Soc Choice Welfare (1995) 12: 181-192

Social Choice and Welfare

Income distribution dynamics: monotone Markov chains make light work

Valentino Dardanoni*, **

Università di Palermo, 90128 Palermo, Italy

Received: 20 December 1993/Accepted: 4 July 1994

Abstract. This paper considers some aspects of the dynamics of income distributions by employing a simple Markov chain model of income mobility. The main motivation of the paper is to introduce the techniques of "monotone" Markov chains to this field. The transition matrix of a discrete Markov chain is called monotone if each row stochastically dominates the row above it. It will be shown that by embedding the dynamics of the income distribution in a monotone Markov chain, a number of interesting results may be obtained in a straightforward and intuitive fashion.

This paper analyses some aspects of the dynamics of income distributions by employing a simple Markov chain model of income mobility. The main motivation of the paper is to introduce the techniques of "monotone" Markov chains to this field. The transition matrix of a discrete Markov chain is called monotone if each row stochastically dominates the row above it. Monotone transition matrices are defined and analyzed in Keilson and Kester (1977) and Conlisk (1990). Even though monotonicity is an ideal assumption to impose on a Markov chain of income mobility, being theoretically plausible, empirically supported and having a number of very convenient mathematical properties, monotone Markov chains do not seem to be widely known and employed by economists. It will be shown that by embedding the dynamics of the income distribution in a monotone Markov chain, a number of interesting results may be obtained in a straightforward and intuitive fashion.

The Markov chain framework is the workhorse of many theoretical and empirical studies of social and economic mobility: see e.g. Kemeny and Snell (1976)

* Address for correspondence: Valentino Dardanoni, Istituto di Scienze Finanziarie, Facoltà di Economia e Commercio, Università di Palermo, 90128 Palermo, Italy, e-mail: vdardano@cere03.cere.pa.cnr.it.

** Thanks are due to Peter Lambert and four anonymous referees for very helpful comments.

for a simple but rigorous introduction to Markov chains, Bartholomew (1982) for a review of Markov chain models of social mobility and Atkinson, Bourguignon and Morrisson (1992) for applications to earnings mobility. Though the results presented in this paper are relevant to various intergenerational and intragenerational mobility contexts, it is convenient for presenting the results to specialize the discussion to the dynamics of the income distribution, where the states of the Markov chain denote income classes arranged in an increasing order. Note however that if the model is to be applied in a social or occupational mobility context, where states are occupational categories, care should be taken to define the states acording to an increasing social status scale. In fact, whenever the Markovian states are not ordered, the stochastic dominance concepts employed here would be meaningless.

It is widely recognized that there are theoretical reasons for believing that the Markovian model may not hold exactly for income and social mobility processes. For example, even if the Markovian property might hold for a given classification of the states, we cannot in general rearrange the classes and retain the Markov property (Kemeny and Snell (1976)). In practice, social and income classes are drawn arbitrarily, and empirical studies (see e.g. Atkinson, Bourguignon and Morisson (1992)) tend to reject the Markov assumption. For example, Shorrocks (1976) rejects the first order Markov property, suggesting that transition probabilities may depend on past history beyond the immediately preceeding period. However, it is possible that rejection of the Markov property depends on the chosen income class boundaries; moreover, by appropriate redefinition of the Markov states, a second (or higher) order chain may be easily transformed to an equivalent first-order chain (see e.g. Billingsley (1961) for details). In short, as argued for example by Bartholomew (1982), it is sufficient that the Markov chain model embodies the main features of the income and social dynamics without being correct in every detail, since the model may be sufficiently near to reality to justify its further use and development.

