Income Mobility: A Robust Approach

by

Frank Cowell
STICERD, London School of Economics and Political Science

and

Christian Schluter
University of Bristol

Discussion Paper
No.DARP/37
July 1998

The Toyota Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020-7955 6678

Partially supported by the ESRC Centre for Analysis of Social Exclusion (CASE), and by ESRC grant R5500XY8. We gratefully acknowledge helpful comments by Dirk Van de Gaer and Maria-Pia Victori-Feser, and research assistance by Chris Soares and Hung Wong.

Abstract

The performance of two broad classes of mobility indices is examined when allowance is made for the possibility of data contamination. Single-stage indices – those that are applied directly to a sample from a multivariate income distribution – usually prove to be non-robust in the face of contamination. Two-stage models of mobility – where the distribution is first discretised and then a transition matrix or other tool is applied – may be robust if the first stage is appropriately specified. We illustrate results using a simple but flexible simulation.

Keywords: Mobility measures; robustness; data contamination.

JEL Nos.: C13, D63.

Contact address: Contact address: Professor Frank Cowell, STICERD, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. Email: f.cowell@lse.ac.uk

1 Introduction

Reliable indicators of mobility are of continuing relevance for theoretical work and policy applications in several important areas, for example, the study of poverty transitions, the modelling of bequest dynamics, the characterisation of earnings or income histories. Because the measurement of income mobility involves the comparison of distributions of income profiles it may inherit some of the practical problems associated with empirical income distributions. The problem of measurement error has long been recognised (Bound et al. 1989, Bound and Krueger 1989), but other difficulties remain. Prominent among these is the problem of contamination:[1] even if one is reasonably confident about a data source, it is obviously inappropriate to assume that the data will automatically give a reasonable picture of the "true" picture of mobility. A researcher may anticipate that, because of miscoding and other types of mistake, some of the observations will be incorrect, and this may have a serious impact upon mobility estimates and comparisons. The purpose of this paper is to examine the performance of some important classes of mobility measures in the presence of contamination.

The central question that we wish to address is whether the properties of mobility indices in conjunction with the characteristics of panel data can give rise to misleading conclusions about income-mobility patterns. Obviously if contamination is in some sense "large" relative to the true data then we cannot expect

[1] The relationship between the two types of approach to imperfections in the data is discussed in Cowell (1998).

to get sensible estimates of mobility indices; but what if the contamination were quite small? Could it be the case that isolated "blips" in the data or extreme values could drive estimates of income mobility? We analyse this problem using methods of robust analysis that have become established in other fields.

There is a special difficulty associated with the problem of data contamination in the present context. Pragmatic approaches that are relatively easy to implement in other income distribution problems may be impractical in applications to issues such as the measurement of mobility. For example, in the analysis of income inequality, it may be appropriate to "trim" data by eye or by algorithm, but the types of rule-of-thumb treatment of outliers that could work well for a univariate problem are likely to be unwieldy in the case of multivariate distributions.

This practical difficulty underlines the importance of understanding the general properties of mobility indices when applied to contaminated data. Our approach is to establish these properties for two broadly-defined types of index using a simple model of data contamination. Section 2 sets out the basic ingredients of the approach; sections 3 and 4 discuss the first of the two principal types of mobility indices; section 6 discusses the second type of index; section 7 concludes.

2 The Fundamentals

We suppose that an income history can be described by a T-dimensional random variable $\mathbf{X}$ where $T \geq 2$. The variate $\mathbf{X}$ may be thought of as a profile of income-events over T discrete periods from which one wishes to estimate income mobility. We write the set of income profiles as $\mathfrak{X} = [\underline{x}, \bar{x}] \times [\underline{x}, \bar{x}] \times ... \times [\underline{x}, \bar{x}]$ where $[\underline{x}, \bar{x}]$ is an interval in $\Re$. Notice that for some approaches to the problem of analysing economic mobility one may wish to restrict $\mathfrak{X}$ to a strict subset of T-dimensional space $\Re^T$ because, for example, one may wish to rule out zero or negative incomes as irrelevant *a priori*; however, unless otherwise specified, we assume that $\underline{x} = -\infty$ and $\bar{x} = \infty$.

We will use the symbol '$\cdot$' to denote the vectorial product (inner product) of two members of $\Re^T$.

2.1 Distributions

Assume that the distribution of income profiles a particular dynamic economy is given by some distribution function $F : \mathfrak{X} \to [0, 1]$. Let $\mathfrak{F}_T$ be the class of all valid T-variate distribution functions. We will find useful a number of derived distributions of linear combinations of the T-variates $\mathbf{X}$. Given a parametric weight vector $\mathbf{w} \in \Re^T$, $\mathbf{w} \cdot \mathbf{1} = 1$ and any $F \in \mathfrak{F}_T$, these derived distributions can be expressed in the form of a functional $\Psi(\cdot; F, \mathbf{w}) : \mathfrak{F}_T \to \Re$ where

$$\Psi(y; F, \mathbf{w}) := \int ... \int_{\{\mathbf{x}:\mathbf{w}\cdot\mathbf{x} \leq y\}} dF(\mathbf{x}) \; . \tag{1}$$

For example the marginal distribution of income in the tth period is $\Psi(\cdot; F, \mathbf{e}_t)$ where $\mathbf{e}_t$ is the tth unit vector $(0, 0, ..., 0, 1, 0., .., 0)$, and the distribution of (unweighted) average income over the T periods is $\Psi\left(\cdot;\ F, T^{-1}\mathbf{1}\right)$.

2.2 Mobility

A *mobility index* M is a real-valued functional defined on the space of T-variate random distributions $M : \mathfrak{F}_T \to \Re$. There are several competing intellectual approaches to the specification of such indices, which need not detain us here. Specific types of mobility indices are discussed in Sections 3 to 6 below; for the moment note that the class of indices M be resolved into two important subclasses:

- *single-stage* indices which attempt to make full use of information in F,
- *two-stage* indices that are based on partial discretisation of the distribution F *a priori.*

For a particular multivariate distribution $F \in \mathfrak{F}_T$ we may express the mobility index as the value of the functional $M(F)$. In many practical applications the "true" distribution will not be known but must be estimated from some dataset.

Let $\mathbf{x} := \mathbf{x}_1, \ldots, \mathbf{x}_n$ denote a sample of size n where each $\mathbf{x}_i \in \mathfrak{X}$ is a realisation of $\mathbf{X}$. An estimator of $M(F)$ is then obtained principally by one of two approaches.

