20) and (43):

$$M_0 = -\frac{1}{n} \sum_{i=1}^{n} \frac{x_1}{\mu_{x_1}} \log \left(\frac{x_0 \, \mu_{x_1}}{x_1 \, \mu_{x_0}} \right) \quad \text{and} \quad M_0' = \frac{1}{2} \text{var}(x_0 - x_1).$$

- We also consider rank mobility measures: (1) $1 - \rho$, where ρ is the Spearman correlation coefficient, and (2) our mobility measure M_0 where incomes are replaced by ranks divided by the number of individuals, as defined in (52), denoted M_0^r hereafter.

Table 6 presents values of these mobility measures in different situations. We consider a three-person world (A, B, C), with always the same incomes in period 0, $x_0 = (e, e^{1.5}, e^2)$, and several scenarios in period 1, with shifted, rescaled and/or reranked incomes.

Elasticity and correlation coefficients are independent of units of measurement of the variables. So mobility indices based on these coefficients respect the scale-independence property. It is clear from Table 6, where scenario 1^a gives a zero value ($1 - \beta = 0$), and scenarios 1^c and 1^d provide the same value ($1 - \beta = 1.5$). Furthermore, the major drawback provided in the introduction is also clear, since zero mobility is obtained with scenarios 1^f or 1^g. It follows that a low value of these measures cannot be associated to low mobility.

The Cowell-Flachaire scale-independent measure M_0 behaves similarly to elasticity and correlation based measures, $1 - \beta$ and $1 - \rho$, but it exhibits non zero values in scenarios 1^f and 1^g. So it shares their unit-free independence property, but it does not share their major drawback.

The Cowell-Flachaire translation-independent measure M_0' behaves similarly to M_0 but it is insensitive to an absolute shift of incomes rather than to a scale factor. It exhibits zero value in 1^b and it provides the same value in 1^c and 1^e.

The Fields-Ok mobility measures are not scale-independent in the sense explained in section 3.4, they have values different from zero in scenario 1^a ($FO_1 = 4.863$ and $FO_2 = 0.693$) and they have different values in 1^c and 1^d. In Table 6, we can see that the same value is given to Fields-Ok measures in scenarios 1^f and 1^g, who share the

[24] No mobility is defined when incomes at both periods are multiplied by the same value.

	period	period						
	0	1^a	1^b	1^c	1^d	1^e	1^f	1^g
A	$e^{1.0}$	$2\,e^{1.0}$	$e^{1.0}+2$	$e^{1.5}$	$2\,e^{1.5}$	$e^{1.5}+2$	e^1	e^1
B	$e^{1.5}$	$2\,e^{1.5}$	$e^{1.5}+2$	$e^{2.0}$	$2\,e^{2.0}$	$e^{2.0}+2$	e^3	e^2
C	$e^{2.0}$	$2\,e^{2.0}$	$e^{2.0}+2$	$e^{1.0}$	$2\,e^{1.0}$	$e^{1.0}+2$	e^2	e^3
Income-mobility measures								
Elasticity	$1-\beta$	**0**	0.312	*1.500*	*1.500*	1.318	**0**	-1.000
Pearson Corr.	$1-r$	**0**	0.001	*1.500*	*1.500*	1.461	0.500	**0**
Cowell-Flachaire	M_0	**0**	0.007	*0.221*	*0.221*	0.164	0.271	0.060
Cowell-Flachaire	M_0'	2.782	**0**	*8.345*	19.471	*8.345*	40.580	22.123
Fields-Ok	FO_1	4.863	2.000	3.114	6.165	3.781	*5.201*	*5.201*
Fields-Ok	FO_2	0.693	0.387	0.667	0.898	0.686	*0.500*	*0.500*
Shorrocks	S_{Theil}	**0**	0.031	0.743	0.679	0.751	0.281	0.069
Shorrocks	S_{Gini}	**0**	**0**	*0.500*	0.459	*0.500*	0.132	**0**
Rank-mobility measures								
Spearman corr.	$1-\rho$	**0**	**0**	*1.500*	*1.500*	*1.500*	0.500	**0**
Cowell-Flachaire	M_0^r	**0**	**0**	*0.251*	*0.251*	*0.251*	0.068	**0**

Table 6: Income and rank mobility measures in different scenarios.

same income values, with the same ranking at the two periods in 1^g and a reranking in 1^f.