The next section of this paper will lay down the formal framework and introduce the main definitions needed in the sequel. Section three will consider the conditions which ensure that a dominance ranking of income distributions is maintained in the future, given that it holds in the present. This analysis may be considered as a discrete counterpart of Kanbur and Stromberg (1988); the discrete framework and the monotonicity assumption will hopefully make the analysis intuitively clear and the conditions themselves easy to derive. Section four analyzes the behavior of the income distribution over time and its relationship with the transition process under operation. Conditions will be given for the income distribution vectors to form a stochastically increasing sequence over time. Section five employs Conlisk's (1985) work on the comparative statics of discrete Markov chains to show how monotonicity may help signing the effect of various changes in the transition matrix on the equilibrium income distribution. Section six considers some aspects of the relationship between the mobility process and lifetime income. It is usually argued that permanent income is a better measure of the differences in opportunities faced by individuals in society than a single observation corresponding to a particular period of time. We consider the conditions which ensure that a reduction in inequality in the static distribution of income, for example owing to a change in the tax system, will imply a reduction in inequality in the distribution of lifetime income. Section seven proves formally the intuitive idea that a more mobile society changes the initial income distribution more rapidly than a less mobile society, using the mobility partial ordering introduced in Dardanoni (1993). The last section contains some concluding remarks.

1. Formal framework and definitions

Consider a discrete Markov chain with constant $n \times n$ transition matrix P; the typical element P_{ij} is the probability that an individual in state i will be in state j in the following period. Clearly, $P \geq 0$ and $P1 = 1$, where 1 is the $n \times 1$ vector of ones (here and hereafter, when an inequality symbol involves vectors and matrices, we mean that the inequality is satisfied elementwise). P is assumed regular, which means that for large enough integer k, P^k is strictly positive (P^k_{ij} denotes the probability that an individual goes from state i to state j in k periods), so that when P is regular, after some number of transitions it is possible to be in any state no matter what the initial state. Regularity of P implies that the $n \times 1$ equilibrium probability vector π^e exists and is the unique solution to $\pi^{e'} = \pi^{e'}P$. We assume that transitions are independent across individuals, and P is constant over time.

It is convenient for presenting our results to specialize the discussion to an income mobility context, with y denoting an $n \times 1$ vector whose components measure income in state $i = 1, 2, \ldots, n$. Adopting the convention that income states are ordered from worst to best, we let $y_1 \leq y_2 \leq \cdots \leq y_n$. The equilibrium distribution of income is the pair $[\pi^e, y]$; in equilibrium an individual chosen at random will have income y_i with probability π^e_i. Clearly, if the model is to be applied in a social or occupational mobily context, care should be taken to rank the "social states" in an increasing order.

The probability vector $\pi(t)$ denotes the (expected) "spot" income distribution at time t, where $\pi(t)_i$ gives the proportion of individuals in state i at time t, and $\pi(t)' = \pi(0)'P^t$. A probability vector π stochastically dominates $\hat{\pi}$ ($\pi \succeq \hat{\pi}$) if $\pi_1 + \pi_2 + \cdots + \pi_k \leq \hat{\pi}_1 + \hat{\pi}_2 + \cdots + \hat{\pi}_k$ for all $k = 1, 2, \ldots, n - 1$. A sequence of spot income distributions is stochastically increasing if $\pi(s + 1) \succeq \pi(s)$ for all $s = 0, 1, 2, \ldots$. We say that a transition matrix P stochastically dominates $\hat{P}$ if $\pi'P \succeq \pi'\hat{P}$ for any income distribution vector π. Abusing notation, we will also denote matrix dominance as $P \succeq \hat{P}$.

A transition matrix is called monotone if each row stochastically dominates the row above it. 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, monotonicity implies that an individual who at time t is in class $i + 1$ faces a better lottery, in terms of stochastic dominance, than an individual in class i. Monotone mobility matrices are also called "order-preserving": if we let v be an $n \times 1$ vector it may be shown that Pv is nondecreasing for all nondecreasing v if and only if P is monotone (see e.g. Keilson and Kester's (1977) Theorem 1). As argued by Conlisk (1990), monotonicity is an ideal assumption to impose on a Markov chain of social mobility, being theoretically plausible and empirically supported. Monotone transition matrices posses several simplifying mathematical properties, reviewed by Keilson and Kester (1977) and Conlisk (1990); and estimated transition matrices are usually either exactly monotone or within sampling errors from being monotone, see for example Dardanoni and Forcina (1993) for a statistical analysis of monotone intergenerational transition matrices.

