ON MOBILITY

By
Esfandiar Maasoumi
Department of Economics, SMU, Dallas, TX, 75275-0496
EMail:maasoumi@mail.smu.edu

This version December 9, 1996

Keywords: Mobility, transition matrices, Markov models, equality, income distribution, PSID, stochastic dominance.

Abstract. There are seemingly two main welfare theoretic approaches to the measurement of mobility. One is based on transition matrices and their reduction to a scalar measure. The other is the so-called Maasoumi-Shorrocks-Zandvakili "inequality reducing" measures that relies on the well known welfare function analysis of inequality indices applied to long run incomes. Both are surveyed in this paper with descriptions of statistical methods for their empirical implementation. It is seen that the two approaches are converging to the same type of welfare orderings. Some popular mobility indices are analyzed in this light. Several empirical examples are briefly described which specially highlight the second method's ability to consider long run income inequality and produce mobility profiles.

Table of Contents.

Introduction

Mobility in any social hierarchy is an indication of opportunity. It is therefore a measure of fairness in any economic system. In a real sense, mobility should be of greater concern to policy makers and analysts than such other important concerns as inequality. It may be understandable and socially tolerable for the young, say, to have less assets or income, but the likelihood and the opportunity to move up and earn in relation to effort is surely a decidedly desirable social goal. What matters in this context is "lifetime equity" rather than instantaneous equality.

There are essentially two types of movements that constitute mobility. One is a "growth" or "structural" mobility which may arise from a general economic movement up, or down. The other is the "latteral type", or "exchange mobility" which obtains when individuals, households, or groups of such units move from one state to another. Since there are at least some temporal limits to such upward movements, mobility, particularly of the latteral type, must be "equalizing". Thus welfare comparison between any two mobility states necessarily requires functionals that reflect degrees of social preference for equity, exchange mobility, and growth (new opportunities).

There are two important questions that need attending to. One is how does one characterize economic status? The other is, how does one evaluate and compare different mobility situations? The latter question poses the issue of why is mobility socially desirable and, if it is, what kinds of Social Welfare Functions (SWFs) represent our preferences for mobility? The first question is all too often resolved by using "income" as a proxy. Such a proxy is evidently more meaningful the more market oriented an economy and a culture, and the less is the level of in-kind and other transfer payments and entitlements.

The reliance on SWFs for evaluating and comparing welfare situations has provided a scientifically useful tool that provides for economy of thought, as well as discipline, since it forces a declaration of principles that are too often implicit. In the instant case, the desirability of both types of mobility suggests functionals that are increasing in incomes (for the growth component), and *ultimately* equality preferring (for the exchange component). While changes in earnings and incomes have been and should be studied with the aim of identifying significant "causal" factors, the evaluation of an existing degree of mobility requires welfare comparisons.

There are currently at least two complementary lines of analyzing mobility. The older approach requires specification of transition probabilities between social states, and a welfare evaluation of the transition matrices that are estimated from the existing data. This clearly requires detailed data, with large panel data being the best source for sufficient cell repetitions which is necessary for reliable statistical inference. There are numerous mobility indices which are mappings from the transition matrices to scalars. Shorrocks (1978a) and Geweke et al (1986) provide systematic discussions of criteria for sensible mobility indices.

In line with a general direction toward unanimous partial ordering in the literature on inequality, useful welfare ranking relations have been developed for transition matrices which rekindle the essential role played by Lorenz and Generalized Lorenz criteria. While the SWF evaluation has begun to be emphasized in this approach, the task of devising consensus SWFs over matrices remains a challenge reminiscent of that faced in the multidimensional analysis of inequality. I provide an account of this line of inquiry in section 3.

The second approach was initiated by Shorrocks (1978b) and generalized by recasting mobility as a multidimensionality question as in Maasoumi (1986a), e.g., see Maasoumi and Zandvakili (1986, 1989, 1990). In this approach inequality indices are computed for multi-period incomes, that is, a type of "permanent income" measured over more than a single period, and mobility indices with a profile of equalization over time are obtained. The latter are directly related to and interpretable by the familiar classes of increasing and equality preferring (Schur concave) SWFs. It will be seen that the more recent welfare theoretic development of the transition matrices approach is converging to the same welfare comparisons, similar notions of long run or "permanent" income, and thus similar preferred mobility indices! Quite general mobility indices are proposed and empirically implemented by Maasoumi and Zandvakili (1989, 1990). This approach is also demanding of data and, like the first approach, it is ideally implemented with plentiful micro data. But the Maasoumi-Shorrocks-Zandvakili (MSZ) measures may be adequately estimated with data grouped by age, income, education, etc. This is useful since much data is made available in this aggregated form, and the approach is particularly focused on "equalization" between income groups in a way that allows controlling for sources of heterogeneity among individuals and households. Such controls with aggregate data appeal to some who may wish to tolerate some diversity due to such things as age or education. I present an account of this approach in section 2.

Our survey makes clear that both approaches share a concern for the distribution of a welfare attribute as well as its evolution. And the convergence in both approaches to rankings by Lorenz-type curves is of econometric significance. There is now a well developed asymptotic inference theory for empirically testing for order relations such as stochastic dominance, Lorenz dominance, and concentration curves. This type of testing is a crucial first step since comparing mobility indices (statistically or otherwise) is of questionable value when, for instance, Generalized Lorenz curves cross. Section 4 contains a brief account of the statistical tools that are available for inference on both the indices of mobility and order relations.

A loosely related strand of econometric research seeks to specify statistical models of earnings or income "mobility". While it is true that such models are focused on explaining "change" and "variation" which are not as meaningful as "mobility", they can shed light on significant *explanations* of earnings changes, as well as account for heterogeneity. This is consequential for policy analysis. Further, there are econometric models that seek to fit transition probabilities. Such studies are directly useful for not only estimating transition matrices, but for explaining the estimated probabilities. We do not delve into this empirically substantive area in this survey. Section 5 concludes with several empirical applications of the "inequality reducing" mobility measures. An insightful survey of income mobility concepts is given in Fields and Ok (1995). Absolute measurement of income mobility and partial ordering of absolute mobility states is treated in Mitra and Ok (1995).

Inequality reducing mobility indices

The MSZ indices

The most common approach to comparing welfare situations is based on indices. For example, two income vectors depicting the respective distribution of households at two points in time, or in different regions, are explicitly and implicitly ranked by scalar measures of such things as inequality, mobility, or poverty. The welfare theoretic underpinnings and limitations of index-based comparisons are now well understood. Notwithstanding these serious limitations, the need for "measurement" has led to the development of many "sensible" indices, supplemented with statistical tools which are based upon the asymptotic distributions of the indices; see Maasoumi (1996a) for a recent survey. This "index-based" approach is limited because of a lack of unanimity on the acceptable index even when there is broad consensus about certain normative principles. But the main difficulty is a consequence of the difficulties of welfare "comparability" as well as the appropriate cardinalization of the class of admissible criterion functions.

In the case of mobility indices, Shorrocks (1978b) and Maasoumi and Zandvakili (1986) argued that a more reliable measure of individual welfare and the distribution of incomes is to be obtained by considering the individual or household's "long run" income. They argued that such a measure of income computed over increasingly longer periods of time would remove transitory and some other life cycle related movements which are picked up by year to year comparisons of income distributions. The annual “snap shots” are incapable of accounting for mobility and returns to investments and/or human capital. The effects of seniority alone may make the notion of income inequality meaningless.

These authors are therefore concerned with dynamics of income distribution, and are thus accounting for, and challenged by, a fundamental lack of *homogeneity* amongst households. The natural labelling of individuals at different points in their life cycle is an essential form of heterogeneity which contradicts the common assumption of symmetry (anonymity) which plays a crucial role in much of the welfare theory that underlies inequality analysis. Shorrocks (1978b) proposed the simple sum of incomes over T periods as the aggregate income. Maasoumi and Zandvakili (1986, 89, 90) proposed more general measures of “permanent income” encompassing the simple average and sum. This author recognized the essential multidimensionality of the mobility analysis and proposed its treatment on the basis of the techniques and the concepts developed in Maasoumi (1986a, 1986b). Consequently, general functions for “permanent incomes” where developed which are Maximum Entropy (ME) aggregators.

The next step in this development is to analyze the inequality in the permanent incomes and compare with single period inequalities. A weighted average of the latter was used by these authors to represent “short run” inequality over any desired number of periods. Clearly, a distribution of permanent incomes is being compared and ranked with a reference short run income distribution. In view of this, all of the rich welfare theory supporting Lorenz and Generalized Lorenz (second order stochastic dominance) dominance relations, as well as the convex inequality measures consistent with such relations, comes at the disposal of the analyst for evaluating mobility profiles and dynamic evolution of the income distribution. At first sight, this appears an un-necessarily restrictive setting for defining the welfare value of mobility. Further reflection suggests that this is not so. Indeed, we will see that the recent development of a welfare theory basis for the alternative of transition matrices has made clear a certain inevitability for the role of the same welfare criteria and, therefore, the same types of mobility measures as the Maasoumi-Shorrocks-Zandvakili indices!

Let X_{it} be the income of the i-th individual in the t-th state (period, say) ; $i \in [1,n]$, and $t \in [1,T]$. Let $S_i(X_i;\alpha,\beta)$ denote the i-th individual’s *permanent* income (living standard?) over a number of periods $k = 1,2,....T$, and $S = (S_1,S_2,....,S_n)$ denote the vector of such incomes for the n households or individuals. The inequality measures which are consistent with a set of axioms to be described below are represented by $I_\gamma(.)$. Let X denote the welfare matrix with the typical element X_{it} and denote its i-th row by X_i, and its t-th column by X^t. The latter is the income vector in the t-th period/state.

The $k-$ *period* long run inequality is given by $I_\gamma(S)$, and short run inequality may be represented by $I_\gamma^k = \sum_{t=1}^k \alpha_t I_\gamma(X^t)$, for $k = 1,2,.....,T$. The vector $\alpha = (\alpha_1,\alpha_2,.....,\alpha_T)$ represents the weights given to the income distribution in different periods, such as the ratio of the mean income in the period, μ_j, and the overall mean of incomes in all the k periods under analysis.

Shorrocks (1978b) proposed $S_i = \sum_{t=1}^k X_{it}$, and $\frac{\mu_t}{\mu}$, the ratio of means just described, as weights, and Maasoumi and Zandvakili (1986) generalized S_i and the weight functions, suggesting the following index of mobility:

$$M_\gamma^k = 1 - \frac{I_\gamma(S)}{I_\gamma^k} \qquad \#$$

$$\text{with } R_\gamma^k = \frac{I_\gamma(S)}{I_\gamma^k}, \text{ as a measure of stability} \qquad \#$$

where it is to be noted that $S(.)$ is measured over the same k periods. When S is quasi-concave $R \in [0,1]$ follows from the convexity of the inequality measures considered. A “mobility profile” is generated by depicting R as k increases over its range. Both Shorrocks (1978b) and Maasoumi and Zandvakili (1986, 89,90) report empirical studies of the US data from the PSID. Section 5 describes some of these studies.

Bounds on R_γ^k may be established using the decompositions of multidimensional measures of inequality given in Maasoumi (1986, Propositions 1 and 2). $P_T = (R_1,R_2,...., R_T)$ is a stability profile which can also reveal the effect of increasing smoothing of the income variable starting from $R_1 = 1$. Other than this smoothing out of the short run effects, R_γ is capable of revealing durable “mobility“ toward equalization in a way that may be obscured by looking at each $I_\gamma(X^t)$, $t = 1,...,T$. This is most clearly seen by considering a situation in which only a permutation of the income vector has occurred between two periods. As is well known, this leaves our *anonymous* (symmetric), relative inequality measures unchanged, but the distribution of the *aggregated* incomes over the two periods will be changed, possibly dramatically (unless there is perfect equality to begin with!). Consequently, $I_\gamma(S)$ and R_γ will measure any mobility over the two periods.