1. For the *non-parametric approach* one replaces F with the empirical distribution: $F^{(n)}(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} \Delta_{\mathbf{x}_i}(\mathbf{x})$ where $\Delta_{\mathbf{y}}$ is a point mass at $\mathbf{y}$. Letting the sample size $n \to \infty$, by the Glivenko-Cantelli theorem the estimator can be written as a functional of the distribution function F, i.e. asymptotically the mobility index $M(F^{(n)})$ becomes $M(F)$ (Victoria-Feser 1998).

2. In the *parametric approach* one assumes *a priori* that $\mathbf{X} \sim F_{\boldsymbol{\theta}}$ where $F_{\boldsymbol{\theta}}$ is a member of a family of distributions characterised by the parameter vector $\boldsymbol{\theta}$. One then finds $\hat{\boldsymbol{\theta}}$ - an estimate of $\boldsymbol{\theta}$ - from the sample $\mathbf{x}$ and estimates mobility using $F_{\hat{\boldsymbol{\theta}}}$.

Here we will assume that a complete set of micro-data is available for the T periods, and we focus upon non-parametric methods.

2.3 Data Contamination

Because in practice a mobility index is usually estimated using a sample one should realistically expect that the data may be subject to contamination: for example the misreporting of weekly as monthly income, or the presence in the sample of data points that have been miscoded by the data transcriber (the classic decimal-point error). If one had reason to suspect that this sort of error were extensive in the data sets under consideration the problem of distributional comparison might have to be abandoned because of unreliability. However, it is possible that there might be a fairly serious problem of comparison even if the amount of contamination were fairly small, so that the data might be considered "reasonably clean".

A standard model of this type of problem is as follows.[2] Suppose that the "true" multivariate distribution for which we wish to estimate mobility is F but, because of the problem of data-contamination, we cannot assume that the data actually observed have really been generated by F. What we actually observe instead of F is a distribution that is in some neighbourhood of it, $F_\varepsilon = (1-\varepsilon)F + \varepsilon H$ where $0 < \varepsilon < 1$ and H is a perturbation distribution. For example, H could be a distribution of discrete masses in $\mathfrak{X}$

$$dH(\mathbf{x}) = \begin{cases} \alpha_1 & \text{if } \mathbf{x} = \mathbf{z}_1 \\ \ldots & \\ \alpha_m & \text{if } \mathbf{x} = \mathbf{z}_m \end{cases} \tag{2}$$

$\forall i$, $\alpha_i \geq 0$, and $\sum \alpha_i = 1$, $\mathbf{z}_1, ..., \mathbf{z}_m \in \mathfrak{X}$. Then F_ε is the mixture model from which an observation has probability $(1-\varepsilon)$ of being generated by F and a probability $\varepsilon\alpha_j$ of being an arbitrary value $\mathbf{z}_j$. The distribution H represents a simple form of *data contamination* at points $\mathbf{z}_1, ..., \mathbf{z}_m$; ε indicates the importance of the contamination; the convex combination F_ε is the observed distribution, and

[2]This approach is based upon the work of Hampel (1968, 1974), Hampel et al. (1986), Huber (1986).

F remains unobservable.[3]

Clearly if ε in (3) were large we could not expect to get sensible estimates of mobility indices; but what if the contamination were very small? To address this question for any given mobility statistic M we can use an elementary version of (2) where a mixture distribution is constructed by combining the "true" distribution F with a single contamination point mass at income $\mathbf{z} \in \mathfrak{X}$:

$$F_{\varepsilon}^{(\mathbf{z})} = [1 - \varepsilon]\, F + \varepsilon H^{(\mathbf{z})} \tag{3}$$

where $H^{(\mathbf{z})}$ is a degenerate distribution defined by:

$$dH^{(\mathbf{z})}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} = \mathbf{z} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The appropriate tool for assessing the impact of an infinitesimal [4] amount of contamination upon the mobility estimate is then given by the *influence function*:

$$\mathrm{IF}(\mathbf{z}; M, F) := \lim_{\varepsilon \to 0} \left[\frac{M\left(F_{\varepsilon}^{(\mathbf{z})}\right) - M(F)}{\varepsilon} \right] \tag{5}$$

[3]Notice that in the multivariate approach of our model it is legitimate to assume that the observations $\mathbf{x}_i$ are iid; the dependence between the *components* of $\mathbf{x}_i$ of course remains. This is by contrast to the problems of robustness in the analysis of time-series data the observations are not iid (Künsch 1984), (Hampel, Ronchetti, Rousseeuw, and Stahel 1986) pp 417ff. However the assumption implies that issues of missing data and attrition in panel data have been appropriatly dealt with.

[4]"Infinitesimal" means here that the probability ε that this contamination occurs tends to zero.

or, when the derivative exists, by $\frac{\partial}{\partial\varepsilon}M\left(F_{\varepsilon}^{(\mathbf{z})}\right)\Big|_{\varepsilon=0}$.

The IF gives the influence on the estimator M of contamination at the point $\mathbf{z}$, and its value will depend upon the position of $\mathbf{z}$ with respect to the position of the majority of the data. The expression (5) indicates whether an estimate of mobility will be stable in the presence of a few "alien" observations in the income profile and, because the IF is the first-order term in the linear expansion of the asymptotic bias of the estimator it will also provide information about the bias of the mobility estimate. If, under the given model of data-contamination (3) IF in (5) is bounded for all $\mathbf{z} \in \mathfrak{X}$, the mobility statistic M is *robust*. Of course it is particularly interesting to know whether IF could in practice be unbounded. Typically, this problem of unboundedness can arise when components of $\mathbf{z}$ approach extreme values such as ∞, $-\infty$ or 0: in this case a single extreme observation in the income profile could drive the mobility estimate by itself.

Clearly it would be useful to know how the influence function will behave for various types of data contamination for a wide class of mobility indices. So in sections 3 to 6 we will examine the problem of characterising IF for certain key types of statistics M.

3 Stability indices

The first subclass of single-stage indices builds upon an extension of inequality analysis. The principal developments are attributable to Shorrocks (1978) and

Maasoumi and Zandvakili (1986, 1990); they proposed a class of mobility measures based on the comparison of the inequality of (weighted) average income to a weighted average of contemporaneous inequalities. Given an inequality index $I : \mathfrak{F}_1 \rightarrow \Re$ and a set of weights $\mathbf{w} := [w_t]$, a typical stability index is

$$M_{\mathrm{S}}(F; I, \mathbf{w}) := 1 - \frac{I(F_{\mathbf{w}})}{\sum_{t=1}^{T} w_t I(F_t)} \tag{6}$$

where $F_{\mathbf{w}}(\cdot) := \Psi(\cdot; F, \mathbf{w})$ and $F_t(\cdot) := \Psi(\cdot; F, \mathbf{e}_t)$ are the distribution of weighted average income and the marginal distribution (see 2.1 above).