Before stating our results, we need to define the summation matrix T that will be crucial in the derivation of much of what follows: T will denote 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' will be lower triangular, and its inverse $(T')^{-1}$ has ones on the main diagonal, minus ones on the first subdiagonal and zeros elsewhere. 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 an $n \times 1$ vector y by T^{-1} differences the elements of y. Use of the summation matrix T implies the possibility of writing many of our definitions and results compactly: for example, the condition that P is monotone can be written as $T^{-1}PT \geq 0$; the condition $\pi \succeq \hat{\pi}$ may be written as $\pi' T \leq \hat{\pi}' T$. Given $n \times 1$ vectors x and y, the identity $x'y = x'TT^{-1}y$, which is a sort of summation by parts, will be exceedingly handy in the demonstration of most of the following results. For example, the well known fact that $\pi' u \geq \hat{\pi}' u$ for all increasing vectors u is equivalent to $\pi' T \leq \hat{\pi}' T$ (utilitarian social welfare is greater under π for any increasing utility vector if and only if $\pi \succeq \hat{\pi}$), admits the following very straight forward proof: $\pi' u \geq \hat{\pi}' u$ can be written as $[(\pi - \hat{\pi})' T][T^{-1}u] \geq 0$; u nondecreasing implies the first $n-1$ elements of $[T^{-1}u] \geq 0$; $\pi \succeq \hat{\pi}$ implies the first $n-1$ elements of $(\pi - \hat{\pi})' T \geq 0$ (the last element is zero); when u equals zero in the first j components and one in the last $n-j$, the first $(n-1)$ elements of the vector $[T^{-1}u]$ are zero everywhere except the jth element which equals minus one. Sufficiency follows immediately, while necessity follows by letting j vary from 1 to n in the above construct.

These simple matrix manipulations will be essential to make the proofs of the results easy to derive and illustrative.

2. Income distribution dominance

Consider two societies following a discrete Markov process with constant $n \times n$ transition matrices P and $\hat{P}$, and let $\pi(t)$ and $\hat{\pi}(t)$ be the spot income distributions at time t. Our first question is the following: Under which conditions is a dominance relation $\succeq$ between two spot distributions $\pi(t)$ and $\hat{\pi}(t)$ preserved at time $t+1$? In other words, are there necessary and sufficient conditions on the transition mechanisms such that $\pi(t) \succeq \hat{\pi}(t)$ implies $\pi(t+1) \succeq \hat{\pi}(t+1)$? A similar question has been analysed by Kanbur and Stromberg (1988) in the context of continuous income distributions under the Lorenz curve ordering. It will be shown that the discrete framework and the monotonicity assumption simplify greatly the derivation and interpretation of the results.

The following Lemmas, due to Keilson and Kester (1977), will be useful for much of what follows:

Lemma 1. *Let $\pi(t+1)' = \pi(t)'P$ and $\hat{\pi}(t+1)' = \hat{\pi}(t)'P$, where P is the common transition matrix. Then the following conditions are equivalent:*

(1) P is monotone;
(2) $\pi(t) \succeq \hat{\pi}(t)$ implies $\pi(t+1) \succeq \hat{\pi}(t+1)$ for all $\pi(t)$ and $\hat{\pi}(t)$.

Proof. [(2) implies (1)]: Rewrite $\pi(t) \succeq \hat{\pi}(t)$ as $\pi(t)' T \leq \hat{\pi}(t)' T$ and $\pi(t+1) \succeq \hat{\pi}(t+1)$ as $\pi(t)' PT \leq \hat{\pi}(t)' PT$. Assume that P is not monotone, i.e. there is at least one element, say the (i,j)th, of $T^{-1}PT$ which is strictly negative. Choose $\hat{\pi}(t) - \pi(t) = e_i - e_{i+1}$, where e_i is the ith unit vector (i.e., the $n \times 1$ vector with one in the ith position and zeros elsewhere). Then $\pi(t)' T - \hat{\pi}(t)' T = e_i$ and the jth element of the vector $(\hat{\pi}(t)-\pi(t))'PT=(\hat{\pi}(t)-\pi(t))'T(T^{-1}PT)$ is negative, a contradiction.