Given that the indices $I_\gamma(.)$ are supported by Schur-concave Social Welfare Functions (SWFs), ordering mobility states by the class of measures in R is consistent with unanimous partial orderings conducted with generalized Lorenz curves and related orderings. For this reason, and to help in the discussion of section 3 below, it is useful to have a brief account of the relevant welfare theory.

The fundamental welfare axioms

Kolm (1969) and Atkinson (1970) provided the modern and influential formalizations of the relationship between SWFs and inequality measures. The need for establishing this important relationship is now recognized widely when analyzing poverty and mobility. However, the axiomatic SWF approach does not by itself identify unique indices even when a particular set of normative properties are consented to by a majority. To appreciate this point, consider the Atkinson family of inequality measures for income x having mean μ_x:

$$A_v = 1 - [\int_0^\infty x^{1-v} dF]^{\frac{1}{1-v}} / \mu_x, \ v > 0, v \neq 1 \qquad \#$$

$$= 1 - exp[\int \log(x/\mu_x) dF], \ v = 1 \qquad \#$$

where F is the *c.d.f* of income. Similarly, the Generalized Entropy (GE) family of indices is given by:

$$I_\gamma(X) = \frac{1}{\gamma(1+\gamma)} \int_0^\infty (x/\mu_x)[(x/\mu_x)^\gamma - 1] dF, \gamma \text{ real} \qquad \#$$

I_γ is ordinally equivalent to the coefficient of variation and the Herfindahl index, and includes the variance of logarithms and Theil's first and second measures, I_0 and I_{-1} , respectively. Also, up to a monotonic transformation, there is a unique member of GE corresponding to each member of the Atkinson family. $v = -\gamma$ is the degree of aversion to relative inequality; the higher its absolute value the greater is the sensitivity of the measure to inequality (transfers) in the tail areas of the distribution.

Employing the functional theory first developed in "information theory" for identifying appropriate measures of divergence between distributions, see Maasoumi (1993), and noting that inequality and many other indices are similar measures of divergence, one puts down an *explicit* set of normative properties (axioms) which inequality indices and/or SWFs must satisfy. Using these axioms as explicit constraints on the function space one then obtains the appropriate inequality index. To exemplify, let us follow Bourguignon (1979) or Shorrocks (1980, 1984) in their discussion of the "fundamental welfare axioms" of Symmetry (Anonymity), Continuity, Principle of Transfers, and additive decomposability. Combined with homogeneity, these axioms identify members of the GE as the desirable family of relative inequality measures.

Axiom 1. (Anonymity) The inequality index (function) is symmetric in incomes.

Axiom 2. (Homogeneity) Invariance to scalar multiplication.

Axiom 3. Continuity.

Axiom 4. Principle of transfers.

This requires that inequality decrease if we redistribute from a single richer individual to a poorer one, leaving their respective ranking and all the other individuals' incomes unchanged.

Axiom 2 is a serious limitation as it restricts attention to "relative" inequality. This is so since this requirement implies mean invariance,—doubling everyone's income would leave inequality unchanged.

The class of functionals satisfying Axioms 1-4 is large! Also, any further axioms are less likely to command unanimity. In fact, any further requirements must be justified by plausible considerations of such things as, heterogeneity amongst households and/or individuals, policy, empirical necessity, and practical interest. The most commonly invoked of such requirements is:

Axiom 5. Additive decomposability.

This requirement, later strengthened as an "aggregation consistency" axiom by Shorrocks (1984), says that total inequality must be the sum of a "between group" component, obtained over group **means**, and an additive component which is a weighted sum of "within group" inequalities. This kind of decomposability is very useful for controlling and dealing with heterogeneity of populations, and as a means of unambiguously identifying the sources of inequality and those that are affected by it. In the context of mobility analysis, this property further serves to identify the contributions of each time interval under consideration.

If the **additive** decomposability requirement of axiom 5 is imposed, such measures as Gini, the correlation coefficient, and variance of logarithms must be excluded. The latter measures provide ambiguous decompositions of overall inequality by population subgroups; see Shorrocks (1984). For the GE family, for instance, a discretized (estimation) formula that helps to demonstrate its decomposability is as follows:

$$I_\gamma = \sum_{r=1}^{R} [X_r / \sum_{j=1}^{n} x_j]^{\gamma+1} (n_r/n)^{-\gamma} I^r + I_\gamma^b \qquad \#$$

where I_γ^b is the between group inequality computed over the group mean incomes, I^r is the inequality within the r-th group, $r = 1,2,.....,R$, n_r is the number of units in group r, and X_r is total income of the $r - th$ group.

Aggregate or permanent income:

As in Maasoumi (1986a), we consider the following weighted generalized entropy measure of divergence between S, on the one hand, and X^t, $t = 1,2,....,T$, we have:

$$D_\beta(S,X;\alpha) = \sum_{t=1}^{T} \alpha_t \{ \sum_{i=1}^{n} S_i [(S_i/X_{it})^\beta - 1] / \beta(\beta+1) \} \qquad \#$$

where $\alpha_j s$ are the weights attached to each period. Minimizing D_β with respect to S_i such that $\sum S_i = 1$, produces the following “optimal” aggregate income functions:

$$S_i \propto (\sum_t^T \alpha_t X_{it}^{-\beta})^{-1/\beta}, \beta \neq 0, -1 \quad \#$$

$$S_i \propto \Pi_t^T X_{it}^{\alpha_t}\ , \beta = 0 \quad \#$$

$$S_i \propto \sum_t \alpha_t X_{it}, \beta = -1 \quad \#$$

These are, respectively, the hyperbolic, the generalized geometric, and the weighted means of the incomes over time. Noting that the “constant elasticity of substitution”-$\sigma = 1/(1+\beta)$, these functional solutions include many of the well known utility functions in economics, as well as some arbitrarily proposed aggregates in empirical applications. For instance, the weighted arithmatic mean subsumes the simple total income discussed earlier, and a popular “composite welfare indicator” based on the principal components of X, when $\alpha_t s$ are the elements of the first eigen vector of the $X^{\prime}X$ matrix; see Ram (1982) and Maasoumi (1989a). The “divergence measure” $D_\gamma(.)$ forces a choice of an aggregate income vector $S = (S_1, S_2, \ldots, S_n)$ with a distribution that is closest to the distributions of its constituent variables. This is desirable when the goal of our analysis is the assessment of income *distribution* and its dynamic evolution. The entropy principle establishes that any other S would be extra distorting of the objective information in the data matrix X. The distribution of the data reflect the outcome of all optimal allocative decisions of all agents in the economy; see Maasoumi (1986b).

The next step in constructing general mobility indices as proposed by Maasoumi and Zandvakili (1986) is the selection of a measure of inequality. The GE index described above was computed for the S_i functions just obtained. It is instructive to analyze this measure in the discrete case:

$$I_\gamma(S) = \sum_{i=1}^n p_i[(S_i^*/p_i)^{1+\gamma} - 1]/\gamma(1+\gamma), \gamma \neq 0, -1 \quad \#$$

$$I_0(S) = \sum S_i^* \log(S_i^*/p_i), \text{ Theil's first index}$$

$$I_{-1}(S) = \sum p_i \log(p_i/S_i^*), \text{ Theil's second index} \quad \#$$

where p_i is the i-th unit’s population share (typically $= \frac{1}{n}$), and S_i^* is S_i divided by the total $K = \sum_{j=1}^n S_j$.

These inequality indices are normalized iso-elastic transformations of the aggregate income functions $S_{i.}$ As such they are “symmetric”, “homogeneous”, and consistent with the Lorenz criterion. They will be homogeneous with respect to every X^t, the t-th column/period in X, if in all the above one works with the matrix of *shares*, $x = (x_{it})$. While this will not change the functional solutions given above, it requires a rather unusual assertion that individual well-being depends on relative incomes. Alternatively, one can impose a general form of scale invariance at the outset; see Maasoumi (1996b).

Useful decomposability properties are possessed by these measures both in population groups and in the time directions.

Theorem (*Decomposability of GE*) *Let* $x_{it} = X_{it}/T_t, T_t = \sum_i X_{it}$, $W_t = T_t/K$, $I_\gamma(X^t)$ = *the GE inequality in the t-th period, and* $\delta_t = \alpha_t/\sum_{f=1}^T \alpha_f$. *Then*:

(i). If $1+\gamma = -\beta$, we have:

$$I_\gamma(S) = \sum_{t=1}^T \delta_t W_t^{1+\gamma} I_\gamma(X^t) + (\sum_t \delta_t W_t^{1+\gamma} - 1)/\gamma(1+\gamma) \quad \#$$

(ii). If the marginal distributions are identical-i.e., $x^t = x^k$, $\forall t \text{ and } k$, we have,

$$I_\gamma(S) = I_\gamma(X^t), any\ t \in [1, T] \quad \#$$

(iii). For Theils first and second measures ($\gamma = 0, -1$), we have:

$$I_0(S) = \sum_{j=1}^m C_t I_0(X^t) - D_{-1}(x, S^*; C) \quad \#$$

where $C_t = \delta_t T_t / \sum_k \delta_k T_k$, and,

$$I_{-1}(S) = \sum_{t=1}^{T} \delta_t I_{-1}(X^t) - D_0(x,S^*;\delta) \qquad \#$$

where by application of L'Hospital's rule to $D_\beta(.)$ defined earlier, we have:

$$D_0(S^*,x;\delta) = \sum_t \delta_t [\sum_i S_i^* \log(S_i^*/x_{it})] \ \geq 0$$

$$\textit{and } S_i \text{ defined at } \beta = 0 \qquad \#$$

and,

$$D_{-1}(S^*,x;C) = \sum_t C_t [\sum_i x_{it} \log(x_{it}/S_i^*)] \ \geq 0$$

$$\textit{and } S_i \text{ defined at } \beta = -1. \qquad \#$$

Proof: See Maasoumi (1986a, Propositions 1-2).

In view of the non-negativity of the D(.) terms in part (iii) of this proposition, it is clear that multiperiod inequality is no more than the weighted average of inequalities in the single periods. This is due to the intertemporal "substitution" effects and a consequence of the convexity of the inequality measures.

Welfare Properties of R_g indices.

In the spirit of the "axiomatic approach" discussed above for inequality measures, Shorrocks (1992) lays down a set of desirable properties that any mobility index may satisfy. These properties are quite suggestive and also useful in any discussion including that of mobility indices based on transition matrices. The desirability of some of the properties may be evident while that of others may be less obvious or compelling. In fact some of the proposed properties are not consistent, so one or more has to be abandoned.

The domain of income structures X over which the measure is well defined is given as follows.

(**A1**) **Universal Domain** : Suppose the feasible set of n-units and T periods for positive incomes is denoted by:

$$\mathsf{X}_{nT} = \{X | \dim X = n \times T; X_{it} > 0\}.$$

A mobility measure should be well defined for all $X \in \mathsf{X}$, where :

$$\mathsf{X} = \cup_{n=2}^{\infty} \cup_{T=2}^{\infty} \mathsf{X}_{nT}$$

However, there are some types of income structures that are eliminated from this set. An example is situations in which, in every period t, all individuals receive the same income μ_t.

The Maasoumi-Shorrocks-Zandvakili (MSZ) indices satisfy this requirement. Other indices may not; see below.

(**A2**) **Continuity**: The degree of mobility varies continuously with the incomes in X.

The MSZ indices satisfy this property.

