JOURNAL OF ECONOMIC THEORY 61, 372–394 (1993)

Measuring Social Mobility

VALENTINO DARDANONI*

Università di Palermo, 90128 Palermo, Italy; and University of California, San Diego, California 92093

Received April 4, 1990; revised September 23, 1992

The paper considers the ranking of mobility matrices in a simple Markov model of social mobility. The approach is the dynamic counterpart ot the "static" inequality ranking of income distributions by the Lorenz curve. The derived partial ordering is motivated by welfare considerations, is shown to be equivalent to some intuitive mobility concepts, and is used to screen some immobility indices. The equivalence of the ranking with the "permanent income" Lorenz ordering gives support to the claim that this approach is the natural extension of Kolm's [The optimal production of social justice, *in* "Public Economics (J. Margolis and H. Guitton, Eds.), MacMillan, London, 1969], Atkinson's [On the measurement of inequality, *J. Econ. Theory* **2** (1970), 244–263], and Dasgupta, Sen, and Starrett's [Notes on the measurement of inequality, *J. Econ. Theory* **6** (1973), 180–187] approaches. *Journal of Economic Literature* Classification Numbers: D31, D63, J62. © 1993 Academic Press, Inc.

1. INTRODUCTION

The theory of inequality measurement is in general concerned with static income distributions, where "snapshots" of the income distribution are the basis of the analysis. In practice, income distributions change over time, under the effect of different transition mechanisms. Transition mechanisms may affect social welfare by changing the shape of the "spot" income distribution. Yet, two societies with the same spot income distributions may have a different level of social welfare depending on the mobility of the populations. For example, as Friedman [10] argues, the income inequality owing to a rigid system where each family stays in the same position each year may be more a cause for concern than the income inequality owing to great mobility and dynamic change associated with equality of opportunity.

* I thank (without implicating) Stefano Demichelis, Peter Lambert, Michael Rothschild, Max Stinchcombe, an Associate Editor, and especially Tilman Börgers, John Conlisk, and Joel Sobel for extremely helpful comments and suggestions. Financial assistance from C.N.R.-Rome is gratefully acknowledged.

0022-0531/93 $5.00

It follows that a proper analysis of the equity implications of public redistribution policies on the income distribution should be complemented by the consideration of the changes in social mobility and in the equality of opportunity.

Mobility studies either make assumptions directly on the various mobility indicators and analyze their properties (Bartholomew [3], Conlisk [5], Geweke, Marshall, and Zarkin [12], Shorrocks [22] and Sommers and Conlisk [24] or focus on the welfare implications of the different mobility structures (Atkinson [2], Conlisk [4], Markandya [18] and Kanbur and Stiglitz [13]).

This paper analyzes the latter problem and considers how economic mobility influences social welfare. Following the approaches of Atkinson [2], Markandya [18], and Kanbur and Stiglitz [13], we consider the welfare prospects of individuals in society by deriving the stream of income distributions which obtains under different mobility structures. A class of Social Welfare Functionals (S.W.F.'s) that aggregates these welfare prospects is then proposed, from which the necessary and sufficient conditions for robust welfare ranking can be derived.

This approach is closely related to the seminal papers by Kolm [16], Atkinson [1], and Dasgupta, Sen, and Starrett [7] on measuring static income inequality, where the welfare of individuals under different income schemes is aggregated into a S.W.F., and a partial ordering of income distributions is created according to the unanimous preference of all S.W.F.'s belonging to a particular class. The fundamental inequality theorem states that the Lorenz curve gives the normatively significant ordering of equal mean income distributions; summary statistical measures of inequality are without much normative significance when Lorenz curves cross. In a similar vein, we do not focus on the derivation of a particular index of mobility, where a transition matrix is reduced to a scalar and a complete ranking is obtained. Rather, we derive 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 gives us a condition for checking a robust welfare recommendation without employing a specific mobility measure.

Summary immobility measures induce a complete order on the set of mobility matrices and have the distinctive advantage of giving definite answers. However, it is clear that there are substantial problems in trying to reduce a matrix of transition probabilities into a single number. Consider the three mobility matrices

$$P_1 = \begin{bmatrix} 0.6 & 0.35 & 0.05 \\ 0.35 & 0.4 & 0.25 \\ 0.05 & 0.25 & 0.7 \end{bmatrix}; \quad P_2 = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}; \quad P_3 = \begin{bmatrix} 0.6 & 0.4 & 0 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.7 \end{bmatrix},$$

where rows denote current state and columns denote future state. Which matrix displays more mobility? Suppose we use some common summary immobility measures to answer this question. From the reviews of Bartholomew [3] and Conlisk [5], we consider the following: the second largest eigenvalue modulus, the trace, the determinant, the mean first passage time, and Bartholomew's measure. These indices are defined and explained in Section 5. Performing the relevant calculations, we find that 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.

	Eigenvalue	Trace	Determinant	Mean first passage time	Bartholomew's measure
Most mobile	P_2	P_1, P_3	P_1	P_3	P_1, P_2, P_3

Our welfare-based partial order is based on the idea that the answer to the above question should depend on an explicit and clear social judgment of the consequences of different mobility structures on social welfare. The comparison of summary measures in inequality for different static income distributions is best understood if the proposed statistical measures are rooted in a well-defined ordering that reflects the welfare views of the society. It is hoped that our analysis will help in the understanding of the welfare properties of the various statistical measures of income mobility.[1]

The next section of this paper motivates and derives the proposed robust welfare ranking of mobility matrices. The third section investigates the role of the discount factor in the derived ordering. Section 4 gives non-welfare interpretations of the ordering in terms of commonly perceived views of what can be considered a more mobile social structure. Section 5 uses the derived ordering to "screen" scalar immobility measures and evaluates whether some commonly employed immobility indices agree with the derived ordering. Section 6 considers the relationship between the partial order derived here and previous orderings of transition matrices and proposes a finer welfare-based partial order. The final section gives concluding remarks.

[1] An Associate Editor correctly reminds me of the distinction between the issue of measuring mobility and the issue of ranking different mobility structures in terms of social welfare. Thus, there is no reason why one should not seek to construct summary immobility measures to capture the intuitive descriptive content of the notion without necessarily going into welfare implications.

2. The Welfare Ranking of Mobility Matrices: Motivation

Consider a discrete Markov chain of income mobility, and assume that there are n income states. Let $P = |p_{ij}|$ such that $p_{ij} \geqslant 0$ and $\sum_j p_{ij} = 1$ be the $(n \times n)$ transition matrix, assumed regular (meaning that for some large enough integer k, P^k is strictly positive), so that the strictly positive equilibrium 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 ith element gives the fraction of the population which is in state i at time $t+1$. We assume that transitions are independent across individuals and P is constant over time. We also adopt the convention that income states are ordered from worst to best.

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 intragenerational or intergenerational; in the latter case we may think of the individuals as dynasties. Let $u = (u_1, u_2, ..., u_n)'$ denote a vector of instantaneous utilities, where u_i denotes the utility value of income state i, and $V^P = (V_1, V_2, ..., V_n)'$ denotes a vector of expected discounted lifetime utilities, whose typical element V_i^P denotes the expected lifetime discounted utility of an individual who starts a life in income class i and is given by the ith element of the vector $V^P = u + \rho P u + \rho^2 P^2 u + \cdots + \rho^\tau P^\tau u$. $0 \leqslant \rho < 1$ denotes the discount factor. 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. Letting τ go to infinity, we have $V^P = [I - \rho P]^{-1} u$. To simplify notation, we normalize V^P as $V^P = (1-\rho)[I - \rho P]^{-1} u$ and denote $(1-\rho)[I - \rho P]^{-1}$ as $P(\rho)$, which is a stochastic matrix, whose typical element $p_{ij}(\rho)$ may be interpreted as the average discounted "lifetime" probability of moving from the initial state i to state j.