[(*1*) implies (*2*)]: $\pi(t)'PT=\pi(t)'T(T^{-1}PT)\leq\hat{\pi}(t)'T(T^{-1}PT)=\hat{\pi}(t)'PT$, where the inequality follows from the fact that $(\pi(t)-\hat{\pi}(t))'PT\leq 0$ and the monotonicity of P. □

Lemma 2. *If P is monotone, so are P^k, $k=0,1,\ldots$*

Proof. $T^{-1}P^kT=T^{-1}PP\ldots PT=(T^{-1}PT)(T^{-1}PT)\ldots(T^{-1}PT)$. But under monotonicity $T^{-1}PT\geq 0$, which implies $T^{-1}P^kT\geq 0$, and the Lemma is proved. □

As argued above, estimated mobility matrices are typically monotone, or within sampling errors from being monotone. Therefore, Lemma 1 tells us that a stochastic dominance relation between two societies which follow a common transition mechanism is likely to be maintained as time unfolds.

Let us consider now the case where P is monotone, but the transition mechanisms are different in the two societies. We need first the following:

Lemma 3. *Let P and $\hat{P}$ be two given transition matrices. Then the following two conditions are equivalent:*

(*1*) $P\succeq\hat{P}$;
(*2*) $PT\leq PT$.

Proof.

[(*1*) implies (*2*)]: Rewrite $\pi'P\succeq\pi'\hat{P}$ for all π as $\pi'PT\leq\pi'\hat{P}T$ for all π and choose π equal to the ith unit vector ($i=1,2,\ldots,n$) to give $PT\leq\hat{P}T$.

[(*2*) implies (*1*)]: Immediate from the definition of $P\succeq\hat{P}$. □

The dominance condition $PT\leq\hat{P}T$ may be given the following interpretation: suppose an individual has to choose either of the two societies P or $\hat{P}$ to live in, and assume he is not allowed to know in advance in which income class he would be. Then when $PT\leq\hat{P}T$, he would face a better income lottery (in terms of stochastic dominance) under P for any initial income class.

We are in a position now to state the following:

Proposition 1. *Let P and $\hat{P}$ be two given transition matrices, and let P be monotone. The following two conditions are equivalent:*

(*1*) $P\succeq\hat{P}$;
(*2*) $\pi(t)\succeq\hat{\pi}(t)$ *implies* $\pi(t)'P\succeq\hat{\pi}(t)'\hat{P}$ *for all* $\pi(t)$ *and* $\hat{\pi}(t)$.

Proof. [(*1*) implies (*2*)]: Rewrite $\pi(t)'P\succeq\hat{\pi}(t)'\hat{P}$ as $\pi(t)'PT-\hat{\pi}(t)'\hat{P}T=(\pi(t)-\hat{\pi}(t))'T(T^{-1}PT)+\hat{\pi}(t)'(P-\hat{P})T$, and note that under monotonicity $(\pi(t)-\hat{\pi}(t))'T(T^{-1}PT)\leq 0$ and $\hat{\pi}(t)'(P-\hat{P})T\leq 0$ using Lemma 3.

[(*2*) implies (*1*)]: Set $\pi(t)=\hat{\pi}(t)$ in (*2*), and use the definition of $P\succeq\hat{P}$. □

The above result tells us that when the transition matrix of the society with a dominating income distribution is monotone, a necessary and sufficient condition for the preservation of the dominance relation in the following snapshot is stochastic dominance of the transition matrix.

Under monotonicity, from Proposition 1 one might expect that when P dominates $\hat{P}$ the steady state income distribution $\pi^{e\prime}=\pi^{e\prime}P$ will also dominate the steady state distribution $\hat{\pi}^{e\prime}=\hat{\pi}^{e\prime}\hat{P}$, because if the dominance relation $\succeq$ holds at each period it will hold in the limit as well. A formal proof that $P \succeq \hat{P}$ implies $\pi^e \succeq \hat{\pi}^e$ when the matrix P is monotone is given in Conlisk's (1992; page 176) .