(**A3**) **Population Symmetry** : If $X^\vee = \Pi X$ for some permutation matrix Π, then X and $X^\vee$ are equally mobile. This requires "rank invariance" to be acceptable since X and $X^\vee$ may not give the same ranking of individuals.

The term "population symmetry" is used because in the analysis of mobility, unlike inequality, we can consider permutation of the time period distributions (the columns of X) as well as permutations of the individual income profiles (the rows of X). One may thus define time symmetry separately:

(**A3**$^\vee$) **Time symmetry**: A mobility measure is time symmetric if X and $X^\vee$ are equally mobile wherever $X = X\Pi$ for some permutation matrix Π.

Broadly speaking, time symmetry suggest that we care about the distribution of an individual's income receipts over time, but not about the time sequence of those receipts. This is more than time symmetry implies, however, since a permutation Π swaps all the incomes in two periods s and t, not just X_{is} with X_{it} for a single person i. But even the weaker idea of time symmetry, as stated, is objectionable since we may not be indifferent to the time sequence of incomes as, for instance, between the situation in which incomes are originally different and then become equal, and the time symmetric equivalent structure in which incomes are initially equal, and then become different.

The MSZ indices satisfy time symmetry, but this may or may not be a desirable quality.

(**A4**) **Population Replication Invariance**: X and $X^{\vee}$ are equally mobile whenever $X^{\vee}$ is a population replication of X.

(**A4**$^{\vee}$)**Time Replication Invariance**: A mobility measure is time replication invariant if X and $X^{\vee}$ are equally mobile whenever $X^{\vee}$ is a time replication of X.

$X^{\vee} \in \mathrm{X}_{rn,T}$ is a population replication of $X \in \mathrm{X}_{nT}$ if r is a positive integer and $X^{\vee}_{jt} = X_{it}$ wherever j = kn + i for some integer k ≥ 0. That is $X^{\vee}$ is the aggregate income structure for sub-populations each having the income structure X. Similarly, $X^{\vee} \in \mathrm{X}_{n,kT}$ is a time replication of $X \in \mathrm{X}_{nT}$ if k is a positive integer and $X^{\vee} = [X, X.....,X]$. Invariance with respect to population replication is the assumption typically used in inequality measurement to compare income distributions for different sized populations.

The MSZ index with additive aggregate income functions exhibit replication invariance with respect to time and the population. However, replication invariance with respect to time may in any case be regarded as a suspect property, for much the same reason as time symmetry. It implies that the degree of mobility is unchanged if the pattern of incomes received in the first T period, say, is exactly repeated for all individuals in the next T periods, and so on. "*But this does not take into account the fact that the distribution of income in period T may be radically different from that in period* 1, *so moving from period T to period* $T+1$ (*and hence back to period* 1 *incomes again*) *may be quite a jolt The desirability of time replication invariance is therefore less than transparent*"; Shorrocks (1992).

(**A5**) **Normalization**: Mobility is a minimum whenever X is completely immobile.

Definition: A structure X is completely immobile if and only if $X_{is}/\mu_s = X_{it}/\mu_t$ for all i,s, and t.

Completely immobile income structures perform a role similar to that played by completely equal distributions in inequality measurement. But it is implicit in the above definition that *relative* mobility is measured since one is only looking at changes in *relative* incomes. This rules out pure exchange mobility. It is a type of rank invariance property that goes beyond requiring that a structure in which all individuals incomes are constant over time is completely immobile.

The MSZ indices satisfy this property. They also satisfy the following stronger normalization property:

(**A6**) **Strong Normalization**: Mobility is a minimum if and only if X is completely immobile.

The other benchmark of interest is perfect mobility. Ideally we would define such an state as one in which the probability of achieving an income level in period $t+1$ is independent of the income received in period t. The concept of perfect mobility is difficult to formulate in terms of *observed income structures* X. A plausible parallel that falls short of "independence" is to require that not only are X_s and X_t uncorrelated, but also any, arbitrarily transformed vectors $\phi_s(X_s)$ and $\phi_t(X_t)$, where $\phi_t(X_t) \equiv (\phi_t(X_{1t}), \phi_t(X_{2t}), ..., \phi_t(X_{nT}))^{\vee}$ for some real valued function ϕ_t. Thus:

Definition: An income structure $X = [X_1, ..., X_T]$ is perfectly mobile if any only if $\phi_s(X_s)$ and $\phi_t(X_t)$ are uncorrelated for all s, t and all real functions ϕ_s, ϕ_t.

An example is given by Shorrocks (1992). Suppose there are J income levels $y_{11}, y_{21}, ..., y_{J1}$ at time 1 and K incomes levels $y_{12}, y_{22}, ..., y_{K2}$ at time 2, and let X consist of the JK income profiles (y_{j1}, y_{k2}) for $j = 1,...,J$ *and* $k = 1,...,K$. Now consider

$$\overline{\phi}_1 = \frac{1}{J}\sum_{j=1}^{J}\phi_1(y_{j1}), \text{ and } \overline{\phi}_2 = \frac{1}{K}\sum_{k=1}\phi_2(y_{k2}), \qquad \#$$

Then the sample covariance of the observed income profile can be shown to be zero.

The associated property is defined as follows:

(**A7**) **Perfect Mobility**: Mobility is a maximum whenever X is perfectly mobile.

Note that this requires that all perfectly mobile structures have the same index value as well as this common value being a maximum. Not all members of the MSZ family satisfy this requirement. See below for a member that partially satisfies (A7). But none of the members satisfy a stronger version of (A7) defined as follows:

(**A8**) **Strong Perfect Mobility**: Mobility is a maximum if and only if X is perfectly mobile.

(**A9**) **Unit Interval Range**: Mobility index should be in [0,1].

This is convenient and satisfied by MSZ and many other indices.

(**A10**) **Scale Invariance**: X and $X^{\vee}$ are equally mobile if $X^{\vee} = \lambda X$, for any scalar $\lambda > 0$.

And,

(**A11**) **Intertemporal Scale Invariance**: X and $X^{\vee}$ are equally mobile whenever $X^{\vee} = X\Lambda$, for any positive diagonal matrix Λ.

Shorrocks' index based on sum of incomes satisfies (A10) but not (A11). But a generalization to a weighted sum aggregate proposed in Maasoumi and Zandvakili (1986, 1990) does satisfy both (A10) and (A11).

How do mobility indices (welfare functions) rank intermediate situations between perfect mobility and total immobility? This requires a careful consideration of transfer sensitivity of mobility measures and has implications for the class of welfare functions. Smoothing transfers that are generally considered as equalizing (therefore preferred) conflict with normalization and change the mean of incomes at the point of transfer. It is difficult to consider them as mobility reducing. Compensating smoothing transfers, on the other hand, preserve the mean incomes (μ_t) but can change cross section distributions; see Shorrocks (1992). Thus the following type of "switches" are considered which can be seen to be mobility enhancing:

Definition: The income structure $X^{\vee}$ is obtained from X by a "simple switch" if, for some i, and j,

$$(X_{i1} - X_{j1})(X_{i2} - X_{j2}) > 0, \text{ and,} \qquad \#$$
$$X^{\vee}_{it} = X_{jt};\ X_{jt} = X_{it}$$
$$X^{\vee}_{ks} = X_{ks} \text{ for } s \neq t \text{ and all } k$$
$$X^{\vee}_{ks} = X_{ks} \text{ for all } s \text{ and all } k \neq i,j$$

Then :

(**A12**) **Atkinson-Bourguignon Condition** (for two periods): The income structure $X^{\vee}$ is more mobile than X whenever $X^{\vee}$ is obtained from X by a simple switch.

This condition implies that if income profiles of the two persons i and j are initially rank correlated, then a switch of incomes in either period enhances mobility. This condition has not been generalized to T≥ 2. The MSZ indices satisfy this condition when income aggregates are weighted averages.

Shorrocks (1992) studied a particular member of the MSZ family which satisfies all of the above properties except (A8), strong perfect mobility. This member is obtained from the measures defined by Maasoumi and Zandvakil (1986,1989) where aggregate income is $\sum_t \alpha_t X_{it}$, $\alpha_t = 1/\mu_t$, and the short run inequality is represented by $\sum \alpha_t \mu_t I_\gamma(X^t)$, and $I_\gamma(.)$ is the coefficient of variation, σ^2/μ. He refers to this as an "ideal" index. Shorrocks (1992) also looked at another mobility index attributed to Hart. The latter index is more conveniently described in relation to indices defined over transition matrices to which we now turn.

Transition Matrices

As was argued before, income distributions change over time under the effect of different transition mechanisms. Transition mechanisms affect social welfare by changing the income distribution. Two societies with the same income distribution at a point in time may have different levels of social welfare depending on the mobility of the populations. This requires welfare functions defined over an expanding time *dimension*.

In a Markov chain model of income generation, Dardanoni (1993) considers how economic mobility influences social welfare by following the approaches of Atkinson [2], Markandya [18], and Kanbur and Stiglitz [13]. He considers the welfare prospects of individuals in society by deriving the discounted stream of income distributions which obtain under different mobility structures. He proposes a class of SWFs *over the aggregates of these welfare prospects*, and derives some necessary and sufficient conditions for unambiguous welfare rankings. Since these aggregates are the discounted stream of incomes, a special case of the aggregates proposed by this author and described in the previous section, the two approaches of this section and the previous one converge when the same welfare functions and the same measures of "permanent" income are used.

The fundamental inequality theorem states that the Lorenz curve gives the normatively significant ordering of *equal mean* income distributions. Inequality indices are difficult to interpret when Lorenz curves cross. In a similar vein, Dardanoni (1993) derives a partial order of social mobility matrices which can be considered as the natural extension of the Lorenz ordering to mobility measurement. The derived ordering may provide conditions for an unambiguous welfare recommendation without employing a specific mobility measure.

Summary mobility measures induce a complete order on the set of mobility matrices and have the advantage of providing intuitive measurements and firm rankings. However, it is clear that there are substantial problems in trying to reduce a matrix of transition probabilities into a single number. This is very much the problem of multidimensional inequality measurement addressed by Maasoumi (1986a), Ebert (1995), and Shorrocks (1995). Dardanoni (1993) offers the following example of three mobility matrices

$$P_{1}=\begin{bmatrix} .6 & .35 & .05 \\ .35 & .4 & .25 \\ .05 & .25 & .7 \end{bmatrix}; P_2=\begin{bmatrix} .6 & .3 & .1 \\ .3 & .5 & .2 \\ .1 & .2 & .7 \end{bmatrix}; P_3=\begin{bmatrix} .6 & .4 & 0 \\ .3 & .4 & .3 \\ .1 & .2 & .7 \end{bmatrix}$$

The rows denote current state and columns denote future state. Suppose we use some common summary immobility measures as proposed and discussed by, for instance, Bartholomew (1982) and Conlisk (1990). Consider the second largest eigenvalue modulus, the trace, the determinant, the mean first passage time, and Bartholomew's measure. These indices are defined below. Any of the three matrices may be considered the most mobile depending on which immobility index is chosen. This is illustrated in the following table, which shows the most mobile mobility matrix according to the different indices.

Indices:	Eigenvalue	Trace	Determinant	mean first passage	Bartholomew
most mobile	P_2	P_1, P_3	P_1	P_3	P_1, P_2, P_3

Different mobility rankings obtain depending on the mobility measure adopted. The welfare-based partial order is similar to the axiomatic welfare analysis exemplified above as it aims to clarify the situation depicted in this example in a sound fashion.