The Shorrocks measures are not scale-independent (scenarios 1^c and 1^d provide different values). In addition, they are sensitive to the choice of the inequality index. Indeed, Table 6 gives very different results with the Theil and Gini indices (S_{Theil}, S_{Gini}). At first sight, the Shorrocks index based on the Gini may appear to be an appropriate measure of rank mobility (Aaberge et al. 2002), because it is equal to zero when no individual position shifts takes place (scenarios $1^a, 1^b$ and 1^g). However, it should not be used to measure rank mobility, because two similar reranking scenarios (1^c and 1^d) give different values of the index (0.5 vs 0.459).

When we turn to rank mobility measures, we can see that our mobility measure behaves similarly to the Spearman correlation based mobility measure. A nice feature of our rank mobility index is that it shares similar foundations as our income mobility measures.

7.2 Empirical application

Chen and Cowell (2017) examine the evidence on rank and income mobility in China during the decades immediately preceding and immediately following the millennium, using data from the China Health and Nutrition Survey.

Table 7 presents some of their results: the transition matrices pre and post millennium, where groupings 1,...,5 are equal-sized twenty percent slices of the distribution. Rank mobility appears to have fallen from the pre-millennium to the post-millennium

		2000							2011				
		1	*2*	*3*	*4*	*5*			*1*	*2*	*3*	*4*	*5*
1989	*1*	**0.29**	0.23	0.21	0.15	0.12	**2000**	*1*	**0.34**	0.26	0.18	0.13	0.08
	2	0.25	**0.25**	0.21	0.17	0.12		*2*	0.24	**0.25**	0.23	0.16	0.13
	3	0.18	0.23	**0.21**	0.23	0.15		*3*	0.18	0.23	**0.22**	0.20	0.17
	4	0.14	0.15	0.21	**0.22**	0.27		*4*	0.13	0.14	0.22	**0.27**	0.24
	5	0.13	0.14	0.16	0.22	**0.34**		*5*	0.11	0.13	0.15	0.24	**0.37**
		$M_0(\Pi) = 0.197$							$M_0(\Pi) = 0.176$				
		[0.177; 0.216]							*[0.157; 0.197]*				

Table 7: Mobility in China: decade rank transition matrices and Cowell-Flachaire summary index. Numbers in brackets are 95%-confidence intervals.

decade, because the values of the diagonal elements increased over time. However, taking into account statistical inference, the authors find two diagonal values significantly different between the two matrices only ($\pi_{11} = 0.29, 0.34$ and $\pi_{44} = 0.22, 0.27$). It is not so easy to compare two matrices with many different values. Such comparisons do not provide always clear results.

A convenient way to capture the mobility implied by a transition matrix is to use a summary statistic. In the case of a quantile transition matrix, the proportion of individuals in each class is the same and our summary index defined in (26) becomes: [25]

$$M_\alpha(\Pi) = \begin{cases} \frac{1}{\alpha(\alpha-1)}\left[\frac{2}{K(K+1)}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\pi_{k\ell}k^{\alpha}\ell^{1-\alpha} - 1\right] & \alpha \neq 0, 1 \\ \frac{-2}{K(K+1)}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\pi_{k\ell}\,\ell\log\left(\frac{k}{\ell}\right) & \alpha = 0 \\ \frac{2}{K(K+1)}\sum_{k=1}^{K}\sum_{\ell=1}^{K}\pi_{k\ell}\,k\log\left(\frac{k}{\ell}\right) & \alpha = 1 \end{cases}$$

We compute this index for $\alpha = 0$, with 95%-confidence intervals: values are given in Table 7.[26] At first sight, rank mobility seems to fall between the two periods, because the values of the index decrease (0.197 vs. 0.176). However, taking into account statistical inference, there is no significant difference between the two values before and after the millennium (confidence intervals intersect).

The transition matrix is a convenient way of providing a simple snapshot of rank-movements in the sample. However, it provides a rather crude snapshot of an income distribution. This illustrates the more general point that, when information is available on the income history of households, it is better to employ the entire information with the appropriate mobility measures.

[25] With a $K \times K$ transition matrix and, thus, K ordered classes, we can define status u_k and v_ℓ by class numbers k and ℓ and we have $\mu_0 = \mu_1 = (K+1)/2$.