Suppose now we want to rank transition matrices according to a real-valued S.W.F. defined on the vector of lifetime expected utilities V^P. 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.

Different mobility processes influence social welfare because of the differences in the implied equilibrium income distributions; this is what sociologists call "structural mobility" and is related to the idea that different mobility structures may imply a different availability of positions in higher income classes. Mobility also influences social welfare through its influence on the 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 by sociologists as "exchange," "circulation," or "pure" mobility, and may be interpreted as the exchange of relative positions in society over time.

To isolate the pure mobility effect, we compare societies with identical steady-state income distributions. In other words, we will consider the case where two societies have within each period an identical spot income distribution, π, but individuals may move from income state to income state differently under the two transition mechanisms. 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 average income.

Given two regular transition matrices P and Q with equal steady-state income distribution vector $\pi' = \pi' P = \pi' Q$, how can we decide which society displays a higher level of social welfare?

As a natural starting point, consider the welfare ranking that corresponds to the class of symmetric and additively separable (i.e., linear) S.W.F. $\sum_i \pi_i V_i^P$ which adds up, for a given u and ρ, the expected lifetime utility of the individuals in the population under the transition matrix P. This is equivalent to the S.W.F. employed in the seminal Atkinson's [1] paper on the inequality ranking of static income distributions. Noting that $\pi' P(\rho) = \pi' Q(\rho)$, we have $\sum_i \pi_i V_i^P = \pi' V^P = \pi' P(\rho) u = \pi' u = \pi' Q(\rho) u = \pi' V^Q = \sum_i \pi_i V_i^Q$, so that given a vector of instantaneous utilities u and a discount factor ρ, any two transition matrices with equal steady-state income distribution will be ranked indifferent by the symmetric additively separable S.W.F.'s.

This result, which is also contained in Atkinson [2] and Kanbur and Stiglitz [13], may at first be surprising. However, as Atkinson explains, it must be remembered that we are ranking here not 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 and what is important is the spot distribution at each period. As Kanbur and Stiglitz put it, the assumptions of additivity of the S.W.F. and additive separable lifetime welfares remove any influence that exchange mobility may have on intertemporal social welfare.

The above example is similar to Diamond's [8] example on the fairness of utilitarianism under uncertainty. In fact, additive S.W.F.'s are often criticized for not taking enough account of fairness considerations. Still, it is our opinion that most of the critiques of the "utilitarian" S.W.F. are based on failures of the symmetry assumption and not of linearity per se. For example, it may be argued that both Diamond's [8] and Sen's [20] examples of the "unfairness" of utilitarian rules could be recast as critiques of the symmetry assumption.

In our context, it may be argued that the perception of fairness associated with each transition matrix does indeed come from rejection of the symmetry of the S.W.F. 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 (linear or not) symmetric ex post S.W.F. defined on the vector of realized utilities will rank the matrices indifferent. Yet, under different transition matrices the composition of people in each income state will be different 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 post Lorenz 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 "past history" in terms of the position of each individual in the past.

In light of these remarks, our proposal is to keep the linearity of the S.W.F., but to abandon the symmetry assumption. The symmetry (or anonymity) assumption is employed to guarantee that all individuals in the society are treated equally regardless to their "labeling." However, in this dynamic context there is a natural "label" for each individual, namely, their starting position in the income ranking. Thus, our specified S.W.F. is a weighted sum of the expected welfares of the individuals, with greater weights to the individuals who start with a lower position in the society, $W(V^P, \lambda) = \sum_i \pi_i \lambda_i V_i^P$, where $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)'$ denotes a nonincreasing nonnegative vector of weights.

The asymmetric treatment corrects for the fact that some individuals start at a lower position. Yet, this 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. We therefore consider the case where the transition matrices are monotone. A transition matrix is called monotone if each row stochastically dominates the row above it. Monotone mobility matrices are defined and analyzed in Keilson and Ketser [14] and Conlisk [5]. 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 by having 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, and, given that $P(\rho)$ is monotone when P is monotone, it follows that under a monotone chain the expected lifetime utility vector will be nondecreasing. Conlisk [5] notes that monotonicity is an ideal assumption to impose on a Markov chain model of income mobility, being theoretically plausible and empirically supported. Estimated transition matrices are either exactly monotone or within sampling errors from being monotone.

If the asymmetric linear S.W.F. with declining weights is adopted for the welfare ranking of the transition matrices, the immediate problem is deciding which is the "right" vector of weights λ. For example, two extreme asymmetric linear S.W.F.'s are found by letting $\lambda_1 = 1$ and $\lambda_i = 0$ for all $i \geqslant 1$, which has a kind of "Rawlsian" flavor or by letting $\lambda_i = 1$ for all i, which of course is the symmetric case. By analogy to the literature on static ranking of income distributions, our objective is to seek 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 nonincreasing positive λ.

Before proving our results, we need to define the summation matrix T that is crucial in the derivation of much of what follows: T denotes the $(n \times n)$ upper triangular matrix with zeros below the main diagonal and ones elsewhere. The inverse T^{-1} has ones on the main diagonal, minus ones on the first superdiagonal, and zeros elsewhere. The transpose T' is lower triangular, and its inverse $(T')^{-1}$ has ones on the main diagonal, minus ones on the first subdiagonal, and zeros elsewhere. Finally, note that postmultiplying P by T transforms each row to a cumulative density, premultiplying P by T' takes the cumulative sum of each column, and premultiplying a $(n \times 1)$ vector y by T^{-1} takes the differences of the elements of y. Let us denote by Π the diagonal matrix with the vector π on the diagonal, so that we can rewrite $W(V^P, \lambda)$ compactly as $\lambda' \Pi P(\rho) u$, and let 1 denote the $(n \times 1)$ vector of ones.

We may now state the following:

THEOREM 1. *Let P and Q be two monotone regular transition matrices obeying $\pi' = \pi' P = \pi' Q$, and let ρ be given. Then the following conditions are equivalent*:

1.i. $W(V^P, \lambda) \geqslant W(V^Q, \lambda)$ *for all nonincreasing λ and nondecreasing u*;

1.ii. $T' \Pi [P(\rho) - Q(\rho)] T \leqslant 0$.

Proof.

1.ii implies 1.i. Consider the obvious identity $\lambda' \Pi [P(\rho) - Q(\rho)] u = \lambda' (T')^{-1} T' \Pi [P(\rho) - Q(\rho)] T T^{-1} u$ and note that: (i) the last row of $T' \Pi [P(\rho) - Q(\rho)] T$ equals $1' \Pi [P(\rho) - Q(\rho)] T = \pi' [P(\rho) - Q(\rho)] T = 0$; (ii) the last column of $T' \Pi [P(\rho) - Q(\rho)] T$ equals $T' \Pi [P(\rho) - Q(\rho)] 1 =$

$T'\Pi(1-1)=0$; (iii) λ nonincreasing is equivalent to the first $n-1$ elements of $\lambda'(T')^{-1} \geqslant 0$ and u nondecreasing is equivalent to the first $n-1$ elements of $T^{-1}u \leqslant 0$. The result follows.