The welfare ranking of mobility matrices

Consider a discrete Markov chain process for income. Let there be n income states. Let $P = [p_{ij}]$, such that $p_{ij} \geq 0$ and $\sum_c p_{ij} = 1$, be the $(n \times n)$ transition matrix, assumed regular (i.e., P^k is strictly positive for sufficiently large integer k), so that the strictly positive steady state probability vector π exists and is the unique solution to $\pi' = \pi' P$. The element p_{ij} is the probability that an individual in state i will be in state j in the following period. $\pi_{t+1} = \pi_t P$ denotes the vector whose i−th element gives the fraction of the population which is in state i at time $t+1$. It is assumed that transitions are independent across individuals and P is constant over time. Strictly speaking P need not be square. Restricting attention to this case exploits the properties of bistochastic matrices. Lastly, income states are in ascending order.

For a given transition matrix, P, we may derive the implied distribution of expected lifetime welfare for the individuals who live in the society whose mobility is governed by P. Consider a society, assumed in equilibrium, consisting of identical individuals who are born simultaneously and live exactly for τ periods. The transition mechanism may be either intra-generational or intergenerational.

Let $u = (u_1, u_2, \ldots, u_n)'$ denote a vector of instantaneous utilities, where u_i denotes the utility value of income in state i, and $V^p = (V_1, V_2, \ldots, V_n)'$ denotes a vector of expected discounted lifetime utilities. The typical element V_i^p denotes the expected lifetime discounted utility of an individual beginning life in income class i, given by the ith element of the vector $V^p = u + \rho P u + \rho^2 P^2 u + \cdots + \rho^\tau P^\tau u$, where $0 \leq \rho < 1$ denotes the discount factor. This typical element is comparable with the S_i functions of the previous section. V^p will in general depend on the vector u, on the transition matrix P, on the discount factor ρ, and on the time period τ. As $\tau \to \infty$, $V^P = [I - \rho P]^{-1} u$ which, for convenience, may be normalized as

$$V^P = (1-\rho)[I-\rho P]^{-1} u = P(\rho) u, \qquad \#$$

say. The typical element of $P(\rho)$, $P_{ij}(p)$,may be interpreted as the average discounted "lifetime" probability of moving from the initial state i to state j.

Dardanoni suggests that transition matrices be ranked according to real-valued SWFs defined over the vector of lifetime expected utilities V^p. This is comparable with the "inequality reducing" rankings of the previous section where Schur-concave welfare functions correspond to inequality measures. There is more a priori structure imposed on the data here by assuming a Markov transition process.

Mobility can occur through general growth in equilibrium income distributions. This is known as "structural mobility", or the "growth component". But there is some intertemporal movement of individuals among the different social classes, for a given equilibrium distribution of the number of individuals in each class; this latter effect is defined as "exchange," or "pure" mobility. Dardanoni (1993) isolates the pure mobility effect by assuming societies with identical steady-state income distributions. In other words two societies have within each period an identical spot income distribution, but individuals may move between income states differently under the two transition mechanisms. Note that under the stated assumptions the distribution of individuals in each state will be given by the equilibrium vector π, with the typical element π_i indicating the proportion of people in income state i. This procedure is the dynamic counterpart to the usual static inequality analysis (e.g., Atkinson [1] and Dasgupta, Sen, and Starrett [7]), where to isolate the pure inequality effect on social welfare one considers societies with equal mean incomes. Lorenz rankings would suffice in that situation.

Allowing for different mean incomes requires consideration of growth. This would be similar to the analysis of inequality on the basis of Generalized Lorenz (GL) curves, proposed by Shorrocks (1983), and second order stochastic dominance. This is the extension considered by Formby, Smith, and Zheng (1995) in the context of mobility matrices.

Following Dardanoni (1993), take two regular transition matrices P and Q with equal steady-state income distribution vector $\pi' = \pi' P = \pi' Q$. He considers the class of symmetric and additively separable (i.e., linear) SWFs $\sum_i \pi_i V_i^P$ which adds up, for a given u and ρ, the expected lifetime utility of the individuals in the population. This is equivalent to the SWFs considered by Atkinson(1970) and Kolm (1969) for the inequality ranking of income distributions. Noting that $\pi' P(\rho) = \pi' Q(\rho)$, we have

$$\sum_i \pi_i V_i^p = \pi' V^p = \pi' P(\rho) u = \pi' u = \pi' Q(\rho) u = \pi' V^Q = \sum_i \pi_i V_{i,}^Q \qquad \#$$

Therefore, given a vector u and a ρ, any two transition matrices with equal steady-state income distributions will be indifferent. This result is given by Atkinson (1983) and Kanbur and Stiglitz (1986) and indicates that we are not ranking mobility as such, but the social welfare implications of each mobility matrix. The symmetric additive social welfare functional implies that movement between income states is irrelevant. What is important is the spot distribution at each period since additive separable lifetime welfares remove any influence that exchange mobility may have on intertemporal social welfare. Thus additive SWFs take inadequate account of fairness considerations. Under the stated assumptions, the equilibrium Lorenz curve of the distribution of income will look identical each period under any transition matrix with equal steady-state distribution, so that any (additive or otherwise) symmetric ex-post SWF defined on the vector of realized utilities will rank the matrices as indifferent. Yet, under different transition matrices the composition of people in each income state will be different in each time period. For example, under the identity transition matrix each individual in the population remains in the same income group as in the initial situation; on the other hand, if transition is governed by a matrix in which each entry is equal to $1/n$, each individual will have the same probability of belonging to any of the n income groups regardless of the initial state. Therefore, though the equilibrium ex-postLorenz curves associated with each of these matrices could look identical for each period, social welfare may well be considered different if we take account of several periods in terms of the position of each individual in the past.

Clearly the natural labelling of welfare units in the context of mobility requires a relaxation of the "symmetry" assumption, such as (A3) above, which are replaced by additional assumptions on "comparability". These assumptions are discussed in, for example, Sen (1971) and Atkinson and Bourguignon (1987). Here the natural "label" for each individual is his/her starting position in the income ranking. Thus one restricted SWF would be the weighted sum of the expected welfares of the individuals, *with greater weights to the individuals who start with a lower position in the society*. That is, $W(V^p, \lambda) = \sum_i \pi_i \lambda_i V_{i,}^P$ where $\lambda = (\lambda_1 \lambda_{2,\ldots,} \lambda_n)'$ denotes a non increasing nonnegative vector of weights. This is a step toward cardinalization; see Maasoumi (1996b). Furthermore, this asymmetric treatment makes sense only if it is a disadvantage to start at a lower position. With no restriction on the mobility matrices, this is not necessarily a disadvantage. There could be a transition matrix such that the lower states are the preferred starting point in terms of lifetime expected utility. Therefore the additional assumption:

Assumption: Transition matrices are *monotone*.

A transition matrix is called monotone if each row stochastically dominates the row above it; see Conlisk (1990). In an intergenerational mobility context, a monotone mobility matrix implies that each child at time t is better off, in terms of stochastic dominance, by having a parent from state $i+1$ than a parent from state i. In an intragenerational mobility context, a monotone mobility matrix implies that an individual who at time t is in state $i+1$ faces a better lottery, in terms of stochastic dominance, than an individual who is in state i. If we let y be a $(n \times 1)$ vector it may be shown that Py is nondecreasing for all nondecreasing y if and only if P is monotone. Since $P(\rho)$ is monotone when P is monotone, the expected lifetime utility vector will be nondecreasing. Estimated transition matrices are often either exactly monotone or within sampling errors from being monotone.

Considering two extreme cases, $\lambda_1 = 1$ and $\lambda_i = 0$ for all $i > 1$, which is "Rawlsian", and $\lambda_i = 1$ for all i, which is the symmetric case, it is seen that there is a need for exploring necessary and sufficient conditions on transition matrices for the unanimous ranking of $W(V^P, \lambda) = \sum_i \pi_i \lambda_i V_i^P$ for *all* non increasing positive λ.

Theorem *Let P and Q be two monotone regular transition matrices such that $\pi' = \pi' P = \pi' Q$, and for a given ρ, the following conditions are equivalent*:

1.*i*. $\quad W(V^P, \lambda) \geq W(V^Q, \lambda),$

1.*ii*. $\quad T' \Pi [(P(p) - Q(p)] T \leq 0,$

Proof: See Dardanoni (1993, Th. 1).

T is the upper triangular summation matrix with its inverse, T^{-1}, as the first "differencing" matrix. For instance, PT transforms each row of P to the cumulative distribution function. Π is a block-diagonal matrix with the typical block being the π vector.

This result is an extended horizon version of the first order stochastic dominance relations obtained by Atkinson (1983).

Further, if we denote by $M(\pi)$ the set of regular monotone transition matrices, condition 1.ii induces (iff) a partial ordering $*_M$, that is reflexive, antisymmetric, and transitive.

If one further assumes monotonicity of the reverse chain-i.e., at each time t an individual in state i has faced a stochastically dominant lottery than an individual in state i-1, the above result would hold for *all* $0 \leq \rho < 1$.

An interesting result concerns the effect of transfers such as smoothing and "simple switches" considered in an earlier section. Dardanoni considers the following "Dynamic Pigou-Dalton" (DPD) transfers: Given integers 0 < i, j, s, k < n, with (i+k) <n, and (j+s) <n , let us decrease the probabilities of the event, "initial state i/lifetime state j", and the event, "initial state (i+k)/lifetime state (j+s)", by a quantity $0 \leq h \leq 1$. Simultaneously, increase by the same h the probabilities of the events, "initial class i/lifetime class (j+s)", and the "initial class (i+k)/lifetime class j", in such a manner as to not violate monotonicity. This transfer is mobility enhancing as it would leave the row and column sums unchanged and improve the lifetime status of a poorer individual. Finally noting that for a more mobile situation there will be smaller covariance between the initial and the lifetime status, the following general result is established:

Theorem *Let P and Q be two transition matrices in* $M(\pi)$ *and let* ρ *be given. Then the following conditions are equivalent*:

(*i*) $P(\rho) *_M Q(\rho)$;
(ii) the Lorenz curve of permanent income for P lies nowhere below that of Q $\forall$ nondecreasing income vectors.
(iii) the covariance between initial status and lifetime status is greater under Q for any nondecreasing score (rank) vectors.
(iv) $P(.)$ can be derived from $Q(.)$ by a finite sequence of DPD exchanges.

Formby et al (1995) extend this result by relaxing the assumption of identical steady states. They note that $T^{\vee}\Pi P(\rho)y$ is the Generalized Lorenz vector of "permanent incomes", and show the following result:

Theorem *Let P and Q be two monotone transition matrices with a given discount factor* ρ. *Then the following conditions are equivalent*:

1.i. $W(V^P,\lambda) \geq W(V^Q,\lambda)$;
1.ii. $T^{\vee}\Pi[(P(p) - Q(p)]T \leq 0$;
1.iii. The Generalized Lorenz curve of permanent income for P lies nowhere below that for Q for all nondecreasing income vectors y ;
1.iv. $P(p)$ can be derived from $Q(p)$ by a finite sequence of DPD exchanges and simple increments.

Proof: By noting that the assumption of steady-state income distribution is not crucial in proving Dardanoni's results as well as in Proposition 1 above, Formby et al (1995) prove the equivalence among conditions 1.i, 1.ii, and 1.iii. Also, 1.iv implies 1.ii. The converse can be similarly proved. Note that each DPD or simple increment leaves all elements other than the (i, j) th of $T^{\vee}\Pi[(P(p) - Q(p)]T$ unchanged.

Formby et al (1995) demonstrate a further result which can be useful in empirical testing based on the generalized concentration curves:

Theorem *Let P and Q be two monotone transition matrices. Then the following conditions are equivalent*:

1.i. $P(p)$ É$_M$ $Q(p)$;
1.ii. $T^{\vee}\Pi Py \leq T^{\vee}\Pi Qy$ for all nondecreasing income vectors y, i.e., the "snap-shot" generalized concentration curve for P lies nowhere below that for Q for all nondecreasing income vectors, if the condition of the monotonicity of the reverse Markov chain is further assumed.