[26] Let us define a n-vector u composed by 20% of each of the following values $1, 2, 3, 4, 5$. We generate a vector $v^\star = \{v_1^\star, \ldots, v_n^\star\}$ where $v_i^\star$ is equal to $1, 2, 3, 4$ or 5 with probabilities $\pi_{u_i,1}, \pi_{u_i,2}, \pi_{u_i,3}, \pi_{u_i,4}, \pi_{u_i,5}$ (row i from the transition matrix). We generate $B = 999$ bootstrap samples $(u, v^\star)$ from which we compute B transition matrices $\Pi^{(1)}, \ldots, \Pi^{(B)}$ and B bootstrap statistics, $M_0^{(1)}, \ldots, M_0^{(B)}$. The 95%-confidence interval is given by the 2.5% and 97.5% empirical quantiles obtained from the set of B bootstrap statistics, $M_0^{(1)}, \ldots, M_0^{(B)}$. We use $n = 2840$ in 1989-2000 and $n = 2600$ in 2000-2011.

	Rank mobility		Income mobility	
	1989-2000	2000-2011	1989-2000	2000-2011
$\alpha = 0$				
overall	0.3776	0.3356	0.4729	0.5555
	[0.3571, 0.3969]	*[0.3153, 0.3558]*	*[0.4408, 0.5155]*	*[0.5076, 0.6205]*
downward	0.1687	0.1464	0.1772	0.1780
	[0.1480, 0.1894]	*[0.1247, 0.1695]*	*[0.1575, 0.2018]*	*[0.1562, 0.2034]*
upward	0.2493	0.2352	0.2689	0.3545
	[0.2269, 0.2701]	*[0.2137, 0.2550]*	*[0.2416, 0.3069]*	*[0.3160, 0.4152]*
$\alpha = 0.5$				
overall	**0.3399**	**0.3077**	**0.4489**	**0.5148**
	[0.3241, 0.3556]	*[0.2907, 0.3241]*	*[0.4230, 0.4812]*	*[0.4762, 0.5603]*
downward	0.1879	0.1591	0.2066	0.1990
	[0.1645, 0.2099]	*[0.1342, 0.1849]*	*[0.1826, 0.2378]*	*[0.1733, 0.2285]*
upward	0.2039	0.1970	**0.2251**	**0.2912**
	[0.1885, 0.2185]	*[0.1816, 0.2121]*	*[0.2053, 0.2529]*	*[0.2641, 0.3336]*
$\alpha = 1.0$				
overall	0.3858	0.3493	0.5110	0.5796
	[0.3641, 0.4056]	*[0.3275, 0.3692]*	*[0.4773, 0.5533]*	*[0.5371, 0.6315]*
downward	0.2437	0.1997	0.2694	0.2432
	[0.2062, 0.2734]	*[0.1622, 0.2328]*	*[0.2313 0.3239]*	*[0.2082, 0.2964]*
upward	0.1901	0.1858	0.2039	0.2643
	[0.1763 0.2037]	*[0.1717, 0.2007]*	*[0.1871, 0.2269]*	*[0.2407, 0.2971]*

Table 8: Mobility in China: Cowell-Flachaire rank and income mobility measures, with 95%-confidence intervals.

Table 8 presents several values of our rank and income mobility measures, with $\alpha = 0, 0.5, 1$, with household data at pre and post millennium, as well as bootstrap 95%-confidence intervals in brackets. Let us consider the case of $\alpha = 0.5$, which gives the same weight to upward and to downward status movements of the same magnitude (see section 4.3). The results suggest that rank mobility decreased and income mobility increased from pre to post millennium (0.3399 vs. 0.3077 and 0.4489 vs. 0.5148). A downward/upward decomposition shows that upward income mobility significantly increased between the two periods (0.2251 vs. 0.2912), not downward income mobility.[27] Similar results are obtained with $\alpha = 0$ and $\alpha = 1$, from which we can see that more weight is given, respectively, to upward and downward movements.

Finally, the use of our mobility measures shows that rank mobility decreased from pre to post millennium. By contrast income mobility has carried on increasing; so has income inequality. The differences in the two stories arise, not from a different structure of the mobility index, but from the use of different status concepts. An upward/downward decomposition for the income-status case shows that the story that emerges is mainly due to upward income movements that significantly increased between the two periods. In the Chinese context, it would be important to look also at rural and urban sub-populations separately as well as together. We refer to the paper of Chen and Cowell (2017) for a detailed empirical study.

8 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 different types of status space also require different types of modelling of the basic mobility ordering.