We assume that the true joint distribution function F is not directly observable, and that we have to work with the contaminated distribution function specified in (3). This of course affects all the derived distributions. For example, under this model of contamination, the observed distribution of weighted average income becomes

$$\Psi\left(y; F_{\varepsilon}^{(\mathbf{z})}, \mathbf{w}\right) = (1-\varepsilon)\Psi(y; F, \mathbf{w}) + \varepsilon\iota\left(\mathbf{w}\cdot\mathbf{z} \,\square\, y\right) \tag{7}$$

where $\iota(c)$ is an indicator function equal to 1 if condition c is met and 0 otherwise.

Furthermore, because the mobility index (6) is defined as a function of the values of an inequality statistic for several derived distributions of F, its influence function will be a function of the influence function for the inequality index implemented for these derived distributions. More precisely, using $F_{\varepsilon}^{(\mathbf{z})}$ in (6) and differentiating yields

$$\mathrm{IF}(\mathbf{z}; M_{\mathrm{S}}, F) = \left.\frac{\partial}{\partial \varepsilon} M_{\mathrm{S}}\left(F_{\varepsilon}^{(\mathbf{z})}; I, \mathbf{w}\right)\right|_{\varepsilon=0} \tag{8}$$

$$= -\frac{\mathrm{IF}(\mathbf{z}; I, F_{\mathbf{w}})}{\sum_{t=1}^{T} w_t I(F_t)} + \frac{I(F_{\mathbf{w}}) \sum_{t=1}^{T} w_t \mathrm{IF}(\mathbf{z}; I, F_t)}{\left[\sum_{t=1}^{T} w_t I(F_t)\right]^2} \tag{9}$$

Whether the influence function of the stability index M_{S} is bounded depends on whether $\mathrm{IF}(\mathbf{z}; I, F_{\mathbf{w}})$ and $\mathrm{IF}(\mathbf{z}; I, F_t)$ are bounded, and on whether the component expressions (9) cancel out. If they do not cancel, $\mathrm{IF}(\mathbf{z}; M_{\mathrm{S}}, F)$ is typically unbounded. A cancellation of terms certainly occurs in two trivial situations. Assume that the weights w_t are deterministic. In the first case, if the inequality measure $I(.)$ belongs to the class of scale-independent measures,[5] a universal multiplication of incomes (including the contamination) between periods $X_t = \delta X_{t-1}$, leaves all inequality measures unaffected. The stability measure (6) is therefore a constant being unaffected by the contamination. Likewise, if $I(.)$ belongs to the class of translation-independent measures, the cancellation is induced by a universal shift in incomes $X_t = X_{t-1} + \delta$; other trivial cases could be found for different types of inequality-neutrality. However, the practical relevance of such cross-sectional behaviour is probably rather slight. The only exception may be a perfectly immobile society in which cross-sections are just replicated.

[5]For any $F \in \mathfrak{F}_1$ let $F^{(+k)}$ be the distribution derived by a translation $k \in \Re$, where $F^{(+k)}(x) = F(x - k)$, and let $F^{(\times k)}$ be the distribution derived from F by transforming the income variable by a scalar multiple $k \in \Re_+$ where $F^{(\times k)}(x) = F\left(\frac{x}{k}\right)$, then scale-independent measures have the property $I\left(F^{(\times k)}\right) = I(F)$ and for translation-independent measures $I\left(F^{(+k)}\right) = I(F)$. See Cowell (1998) for a discussion of these and related concepts.

It appears that in practice one might expect the influence function of the stability index to be unbounded. However two issues remain to be resolved. The first is whether this will actually occur (outside trivial income profiles) for standard implementations of stability measures given deterministic weights $\mathbf{w}$: not all inequality indices are inherently robust (Cowell and Victoria-Feser 1996). The second is that we also need to consider cases where weights are stochastic and themselves are subject to contamination: it is possible that reweighting income within the concept of aggregated income $\mathbf{w} \cdot \mathbf{X}$ and reweighting average inequality in the denominator of (6) may have a non-trivial impact upon the robustness property of the stability index. These issues will now be examined for two principal types of inequality index and their associated stability indices.

3.1 The generalised entropy index

A popular inequality index for use in the mobility measure (6) is the generalised entropy index. For any $G \in \mathfrak{F}_1$ this is given by

$$I_{\mathrm{GE}(\alpha)}(G) = \frac{1}{\alpha^2 - \alpha} \left[\frac{\mu_\alpha(G)}{\mu(G)^\alpha} - 1 \right] \tag{10}$$

where μ_α is the functional

$$\mu_\alpha(G) = \int x^\alpha dG(x), \tag{11}$$

μ is the mean $\mu(G) := \mu_1(G)$and $\alpha \in \Re$ is the sensitivity parameter of the index.[6] However the principal difficulty with (10) from the point of view of practical application to the construction of a stability index is that the inequality index is inherently non-robust. The primary reason for this lies in the behaviour of the integral in (11): consider a one-dimensional version of (3) where there is point contamination at $x = \bar{z}$; then for the mixture distribution $G_\varepsilon^{(\bar{z})}$ the integral in (11) becomes

$$\int x^\alpha dG_\varepsilon^{(\bar{z})}(x) = (1-\varepsilon)\mu_\alpha(G) + \varepsilon \bar{z}^\alpha \; ; \tag{12}$$

the last term in (12) can be arbitrarily large when suffering from a large outlier if $\alpha > 0$ or an outlier near zero if $\alpha < 0$. For the second reason for the nonrobustness of the GE index consider its influence function which is given by[7]

$$\mathrm{IF}(\bar{z}; I_{\mathrm{GE}(\alpha)}, G) = A + B\bar{z}^\alpha + C\bar{z} \tag{13}$$

[6] For α large and positive the index is sensitive to changes at the top of the income distribution, for α negative the index is sensitive to changes at the bottom of the distribution. At $\alpha = 0$ and $\alpha = 1$ (10) adopts the form of the MLD and the Theil index respectively. (Cowell 1998)