This result establishes that if the generalized concentration curve for P lies nowhere below that for Q, then the generalized Lorenz curve of the "lifetime" for P also lies above that for Q. If one adopts Dardanoni's assumption of identical steady states (no growth mobility), the above result reduces to the ordinary concentration curve dominance which can be used as an intuitive tool in ranking "pure" mobility.

This result is empirically significant as it suggests the concentration curve as the dynamic counterpart of the standard Lorenz curve in terms of social welfare. If the generalized concentration curve of P dominates that of Q, the social welfare of *permanent income* under P will be no less than under Q for all *Schur* -concave symmetric SWFs.

Two main points emerge. One is that, both the inequality reducing and transition matrix measures are supported by the same type of welfare functions in terms of "lifetime incomes". Secondly, the empirical techniques for testing Lorenz type dominance, or stochastic dominance, are all made available. For instance, Bishop, Chow and Formby (1994) show that matched pairs of estimates of generalized Lorenz and concentration curve ordinates have a joint normal distribution and this sampling property is distribution-free. In general they have to be applied to the same statistics which are, however, defined over the types of aggregate income functions proposed by Maasoumi and Zandvakili (1986, 1989, 1990). Further description of some statistical techniques is given in the next section.

Some mobility indices

We close this section with a definition of some measures of immobility based on transition matrices. The first is attributed to S. Prais and is called the trace index:

$$Trace = [trace(P) - 1]/(n-1) \qquad \#$$

This index has been criticized for ignoring the off-diagonal transition probabilities. It also violates transfer properties such as A(12) above, for instance when DPD transfers increase any element along the diagonal of P.

$$\text{Determinant} = |P|^{1/(n-1)} \qquad \#$$

This is an index of immobility such as 1-R_γ. Unfortunately, this index obtains its perfectly mobile limit whenever *any* two rows or columns of P are identical! It also violates the ranking condition of the above theorems.

$$\text{Second Largest Eigenvalue modulus } |\lambda_2| \qquad \#$$

This measure of immobility is in the unit interval (A(9)) and measures the speed of escape from the initial state; see Theil (1972). But it too is incoherent with the ranking conditions for DPD type exchanges.

$$Bartholomew's = \frac{1}{n-1}\sum_i \sum_j \pi_i p_{ij}|i-j| \qquad \#$$

This immobility index measures the expected number of crossings between periods in the steady state. It satisfies the normalization and the unit interval conditions. This measure is also coherent with DPD rankings.

$$\textit{The} \text{ mean first passage time} = \pi' M^P \pi \qquad \#$$

where M^P is the mean first passage matrix. This index measures the expected number of periods before the "first" individual reaches the state of the second individual. This measure of immobility is not coherent with DPD rankings.

As we have seen, all of the MSZ measures in the previous section, or indeed any reasonable measure corresponding to Schur-concave SWFs, and based on "permanent income" is coherent with DPD rankings and the results of the above theorems. Shorrocks (1992) considered an interesting measure due to Hart (1976a). Let the Galtonian model of income evolution be written in terms of the geometric mean incomes, m_t at time t, as follows:

$$\ln(x_{t+1}/m_{t+1}) = \beta_t \ln(x_t/m_t) + \varepsilon_t \qquad \#$$

This model can be used to analyze both income movements over time as well as the effect of mobility on the distribution of income. β_t measures the extent to which incomes regress toward the geometric mean. The case of a "unit root" corresponds to Gibrat's law of proportional effect: changes in relative incomes are independent of current income. This simple model of mobility may be extended by further modelling the ε_t in terms of individual specific characteristics and/or time varying effects. An example is Lillard and Willis (1978) where panel data are used. Of course, using the techniques of limited dependent variable models, transition probabilities can be similarly modelled in terms of individual specific and time varying components. A survey of several applications is given in Creedy (1985). Alternative models of diffusion describing the evolution of income have been proposed which derive its steady state distribution forms. An example is Sargan (1957). Interestingly, the focus seems to have shifted to analyzing the properties of the equilibrium distribution and instantaneous inequality and poverty, rather than the dynamic welfare implications of the evolution mechanism. Mobility analysis is thus a return to first principles.

Shorrocks defined the following function of Hart's mobility index, see Hart (1976a, 1976b, 1981, 1983), which is derived from the Galtonian model above:

$$H(x_t, x_{t+1}) = 1 - \frac{cov(\ln x_t, \ln x_{t+1})}{\sigma(\ln x_t)\sigma(\ln x_{t+1})} \qquad \#$$

This is related to the simple correlation coefficient. In the Galtonian model, $\beta_t < 1$ will reduce inequality while $\sigma(\varepsilon_t)$ would increase it. Shorrocks (1992) shows that H(.) is consistent with A(2)-A(5), A(10)-A(12), and partially with A(7) and A(9). It fails the universal domain, strong normalization, and strong perfect mobility properties. Also, it satisfies time symmetry but not necessarily time replication invariance. It is coherent with DPD type rankings. As noted above, the MSZ indices are consistent with almost all of (A1)-(A12) properties and thus are superior to H(.).

Recently, Chakravarty (1995) proposed the Kullback's minimum discrimination statistic as a measure of mobility. This is given by:

$$K = \sum_i \pi_i^t \, ln(\frac{\pi_i^t}{\pi_i^{t+1}}) \qquad \#$$

where π_i^t stands for proportion of population at time t in state i.

This is a natural entropy measure of divergence between two distributions. Indeed it may be extended so that we may measure the divergence between the distribution of the "permanent income" and the distribution of income at any desired point. Examples of the latter are, the "short run" income distribution of MSZ indices, perfect mobility distribution, and complete immobility distributions.

The first property to note about K is that it is not a metric as it violates the triangularity rule. There are other entropy measures that have similar properties and are "metric"; See Maasoumi (1993). Chakravarty (1995) notes that K satisfies many of the useful properties discussed in Shorrocks (1978a) and summarized above. But it fails to satisfy the "monotonicity" property. We note that the ordering relation discussed by Dardanoni (1993) is coherent with perfect mobility, as is K, but does not imply monotonicity.

Chakravarty (1995) points out that the well-known asymptotic χ^2 distribution of $2K$ may be used to test some very interesting hypotheses about mobility.

Shorrocks (1976) gives a good account of the properties of mobility indices based on transition matrices. Dardanoni's (1993) account of the same also leaves one with the conclusion that at least some members of the MSZ family are "ideal".

Statistical inference

There have been several significant advances in the development of statistical inference tools in the area of income inequality. These are generally applicable to inference on mobility indices and on ranking distributions.

For mobility indices such as the MSZ the connection is rather immediate. Inequality indices are estimated by the Method of Moments (MM) estimators since they are functions of population moments. Explicit formulae are derived for derivatives that are required in the delta method which extends the well known theory of MM asymptotic distributions to that of inequality indices. This is surveyed in Maasoumi (1996a) which contains an extensive citation to original sources. The extension to mobility indices requires thinking in terms of "long run" incomes and the inequality in their distributions. Trede (1995) gives the details for the asymptotic distribution of some of the MSZ measures, such as those based on the Atkinson family and Theil's inequality indices, but where aggregate income is the simple sum of incomes analyzed by Shorrocks (1978b). Extension to weighted sum function is immediate, but the statistical theory for the more complicated aggregate functions is developed in Maasoumi and Trede (1997). Trede (1995) also gives the asymptotic distributions of the mobility indices that are based on transition matrices. Some of these measures were discussed earlier. An application to German data is reported in Trede (1995) which analyzes earnings mobility for different sexes. A program written in GAUSS code is made available by Trede.

But inference about indices derived in either of the two approaches described in this paper may be inconclusive, if not bewildering, when Lorenz-type curves cross. Therefore it is desirable to test for rank relations of the type described above and in Maasoumi (1996a, 1996b). This type of testing is now possible and is inspired by testing for inequality restrictions in econometrics and statistics. A brief account of some of the available techniques for stochastic dominance would be helpful and follows.

Tests for Stochatic Dominance

Stochastic Dominance(SD) relations and comparisons of distributions on the basis of their Lorenz and Generalized Lorenz (GL) curves are intended to avoid the "index choice" problem. As we have seen, the SWF rankings of mobility structures can essentially follow the same path as that of static inequality analysis but in terms of lifetime incomes. In practice, however, numerical SD rankings often encounter a predictable difficulty since many distributions and (Lorenz) curves cross, making it impossible to be decisive. But the realization that all such comparisons are based on sample based estimates of distribution functions (or curves) suggests that such comparisons should be conducted statistically and tested accordingly. The statistical approach is both sound and able to deliver more clear-cut statistical decisions!

The basic characteristic of tests for rankings is that of ordered populations and inequality restrictions. Starting with the work of Lehmann (1959), and Bartholomew (1959), likelihood ratio and Wald-type tests have been and are being developed for such hypotheses. These tests supplement other well known procedures based on one-sided Wilcoxan rank, and the multivariate versions of the Kolmogorov-Smirnov tests. See Maasoumi (1996a) for a recent selective survey.

In the area of income distributions and tax analysis, initial developments focused on tests for Lorenz curve comparisons as in Beach and Davidson (1983), Bishop, Formby, and Thistle (1989), and Bishop et al (1991) . In practice, a finite number of ordinates of the desired curves or functions are compared. These ordinates are typically represented by quantiles and/or conditional interval means. Thus, the distribution theory of the proposed tests are typically derived from the existing asymptotic theory for ordered statistics and quantiles. Recently Beach, Davidson, and Slotve (1995) have outlined the asymptotic distribution theory for cumulative/conditional means and variances which are useful for statistically comparing Lorenz and GL curves. This theory is particularly useful for third order stochastic dominance (TSD) ranking of crossing GL curves when a "transfer sensitivity" condition is assumed; see the definition of TSD below. To control for the size of a sequence of tests at several points the Union Intersection (UI) test and Studentized Maximum Modulus technique for multiple comparisons is generally favored in this area.

Some alternatives to these multiple comparison techniques have been suggested which are typically based on Wald type joint tests of equality of the same ordinates, see Bishop et al (1994) and Anderson (1994). These alternatives are sometimes problematic when their implicit null and alternative hypotheses are not a satisfactory representation of the **inequality** (order) relations that need to be tested. Xu et al (1995), and Xu (1995) take proper account of the inequality nature of such hypotheses and adapt econometric tests for inequality restrictions to testing for FSD and SSD, and to GL dominance, respectively. Their tests follow the work in econometrics of Gourieroux et al (1982) Kodde and Palm (1986), and Wolak (1988, 1989), which complements the work in statistics exemplified by Kudô (1963), Perlman (1969), Robertson and Wright (1981), and Shapiro (1985). The asymptotic distributions of these $\chi - bar$ squared tests are mixtures of chi-squared variates with probability weights which are generally difficult to compute. This leads to bounds tests involving inconclusive regions and conservative inferences. In addition, the computation of the χ *bar* squared statistic requires Monte Carlo or Bootstrap estimates of covariance matrices, as well as inequality restricted estimation which requires optimization with quadratic linear programming.

In contrast, Maasoumi et al (1996) propose a direct bootstrap approach that bypasses many of these complexities while making less restrictive assumptions about the underlying processes. They offer an empirical application for ranking US income distributions from the CPS and the PSID data. Their chosen statistic is the Kolmogorov-Smirnov (KS) as characterized by McFadden (1989), Klecan et al (1991), and Kaur et al (1994).