1.i implies 1.ii. Assume that the (i, j)th element of $T'\Pi[P(\rho)-Q(\rho)]\,T$ is positive. Then choose $\lambda_i=(1, 1, ..., 1, 0, ..., 0)$ and $u_j=(0, 0, ..., 0, 1, ..., 1)$ so that $\lambda_i' T^{-1}$ equals the $(1 \times n)$ vector with one in the ith position and zeros elsewhere and $T^{-1}u_j$ equals the $(n \times 1)$ vector with minus one in the jth position and zeros elsewhere, and we obtain the desired contradiction. Q.E.D.

The ranking condition 1.ii involves the comparison of the cumulative sums of the matrices $\Pi P(\rho)$ and $\Pi Q(\rho)$. Following Kemeney and Snell [15], we may call these matrices the "lifetime exchange matrices." $\Pi P(\rho)$ and $\Pi Q(\rho)$ array the equilibrium joint density of initial state and lifetime state, and their typical element (i, j) gives the probability of the event (starting position i, lifetime position j). Condition 1.ii requires the comparison of each element of the cumulative sum of the two lifetime exchange matrices, where the cumulative sum is taken from the top left element. When the condition is satisfied, then $\Pi P(\rho)$ displays more mobility than $\Pi Q(\rho)$ in the following sense: the cumulative probability that an individual who starts in class k or lower will stay in lifetime class j or lower is greater under $\Pi Q(\rho)$ for all k and j.

If we restrict our attention to monotone Markov chains with equal steady-state income distribution and assume that lifetime welfare is reflected by discounted expected utility, then when condition 1.ii is satisfied we may deduce that social welfare is superior under the mobility matrix P for any additive asymmetric S.W.F. that gives greater weights to the individuals who start at a lower position in the society. Conversely, if we agree that the social welfare function should belong to the above class, without agreeing on its precise form, then to say that social welfare is higher under the matrix P than under Q implies that P and Q stand in the relation given by 1.ii.

The above necessary and sufficient condition is in effect a first order stochastic dominance result and may be considered the infinite horizon extension to Atkinson's [2] ordering of bistochastic transition matrices for a two period society. Denote by $\mathcal{M}(\pi)$ the set of regular monotone transition matrices with equilibrium vector π. Condition 1.ii induces a (reflexive, antisymmetric, and transitive) partial order on $\mathcal{M}(\pi)$. Given two matrices P and Q in $\mathcal{M}(\pi)$, we denote by $\succcurlyeq_M$ the mobility order induced by 1.ii:

DEFINITION. $P \succcurlyeq_M Q$ on $\mathcal{M}(\pi)$ if and only if $T'\Pi(P-Q)\,T \leqslant 0$ and P, Q belong to $\mathcal{M}(\pi)$.

To get a further intuitive appreciation for the derived ordering $\succcurlyeq_M$ on the set $\mathcal{M}(\pi)$, we can consider the "minimal" and "maximal" elements of this partially ordered set.

THEOREM 2. *For any transition matrix P in $\mathcal{M}(\pi)$ the following is true:*

2.i. $T'\Pi[P-I]\,T \leqslant 0$;

2.ii. $T'\Pi[P-1\pi']\,T \geqslant 0$.

Proof.

2.i. Note that the last row and column of $T'\Pi[P-I]\,T$ consist of zeros; consider then the typical element of $T'\Pi[P-I]\,T$ when $i \leqslant j < n$,

$$\sum_{t=1}^{i} \pi_t \left(\sum_{s=1}^{j} p_{ts} \right) - \sum_{t=1}^{i} \pi_t = \sum_{t=1}^{i} \pi_t \left(1 - \sum_{s=j+1}^{n} p_{ts} \right) - \sum_{t=1}^{i} \pi_t$$
$$= - \sum_{t=1}^{i} \sum_{s=j+1}^{n} \pi_t p_{ts} \leqslant 0,$$

while, when $j < i < n$, we have,

$$\sum_{s=1}^{j} \left(\sum_{t=1}^{i} \pi_t p_{ts} \right) - \sum_{s=1}^{j} \pi_s = \sum_{s=1}^{j} \left(\sum_{t=1}^{n} \pi_t p_{ts} - \sum_{t=i+1}^{n} \pi_t p_{ts} \right) - \sum_{s=1}^{j} \pi_s$$
$$= \sum_{s=1}^{j} \left(\pi_s - \sum_{t=i+1}^{n} \pi_t p_{ts} \right) - \sum_{s=1}^{j} \pi_s$$
$$= - \sum_{s=1}^{j} \sum_{t=i+1}^{n} \pi_t p_{ts} \leqslant 0.$$

2.ii. Consider the following inequalities: $T'\Pi PT = (T'\Pi T)(T^{-1}PT) \geqslant (T'\Pi 1\pi' T)(T^{-1}PT) = T'\Pi 1\pi' PT = T'\Pi 1\pi' T$. Here use is made of: P monotone implies $T^{-1}PT \geqslant 0$; $T'\Pi T \geqslant T'\Pi 1\pi' T$ from 2.i; $\pi' = \pi' P$. Q.E.D.

The mobility ordering $\succcurlyeq_M$ thus agrees with the often-argued view (e.g., Shorrocks, [22]) that the identity matrix should be considered as displaying at least as much immobility as any other transition matrix. On the other side of the spectrum, the ordering agree with Prais' [19] view of a "perfectly mobile" society; the matrix $1\pi'$ has identical rows, with all elements equal to π', thus indicating equality of opportunity and origin independence.

3. THE DISCOUNT FACTOR

Condition 1.ii is crucially dependent on the discount factor employed in the analysis. It is clearly possible that an ordering derived for a given ρ

might be upset under a different discount factor. In this section we investigate this problem and provide conditions under which a definite ranking may be obtained for all the discount factors $0 \leqslant \rho < 1$.

Let P be the transition matrix for a regular Markov chain, with $\pi' = \pi' P$. The *reverse Markov chain* for P is a Markov chain with transition matrix given by $\Pi^{-1} P' \Pi$ (see Kemeny and Snell [15] for a review of properties of reverse chains). A reasonable assumption which is likely to be satisfied by estimated transition matrices is monotonicity of the reverse chain, which implies that at each time, t, an individual in class i has faced a better lottery (in terms of stochastic dominance) than an individual in class $i-1$. We have the following:

THEOREM 3. *Let P and Q be two transition matrices in $\mathcal{M}(\pi)$, and assume that the reverse chain for Q is monotone. Then the following conditions are equivalent*:

3.i. $P \succcurlyeq_M Q$;

3.ii. $P(\rho) \succcurlyeq_M Q(\rho)$ *for all* $0 \leqslant \rho < 1$.

Proof.

3.i. implies 3.ii. Rewrite $T'\Pi[P(\rho) - Q(\rho)]\, T \leqslant 0$ as $T'\Pi[\sum_{t=0}^{\infty} \rho^t P^t - \sum_{t=0}^{\infty} \rho^t Q^t]\, T \leqslant 0$. Given $T'\Pi P T \leqslant T'\Pi Q T$, to prove the result it suffices to show that if $T'\Pi[P^t - Q^t]\, T \leqslant 0$, then $T'\Pi[P^{t+1} - Q^{t+1}]\, T \leqslant 0$. Consider then the following inequalities, where use is made of the fact that under monotonicity of the reverse transition matrix $\Pi^{-1} Q' \Pi$, we have $T'\Pi Q^k \Pi^{-1} T^{-1'} \geqslant 0$ for all k:

$$
\begin{aligned}
T'\Pi P^{t+1} T &= (T'\Pi P^t T)(T^{-1} P T) \leqslant (T'\Pi Q^t T)(T^{-1} P T) = T'\Pi Q^t P T \\
&= (T'\Pi Q^t \Pi^{-1} T^{-1'})(T'\Pi P T) \leqslant (T'\Pi Q^t \Pi^{-1} T^{-1'})(T'\Pi Q T) \\
&= T'\Pi Q^{t+1} T.
\end{aligned}
$$

3.ii implies 3.i. Let $H(\rho) = T'\Pi[P(\rho) - Q(\rho)]\, T$ denote a matrix valued function of ρ. Note that $H(0) = 0$, and $H(\rho) \leqslant 0$ for all $\rho > 0$. Thus, it follows that $(dH(\rho)/d\rho)|_{\rho=0}$ must be non-positive. But $(dH(\rho)/d\rho)|_{\rho=0} = T'\Pi[(I - \rho P)^{-2} P - (I - \rho Q)^{-2} Q]\, T|_{\rho=0} = T'[P - Q]\, T$. Q.E.D.