[7] Using (10) the generalised entropy index for the contaminated income distribution is

$$I_{\mathrm{GE}(\alpha)}(G_\varepsilon^{\bar{z}}) = \frac{1-\varepsilon}{\alpha^2-\alpha}\left[\int \left[\frac{x}{\mu(G_\varepsilon^{\bar{z}})}\right]^\alpha dG(x) - 1\right] + \frac{\varepsilon}{\alpha^2-\alpha}\left[\left[\frac{\bar{z}}{\mu(G_\varepsilon^{\bar{z}})}\right]^\alpha - 1\right]$$

and so, differentiating with respect to ε in the neighbourhood of $\varepsilon = 0$, the influence function is

$$\frac{1}{\alpha^2-\alpha}\left[\left[\frac{\bar{z}}{\mu(G)}\right]^\alpha - \int\left[\frac{x}{\mu(G)}\right]^\alpha dG(x)\right] - \frac{1}{\alpha-1}\int\left[\frac{x}{\mu(G)}\right]^\alpha dG(x)\left[\frac{\bar{z}}{\mu(G)} - 1\right]$$

which gives (13).

where $A := I_{\mathrm{GE}(\alpha)}(G)[\alpha-1]+1/\alpha$, $B := \mu(G)^{-\alpha}\left[\alpha^2-\alpha\right]^{-1}$, $C := \frac{1}{\mu(G)}\left[\alpha I_{\mathrm{GE}(\alpha)}(G)+\frac{1}{\alpha-1}\right]$ are expressions independent of the contamination. Clearly, the function (13) is unbounded for all values of α for sufficiently large $\bar{z}$ because of the last term in parentheses which comes from the impact of contamination upon the mean $\mu\left(G_\varepsilon^{(\bar{z})}\right)$ (Cowell and Victoria-Feser 1996).

Then, by setting $G = F_t$ and $G = F_{\mathbf{w}}$, we can see that this behaviour of the influence function will apply to all the component inequality indices in the stability index (6): all the component inequality indices are nonrobust.[8] To see whether this implies nonrobustness of the associated mobility index the influence functions for M_{S} can now be derived for the generalised-entropy implementation. Assume first that the set of weights $\mathbf{w}$ in (6) is deterministic. Then the function (9) becomes

$$\mathrm{IF}(\mathbf{z}; M_{\mathrm{S}}, F) = \frac{\sum_{t=1}^{T} w_t\left[I_{\mathrm{GE}(\alpha)}(F_{\mathbf{w}})\mathrm{IF}(\mathbf{z}; I_{\mathrm{GE}(\alpha)}, F_t) - I_{\mathrm{GE}(\alpha)}(F_t)\mathrm{IF}(\mathbf{z}; I_{\mathrm{GE}(\alpha)}, F_{\mathbf{w}})\right]}{\left[\sum_{t=1}^{T} w_t I_{\mathrm{GE}(\alpha)}(F_t)\right]^2} \tag{14}$$

Substituting from (13) into (14) we see that the terms involving the unbounded contamination terms z_t and z_t^α vanish only if (a) true inequality $I_{\mathrm{GE}(\alpha)}(F_t)$ in each cross-section t equals $I_{\mathrm{GE}(\alpha)}(F_{\mathbf{w}})$ and (b) z_t is proportional to z_{t-1} for all t. This can only occur if every income profile is a straight line with identical slope and it follows from the fact that the generalised entropy index is a scale indepen-

[8] This can be seen from the above argument by for putting $\bar{z} = \mathbf{w}\cdot\mathbf{z}$ for $G = F_{\mathbf{w}}$ and $\mathbf{w} = \mathbf{e}_t$ for $G = F_t$.

dent inequality measure. Except for this trivial situation (14) is unbounded as $z_t \to \infty$: the mobility index is not robust.

In the case where the set of weights $\mathbf{w}_t$ is endogenous a conventional parameterisation is

$$w_t = \frac{\mu(F_t)}{\mu(\bar{F})} \tag{15}$$

where $\bar{F}(\cdot) = \Psi\left(\cdot;\ F, T^{-1}\mathbf{1}\right)$ is the distribution of T-period average income. This modification is easily incorporated in the denominator of (10), leaving the subsequent derivation essentially unchanged: the mobility index M_{S} for the Generalised Entropy class is not robust whether or not weights are exogenous.

3.2 The Gini coefficient

For any $G \in \mathfrak{F}_1$ the Gini coefficient can be written as the functional

$$I_{\mathrm{Gini}}(G) = 1 - 2\frac{\int_0^1 C(G;q)dq}{\mu(G)} \tag{16}$$

where $C(G;q) := \int_{\underline{x}}^{Q(G;q)} xdG(x)$, $Q(G;q) := \inf\{x : G(x) \geq q\}$.$C$ and Q are the cumulative-income and quantile functionals respectively. Now from Cowell and Victoria-Feser (1996) and Monti (1991) we may derive

$$\begin{aligned}\mathrm{IF}(z; I_{\mathrm{Gini}}, G) \quad &= \quad [1 - I_{\mathrm{Gini}}(G)]\left[\frac{z}{\mu(G)} + 1\right] \\ &\quad - \left[\frac{\mu(G) - z}{\mu(G)}\left(1 - G(\mu(G) - z)\right) + \frac{\mu(G) + z}{\mu(G)}\left(1 - G(\mu(G) + z)\right)\right]\end{aligned}$$

$$-\frac{1}{\mu(G)}\left[\int_{-\infty}^{\mu(G)-z} udG(u) + \int_{-\infty}^{\mu(G)+z} udG(u)\right] \tag{17}$$

which may be shown to be unbounded as $z \to \pm\infty$. The influence functions for Gini-inequality of aggregated income and of individual cross-sectional distributions are derived from application of (17).

With endogenous weights (15) the mobility index is found by using (15) and $I_{\text{Gini}}(G)$ in (6)

$$M_{\text{S}}(F; I, \mathbf{w}) = 1 - \frac{\mu(\bar{F}) - \int_0^1 C(\bar{F}; q)dq}{\sum_{t=1}^{T}\left[\mu(F_t) - \int_0^1 C(F_t; q)dq\right]} \tag{18}$$

The influence function of the mobility index is then given as a special case of (8) by applying (16); it is immediate from (17) that $\text{IF}(\mathbf{z}; M_{\text{S}}, F)$ is also unbounded for the Gini index with endogenous weights.

4 "Distance" and related measures

A second principal subclass of single-stage indices interprets mobility in terms of "distributional change" (Cowell 1985) and typically focuses upon measures that incorporate a concept of distance between incomes. As far as the measures' properties in the face of contaminated data are concerned they can be treated in the same manner as the approach of section 3. The distributional-change approach requires restriction to a two-period interpretation of mobility: we will label the two periods $(t-1, t)$ and $\mathbf{x} := (x_{t-1}, x_t)$.