McFadden (1989) and Klecan, McFadden, and McFadden (1991) have proposed tests of first and second order "maximality" for stochastic dominance which are extensions of the Kolmogorov-Smirnov statistic. McFadden (1989) assumes i.i.d. observations and independent variates, allowing him to derive the asymptotic distribution of his test, in general, and its exact distribution in some cases (see Durbin (1973, 1985). Klecan et al generalize this earlier test by allowing for weak dependence in the processes both across variables and observations. They demonstrate with an application for ranking investment portfolios. The asymptotic distribution of these tests cannot be fully characterized, however, prompting Monte Carlo and bootstrap methods for evaluating critical levels. In the following section some definitions and results are summarized which help to describe these tests.

Definitions and Tests

Let X and Y be two income variables at either two different points in time, or two lifetime income vectors. Let $X_1, X_2, ..., X_n$ be the not necessarily i.i.d observations on *X*, and $Y_1, Y_2, ..., Y_m$ be similar observations on *Y*. Let U_1 denote the class of all utility functions u such that $u' \geq 0$, (increasing). Also, let U_2 denote the class of all utility functions in U_1 for which $u'' \leq 0$ (strict concavity).Let $X_{(i)}$ and $Y_{(i)}$ denote the i-th order statistics, and assume F(x) and G(x) are continuous and monotonic cumulative distribution functions (cdf,s) of X and Y, respectively.

Quantiles $q_x(p)$and $q_y(p)$ are implicitly defined by, for example, $F[X \leq q_x(p)] = p$.

Definition (**FSD**): X First Order Stochastic Dominates Y, denoted X FSD Y, if and only if any one of the following equivalent conditions holds:

(1) $E[u(X)] \geq E[u(Y)]$ for all $u \in U_1$, with strict inequality for some u.

(2) $F(x) \leq G(x)$ for all x in the support of X, with strict inequality for some x.

(3) $q_x(p) \geq q_y(p)$ for all $0 \leq p \leq 1$.

Definition (**SSD**): X Second Order Stochastic Dominates Y, denoted X SSD Y, if and only if any of the following equivalent conditions holds:

(1) $E[u(X)] \geq E[u(Y)]$ for all $u \in U_2$, with strict inequality for some u.

(2) $\int_{-\infty}^{x} F(t)dt \leq \int_{-\infty}^{x} G(t)dt$ for all x in the support of X and Y, with strict inequality for some x.

(3) $\int_0^p q_x(t)dt \geq \int_0^p q_y(t)dt$, for all $0 \leq p \leq 1$, with strict inequality for some value(s) p.

The tests of FSD and SSD are based on empirical evaluations of conditions (2) or (3) in the above definitions. Mounting tests on conditions (3) typically relies on the fact that quantiles are consistently estimated by the corresponding order statistics at a finite number of sample points. Mounting tests on conditions (2) requires empirical cdfs and comparisons at a finite number of observed ordinates. Also, from Shorrocks (1983) or Xu (1995) it is clear that condition (3) of SSD is equivalent to the requirement of Generalized Lorenz (GL) dominance. FSD implies SSD.

Noting the usual definition of the Lorenz curve of, for instance, X as $L_x(x) = \frac{1}{\mu_x} \int_{-\infty}^{x} X \times dF(t)$, and its GL (x) = $\mu_x L_x(x)$, some authors have developed tests for Lorenz and GL dominance on the basis of the sample estimates of conditional interval means and cumulative moments of income distributions; e.g. see Bishop et al (1989), Bishop et al (1991), Beach et al (1995), and Maasoumi (1996a) for a general survey of the same. The asymptotic distributions given by Beach et al (1995) are particularly relevant for testing for Third Order Stochastic Dominance (TSD). The latter is a useful criterion when Lorenz or GL curves cross at several points and the investigator is willing to adopt "transfer sensitivity" of Shorrocks and Foster (1987), that is a relative preference for progressive transfers to poorer individuals. When either Lorenz or Generalized Lorenz Curves of two distributions cross unambiguous ranking by FSD and SSD is not possible. Whitmore (1970) introduced the concept of third order stochastic dominance (TSD) in finance. Shorrocks and Foster (1987) showed that the addition of the "transfer sensitivity" requirement leads to TSD ranking of income distributions. This requirement is stronger than the Pigou-Dalton principle of transfers and is based on the class of welfare functions U_3 which is a subset of U_2 with $u''' \geq 0$. TSD is defined as follows:

Definition (**TSD**): X Third Order Stochatic Dominates Y, denoted X TSD Y, if and only if any of the following equivalent conditions holds:

(1) $E[u(X)] \geq E[u(Y)]$ for all $u \in U_3$, with strict inequality for some u.

(2) $\int_{-\infty}^{x} \int_{-\infty}^{v} [F(t) - G(t)]dt\, dv \leq 0$, for all x in the support, with strict inequality for some x,

with the end-point condition:

$\int_{-\infty}^{+\infty} [F(t) - G(t)]dt \leq 0.$

(3) When E[X] = E[Y], X TSD Y iff $\lambda_x^2(q_i) \leq \lambda_y^2(q_i)$, for all Lorenz curve crossing points $i = 1,2,\ldots\ldots, (n+1)$; where $\lambda_x^2(q_i)$ denotes the "cumulative variance" for incomes upto the *ith* crossing point. See Davies and Hoy (1995).

When $n = 1$, Shorrocks and Foster (1987) showed that X TSD Y if (a) the Lorenz curve of X cuts that of Y from above, and (b) Var(X) ≤ Var(Y). This situation revives the coefficient of variation as a useful statistical index for ranking distributions.

Kaur et al (1994) assume i.i.d observations for independent prospects X and Y. Their null hypothesis is condition (2) of SSD *for each x* against the alternative of strict violation of the same condition *for all x*. The test of SSD then requires an appeal to union intersection technique which results in a test procedure with maximum asymptotic size of α if the test statistic at each x is compared with the critical value Z_α of the standard Normal distribution.

McFadden offers a definition of "maximal" sets, as follows:

Definition (**Maximality**): Let $Æ = \{X_1, X_2, \ldots, X_K\}$ denote a set of K distinct random variables. Let F_k denote the cdf of the k-th variable. The set Æ is *first* (*second*) *order maximal* if no variable in Æ is first (second) order weakly dominated by another.

Let $X_{.n} = (x_{1n}, x_{2n}, \ldots, x_{Kn})$, n = 1, 2,....,N , be the observed data. Assume $X_{.n}$ is strictly stationary and α – *mixing*, and assume $F_i(X_i)$, i = 1,2,...,K, are *exchangeable* random variables, so that resampling estimates of the test statistics converge appropriately. This is less demanding than the assumption of independence which is not realistic in many applications (as in mobility analysis, and before and after tax scenarios). In general F_k is unknown and estimated by the empirical distribution function $F_{kN}(X_k)$. Finally, if we adopt Klecan et al's mathematical regularity conditions pertaining to von Neumann-Morgenstern (VNM) utility functions that generally underlie the expected utility maximization paradigm, the following theorem defines the tests and the hypotheses being tested:

Theorem *Given the mathematical regularity conditions*;

(*a*) *The variables in Æ are first-order stochastically maximal*; *i.e.*,

(1) $d = \min_{i \neq j} \max_{x} [F_i(x) - F_j(x)] > 0,$

if and only if for each i and j, there exists a continuous increasing function u such that $E\, u(X_i) > E\, u(X_j)$.

(*b*) *The variables in Æ are second order stochastically maximal*; *i.e.*,

(2) $S = \min_{i \neq j} \max_x \int_{-\infty}^{x} [F_i(\mu) - F_j(\mu)]d\mu > 0,$

if and only if for each i and j, there exists a continuous increasing and strictly concave function u such that $Eu(X_i) > Eu(X_j)$.

(*c*) *Assuming the stochastic process* $X_{.n}$, $n = 1, 2, \ldots$, *to be strictly stationary and* α – *mixing with* $\alpha(j) = O(j^{-\delta})$, *for some* $\delta > 1$, *we have*:

$d_{2N} \rightarrow d$, *and* $S_{2N} \rightarrow S$, *where* d_{2N} *and* S_{2N} *are the empirical test statistics defined as* :

$$(3)\ d_{2N} = \min_{i \neq j} \max_{x} [F_{iN}(x) - F_{jN}(x)]$$

and,

$$(4)\ S_{2N} = \min_{i \neq j} \max_{x} \int_0^x [F_{iN}(\mu) - F_{jN}(\mu)]d\mu$$

[Proof] See Theorems 1 *and* 5 *of Klecan et al* (1991).

The null hypotheses tested by these two statistics is that, respectively, Æ is *not* first (second) order maximal— i.e., X_i FSD(SSD) X_j, for some *i* and *j*. We reject the null when the statistics are positive and large. Since the null hypothesis in each case is composite, power is conventionally determined in the least favorable case of identical marginals $F_i = F_j$. Thus, as is shown in Kaur et al (1994) and Klecan et al (1991), tests based on d_{2N} and S_{2N} are consistent. Furthermore, the asymptotic distribution of these statistics are non-degenerate in the least favorable case, being Gaussian (see Klecan et al (1991), Theorems 6-7).

The statistic S_{2N} has, in general, neither a tractable distribution, nor an asymptotic distribution for which there are convenient computational approximations. The situation for d_{2N} is similar except for some special cases—-see Durbin (1973, 1985), and McFadden (1989) who assume i.i.d. observations (not crucial), and independent variables in Æ (consequential). Unequal sample sizes may be handled as in Kaur et al.

Klecan et al (1991) suggest Monte Carlo procedures for computing the significance levels of these tests. This forces a dependence on an assumed parametric distribution for generating MC iterations, but is otherwise quite appealing for very large iterations. Maasoumi et al (1996) employ bootstrap methods to obtain the empirical distributions of the test statistics and confidence intervals. They report an empirical examination of the US income distribution based on the CPS and PSID data. Their methods are directly applicable to ranking of mobility structures described previously.

Some empirical examples

Creedy (1985) contains detailed descriptions of empirical studies which implement the transition matrix and other model-based techniques. Shorrocks (1976, 1978a) and Lillard and Willis (1978) also implement the transition matrix method using some of the same US panel data which I will describe below.

The MSZ index method has been implemented by Shorrocks, Maasoumi, Zandvakili, Trede and others. Trede's work is based on German panel data and, as mentioned earlier, reports statistical tests of significant change in mobility. The first three authors use the Michigan Panel data. We now exemplify some of these latter studies:

Mobility and gender

The MSZ family of mobility measures was investigated by Maasoumi and Zandvakili (1990) using the Michigan Panel Study of Income Dynamics (PSID). These measures are decomposed in order to learn about components that are due to differences in *gender and income* groups, on the one hand, and within group components which are free of such group characteristics. Several aggregator functions were used to compute the "Permanent income" variable. Their justification and role in robustifying inferences was investigated.

"Household" income data for the period 1969-81 were taken from the PSID. Household's income (head and spouse, if any) consists of the following: income from wages, salaries, rents, dividends, interest, business, bonuses, commissions, professional practice, aid to dependent children, social security, retirement pay, pension or annuities, unemployment compensation, child support, and other transfer payments. Real total income is obtained using the current consumer price index. Income is adjusted for family size (in 1975) to provide a better measure of family income since family members effectively pool their incomes; see Kakwani (1984) and Rosen (1984). We refer to this adjusted income as the "Per Capita Family Income" (PCFI).

In computing the permanent incomes three different schemes were used in order to weight income at different times. These α_t weights are, (i) equal weights for all years, (ii) the ratio of mean income at time *t* to the mean income over the entire *T* periods (MIW in tables), and (iii) the normalized elements of the eigen vector corresponding to the first principal component of the $X^{\vee}X$ matrix. We did not find any qualitative differences in our results between these three cases, and thus report only the computations based on ratio-of-means weights. The other two cases are reported in Zandvakili (1987).