The result is useful in the following sense: if a "one shot" transition matrix P dominates another matrix Q and $\Pi^{-1} Q' \Pi$ is monotone, then we do not need to carry out the relevant computations for the set of discount factors of interest. Conversely, if P does not dominate Q, we know that there exist some values of the discount factor such that the lifetime matrix $P(\rho)$ will not dominate $Q(\rho)$. Therefore, the researcher may conduct a grid search for the dominance condition at various values of ρ.

4. Non-Welfare Interpretations

In this section we investigate the following question: suppose we have two matrices P and Q which can be ordered according to the welfare-based ordering $\succcurlyeq_M$. Are there any non-welfare mobility interpretations of the derived ordering?

In the literature on inequality comparisons, a fundamental role is played by the "Pigou-Dalton" principle of transfers, which says that a rich to poor income transfer should decrease measured inequality. This is considered by many the "key" economic assumption that should be satisfied by an inequality index. In the mobility context, consider the following: a lifetime exchange matrix $\Pi P(\rho)$ may be thought of as arraying the joint distribution of the event (initial class i, lifetime class j). Denote by D.P.D. a *dynamic Pigou-Dalton* transfer of the following kind: given integers $0 < i, j, s, k < n$ with $i + k \leqslant n$ and $j + s \leqslant n$, decrease the probabilities of the events (initial class i, lifetime class j) and (initial class $i + k$, lifetime class $j + s$) by a quantity $0 \leqslant h \leqslant 1$ and increase the probabilities of the events (initial class i, lifetime class $j + s$) and (initial class $i + k$, lifetime class j) by h. Note that this transformation leaves row and column sums unchanged and assume that h is chosen in a manner that preserves monotonicity. In terms of stochastic dominance, a D.P.D. transfer shifts probability mass to the left, inducing a worst "lifetime lottery" for an individual who starts in class $i + s$, and shifts probability mass to the right, inducing a better lifetime lottery for an individual who starts in class i. This implies that in the monotone Markov chain context, a D.P.D. decreases the lifetime disadvantage of poorer individuals.

Another non-welfare interpretation may be gained by considering the Lorenz curve for the distribution of permanent income. Consider a society which follows a Markov chain with steady-state income distribution π and transition matrix P. Let y denote the income state vector, assumed increasing, and let $Y^P = P(\rho)\, y$ denote the permanent income vector. Under monotonicity, the Lorenz curve for the permanent income distribution $[\pi, Y^P]$ has horizontal coordinates given by $\pi_1, \pi_1 + \pi_2, \ldots, 1$ and vertical coordinates given by $[\pi_1 Y_1^P, \pi_1 Y_1^P + \pi_2 Y_2^P, \ldots, \pi' Y^P]/[\pi' Y^P]$. Exchange mobility implies that the Lorenz curve for the permanent income distribution will lie everywhere above the static Lorenz curve.[2] Assume we observe two societies P and Q with the same steady-state income distribution $\pi' = \pi' P = \pi' Q$. Then the "static" Lorenz curve for P will be identical to that for Q. However, intuitively we would expect that if society P displays more mobility than Q the permanent income Lorenz curve for P will lie above that for Q. This turns out to be so (see below).

[2] This is an obvious implication of Theorems 2 and 4.

A final interpretation of the ranking $\succcurlyeq_M$ may be obtained by the following construct: let f be a $(n \times 1)$ nondecreasing vector of status scores for the initial position, so that f_i measures the status of the ith starting position, and let g be a $(n \times 1)$ nondecreasing vector of status scores for the lifetime position, so that g_j measures the status of the jth lifetime position. Assume that f and g are normalized to have zero mean. When the Markov chain is regular and is in equilibrium, we define the covariance between the initial and lifetime status by $\mathrm{Cov}[f_{k(0)}, g_{k(1)}]$, where the random indices $k(0)$ and $k(1)$ denote the initial and lifetime states of the chain. For a regular transition matrix P with $\pi' = \pi' P$, the (i, j)th element of $\Pi P(\rho)$ gives the probability that $[f_{k(0)}, g_{k(1)}] = (f_i, g_j)$, and thus $\mathrm{Cov}[f_{k(0)}, g_{k(1)}] = f' \Pi P(\rho)\, g$. Under monotonicity, $\mathrm{Cov}[f_{k(0)}, g_{k(1)}]$ will be nonnegative; intuitively we would expect that if P is a more mobile society than Q, P will display a lower covariance between initial and lifetime status.

The following result establishes the desired non-welfare interpretations of $\succcurlyeq_M$:

THEOREM 4. *Let P and Q be two transition matrices in $\mathcal{M}(\pi)$ and let ρ be given. Then the following conditions are equivalent*:

4.i. $P(\rho) \succcurlyeq_M Q(\rho)$;

4.ii. *the permanent income Lorenz curve for P lies nowhere below that for Q for all nondecreasing income vectors y*;

4.iii. *the covariance between initial status and lifetime status is greater under Q for any nondecreasing status score vectors f and g*;

4.iv. *$P(\rho)$ can be derived from $Q(\rho)$ by a finite sequence of D.P.D. transfers.*

Proof. Condition 4.iii may be written as $f' \Pi P(\rho)\, g \leqslant f' \Pi Q(\rho)\, g$ and, under monotonicity, condition 4.ii may be written as $T' \Pi [P(\rho) - Q(\rho)]\, y \leqslant 0$ for all nondecreasing y. Then the equivalence between 4.i, 4.ii, and 4.iii follows by the same line of reasoning as in Theorem 1. That 4.iv implies 4.i is obvious. The converse can be proved in the following way: Let $h_{ij} \geqslant 0$ equal the (i, j)th element of $T' \Pi [Q(\rho) - P(\rho)]\, T$. Then for each $i = 1, 2, ..., n-1$ and $j = 1, 2, ..., n-1$ operate $\Pi Q(\rho)$ with the following D.P.D.: $\pi_i q_{ij} - h_{ij}$; $\pi_i q_{i,j+1} + h_{ij}$; $\pi_{i+1} q_{i+1,j} + h_{ij}$; $\pi_{i+1} q_{i+1,j+1} - h_{ij}$ and note that each D.P.D. leaves all elements [exept the (i, j)th] of $T' \Pi Q(\rho)\, T$ unchanged. Q.E.D.