The aggregation of changes in status involves the application of standard principles to individual histories. 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 each individual's contribution to mobility, where each of these contributions depends on status in the two periods, but

[27] The overall mobility index can be computed with the upward and downward mobility indices using (27). If we consider upward income movements defining the first group ($k = 1$) and downward income movements defining the second group ($k = 2$), in 2000-2011, we also have $p_1 = 0.7818898$, $\mu_{u,1} = 8417.185$, $\mu_{v,1} = 27815.95$, $p_2 = 0.2181102$, $\mu_{u,2} = 15378.68$, $\mu_{v,2} = 7071.294$, $\mu_u = 9935.559$, $\mu_v = 23291.33$.

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 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, where the α values enable the oberver to apply his/her own judgment as to whether greater weight should be placed on upward or downward movement.

We have also shown that the empirical performance of the mobility measures in our M_α-family accords well with intuitive understanding of mobility, whereas some commonly-used mobility measures in the literature do not (for some measures a low value of measured mobility does not, in fact, mean low mobility). Furthermore, because our approach can capture income mobility and rank mobility within the same framework, it becomes possible to examine side-by-side mobility comparisons for each of the two underlying status concepts. As illustrated by the example of China around the millennium, this enables us to see more clearly the contrasting patterns of mobility through time with different interpretations according to the type of mobility (up or down) and according to the status concept.

References

Aaberge, R., A. Björklund, M. Jäntti, P. J. Pedersen, N. Smith, and T. Wennemo (2002). Income inequality and income mobility in the Scandinavian countries compared to the United States. *Review of Income and Wealth 48*, 443.

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.

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.

Bhattacharya, D. and B. Mazumder (2011). A nonparametric analysis of black-white differences in intergenerational income mobility in the United States. *Quantitative Economics 2*, 335–379.

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.

Chen, Y. and F. A. Cowell (2017). Mobility in China. *Review of Income and Wealth 63*, 203–218.

Corak, M. (2006). Do poor children become poor adults? Lessons for public policy from a cross country comparison of generational earnings mobility. *Research on Economic Inequality 13*.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives 27*, 79–102.

Corak, M., M. J. Lindquist, and B. Mazumder (2014). Do poor children become poor adults? Lessons for public policy from a cross country comparison of generational earnings mobility. a comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics 30*, 185–200.

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

Cowell, F. A. (2011). *Measuring Inequality* (Third ed.). Oxford: Oxford University Press.

Cowell, F. A. and E. Flachaire (2015). Statistical methods for distributional analysis. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 2. New York: Elsevier.

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 (2012). Inequality adjusted income growth. *Economica 79*, 747–765.

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.

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

Jäntti, M. and S. P. Jenkins (2015). Income mobility. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Volume 2, pp. 807–935. Elsevier.

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

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

Polyanin, A. D. and V. F. Zaitsev (2004). *Handbook of Nonlinear Partial Differential Equations*. Boca Raton: Chapman and Hall / CRC Press.

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 57*, 1–22.

Shorrocks, A. F. (1978a). Income inequality and income mobility. *Journal of Economic Theory 19*, 376–393.

Shorrocks, A. F. (1978b). 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.

Appendix: Proofs

Proof. [**Theorem 1**].In both the case where Z is a connected subset of $\mathbb{R} \times \mathbb{R}$ and the case where Z is $\mathbb{Q}_+ \times \mathbb{Q}_+$ Theorem 5.3 of Fishburn (1970) of can be invoked to show that axioms 1 to 3 imply that $\succeq$ can be represented as

$$\sum_{i=1}^{n} \phi_i(z_i), \forall \mathbf{z} \in Z^n \tag{65}$$

where ϕ_i is continuous, defined up to an affine transformation and, by Axiom 2 is increasing in v_i if $v_i > u_i$ and *vice versa*. Using Axiom 4 in (65) we have

$$\phi_i(u_i, u_i) = \phi_i(u_i + \delta, u_i + \delta). \tag{66}$$

where $\delta := u_i' - u_i$. Equation (66) implies that ϕ_i must take the form $\phi_i(u, u) = a_i + b_i u$. Since ϕ_i is defined up to an affine transformation we may choose $a_i = 0$ and so we have

$$\phi_i(u, u) = b_i u. \tag{67}$$

■

Proof. [**Theorem 2**]. The proof proceeds by considering two cases of (λ_0, λ_1).