In our computations the substitution parameter β is restricted by the relations $-\gamma = 1 + \beta$. We computed four different aggregator functions corresponding to four inequality measures with $-\gamma = \nu = (2, 1, .5, .0)$. $\nu = 0.0$ and 1.0 correspond to Theil's first and second inequality measures, respectively combined with the linear and the Cobb-Douglas forms of the aggregator function. Tables 1-3 are from Maasoumi and Zandvakili (1990) which provide, respectively, the annual short-run inequalities, the inequalities in the aggregated (long-run) incomes, and the income stability measure R_γ. Decomposition of each based on gender is also given. Note that as one moves toward 1981 the number of periods over which S_i, $I_\gamma(S)$ and R_γ are calculated is increasing from one to 13. The results for every other year are reported to save space.

[Tables 1-3 go about here]

Short-run inequality in Table 1 has generally increased. As expected, inequality is greater with larger degrees of relative inequality aversion (v). There are 1776 male and 529 female headed households in the sample. The "within-group" component of short-run inequalities is dominant. The column heading "within" refers to weighted average of within group inequalities as in the formulae previously discussed in this paper. The absolute value of the "between-group" component, however, has increased over the 13 years. For both sexes annual inequalities have a rising trend (less uniformly so for $v = 2$). But short-run inequality amongst female headed households is always greater than amongst men. These annual values, however, contain many transitory components which are partially removed from the aggregated values in Table 2.

Table 2 values exhibit much less volatility. After a decline in the initial years, $I_\gamma(S)$ has increased back to about its original value. Also, long-run inequality is always smaller than the corresponding short-run inequality. Once again, inequality among women is greater than among men, and within group inequality is several times the between group component. These relative values are somewhat sensitive to the family size adjustment of incomes. For instance, the between group component increases to 15-25% of overall inequality for unadjusted incomes; see Zandvakili (1987). This is partly due to a larger proportion of two income earners being among the male headed families.

The corresponding stability measures are presented in Table 3. Again, seven of the 13 possible values are reported without any qualitative loss. $0 < R_\gamma < 1$ in all cases. The following may be concluded from Table 3 :

(i) There is a tendency for the profiles to fall and then level off as the number of aggregated years increases from one to 13.

(ii) The profiles for households headed by men fall faster and further than those of women headed households.

(iii) These patterns are robust with respect to the choice of aggregation function, family size adjustment, and inequality measure.

The fact that the profiles are becoming flatter is an indication that, although there have been some transitory movements in the size distribution of income, there is a lack of any permanent equalization. Further, while some equalization has taken place within each group of households, inequality between men and women headed households has increased in absolute value.

Mobility and Income Level.

Maasoumi and Zandvakili (1989) give inequality and mobility decompositions by age, education, and race. Similar decompositions by income level can reveal the aggregate impact of all such non-income characteristics (including gender). It is anticipated that if the major causes of variation in incomes are transitory in nature, the length of time spent in any income class will be short. "Permanent" income inequality changes will be very revealing in this context.

The total sample is divided into seven income groups (G1-G7). The assignment to groups is on a one time basis and according to the simple arithmetic mean income of the individual household over the thirteen year period. These real income levels begin with mean incomes of less than $4,999, and increasing in increments of $5,000. The last group contains mean incomes of $35,000 or more.

Short-run inequalities and their decompositions based on income level are given in Table 4. All the tables and figures in this section are taken from Maasoumi and Zandvakili (1990). There are several recognizable patterns. The "between group" inequality has increased steadily over this period. The "within group" component of inequality fluctuates around a relatively constant mean value. The observed patterns suggest that the non-income differences do contribute to the increase in between group inequality. Over 70% of women headed households earn less than $15,000. Of course, this is confounded by the differential impact of inflation on different income groups (we use real incomes).

[Tables 4-6 go about here]

In Table 5 long run inequality levels have risen after an initial decline. Decomposition by income level shows that the between group component of $I_\gamma(S)$ has increased uniformly. At the same time the within group inequality has *decreased* steadily. This change has been dramatic so that in the later years the between group component is larger than the within component. These changes include the well known life cycle and human capital effects, and are not inconsistent with the cumulative effects predicted by discrimination theories.

The long run *within group* inequalities reveal a falling trend for each of the seven income groups. This is anticipated since transitory components are smoothed out and individual incomes have approached group mean incomes in the long run. These long run grouping observations are somewhat sensitive to the family size. Within group aggregate income inequalities are noticeably smaller when income is not adjusted for family size, and there is generally less inequality within the higher income groups.

Table 6 reports the stability profiles which reveal much higher degrees of permanent equalization *within* income groups than was observed for the gender groups of the last section. Note that as the stability profiles of the whole sample flatten, the corresponding within group profiles continue to fall. In our view some equalization has occurred, but this is mostly confined to within income groups.

On the basis of the approximately 2300 households which remained in the Michigan panel over the period 1969-81, Maasoumi and Zandvakili (1990) conclude that: (i) there is not a great deal of inequality between the men and women-headed households; (ii) the dominant within group component of inequality is either increasing over this period or, when incomes are smoothed by time aggregation, relatively stable. (iii) This larger within group component of inequality is due to high levels of inequality within lower income groups (such as women headed households). (iv) Grouping by real income brackets leads predictably to very large between group inequality values. (v) Some equalization of real incomes has occurred over time *within* most *income* groups, but this is very hard to judge by a comparison of annual inequality measures and most clearly revealed by using our "permanent income" distributions. (vi) Modest levels of mobility are recorded as the aggregation interval is expanded, but the corresponding profiles flatten-out after about eight or nine years.

We close this subsection with Figure 1 which summarize the evolution of the income distribution for this sample with the graph of the stability profile P_t.

[Figure 1 goes here]

Mobility by education, age, and race

Maasoumi and Zandvakili (1989) is based on the same data as the previous section, but the role of years of schooling of the head of household, his/her age, and race were examined through decomposition of the inequality/mobility measures. Tables 7-15 are from that source.

Table 7 is a summary of short run, long run, and the stability values for all the thirteen years. Tables 8-10 provide decompositions by educational attainment which was indicated by the years of completed schooling by the head. The increase in overall short-run inequality is primarily due to increases in within group inequalities. Long run inequality is quite stable. the *R* measure declines over longer periods. This indicates that while there is much short run mobility (change), this does not change permanent income inequality. Note that this phenomenon may be partly due to the anonymity of our measures which are invariant to short run *switching of positions* by individuals.

It should be noted that education is both a capital good and a provider of a stream of consumption. It has different values for different individuals. This heterogeneity effect is here controlled for leading to conditional inferences. For a discussion of these issues and a multidimensional treatment in which education is regarded as a distinct attribute (with income and wealth) see Maasoumi and Nickelsburg (1988).

[Tables 7-10 go about here]

Tables 8-10 indicate that the greatest inequality is within the group with fewest schooling years. Indeed, within group inequality declines steadily with educational attainment: education is an equalizer (some might argue it is a restraint over unusual earnings!). Between group inequality is rising somewhat over these years but is about one quarter of total inequality, and declining proportionately.

Long run inequality is much more stable over time and is smaller than short run inequalities. Looking at these figures a policy maker is less likely drawn to quick reactions to transitory phenomenon, and more likely to focus on stable features, for instance, the fact that some mobility is experienced in the early part of this period, but has ceased in the latter part of the sample. Similarly, we note that between group *long run* inequality has risen consistently, suggesting that the expected returns to schooling has materialized. In fact the inequality gap between the educational groups here has widened by 50% over time; see John et al (1993).

[Tables 11-13 go here]

Tables 11-13 have the same structure as before but focus on the impact of age. These tables suggest that between group inequality, both short and long run, has increased dramatically over time. Seniorage matters! Within group inequality is larger the older the group. This is also due to accumulation of returns to different investments, opportunities, and attainments. Short run inequalities increase within groups, while long run inequalities are stable with a moderate increase toward the end of this period. Maasoumi et al (1996) find these trends have continued. Finally we note that these figures are based on per capita incomes. Since family size and composition changes over time these figures show greater volatility than the authors found with total family incomes unadjusted for family size; see Maasoumi and Zandvakili (1989, Appendix).

[tables 14-15 go here]

Tables 14-15 provide decompositions by race, noting that the "non-white" group includes all heads not classified as "white". This explains the large within group inequality. The number of households in each group is given in the last column. Inequality amongst non-whites has increased faster than amongst whites. Short run inequality has increased within both groups, and somewhat increased in the aggregated incomes. Between group inequality in the short run distributions declined somewhat in the first half of the period and increased again in the last 4-5 years of the sample. For the long run incomes, between group inequalities are rather stable with a slight decline over time. It would appear that within group characteristics are controlling of the degree of inequality in this sample and for this decomposition. Other grouping criteria that are more race specific than "non-white" are known to indicate greater between group inequality. See John et al (1991).

Experimentation over the members of the GE family, as well as with different sets of weights given to incomes at different points in time, represent an attempt to robustify summary findings. This is an important element of empirical work in this area since unanimity with respect to weights and degree of aversion to inequality is not likely. Of course, an interpretation of this "robustification" technique is that it is an empirical substitute for unanimous ranking by Lorenz-type comparisons over plausible ranges of parameter values. This is useful when such curves cannot be statistically ordered when they cross only at extreme parametric values.

Several other applications to US and UK data are reported in Shorrocks (1978b, 1981). A good deal more is now possible given the dominance testing techniques of section 4, and the asymptotic distribution theory summarized in Maasoumi (1996a). The bootstrap alternative appears very promising, as demonstrated by Mills and Zandvakili (1996).

Anderson, G.J. (1994),“Nonparametric tests of stochastic dominance in income distributions,” Working paper (mimeo.), University of Toronto.

Atkinson, A. (1970),“On the Measurement of Inequality,” **Journal of Economic Theory**, II, 244-63.

________(1983),“The measurement of economic mobility,” in **Social Justice and Public Policy**, MIT Press, Cambridge, MA.

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

_____________________________ (1987),“Income Distribution and Differences in Needs,” in G.R. Feiwel (ed.), **Arrow and the Foundations of the Theory of Economic Policy**, Macmillan: London.

____________________________ (1989),“The design of Direct Taxation and Family Benefits,” **Journal of Public Economics**, 41, 3-29.

Atkinson, A., F., Bourguignon, F., and C. Morrison (1992). **Empirical studies of earnings mobility**, Fundamentals of Pure and Applied Economics, vol. 52, Distribution Section, Harwood Academic.

Bartholomew, D.J. (1959),“A test of homogeneity for ordered alternatives,” **Biometrica**, 46, 36-48.

____________(1982), **Stochastic models for social processes**, J. Wiley, New York (3rd. ed).

Beach, C.M. and R. Davidson (1983),“Distribution-free statistical inference with Lorenz curves and income shares,” **Review of Economic Studies**, 50, 723-735.

_________and J. Richmond (1985),“Joint confidence intervals for income shares and Lorenz curves,” **International Economic Review**, 26, 439-450.

Beach, C.M., R. Davidson, and G.A. Slotsve (1995),“Distribution-free statistical inference for inequality dominance with crossing Lorenz curves,” Unpublished paper N95A03, GREQAM, Universites *d'Aix*-Marseille .

Bishop, J. A., J. P. Formby, and P. D. Thistle, (1989),“Statistical inference, income distributions, and social welfare,” **Research on Economic Inequality**, 1, 49-82.

Bishop, J.A.,K.V.Chow, and J.P.Formby (1994),“Testing for marginal changes in income distributions with Lorenz and Concentration curves,” **International Economic Review**, 35, 2, 479-488.