By analogy with the welfare interpretation given in Section 2, it is worth noting that the Theorem shows that the proposed ordering makes sense only if the transition matrices under consideration are monotone. In fact, if the chain is not monotone, it is clear from the discussion above that it is possible that a D.P.D. might imply a "poor to rich" transfer of lifetime resources. Consider the following two transition matrices P and Q:

$$P = \begin{bmatrix} 0.246154 & 0.223077 & 0.530769 \\ 0.223077 & 0.311538 & 0.465385 \\ 0.530769 & 0.465385 & 0.003846 \end{bmatrix}; \qquad Q = \begin{bmatrix} 0.25 & 0.25 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{bmatrix}.$$

Note that both P and Q are bistochastic (so that they have a common steady-state vector $\pi = (1/3, 1/3, 1/3)'$), but clearly they are not monotone. Assume that the income state vector $y = (10, 95, 100)'$ and that the discount factor $\rho = 0.5$, and calculate the permanent income vectors $Y^P = (40.71, 83.82, 80.47)'$ and $Y^Q = (40.71, 83.87, 80.72)'$. The horizontal coordinates of the Lorenz curve for Y^P are (0.1986, 0.5911, 1), while for Y^Q are (0.1986, 0.5923, 1); thus the permanent income Lorenz curve for Q lies nowhere below that for P. However P is "welfare superior" according to condition 1.ii:

$$T'[P(0.5) - Q(0.5)]\, T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -0.05 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$P(0.5)$ can be derived from $Q(0.5)$ by a transfer of lifetime resources from individuals starting in class 3 to individuals starting in class 2. Given that lifetime income is greater when starting in class 2, this creates greater lifetime inequality.

In this section we have given additional mobility interpretations to the ordering $\succcurlyeq_M$. A final understanding of the properties of the derived ordering may be grained by exploiting the well-known equivalence between the Lorenz curve ordering and other welfare orderings; for example, from Dasgupta, Sen, and Starrett [7], it follows that the partial ordering $\succcurlyeq_M$ is equivalent to the welfare ordering derived from the class of S-concave S.W.F. defined on permanent income for any nondecreasing income vectors and is also equivalent to the welfare ordering derived from the class of concave symmetric additively separable S.W.F. defined on permanent income, for any nondecreasing income vector.

5. The Relationship with Summary Immobility Indices

As we have already seen in the introduction, there are inherent problems in reducing a transition matrix into a single scalar index of immobility. To sort out the desirable properties of various immobility measures, the mobility ordering $\succcurlyeq_M$ might be used as a screening device in the following fashion: A real-valued immobility index $I(\cdot)$ defined on the set of mobility matrices is said to be *coherent* with $\succcurlyeq_M$ if $I(P) \leqslant I(Q)$ for any P and Q in $\mathscr{M}(\pi)$ such that $P \succcurlyeq_M Q$. This approach is similar in spirit to the use of the Lorenz curve ordering to screen "static" inequality measures (Atkinson [1]).

In the sequel, use is be made of the following two bistochastic regular monotone transition matrices:

$$A = \begin{bmatrix} 0.6 & 0.31 & 0.09 \\ 0.21 & 0.68 & 0.11 \\ 0.19 & 0.01 & 0.8 \end{bmatrix}; \qquad B = \begin{bmatrix} 0.6 & 0.32 & 0.08 \\ 0.22 & 0.66 & 0.12 \\ 0.18 & 0.02 & 0.8 \end{bmatrix}.$$

Individuals who start at the lowest initial position face a better income lottery under A, while individuals who start at the highest initial position face a better income lottery under B. Given $T'[A-B]\,T \leqslant 0$, an immobility index $I(\cdot)$ is not coherent when $I(A) > I(B)$.

Trace. The trace of transition matrix is sometimes employed as an immobility measure in the form $(\text{trace}(P)-1)/(n-1)$. It is criticized on the grounds that it pays no attention to the off-diagonal elements (Sommers and Conlisk [24]). In fact, this measure is not coherent, as follows by operating a regular monotone transition matrix with a D.P.D. transfer which increases any element along the main diagonal.

Determinant. The determinant of a transition matrix has been proposed as an immobility measure in the form $|P|^{1/(n-1)}$. Critics of this measure point out that it gives the completely mobile value when any two rows or columns are identical (Shorrocks [22]). This measure is also not coherent, given $|A|^{1/2} = 0.519 > |B|^{1/2} = 0.506$.

Second Largest Eigenvalue. The second largest eigenvalue modulus of a transition matrix has been proposed as an immobility measure and has been given interpretations in terms of the speed of escape from the initial conditions and in terms of regression to the mean (Theil [25], Shorrocks [22], and Sommers and Conlisk [24]). Conlisk [5] shows that the second largest eigenvalue modulus of a monotone transition matrix, P, denoted $E(P)$, is real and $0 < E(P) < 1$. The second largest eigenvalue for A, $E(A) = 0.691$, which is greater than the value for B, $E(B) = 0.685$. Thus, the index is not coherent.

Sommers and Conlisk [24] propose a closely related variant to $E(P)$, namely the second largest eigenvalue modulus of a "symmetrized" version of P, $\hat{P} = (1/2)(P + \Pi^{-1}P'\Pi)$. $\hat{P}$ is a transition matrix with the same equilibrium and immobility correlation as P, but with a simpler eigenvalue structure and better behaved near the extreme of perfect mobility. Sommers and Conlisk gave intergenerational correlation interpretations to $E(\hat{P})$, computed the two measures $E(P)$ and $E(\hat{P})$ for a set of 20 estimated intergenerational mobility matrices, and found that in each case the difference between the two indices was minimal. This is due to the approximate symmetry of ΠP for near all the estimated transition matrices. In Appendix 1 we show that coherence of $E(\hat{P})$ follows under monotonicity of the reverse chains and under some mild additional assumptions. Thus the modified second eigenvalue index will typically be coherent.

Bartholomew's Immobility Measure. Bartholomew's [3] index of mobility may be written as $\sum_i \sum_j \pi_i p_{ij} |i-j|$ and may be interpreted as the expected number of class boundaries crossed from one time to the next when the chain is in its steady state. The immobility index $[1/(n-1)](\sum_i \sum_j \pi_i p_{ij} |i-j|)$ is normalized to take values between zero and one. A D.P.D. transfer has a nonincreasing effect on $\sum_i \sum_j \pi_i p_{ij} |i-j|$. By Theorem 4, coherence follows.

Mean First Passage Time. Consider a steady-state Markov chain and let two individuals from the population be chosen at random. Conlisk [5] considers as an immobility measure the expected number of periods which must pass before the first individual achieves the state of the second individual. Letting M^P denote the mean first passage matrix (see, e.g., Kemeny and Snell [15]), the immobility measure is $\pi' M^P \pi$. $\pi' M^P \pi$ is not coherent, as shown by $\pi' M^A \pi = 0.652 > \pi' M^B \pi = 0.641$.

Lifetime Inequality Measures. Economists (e.g., Friesen and Miller [11]) have suggested the use of static inequality indices applied to the distribution of permanent income as measures of intertemporal equality of opportunities. From Theorem 4 it follows that any "reasonable" (i.e., S-convex) inequality measure will be coherent.

6. Related Orderings

6.1. *Welfare-Based Orderings*

Stochastic dominance rules specify unanimous preference for a given class of S.W.F. By considering different classes of S.W.F.'s, alternative stochastic dominance concepts may be obtained that imply a trade-off between the class of admissible S.W.F.'s and the strength of the conditions on the mobility matrices. A partial order defined on the set of monotone regular transition matrices with equal steady-state distribution is *finer* than $\succcurlyeq_M$ if $\succcurlyeq_M$ implies it (that is, if it orders all the mobility matrices that $\succcurlyeq_M$

orders). A natural way to obtain orderings finer than $\succcurlyeq_M$ is to consider a subset of our S.W.F. class.