Theorem 1 of Cowell (1985) establishes that subgroup-decomposable continuous measures of distributional change must take the form

$$M_{\text{Dist}}(F) := \phi\left(\int D(\mathbf{x})\, dF(\mathbf{x}),\ \boldsymbol{\mu}(F)\right) \tag{19}$$

where the function $D : \mathfrak{X} \times \mathfrak{X} \to \Re$ embodies the concept of distance. A principal example of (19) is derived in the case where the function D is homothetic, in which case the measure takes the form of a generalised conditional entropy index:

$$\frac{1}{\alpha^2 - \alpha} \int \int \left[\left[\frac{x_{t-1}}{\mu\left(F_{t-1}\right)} \right]^{1-\alpha} \left[\frac{x_t}{\mu\left(F_t\right)} \right]^{\alpha} - 1 \right] dF(\mathbf{x}) \tag{20}$$

where α is a sensitivity parameter (Cowell 1980). In the light of the argument in subsection 3.1 it is immediate that (20) is nonrobust for all α. However there are other commonly-used indices that are either examples of (19) or that employ a similar notion of aggregating the "distance" between individuals' incomes in the two distributions.

4.1 The Hart index

The Hart index incorporates the concept of distance that is implicit in the use of the variance of logarithms:

$$M_{\text{Hart}}(F) := 1 - r(\log x_{t-1}, \log x_t) \tag{21}$$

where $r(.)$ is the correlation coefficient (Shorrocks 1993). It may be defined equivalently as

$$M_{\text{Hart}}(F) := 1 - \frac{A(F)}{B(F)} \tag{22}$$

where

$$A(F) := \int\int [\log x_{t-1}] [\log x_t] \, dF - \mu(F_{t-1}^*)\mu(F_t^*)$$

$$B(F) := \sqrt{\left[\mu_2(F_{t-1}^*) - \mu(F_{t-1}^*)^2\right] \left[\mu_2(F_t^*) - \mu(F_t^*)^2\right]}$$

and $\log \mathbf{x} \sim F^*$. The index (21) is thus composed entirely of linear functionals, each of which is nonrobust; its influence function then is $\frac{B(F)\,\mathrm{IF}(\mathbf{z};A,F) - A(F)\,\mathrm{IF}(\mathbf{z},B,F)}{B(F)^2}$. Moreover, the contaminations do not cancel out, leading us to conclude that M_{Hart} is not robust.[9]

[9] Examine for instance one term of B, such as $\mu_2(F_{t-1}^*) - \mu(F_{t-1}^*)^2$. For the contaminated distribution $F_\varepsilon^{(\mathbf{z})}$ this becomes

$$(1-\varepsilon)\mu_2(F_{t-1}^*) + \varepsilon z_{t-1}^2 + (1-\varepsilon)^2\mu(F_{t-1}^*)^2 + \varepsilon^2 z_{t-1}^2 - 2(1-\varepsilon)\varepsilon\mu(F_{t-1}^*)z_{t-1}.$$

The observation that no terms cancel out is then immediate.

4.2 The Fields-Ok Index

Recently Fields and Ok (1997) have proposed a mobility index which can be seen as a special case of (19) where the distance concept is based on the absolute differences of logarithms:

$$M_{\mathrm{FO}}(F) = c \int \int |\log x_{t-1} - \log x_t| \, dF(x_{t-1}, x_t). \tag{23}$$

In order to establish the properties of (23), since $|a - b| \geq (a - b)$, it is helpful to examine the behaviour of the following functional:

$$m(F) := c \int \int (\log x_{t-1} - \log x_t) \, dF(x_{t-1}, x_t). \tag{24}$$

In the presence of the contamination given in equation (3), this becomes

$$m\left(F_{\varepsilon}^{(\mathbf{z})}\right) = (1 - \varepsilon) m(F) + \varepsilon c (\log z_{t-1} - \log z_t) \tag{25}$$

The influence function is thus

$$\mathrm{IF}(m; F, z) = -m(F) + c\left[\log z_{t-1} - \log z_t\right] \tag{26}$$

As $z_{t-1} \to \infty$ it is clear that (26) is unbounded from above unless $z_{t-1} = z_t$. Since $M_{\mathrm{FO}}(F) \geq m(F)$ this implies that the index (23) is also non-robust. A similar argument applies if $z_t \to 0$; so M_{FO} is non-robust for this form of contamination. For the reverse contamination profile, exchange x_{t-1} and x_t in (24): we may conclude that the index M_{FO} is generally non-robust.

4.3 The King Index

King's (1983) welfarist approach to the measurement of mobility differs in two aspects from other measures incorporating a distance concept. Following Atkinson (1970), he derives axiomatically a social-welfare function consistent with the proposed mobility measure. Second, he refers to social mobility as changing ranks within distributions. The index can be expressed as

$$M_{\text{King}}(F) = 1 - \left[\frac{\int\int \left(x_t e^{\gamma s(F,\mathbf{x})} \right)^k dF(\mathbf{x})}{\mu_k(F_t)} \right]^{\frac{1}{k}} \quad k \,\square\, 1,\ k \neq 0,\ \gamma \geq 0 \tag{27}$$

where $s(F;\mathbf{x}) := \frac{|x_t - Q(F_t; F_{t-1}(x_{t-1}))|}{\mu(F_t)}$ is the "scaled order statistic" which captures reranking.

To examine its robustness properties write (27) as $1 - [A(F)/B(F)]^{1/k}$ where

$$A(F) := \int\int \left(x_t e^{\gamma s(F,\mathbf{x})} \right)^k dF(\mathbf{x}) \tag{28}$$

and $B(F) := \mu_k(F_t)$. It suffices to examine the expressions $A(F)$ and $B(F)$

separately. With point mass contamination at $\mathbf{z}$ it follows that

$$\left.\frac{\partial A(F_\varepsilon)}{\partial\varepsilon}\right|_{\varepsilon=0} = -A(F)+\left[z_t e^{\gamma\iota(s(F,(x_{t-1},z_t)))}\right]^k + \mathrm{IF}(\mathbf{z}; s, F)\int\int k\left(x_t e^{\gamma s(F,\mathbf{x})}\right)^k \gamma dF(\mathbf{x}) \tag{29}$$

where $\iota(s)$ is an indicator equal to zero if $s = 0$ and $O(z_t)$ otherwise. Similarly

$$\left.\frac{\partial B(F_\varepsilon)}{\partial\varepsilon}\right|_{\varepsilon=0} = -B(F) + z_t^k. \tag{30}$$

The term $z_t e^{\gamma\iota(s)}$ is unbounded for arbitrarily large contamination and so it is clear that the (29) and (30) are unbounded if $z_t \to \infty$ in the case where $k > 0$;[10] in the case where $k < 0$ the expressions are unbounded if $z_t \to 0$. We may conclude that the influence function of King's index is unbounded irrespective of the mobility pattern (i.e. the value of s) and even if the mean is deterministic.