Case 1: $\lambda_0 = \lambda_1 = \lambda > 0$.

Theorem **1** implies that if $\mathbf{z} \sim \mathbf{z}'$ then

$$\sum_{i=1}^{n} \phi_i(z_i) = \sum_{i=1}^{n} \phi_i(z_i'). \tag{68}$$

Axiom 5 further implies that

$$\sum_{i=1}^{n} \phi_i(\lambda z_i) = \sum_{i=1}^{n} \phi_i(\lambda z_i').$$

These two equations imply that the function (11) is homothetic so that we may write

$$\sum_{i=1}^{n} \phi_i(\lambda z_i) = \theta\left(\lambda, \sum_{i=1}^{n} \phi_i(z_i)\right), \tag{69}$$

where $\theta : \mathbb{R} \to \mathbb{R}$ is increasing in its second argument. Consider the case where, for arbitrary distinct values j and k, we have $v_i = u_i = 0$ for all $i \neq j, k$. This implies that $\phi_i(u_i, v_i) = 0$ for all $i \neq j, k$ and so, for given values of v_j, v_k, λ, (69) can be written as the functional equation:

$$f_j(u_j) + f_k(u_k) = h(g_j(u_j) + g_k(u_k)), \tag{70}$$

where $f_i(u) := \phi_i(\lambda u, \lambda v_i)$, $g_i(u) := \phi_i(u, v_i)$, $i = j, k$ and $h(x) := \theta(\lambda, x)$. Alternatively, for given values of u_j, u_k, λ, (69) can be written as the functional equation

$$f_j(v_j) + f_k(v_k) = h(g_j(v_j) + g_k(v_k)), \tag{71}$$

with $f_i(v) := \phi_i(\lambda u_i, \lambda v)$, $g_i(v) := \phi_i(u_i, v)$, $i = j, k$ and $h(x) := \theta(\lambda, x)$. Take first the functional equation (70): it has the solution

$$\begin{aligned} f_i(u) &= a_0 g_i(u) + a_i, \ i = j, k, \\ h(x) &= a_0 x + a_j + a_k, \end{aligned}$$

where a_0, a_j, a_k, are constants that may depend on λ, v_j, v_k (Polyanin and Zaitsev 2004, Supplement S.5.5). Therefore:

$$\phi_j(\lambda u_j, \lambda v_j) = a_0(\lambda, v_j, v_k)\phi_j(u_j, v_j) + a_j(\lambda, v_j, v_k) \tag{72}$$

$$\phi_k(\lambda u_k, \lambda v_k) = a_0(\lambda, v_j, v_k)\phi_k(u_k, v_k) + a_k(\lambda, v_j, v_k). \tag{73}$$

Since j and k are arbitrary, we could repeat the analysis for arbitrary distinct values j and ℓ and $v_i = u_i = 0$ for all $i \neq j, \ell$, where $\ell \neq k$; then we would have

$$\phi_j(\lambda u_j, \lambda v_j) = a_0'(\lambda, v_j, v_k)\phi_j(u_j, v_j) + a_j'(\lambda, v_j, v_\ell) \tag{74}$$

$$\phi_k(\lambda u_\ell, \lambda v_\ell) = a_0'(\lambda, v_j, v_k)\phi_\ell(u_\ell, v_\ell) + a_\ell'(\lambda, v_j, v_\ell). \tag{75}$$

where a_0', a_j', a_ℓ, are constants that may depend on λ, v_j, v_ℓ. The right-hand sides of (72) and (74) are equal and so a_j must be independent of v_j and a_0 must be independent of v_j, v_k. Therefore, because j and k are arbitrary we have

$$\phi_i(\lambda u_i, \lambda v_i) = a_0(\lambda)\phi_i(u_i, v_i) + a_i(\lambda, v_i), \ i = 1, ..., n. \tag{76}$$

In the case where $v_i = u_i$, (67) and (76) yield

$$b_i \lambda v_i = a_0(\lambda) b_i v_i + a_i(\lambda, v_i)$$

so that

$$a_i(\lambda, v_i) = [\lambda - a_0(\lambda)] b_i v_i$$

and (76) can be rewritten

$$\phi_i'(\lambda u_i, \lambda v_i) - b_i v_i = a_0(\lambda)\phi_i'(u_i, v_i), \ i = 1, ..., n. \tag{77}$$

where $\phi_i'(u_i, v_i) := \phi_i(\lambda u_i, \lambda v_i) - b_i v_i$. From Aczél and Dhombres (1989), page 346 there must exist $\beta \in \mathbb{R}$ and a function $h : \mathbb{R}_+ \to \mathbb{R}$ such that $\phi_i'(u_i, v_i) = u_i^\beta h_i(v_i/u_i)$, so that

$$\phi_i(u_i, v_i) = u_i^\beta h_i\left(\frac{v_i}{u_i}\right) + b_i u_i. \tag{78}$$