In the literature on the measurement of income inequality, partial orderings finer than the Lorenz ordering are obtained by considering some kind of "transfer sensitivity axiom," which requires a given rich to poor income transfer to be more inequality reducing if performed at the lower end of the income distribution. This axiom was introduced by Kolm [17], who calls it the "principle of diminishing transfer" (see also Shorrocks and Foster [23] and Dardanoni and Lambert [6] for applications to inequality measurement). In this dynamic setting, a natural subset of the additive asymmetric S.W.F. with nonincreasing weights is obtained by considering those S.W.F.'s which: (i) satisfy Kolm's principle of diminishing transfer that, in our context, implies that greater weight is given to greater mobility at the lower levels so that the system of weights is decreasing at an increasing rate and (ii) insist that the utility vector u should increase at a decreasing rate. Our next theorem seeks the conditions on the transition matrices which ensure unanimous preference by all S.W.F.'s in this class:

THEOREM 5. *Let P and Q be two transition matrices in $\mathcal{M}(\pi)$ and let ρ be given. Then the following conditions are equivalent*:

5.i. $W(V^P, \lambda) \geqslant W(V^Q, \lambda)$ *for all nonincreasing λ with $\lambda' T^{-1'}$ nondecreasing and nondecreasing u with uT^{-1} nonincreasing*;

5.ii. $T'^2 \Pi[P(\rho) - Q(\rho)]\, T^2 \leqslant 0$.

Proof.

5.ii implies 5.i. Consider the obvious identity $\lambda' \Pi[P(\rho) - Q(\rho)]\, u = \lambda'(T')^{-2}\, T^{2'} \Pi[P(\rho) - Q(\rho)]\, T^2 T^{-2} u$ and note that: (i) the last two rows and columns of $T^{2'} \Pi[P(\rho) - Q(\rho)]\, T^2$ are equal to each other; (ii) the sum of the last two elements of $\lambda' T^{-2'}$ equals $(\lambda_{n-1} - \lambda_n) \geqslant 0$ and the sum of the last two elements of $T^{-2} u$ equals $(u_{n-1} - u_n) \leqslant 0$; (iii) the first $n-2$ elements of $\lambda' T^{-2'}$ are nonnegative and the first $n-2$ elements of $T^{-2} u$ are nonpositive. The result follows.

5.i implies 5.ii. Assume that the (i, j)th element of $T^{2'} \Pi[P(\rho) - Q(\rho)]\, T^2$ is positive. Then choose $\lambda_i = (i, i-1, \ldots, 1, 0, \ldots, 0)$ and $u_j = (-j, -j+1, \ldots, -1, 0, \ldots, 0)$ so that $\lambda_i'(T')^{-2}$ equals the $(1 \times n)$ vector with one in the ith position and zeros elsewhere and $T^{-2} u_j$ equals the $(n \times 1)$ vector with minus one in the jth position and zeros elsewhere, and we obtain the desired contradiction. Q.E.D.

By imposing further restrictions on the admissible class of S.W.F.'s, the above theorem provides a broader condition for unanimous ranking of transition matrices. In effect, condition 5.ii resembles a second order stochastic dominance ranking and may be compared with Atkinson's [2] results for the case of two period societies. In the two period case the condition requires restrictions on the sign of third and fourth order cross partial derivatives of the two argument utility function of the representative individual. In this formulation we require the joint restriction that the system of weights should satisfy Kolm's principle of diminishing transfer and the utility vector should decrease at an increasing rate.

Alternatively, one could consider weaker classes of S.W.F. to obtain orderings that imply $\succcurlyeq_M$. A weaker class than the one considered above is analyzed by Kanbur and Stiglitz [13], where given two bistochastic matrices P and Q, P is preferred to Q when $W(V^P) \geq W(V^Q)$ for all symmetric and quasi-concave real-valued $W(\cdot)$ and all utility vectors u. Under monotonicity, Kanbur and Stiglitz's ordering is equivalent to the ranking obtained with the additive asymmetric S.W.F. class when it is not required that the utility vector be nondecreasing.[3] Given that our class considers only a subset of all vectors u (i.e., only those nondecreasing), our class is more restrictive and therefore our condition $\succcurlyeq_M$ is weaker than Kanbur and Stiglitz's. However, it seems natural to insist on the unanimous preference of only nondecreasing vectors u, and therefore it may be argued that we have obtained a finer partial order at almost no cost.

6.2 *Non-Welfare-Based Orderings*

Shorrocks' [22] and Conlisk's [5] treatments of the measurement of mobility propose intuitively reasonable mobility criteria that induce partial orders over the set of transition matrices. While the motivation for these orderings is quite different from that of the welfare-based ordering $\succcurlyeq_M$, it is interesting to compare the relationship between their orderings and $\succcurlyeq_M$. According to Shorrock's *monotonicity axiom*, given two transition matrices P and Q of equal size, if $p_{ij} \geq q_{ij}$ for all $i \neq j$ (with strict inequality for some $i \neq j$), then any scalar measure of immobility $I(\cdot)$ should declare $I(P) < I(Q)$. In Appendix 2 we show that Shorrocks' ordering implies $\succcurlyeq_M$ on the set of transition matrices with an equal steady-state vector. That Shorrocks' monotonicity axiom is not implied by $\succcurlyeq_M$ can be seen by the following example, where P and Q are monotone bistochastic matrices:

$$P = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/6 & 2/3 & 1/6 \\ 1/6 & 1/6 & 2/3 \end{bmatrix}; \qquad Q = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix}.$$

[3] This may be shown by noting that Kanbur and Stiglitz's ranking is equivalent to Sherman's [21] partial ordering of bistochastic matrices: given two bistochastic matrices P and Q, Px majorizes Qx for all $x \in \Re^n$ if and only if there is a bistochastic matrix R such that $P = RQ$. In our context, under monotonicity we may rewrite the majorization condition as $T'[P(p) - Q(p)]\, u \geq 0$ for all u; if we insist that u be nondecreasing we get 1.ii.

Poorer individuals face a better lottery under Q, and richer individuals face a better lottery under Q than under P. P is welfare superior according to $\succcurlyeq_M$. However P and Q cannot be ordered by Shorrock's partial ordering.

Building on Shorrocks' framework, Conlisk [5] proposes a new mobility criterion, called the *D-criterion*, which induces a partial order on the set of monotone transition matrices. Denote by $D(P)$ the $(n-1)\times(n-1)$ matrix gotten by deleting the last row and column of $T^{-1}PT$. According to Conlisk's D-criterion, given two monotone transition matrices P and Q, Q displays at least as much immobility as P if $D(Q)\geqslant D(P)\geqslant 0$. Thus if the D-criterion holds, the matrix Q is " more monotone" than the matrix P; this intuitively means that the "differential advantage" (in terms of stochastic dominance) of an individual who is in class $i+1$ over an individual in class i is greater under Q for all income classes $i=1, ..., n-1$.

In Appendix 2 we show that Conlisk's D-criterion implies $\succcurlyeq_M$ on $\mathcal{M}(\pi)$; however, Conlisk's ordering is not implied by $\succcurlyeq_M$, as shown by the following example, where P and Q are monotone bistochastic matrices and ε is a small positive scalar:

$$P=\begin{bmatrix} 1/3+\varepsilon & 1/3 & 1/3-\varepsilon \\ 1/3 & 1/3 & 1/3 \\ 1/3-\varepsilon & 1/3 & 1/3+\varepsilon \end{bmatrix}; \qquad Q=\begin{bmatrix} 1-2\varepsilon & \varepsilon & \varepsilon \\ \varepsilon & 1-2\varepsilon & \varepsilon \\ \varepsilon & \varepsilon & 1-2\varepsilon \end{bmatrix}.$$

Here, when ε tends to zero P tends to the perfect mobility matrix and Q to the identity matrix, and $P\succcurlyeq_M Q$. However, P and Q cannot be ordered by Conlisk's partial ordering, given $D(P)=\left[\begin{smallmatrix}\varepsilon & \varepsilon \\ \varepsilon & \varepsilon\end{smallmatrix}\right]$ and $D(Q)=\left[\begin{smallmatrix}1-3\varepsilon & 0 \\ 0 & 1-3\varepsilon\end{smallmatrix}\right]$.