5 Simulation

We have seen that most of the single stage measures introduced in sections 3 and 4 are non-robust. *In principle* they might be extraordinarily sensitive in that an infinitesimal amount of contamination in the wrong place could cause the value of the index to be biased away from the value it would adopt for the uncontaminated distribution. It remains to establish how important this issue is

[10] As $z_t \to \infty$ there are two possibilities for the term $s(F, (x_{t-1}, z_t))$ in (29): either (i) it diverges to infinity, or (ii) it vanishes in the case where $F_{t-1}(x_{t-1}) = 1$ so that in the limit $Q(F_t; F_{t-1}(x_{t-1})) = z_t$. Given that $\gamma \geq 0$, the result follows.

likely to be in practice.

To investigate this we could have taken a set of panel data and manipulated some of the observations. However, there is always the danger that some results may be specific to the dataset chosen, and it would clearly be more illuminating to be able to examine systematically the sensitivity of the simulation results to changes in the characteristics of the underlying distribution. Given that our purpose is to examine the behaviour of practical tools, rather than to discuss case studies of particular examples of income mobility, it makes sense to use a "dataset" over which one has some control @@

We therefore carried out a simulation on an artificial distribution that has characteristics similar to actual data. Our baseline distribution was a bivariate lognormal with parameters that would be of the same order of magnitude as empirical estimates for the Michigan Panel Study of Income Dynamics:[11] this suggested simulated data where marginal distributions were given by $\Lambda(10.25, 0.5)$;[12] a number of values for the correlation coefficient on log-income were used in the experiment

There are two main types of contamination that may then be modelled within this bivariate framework. Type 1 is that of the "rogue profile": both components of the income profile (x_{t-1}, x_t) are simultaneously contaminated for particular

[11] The PSID income concept used was log annual, unequivalised, real, post-tax, post-benefit income in 1989.

[12] We also calculated results for larger values of the scale parameter, but the qualitative results remain intact.

	correlation=0.50				correlation=0.75			
contam:	*2.5%*	*5%*	*7.5%*	*10%*	*2.5%*	*5%*	*7.5%*	*10%*
"Stability" indices								
GE(-1)	0.9575	0.9344	0.9223	0.9133	0.9746	0.9615	0.9539	0.9488
	(0.0131)	(0.0117)	(0.0108)	(0.0096)	(0.0093)	(0.0081)	(0.0073)	(0.0068)
GE(0)	0.9328	0.9042	0.8908	0.8816	0.9614	0.9456	0.9377	0.9327
	(0.0123)	(0.0094)	(0.0076)	(0.0061)	(0.0079)	(0.0060)	(0.0047)	(0.0040)
GE(1)	0.9055	0.8811	0.8726	0.8677	0.9456	0.9326	0.9272	0.9245
	(0.0156)	(0.0109)	(0.0089)	(0.0075)	(0.0100)	(0.0073)	(0.0058)	(0.0052)
GE(2)	0.8954	0.8856	0.8846	0.8849	0.9374	0.9337	0.9315	0.9316
	(0.0350)	(0.0304)	(0.0295)	(0.0255)	(0.0255)	(0.0238)	(0.0205)	(0.0196)
Gini	0.9626	0.9437	0.9341	0.9275	0.9803	0.9706	0.9655	0.9622
	(0.0072)	(0.0054)	(0.0042)	(0.0033)	(0.0042)	(0.0031)	(0.0024)	(0.0020)
"Distance"-based indices								
King	1.2146	1.2361	1.2383	1.2386	1.3805	1.4718	1.4843	1.4870
	(0.0632)	(0.0178)	(0.0128)	(0.0104)	(0.1839)	(0.0853)	(0.0592)	(0.0514)
Hart	0.8019	0.6655	0.5811	0.5112	0.7982	0.6643	0.5765	0.5106
	(0.0552)	(0.0478)	(0.0425)	(0.0360)	(0.0642)	(0.0542)	(0.0457)	(0.0414)
Fields-Ok	1.0017	0.9995	1.0008	1.0001	1.0005	0.9999	0.9999	1.0002
	(0.0334)	(0.0329)	(0.0333)	(0.0331)	(0.0347)	(0.0347)	(0.0342)	(0.0349)

Table 1: **Bias in mobility indices resulting from type-1 contamination**

observations in the data-set. Type-2 contamination may be thought of as the "blip" problem: contamination may afflict individual components of the profile. The experiment reported in Table 1 models "decimal-point contamination"[13] of the first type in a sample of size 500 where the contaminated observations range from 2.5% to 10% of the sample.

shows the contaminated mobility estimate as a ratio of the true value (so an unbiased entry would have the value 1.0000). The figures in parentheses show the standard errors of the estimate. As the top part of the table shows the

[13]This means that a proportion of the observations are recorded as being 10 times larger (in our case) or smaller than they should be: it is one of several typical manual recording errors found in practice.

stability indices based on GE-measures or the Gini index can exhibit substantial downward bias (4 to 13 percent) if the correlation coefficient of the log-income process is low; if the correlation is higher, the bias is reduced (the bias worsens with a reduction in the lognormal dispersion parameter). The lower part of the table shows that the bias for two of the distance-related measures can be very large: the King index is biased upwards and the Hart index downwards. This phenomenon persists even where the underlying log-income correlation is high.

The Fields and Ok index appears to perform extremely well in this case, but in a "blip" experiment it performs as badly, or worse than, the King index - see Table 2. Inspection of (23) reveals why this is the case: simultaneous similarly-sized perturbations of x_{t-1} and x_t will effectively cancel each other out, a phenomenon that is absent from the "blip" model.

6 Transition matrices and related techniques

Income mobility is inherently a complex process, and the attempts at measuring mobility usually involve some attempt at simplifying the underlying model of the process; this *a priori* simplification then has consequences for the way in which sample data are to be handled. The simplifications usually involve discretisation of the process, in one or both of two aspects - in state space and in terms of time. The time discretisation is implicit in the discussion of Section 2.