From (67) we see that (78) implies $h_i(1) = 0$. Now return to the alternative functional equation (71): following the same argument this must have a solution of the form

$$\phi_i(u_i, v_i) = u_i^{\beta'} h_i'\left(\frac{v_i}{u_i}\right) + b_i v_i \tag{79}$$

Case 2: $\lambda_0 = 1,\ \lambda_1 = \lambda \neq 1.$

Again if $\mathbf{z} \sim \mathbf{z}'$ then (68) holds. Now Axiom 5 implies

$$\sum_{i=1}^{n} \phi_i (u_i, \lambda v_i) = \sum_{i=1}^{n} \phi_i (u_i, \lambda v_i') . \tag{80}$$

Equations (68) and (80) imply that the function (11) is homothetic in v so that we may write

$$\sum_{i=1}^{n} \psi_i (\lambda v_i) = \theta \left(\lambda, \sum_{i=1}^{n} \psi_i (v_i) \right), \tag{81}$$

where $\psi_i (v) := \phi_i (u_i, v)$ and $\theta : \mathbb{R} \to \mathbb{R}$ is increasing in its second argument. By the same argument as before we have

$$\psi_i (\lambda v_i) \quad = \quad a_0 (\lambda) \psi_i (v_i) + a_i (\lambda) ,\ i = 1, ..., n. \tag{82}$$

Putting $v_i = 0$ in (82) we see that

$$a_i (\lambda) = \psi_i (0) [1 - a_0 (\lambda)]$$

and so we may rewrite(82) as

$$\psi_i' (\lambda v) \quad = \quad a_0 (\lambda) \psi_i' (v) , \text{ where} \tag{83}$$

$$\psi_i' (v) \quad := \quad \psi_i (v) - \psi_i (0) . \tag{84}$$

Equation (83) can be expressed as $f (x + y) = g(y) + f(x)$ where$f (\centerdot) := \log (\psi_i' (\centerdot))$, $g (\centerdot) := \log (a_0 (\centerdot))$, $x = \log v$, $y = \log \lambda$. This Pexider equation has the solution $f (x)$=$bx + c$, $g (y) = \alpha y$

$$\log \psi_i' (v) = a + b \log v,\ \log (a_0 (\lambda)) = b (\log \lambda)$$

where the constant a may depend on i and u_i. This implies

$$\phi_i (u_i, v_i) = A_i (u_i) v_i^b + \phi_i (u_i, 0) . \tag{85}$$

where $A_i (u_i) = \exp (a)$. Putting $v_i = u_i$ in (78) and (85) we find

$$A_i (u_i) u_i^b + \phi_i (u_i, 0) = b_i u_i.$$

since the RHS is linear in u_i we must have $A_i (u_i)$ proportional to u_i^{1-b}. Therefore

$$\phi_i (u_i, v_i) = c_i v_i^b u_i^{1-b} + \phi_i (u_i, 0) \tag{86}$$

Now combine the results from the two cases. Since (78), (79) and (86) are true for arbitrary u_i, v_i this implies that

$$\phi_i (u_i, v_i) = c_i u_i^{\alpha} v_i^{1-\alpha} + c_i' u_i + c"_i v_i \tag{87}$$

where $\alpha := 1 - b$. Differentiating (87) we have

$$\frac{\partial \phi_i (u_i, v_i)}{\partial u_i,} = \alpha c_i u_i^{\alpha-1} v_i^{1-\alpha} + c_i' \tag{88}$$

$$\frac{\partial \phi_i (u_i, v_i)}{\partial v_i,} = [1 - \alpha] c_i u_i^{\alpha} v_i^{-\alpha} + c"_i \tag{89}$$

In view of Axiom 2 (88) and (89) must be zero when $v_i = u_i$: this requires $c_i' = -\alpha c_i$ and $c"_i = -[1 - \alpha] c_i$. This in turn implies (12) . ■