Shorrocks' and Conlisk's orderings are very intuitive and are applicable to transition matrices which do not necessarily have a common equilibrium vector. It is worth noting that because the welfare-based ordering $\succcurlyeq_M$ is finer than both these orderings, it will be more selective as a screening device for immobility measures. For example, it is easy to check that the trace is a coherent immobility index with both Shorrocks' and Conlisk's criteria, while it is not coherent under $\succcurlyeq_M$, as shown in the previous section.

7. Concluding Remarks

(1) In this paper we have considered the ranking of mobility matrices by deriving the lifetime welfare prospects under different transition mechanisms and aggregating them with a weighted linear S.W.F. which gives greater weight to individuals starting at a lower position. By considering the unanimous preference for this class of S.W.F.'s we have

derived a robust partial ordering which emphasizes intertemporal equality of opportunity.

The multiperiod framework allows the consideration of both intergenerational and intragenerational mobility. The linearity of the S.W.F. and the monotonicity assumption make the proofs of the theorems transparent; the derived robust rankings are very easy to apply in practice, essentially involving only simple matrix multiplication. The linearity of the problem has the further advantage that the rankings obtained are dynamically consistent and identical to those obtained considering the expected value of an intertemporal ex post S.W.F. defined on realized income distributions. Therefore, this approach is consistent with both ex ante and ex post approaches to social decision making under uncertainty.

(2) This approach may be considered as the intertemporal counterpart to the static inequality ranking of income distributions by the Lorenz curve. The seminal papers by Kolm [16], Atkinson [1], and Dasgupta, Sen, and Starrett [7] justify the Lorenz partial ordering with welfare-based considerations, show the equivalence of the ordering with some intuitive concepts of what can be regarded as a more unequal income distribution, and use the ordering to screen commonly employed inequality measures. Similarly, we have motivated our ordering by welfare considerations, shown the equivalence of the derived ordering with some intuitive concepts of what can be regarded as a more mobile society in terms of intertemporal fairness, and employed the derived ordering to screen immobility measures for coherence with the ordering. The equivalence of our ranking with the "permanent income" Lorenz ranking, as shown in Theorem 4, gives support to the claim that this approach is the natural extension of Kolm's [16], Atkinson's [1], and Dasgupta, Sen, and Starrett's [7] approaches.

(3) We have considered societies which have a common steady-state income distribution vector. In this case, societies display identical snapshot inequality, but social welfare is influenced by the exchange of relative positions over time. This approach ranks pure mobility, abstracting to what is called "structural mobility," which refers to the change of available positions in the social ladder over time. Again, this approach is identical in spirit to the static Lorenz ranking, where to abstract from efficiency considerations one compares income distributions with equal average income.

(4) In practice, actual static income distributions are likely to have different average incomes. Analogously, actual societies may not be in steady state and different mobility matrices might imply different equilibrium distributions. Lorenz curves for static income distributions are routinely calculated in the applied income inequality literature, even if the

welfare interpretation of the derived ranking may be doubious. The justification for the use of the Lorenz ranking in the case of different mean incomes is often given in terms fo unequal *relative* shares of the available resources. Similarly, to concentrate on the exchange of relative positions over time it is often suggested to consider *fractile* Markov chains (e.g., Geweke, Marshall, and Zarkin [12] and Kanbur and Stiglitz [15]). A Markov chain is fractile if for all t and all $i = 1, 2, ..., n$ we have $\pi_{ti} = n^{-1}$. The transition matrix for a regular fractile chain is bistochastic and has unique equilibrium vector with equal number of individuals in each income state. An individual who at time t is in state i is in the poorest ith fractile of the population; thus the values of the states of the chain will have to be interpreted as "relative" position in society.

(5) Alternatively, to eliminate the effect of structural mobility on the static income distribution over time, one could consider continuous time Markov chains, where the transition probabilities are governed by the system of differential equations $dP(t)/dt = RP(t)$, $P(0) = I$, where $p_{ij}(t)$ denotes the probability that an individual starting in class i will be in class j at time t, and R is the $(n \times n)$ "intensity matrix" (see Doob [9] for a general treatment of continuous time Markov processes and Geweke, Marshall, and Zarkin [12] for an application to mobility measurement. Continuous time Markov chains are more appropriate for the analysis of intragenerational, rather than intergenerational, mobility). The mobility ordering $\succcurlyeq_M$ proposed here may be employed to rank intensity matrices of different societies. The ordering must then be reinterpreted in terms of the "instantaneous" equalizing effect of exchange mobility on social welfare, on the Lorenz curve, and so on.

Appendix 1

Under monotonicity of the reverse chains, $\hat{P}$ and $\hat{Q}$ are monotone, and from $T'\Pi(P-Q)T \leqslant 0$ it follows that $T'\Pi(\hat{P}-\hat{Q})T \leqslant 0$. Thus, by Theorem 4, $\hat{P}$ may be obtained from $\hat{Q}$ by a finite sequence of D.P.D. transfers. Consider the following differentiable parametrization of $\hat{Q}$, $\hat{Q}(s) = \hat{Q} + sH$, where s is a positive scalar and H arrays a D.P.D. transfer. Denote by $E[\hat{Q}(s)]$ the second largest eigenvalue modulus of $\hat{Q}(s)$ such that $E[\hat{Q}(0)] = E(\hat{Q})$, and let $x(s)$ denote the eigenvector corresponding to $E[\hat{Q}(s)]$, normalized so that $x'(s)\,x(s) = 1$, and let x denote the eigenvector of $\hat{Q}$ corresponding to $E(\hat{Q})$, so that $x = x(0)$. Then coherence follows if $E[\hat{Q}(s)]$ is increasing in s for small s.

For small s, differentiate $E[\hat{Q}(s)] = x'(s)\,\hat{Q}(s)x(s)$ to get

$$\frac{dE[\hat{Q}(s)]}{ds} = \frac{dx'(s)}{ds}\hat{Q}(s)\,x(s) + x'(s)\frac{d\hat{Q}(s)}{ds}x(s) + x'(s)\,\hat{Q}(s)\frac{dx(s)}{ds}$$

and differentiate the normalization condition $x'(s)\,x(s)=1$ to get the identity

$$\frac{dx'(s)}{ds}\,x(s)+x'(s)\,\frac{dx(s)}{ds}=0.$$

Rearranging and using the equations $\hat{Q}(s)\,x(s)=E[\hat{Q}(s)]\,x(s)$ and $x'(s)\,\hat{Q}(s)=E[\hat{Q}(s)]\,x'(s)$, we get

$$\frac{dE[\hat{Q}(s)]}{ds}=x'(s)\,\frac{d\hat{Q}(s)}{ds}\,x(s).$$

At $s=0$, we have $dE[\hat{Q}(s)]/ds|_{s=0}=x'Hx'$.

Conlisk [5] shows that monotonicity forces the eigenvector corresponding to the second largest eigenvalue modulus to be strictly increasing under the mild additional assumptions that both Q and $D(Q)$ are primitive. Conlisk notes that these conditions virtually always hold for mobility matrices. Given that x is strictly increasing, from the definition of the perturbation matrix H, it follows that $x'Hx>0$. Thus, under monotonicity of the reverse chain (along with the additional conditions from above) we establish coherence of the modified second eigenvalue index.