Two-stage mobility indices involve discretisation of the state space. The tran-

	correlation=0.50				correlation=0.75			
contam:	*2.5%*	*5%*	*7.5%*	*10%*	*2.5%*	*5%*	*7.5%*	*10%*
"Stability" indices								
GE(-1)	0.9968	1.0090	1.0240	1.0402	1.0130	1.0402	1.0663	1.0929
	(0.0140)	(0.0136)	(0.0140)	(0.0137)	(0.0109)	(0.0118)	(0.0133)	(0.0130)
GE(0)	0.9919	0.9937	0.9978	1.0020	1.0155	1.0330	1.0463	1.0586
	(0.0124)	(0.0114)	(0.0104)	(0.0096)	(0.0087)	(0.0081)	(0.0079)	(0.0078)
GE(1)	1.0090	1.0063	1.0019	0.9960	1.0501	1.0640	1.0665	1.0662
	(0.0144)	(0.0150)	(0.0149)	(0.0146)	(0.0109)	(0.0123)	(0.0131)	(0.0131)
GE(2)	1.0795	1.0707	1.0543	1.0353	1.1592	1.1616	1.1468	1.1289
	(0.0195)	(0.0160)	(0.0170)	(0.0173)	(0.0228)	(0.0159)	(0.0172)	(0.0172)
Gini	0.9904	0.9833	0.9789	0.9745	1.0021	1.0037	1.0041	1.0045
	(0.0090)	(0.0091)	(0.0088)	(0.0082)	(0.0063)	(0.0069)	(0.0068)	(0.0070)
"Distance"-based indices								
King	1.0718	1.0593	1.0658	1.0584	1.2522	1.2503	1.2458	1.2287
	(0.1130)	(0.1114)	(0.1088)	(0.1070)	(0.1623)	(0.1500)	(0.1525)	(0.1534)
Hart	1.1048	1.1864	1.2426	1.2821	1.3138	1.5533	1.7123	1.8551
	(0.0716)	(0.0750)	(0.0780)	(0.0785)	(0.0965)	(0.1059)	(0.1161)	(0.1232)
Fields-Ok	1.0750	1.1534	1.2289	1.3073	1.1159	1.2382	1.3525	1.4772
	(0.0343)	(0.0348)	(0.0355)	(0.0352)	(0.0353)	(0.0351)	(0.0378)	(0.0375)

Table 2: **Bias in mobility indices resulting from type-2 contamination**

sition matrix approach is a standard example of the two-stage approach and permits discussion of a richer pattern of income mobility than can be embodied within a single class of stability or distance-based indices. It might be thought that, as with the distance-based single-stage measures, the two-stage approach makes sense only for cases where $T = 2$; but there is no reason *a priori* why this should be so.[14]

The essential components of the approach are as follows. One specifies a set of income classes (or "bins") into which observations from an empirical distribution are sorted

$$B_i(F) := [b_i(F), b_{i+1}(F)),\ i = 1, ..., \tau$$

such that $b_1 = \underline{x}$, $b_{\tau+1} = \bar{x}$. For simplicity we assume that the set of bins is the same for both periods. The *transition matrix* is $\mathbf{P}(F) := [p_{ij}(F)]$ where the *transition probabilities* $p_{ij}(F) := \Pr(x_t \in B_j(F) \mid x_{t-1} \in B_i(F))$ may then be expressed as [15]

$$\frac{F\left(b_{i+1}(F), b_{j+1}(F)\right) - F\left(b_i(F), b_j(F)\right)}{F\left(b_{i+1}(F), \bar{x}\right) - F\left(b_i(F), \underline{x}\right)}\ . \tag{31}$$

[14]One of the few authors who has attempted to deal with multiperiod generalisations of the two-stage concept is Hills (1998). The modification of the approach to continuous time is discussed in Geweke et al. (1986).

[15]The maximum likelihood estimator of (31) is: $p_{ij}\left(F^{(n)}\right) = \frac{n_{ij}}{n} / \sum_j \frac{n_{ij}}{n}$.

The mobility index is then expressed as a function - of the transition matrix[16]

$$M_{\text{trans}}(F; -) = - \left(\mathbf{P}(F)\right). \tag{34}$$

The issues that concern us here fall roughly into two groups: the general characteristics of the function - and the specification of the bins. This is easily seen if we evaluate the influence function for this general class of measures. If we assume that - in (34) is differentiable with bounded slope for all $p_{ij} \in [0, 1]$ then we have:

$$\begin{aligned} \mathrm{IF}(\mathbf{z}; M_{\text{trans}}, F) &= \sum_{i,j=1}^{\tau} -_{ij}(F)\mathrm{IF}(\mathbf{z}; p_{ij}, F) & (35) \\ &= \sum_{i,j=1}^{\tau} -_{ij}(F) \left.\frac{\partial}{\partial \varepsilon} p_{ij}\left(F_{\varepsilon}^{(\mathbf{z})}\right)\right|_{\varepsilon=0} & (36) \end{aligned}$$

where $-_{ij} := \partial - \left(\mathbf{P}(F)\right) / \partial p_{ij}$.

Exogenous bins. We need to focus upon the differential in (36). Letting $b_i(F) = b_i^*$ for all $i = 1, ..., \tau$ and assuming that $z_t \in B_i(F)$ and $z_{t-1} \in B_j(F)$,[17]

[16]Two commonly-used examples of - are the Prais index, defined as

$$M_{\text{trans}}(F; \mathrm{tr}) = \frac{n - \mathrm{tr}(\mathbf{P})}{n-1} = \frac{n - \sum_i \lambda_i}{n-1} \tag{32}$$

where $\mathrm{tr}(\mathbf{P})$ is the trace of the $n \times n$ transition matrix $\mathbf{P}$, and λ_j its jth ordered eigenvalue. The eigenvalue index is given by

$$\frac{n - \sum_i |\lambda_i|}{n-1} \tag{33}$$

which captures the speed of convergence of the underlying Markov process since all eigenvalues of the stochastic matrix are bounded by one. The eigenvalue index equals the Prais index if the eigenvalues of $\mathbf{P}$ are all real and non-negative.