Blackorby, C. and D. Donaldson (1978),“Measures of relative equality and their meaning in terms of social welfare,” **Journal of Economic Theory**, 18, 59-80.

Bourguignon, F. (1979),“Decomposable Income Inequality Measures,” **Econometrica**, 47, 901-20.

Chakravarty, S.R. (1995),“A note on the measurement of mobility,” **Economics Letters**, 48, 33-36.

Conlisk, J. (1989),“Ranking mobility matrices,” **Economics Letters**, 29, 231-235.

________(1990),“Monotone mobility matrices,” **Journal of Mathematical Sociology**, 15, 173-191.

Creedy, J. (1985), **Dynamics of income distribution**, Oxford: Basil Blackwell.

Davies, James, and Michael Hoy (1994),“The Normative significance of using third-degree stochastic dominance in comparing income distributions,” **Journal of Economic Theory**, 64, 520-30.

_________ (1995),“Making inequality comparisons when Lorenz curves intersect,”**American Economic Review**, 85, 4, 980-986.

Deshpande,J.V., and S.P.Singh (1985),“Testing for second order stochastic dominance,”**Communications in Statistics-Theory and Methods**, 14, 887-893.

Dardanoni, V (1992),“On multidimensional inequality measurement,” University of Perugia, Italy (Unpublished).

___________(1993),“On measuring social mobility,” **Journal of Economic Theory**, 61, 372-394.

Durbin, J.(1973), **Distribution theory for tests based on the sample distribution function**, SIAM, Philadelphia.

Durbin, J. (1985),“The first passage density of a continuous Gaussian process to a general boundary,” **Journal of Applied Probability**, 22, 99-122.

Dykstra, R.L., Kochar, S., and T. Robertson (1995),“Inference for likelihood ratio ordering in the two sample problem,” **Journal of American Statistical Association**, 90, 1034-40.

Ebert, U.(1995a),“Income inequality and differences in household size,” **Mathematical Social Sciences**, 30, 37-55.

______(1995b),“Using equivalence scales to compare income distributions when households are different,” October 1995, unpublished paper, University of Oldenburg, Germany.

Fields, G., and E.A. Ok (1995),“The meaning and the measurement of income mobility,” In J. Silber (ed) **Income inequality measurement**: **From theory to practice**, Kluwer, forthcoming.

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

Gourieroux, C., Holly, A., and A. Monfort (1982),“Likelihood ratio test, Wald test and Kuhn-Tucker test in linear models with inequality constraints in the regression parameters,” **Econometrica**, 50, 63-80.

Hadar, J. and W.R. Russell (1969),“Rules for ordering uncertain prospects,”American Economic Review, **59**, **25-34**.

Hadar,J. and W.R. Russell (1971),“Stochastic dominance and diversification,” **Journal of Economic Theory**, 3, 288-305.

Hanoch,G. and H. Levy (1969),“The efficiency analysis of choices involving risk,” **Review of Economic Studies**, 36,

335-346.
Hart, P.E. (1976a),“The Dynamics of Earnings, 1963-1973,” **Economic Journal**, 86, 541-65.
________(1976b),“The comparative statics and dynamics of income distributions,” **Journal of the Royal Statistical Society**, Series A, 139, 108-125.
________(1981),“The statics and dynamics of income distributions: A survey,” in Klevmarken and Lybeck (eds.), cited below.
_________(1983),“The size mobility of earnings,” **Oxford Bulletin of Economics and Statistics**, 45, 181-193.
John, C. , K.M. Murphy, and B. Pierce (1991),“Accounting for the slowdown in Black-White wage convergence,” In **Workers and Their wages**: **Changing Patterns in the United States**., Marvin H. Kosters (Ed.) Washington: Americam enterprise Institute Press.
John, C., K.M. Murphy, and B. Pierce (1993),“Wage inequality and the rise in returns to skill,” **Journal of Political Economy**, 101, no.3., 410-442.
Kanbur, S.M.R., and J.E. Stiglitz (1986),“Intergenerational mobility and dynastic inequality,” Woodrow Wilson discussion paper No. 111, Princeton Univ.
Kaur, A., B.L.S. Prakasa Rao, and H. Singh (1994),“Testing for second-order stochastic dominance of two distributions,” **Econometric Theory**, 10, 849-866.
Klecan, L., R. McFadden, and D. McFadden (1991),“A robust test for stochastic dominance,” Working paper, Economics Dept., MIT.
Klevmarken, .A., and J.A. Lybeck (1981), **The statics and dynamics of income**, Clevedon: Tieto limited.
Kodde, D.A., and F.C. Palm (1986),“Wald criteria for jointly testing equality and inequality restrictions,” **Econometrica**, 50, 1243-1248.
Kolm, S. C. (1969),“The optimal production of social justice,” in **Public Economics**, J. Margolis and H. Guitton (eds.), MacMillan: London, 145-200; Also presented at IEA conference on Public Economics, Biaritz, 1966.
Kolm, S.C. (1976a),“Unequal inequalities : I,” **Journal of Economic Theory**, 12, 416-442.
___________(1976b),“Unequal inequalities : II,” **Journal of Economic Theory**, 13, 82-111.
Lehmann, E.L. (1959), **Testing statistical hypotheses**, J. Wiley and Sons, N.Y.
Lillard, L., and Willis, R.(1978),“Dynamic Aspects of Earnings Mobility,” **Econometrica**, 46, 985-1012.
Maasoumi, E.(1986),“The Measurement and Decomposition of Multi-Dimensional Inequality,” **Econometrica**, 54, 991-97.
Maasoumi, E.(1989a),“Composite Indices of Income and Other Developmental Indicators : A General Approach,” **Research on Economic Inequality**, Vol. 1, 269-286.
_____________(1989b),“Continuously distributed attributes and measures of multivariate inequality,” **Journal of Econometrics**, 42, 1, 131-144.
___________. (1991), **Measurement of Welfare and Inequality**, (Guest ed.), Annals of Econometrics, **Journal of Econometrics**, Vol. 50, No. 1/2.
___________. (1993),“A Compendium to Information Theory in Economics and Econometrics,” **Econometric Reviews**, Vol 12, 3, 1-49.
Maasoumi, E. (1998), “Empirical Analyses of Inequality and Welfare”, in P.Schmidt and M.H. Pesaran (eds.) **Handbook of Applied Microeconometrics**, Blackwell publishers.
__________(1999),“Multidimensioned approaches to welfare analysis,” Chapter 15, in J. Silber (ed.), **Handbook of Income Inequality Measurement**, Kluwer Publishers, Forthcoming.
Maasoumi, E., and J. Nickelsburg (1988),“Multivariate measures of well being and an analysis of Inequality in the Michigan data,” **Journal of Business and Economic Statistics**, 6, 3, 327-334.
Maasoumi, E., and M. Trede (1997),”Comparing Income Mobility in Germany and the US using Generalized Entropy Mobility Measures”, with Mark Trede, **The Review of Economics and Statistics**, forthcoming.
Maasoumi, E., and S. Zandvakili (1986),“A Class of Generalized Measures of Mobility with Applications,” **Economics Letters**, 22, 97-102.
_________________,(1989),“Mobility Profiles and Time Aggregates of Individual Incomes,” **Research on Economic Inequality**, 195-218..
__________________,(1990),“Generalized entropy measures of mobility for different sexes and income levels,” **Journal of Econometrics**, 121-133.
Maasoumi, E., J. Mills, and S. Zandvakili (1996),“Consensus ranking of US income distributions: A bootstrap application of tests for stochastic dominance,” Mimeo, Department of Economics, SMU, Dallas, TX..
Markandya, A. (1984),“The welfare measurement of changes in economic mobility,” **Economica**, 51, 457-471.
McFadden, D. (1989),“Testing for stochastic dominance,” in Part II of T. Fomby and T.K. Seo (eds.) **Studies in the Economics of Uncertainty** (in honor of J. Hadar), Springer-Verlag.
Mills, J. and S. Zandvakili (1996),“bootstraping inequality measures,” Working paper, University of Cincinnati, Economics.
Mitra, T., and E.A. Ok (1995),“The measurement of income mobility: A partial ordering approach,” Cornell University,

April.

Perlman, M.D. (1969),“One-sided testing problems in multivariate analysis,” **Annals of Math. Statistics**, 40, 549-62.

Prais, S.J. (1955),“Measuring social mobility,” **J. of Royal Statistical Society**, Ser. A., 56-66.

Robertson, T., and F.T. Wright (1981),“Likelihood ratio tests for and against stochastic ordering between multinomial populations,” **Annals of Statistics**, 9, 1248-57.

Robertson, T., Wright, F.T., and R. Dykstra (1982?), **Order restricted statistical inference**, J. Wiley and Sons, New York, N.Y.

Sargan ,J.D. (1957),“The distribution of wealth,” **Econometrica**, 25, 568-90, reprinted as Chapter 3, in E. Maasoumi (ed.) Contributions to Econometrics : J.D. Sargan, Vol. 1, (1988), Cambridge University Press.

Sen, A. (1970) **Collective choice and social welfare**, Holden Day: San Francisco, (reprinted, North-Holland, 1979).

(1988),“Towards a unified theory of inequality constrained testing in multivariate analysis,” **International Statistical Review**, 56, 49-62.

Shorrocks, A.F.(1976),“Income Mobility and the Markov Assumption,” **Economic Journal**, 86, 566-77.

___________(1978a),“The measurement of mobility,” **Econometrica**, 46, 1013-1024.

______(1978b),“Income Inequality and Income Mobility,” **Journal of Economic Theory**, 19, 376-93.

_______(1980),“The class of additively decomposable inequality measures,” **Econometrica**, 48, 613-625.

_______(1981),“Income Stability in the United States,” Ch. 9 in **The Statics and Dynamics of Income**, Klevmarken, N.A., and J.A. Lybeck (eds.), Tieto Ltd, Bristol.

_______(1983),“Ranking income distributions,” **Economica**, 50, 3-17.

_______(1984), “Inequality Decomposition by population Subgroups,” **Econometrica**, 52, 1369-85.

________, and J. Foster (1987),“Transfer sensitive inequality measures,” **Review of Economic Studies**, 54, 485-497.

_____________(1992),“On the Hart measure of income mobility,” in , Industrial concentration and economic inequality, Festschrift for Peter Hart, forthcoming.

__________(1995),“Inequality and welfare evaluation of heterogeneous income distributions,” Unpublished paper, University of Essex, UK.

Theil, H. (1972), **Statistical Decomposition Analysis**, North-Holland, Amsterdam.

Trede, M. (1995),“Statistical inference in mobility measurement: Sex differences in earnings mobility,” University of Koln, Germany (unpublished).

U.S. Bureau of the Census, Current Population Reports : Consumer Income, Series $P-60$ (various years).

Wolak, F.A. (1988),“Duality in testing multivariate hypothesis,” **Biometrica**, 75, 611-5.

_______(1989),“Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models,” **Econometric Theory**, 5, 1-35.

__________(1991),“The local nature of hypothesis tests involving inequality constraints in nonlinear models,” **Econometrica**, 981-95.

Xu, K., G. Fisher, and D. Wilson (1995),“New distribution-free tests for stochastic dominance,” Working paper No. 95-02, February, Dept. of Economics, Dalhousie University, Halifax, Nova Scotia.

Xu, K. (1995),“Asymptotically distribution-free statistical test for generalized Lorenz curves:An alternative approach,” Dept. of Econ., Dalhousie University, Halifax, Nova Scotia, presented at the 7th World Congress of Econometric Society, Japan.

Zandvakili, S. (1987), **The Measurement of Mobility and Income Stability** : **An Information Theoretic Approach**, Ph.D. Thesis, Indiana University, Bloomington.