Appendix 2

Let P and Q be two transition matrices in $\mathcal{M}(\pi)$. To demonstrate that Shorrocks' ordering implies $\succcurlyeq_M$, consider the (i,j)th element of $T'\Pi PT$ when $i\leqslant j<n$,

$$\sum_{t=1}^{i}\pi_t\left(\sum_{s=1}^{j}p_{ts}\right)=\sum_{t=1}^{i}\pi_t\left(1-\sum_{s=j+1}^{n}p_{ts}\right)=\sum_{t=1}^{i}\pi_t-\sum_{t=1}^{i}\sum_{s=j+1}^{n}\pi_t p_{ts},$$

while, when $j<i<n$, we have

$$\sum_{s=1}^{j}\left(\sum_{t=1}^{i}\pi_t p_{ts}\right)=\sum_{s=1}^{j}\left(\sum_{t=1}^{n}\pi_t p_{ts}-\sum_{t=i+1}^{n}\pi_t p_{ts}\right)=\sum_{s=1}^{j}\pi_s-\sum_{s=1}^{j}\sum_{t=i+1}^{n}\pi_t p_{ts}.$$

When $i\leqslant j<n$, the (i,j)th element of $T'\Pi(P-Q)\,T$ equals $\sum_{t=1}^{i}\sum_{s=j+1}^{n}\pi_t(q_{ts}-p_{ts})$, while it equals to $\sum_{s=1}^{j}\sum_{t=i+1}^{n}\pi_t(q_{ts}-p_{ts})$ when $j<i<n$. Note that the last row and column of $T'\Pi(P-Q)\,T$ consist of zeros. Thus, it follows that if $p_{ij}\geqslant q_{ij}$ for all $i\neq j$ we have $T'\Pi(P-Q)\,T\leqslant 0$.

To demonstrate that Conlisk's ordering implies $\succcurlyeq_M$ on $\mathcal{M}(\pi)$, consider first the identity

$$(P-Q)\,T = T\begin{bmatrix} [D(P)-D(Q)] & 0 \\ [(P_n-Q_n)\,T]_{n-1} & 0 \end{bmatrix},$$

where P_n and Q_n denote the last row of P and Q, respectively, and $[(P_n-Q_n)\,T]_{n-1}$ denotes the first $n-1$ elements of $(P_n-Q_n)\,T$. Premultiplying this identity by π' and using $\pi'=\pi'P=\pi'Q$ we get

$$\pi'(P-Q)\,T = \pi'T\begin{bmatrix} [D(P)-D(Q)] & 0 \\ [(P_n-Q_n)\,T]_{n-1} & 0 \end{bmatrix} = 0,$$

from which we derive $[(P_n-Q_n)\,T]_{n-1} = -\pi'_{n-1}T_{n-1}[D(P)-D(Q)]$, where π_{n-1} denotes the first $n-1$ elements of π and T_{n-1} denotes the first $n-1$ rows and columns of T.

That $[D(P)-D(Q)]\leqslant 0$ implies $T'\Pi(P-Q)\,T\leqslant 0$ is shown by the inequalities

$$\begin{aligned} T'\Pi(P-Q)\,T &= T'\Pi T T^{-1}(P-Q)\,T \\ &= \begin{bmatrix} T'_{n-1}\Pi_{n-1}T_{n-1} & T'_{n-1}\pi_{n-1} \\ \pi'_{n-1}T_{n-1} & 1 \end{bmatrix}\begin{bmatrix} [D(P)-D(Q)] & 0 \\ -\pi'_{n-1}T_{n-1}[D(P)-D(Q)] & 0 \end{bmatrix} \\ &= \begin{bmatrix} (T'_{n-1}\Pi_{n-1}T_{n-1} - T'_{n-1}\pi_{n-1}\pi'_{n-1}T_{n-1})[D(P)-D(Q)] & 0 \\ 0 & 0 \end{bmatrix} \\ &\leqslant 0, \end{aligned}$$

where Π_{n-1} denotes the first $n-1$ rows and columns of Π and the last inequality may be deduced from Theorem 2.

References

1. A. B. Atkinson, On the measurement of inequality, *J. Econ. Theory* **2** (1970), 244–263.
2. A. B. Atkinson, The measurement of economic mobility, *in* "Social Justice and Public Policy," MIT Press, Cambridge, MA, 1983.
3. D. J. Bartholomew, "Stochastic Models for Social Processes," 3rd ed., Wiley, New York, 1982.
4. J. Conlisk, Ranking mobility matrices, *Econ. Letters* **29** (1989), 231–235.
5. J. Conlisk, Monotone mobility matrices, *J. Math. Sociol.* **15** (1990), 173–191.
6. V. Dardanoni and P. J. Lambert, Welfare ranking of income distributions: A role for the variance and some insights for tax reform, *Soc. Choice Welf.* **5** (1988), 1–17.
7. P. Dasgupta, A. K. Sen, and D. Starrett, Notes on the measurement of inequality, *J. Econ. Theory* **6** (1973), 180–187.
8. P. A. Diamond, Cardinal welfare, individualistic ethics and interpersonal comparisons of utility: A comment, *J. Polit. Econ.* **75** (1965), 765–766.

9. J. L. Doob, "Stochastic Processes," Wiley, New York, 1953.
10. M. Friedman, "Capitalism and Freedom," Univ. Chicago Press, Chicago, 1962.
11. P. H. Friesen and D. Miller, Annual inequality and lifetime inequality, *Quart. J. Econ.* **98** (1983), 139–155.
12. J. Geweke, R. C. Marshall, and G. A. Zarkin, Mobility indices in continuous time Markov chains, *Econometrica* **54** (1986), 1407–1423.
13. S. M. R. Kanbur and J. E. Stiglitz, "Intergenerational Mobility and Dynastic Inequality," Woodrow Wilson Discussion paper No. 111, Princeton University, 1986.
14. J. Keilson and A. Kester, Monotone matrices and monotone Markov chains, *Stochastic Process. Appl.* **5** (1977), 231–241.
15. J. G. Kemeny and J. L. Snell, "Finite Markov Chains," Springer, New York, 1976.
16. S.-Ch. Kolm, The optimal production of social justice, *in* "Public Economics" (J. Margolis and H. Guitton, Eds.), MacMillan, London, 1969.
17. S.-Ch. Kolm, Unequal inequalities II, *J. Econ. Theory* **13** (1976), 82–111.
18. A. Markandya, The welfare measurement of changes in economic mobility, *Economica* **51** (1984), 457–471.
19. S. J. Prais, Measuring social mobility, *J. Roy. Statist. Soc. Ser. A* **118** (1955), 56–66.
20. A. K. Sen, "On Economic Inequality," Clarendon Press, Oxford, 1973.
21. S. Sherman, On a conjecture concerning doubly stochastic matrices, *Proc. Amer. Math. Soc.* **3** (1952), 350–353.
22. A. F. Shorrocks, The measurement of mobility, *Econometrica* **46** (1978), 1013–1024.
23. A. F. Shorrocks and J. Foster, Transfer sensitive inequality measures, *Rev. Econ. Stud.* **54** (1987), 485–497.
24. P. M. Sommers and J. Conlisk, Eigenvalue immobility measures for Markov chains, *J. Math. Sociol.* **6** (1978), 253–276.
25. H. Theil, "Statistical Decomposition Analysis," North-Holland, Amsterdam, 1972.