[17]The other, easier cases can be derived immediately.

we have

$$p_{ij}\left(F_{\varepsilon}^{(\mathbf{z})}\right)=\frac{[1-\varepsilon]\left[F\left(b_{i+1}^{*},b_{j+1}^{*}\right)-F\left(b_{i}^{*},b_{j}^{*}\right)\right]+\varepsilon}{[1-\varepsilon]\left[F\left(b_{i+1}^{*},\bar{x}\right)-F\left(b_{i}^{*},\underline{x}\right)\right]+\varepsilon} \tag{37}$$

So we have

$$\left.\frac{\partial}{\partial\varepsilon}p_{ij}\left(F_{\varepsilon}^{(\mathbf{z})}\right)\right|_{\varepsilon=0} = \frac{1-F\left(b_{i+1}^{*},b_{j+1}^{*}\right)+F\left(b_{i}^{*},b_{j}^{*}\right)}{F\left(b_{i+1}^{*},\bar{x}\right)-F\left(b_{i}^{*},\underline{x}\right)} - \frac{\left[F\left(b_{i+1}^{*},b_{j+1}^{*}\right)-F\left(b_{i}^{*},b_{j}^{*}\right)\right]\left[1-F\left(b_{i+1}^{*},\bar{x}\right)+F\left(b_{i}^{*},\underline{x}\right)\right]}{\left[F\left(b_{i+1}^{*},\bar{x}\right)-F\left(b_{i}^{*},\underline{x}\right)\right]^{2}} \tag{38}$$

which is clearly bounded because it is independent of $\mathbf{z}$.

Endogenous bins. It is quite common to link the bin boundaries b_i to a proportion of some statistic of the distribution, for example to a proportion of the mean or to one of the quantiles. Clearly the expression (37) will now involve additional terms of the form $(1-\varepsilon)F\left(b_{i+1}(F_{\varepsilon}^{(\mathbf{z})}),b_{j+1}(F_{\varepsilon}^{(\mathbf{z})})\right)$. Differentiating this term with respect to ε in the neighbourhood of 0 gives

$$-F\left(b_{i+1},b_{j+1}\right)+\frac{\partial F\left(b_{i+1},b_{j+1}\right)}{\partial b_{i+1}(F)}\mathrm{IF}(\mathbf{z};b_{i+1},F)+\frac{\partial F\left(b_{i+1},b_{j+1}\right)}{\partial b_{j+1}(F)}\mathrm{IF}(\mathbf{z};b_{j+1},F)\ . \tag{39}$$

Thus, unless the bin boundaries are parametrised as robust statistics such as functions of quantiles, the transition probabilities estimator suffers from an unbounded influence function. However, the positive result is that transition matrices computed on the basis of deciles or other quantiles are indeed robust.

The robust choice of income classes then implies robust estimates of the transition probabilities. The choice of the mobility index from this class of indices is irrelevant from the view point of robustness, and should be guided by other considerations.

7 Concluding Remarks

We have seen that in the presence of data contamination commonly used "single-stage" mobility measures usually behave rather differently from appropriately designed two-stage models of mobility. Why do single-stage models go wrong? These measures are typically expressible in the form

$$M(F) = A\left(L_1(F), L_2(F), ...\right) \tag{40}$$

where

$$L_i(F) := \int E_i(\mathbf{x}) W_i\left(F(\mathbf{x})\right) dF(\mathbf{x}), i = 1, 2... , \tag{41}$$

$E_i : \mathfrak{X} \to \Re$ is an income evaluation function, $W_i : \mathfrak{F}_1 \to \Re$ is a linear or constant weighting functional. The form (40) typically exhibits the following characteristics:

- The linear functionals L_i of the distribution are defined over all of $\mathfrak{X}$.
- The integrand in (41) diverges to infinity for some $\mathbf{x} \in \mathfrak{X}$.
- A nonlinear aggregation function A.

The problem with single-stage indices comes partly from integrating over the whole domain, partly from the form of the sensitivity of the E or W that causes the integrand to diverge, and partly from the fact the components of the impact of contamination do not cancel because of nonlinearity of the aggregation function.

The two-stage approach deals with these things separately. In stage 1 we process information: a non-linear function filters out information from parts of the domain $\mathfrak{X}$; in particular extreme values may be filtered if the data "bins" are function of robust statistics of the distribution. In stage 2 the evaluation and weighting jobs performed by the functions E and W in (41) are achieved by appropriate specification of the function - in (34).

The analysis of robustness has an important role to play in the specification and selection of income-mobility indices. Unlike the case of inequality measures or Social-Welfare Functions there is not really a good *a priori* case for one mobility index rather than another or one class of indices rather than another. Instead, most commonly-used mobility measures are essentially pragmatic. Robustness properties can be one good guide to the choice of a pragmatic index.

References

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory 2*, 244–263.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1989). Measurement error in cross-sectional and longitudinal labor market surveys: Results from two validation studies. Working Paper 2884, NBER.

Bound, J. and A. B. Krueger (1989). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right ? Working Paper 2885, NBER.

Cowell, F. A. (1980). Generalized entropy and the measurement of distributional change. *European Economic Review 13*, 147–159.

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

Cowell, F. A. (1998). Measurement of inequality. In A. B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution*, Chapter 2. Amsterdam: North Holland.

Cowell, F. A. and M.-P. Victoria-Feser (1996). Robustness properties of inequality measures. *Econometrica 64*, 77–101.

Fields, G. S. and E. A. Ok (1997, January). A subgroup decomposable measure of relative income mobility. Working paper, Cornell University.

Geweke, J., R. C. Marshall, and G. A. Zarkin (1986). Mobility indices in continuous time Markov chains. *Econometrica 54*(6), 1407–1423.

Hampel, F. R. (1968). *Contribution to the Theory of Robust Estimation.* Ph. D. thesis, University of California, Berkeley.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association 69*, 383–393.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel (1986). *Robust Statistics: The Approach Based on Influence Functions.* New York: John Wiley.

Hills, J. R. (1998). Does income mobility mean that we do not have to worry about poverty? In A. B. Atkinson and J. R. Hills (Eds.), *Exclusion, Employment and Opportunity*, Chapter 3, pp. 31–55. LSE, Houghton Street, London: Centre for Analysis of Social Exclusion, STICERD.

Huber, P. J. (1986). *Robust Statistics.* New York: John Wiley.

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

Künsch, H. (1984). Infinitesimal robustness for autoregressive processes. *The Annals of Statistics 12*, 843–863.

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

Maasoumi, E. and S. Zandvakili (1990). Generalized entropy measures of mobility for different sexes and income levels. *Journal of Econometrics 43*, 121–133.

Monti, A. C. (1991). The study of the Gini concentration ratio by means of the influence function. *Statistica 51*, 561–577.

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

Shorrocks, A. F. (1993). On the Hart measure of income mobility. In M. Casson and J. Creedy (Eds.), *Industrial Concentration and Economic Inequality*. Edward Elgar.

Victoria-Feser, M.-P. (1998). The sampling properties of inequality indices: Comment. In J. Silber (Ed.), *Income Inequality Measurement: From Theory to Practice*. Kluwer, Dewenter.