A further appreciation of the properties of monotone chains may be obtained by the following example: let P and $\hat{P}$ be the following (non monotone) transition matrices:

$$P = \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.3 & 0.4 & 0.3 \\ 0.5 & 0.4 & 0.1 \end{bmatrix}; \qquad \hat{P} = \begin{bmatrix} 0.1 & 0.5 & 0.4 \\ 0.3 & 0.5 & 0.2 \\ 0.5 & 0.5 & 0.0 \end{bmatrix}.$$

The equilibrium vectors are $\pi = (0.3, 0.4, 0.3)$ and $\hat{\pi} = (0.286, 0.5, 0.214)$; thus π does not dominate $\hat{\pi}$, even though $P \succeq \hat{P}$. Failure of monotonicity may imply the rather paradoxical result that even though in a society at every time period each income class faces a better income lottery than in another society, still in equilibrium the income distribution of the first society does not dominate that of the second one.

3. The sequence of income distribution vectors over time

We consider now the behavior of the spot distribution $\pi(t)$ over time, and in particular the conditions under which $\pi(k+1) \succeq \pi(k)$ for all k, so that the income distribution vectors form a stochastically increasing sequence over time. The following result gives conditions for intertemporal dominance:

Proposition 2. *Let $\pi(k)' = \pi(0)' P^k$, with P monotone. Then $\pi(k+1) \succeq \pi(k)$ for all $k = 0, 1, \ldots$ if and only if $\pi(1) \succeq \pi(0)$.*

Proof. Sufficiency follows immediately from Lemmas 1 and 2; necessity is trivial. □

Under monotonicity, a simple condition for the sequence of "spot" income distributions to be stochastically increasing is that $\pi(1)$ dominates $\pi(0)$. Intuitively, the chain will be monotonically increasing over time if it does not start with a "too rich" income distribution. For example, under monotonicity it follows that when the chain starts with all the population belonging to the poorer (richer) class, society's income distribution will get stochastically richer (poorer) over time.

A more precise characterization on the initial income distribution which ensures a stochastically increasing sequence of income distributions over time may be obtained, following Keilson and Kester (1977), by linking the initial income vector to the steady state income distribution. We need now the following definition: A transition matrix P is tridiagonal if $P \geq 0$, $P1 = 1$ and in addition $P_{ij} = 0$ for $|i - j| > 1$. Tridiagonal transition matrices arise quite naturally when the transition period is short, and each individual can move (up or down) each period only to adjacent states. Tridiagonal transition matrices are extensively studied under what are called "birth and death" stochastic processes. Denote by $D(\pi^e)$ the diagonal matrix with π^e on the diagonal. We have the following:

Proposition 3. *Let P be a regular monotone transition matrix. A sufficient condition for $\pi(k+1) \succeq \pi(k)$ for all $k = 0, 1, \ldots$ is that $\pi(0)_i / \pi_i^e$ is nonincreasing in $i = 1, 2, \ldots, n$. If P is in addition tridiagonal, the condition is also necessary.*

Proof. [Sufficiency]: From Proposition 2, it follows that it suffices to show that $\pi(0)_i / \pi_i^e$ nonincreasing in i implies $\pi(1) \succeq \pi(0)$, i.e. $\pi(0)' PT \leq \pi(0)' T$. $\pi(0)_i / \pi_i^e$

nonincreasing may be rewritten as $\pi(0)'D(\pi^e)^{-1}$ nonincreasing, which is equivalent to $\pi(0)'D(\pi^e)^{-1}(T')^{-1} \geq 0$. Given the identities $\pi(0)'PT = \pi(0)'D(\pi^e)^{-1}(T')^{-1}T'D(\pi^e)PT$ and $\pi(0)'T = \pi(0)'D(\pi^e)^{-1}(T')^{-1}T'D(\pi^e)T$, to show $\pi(0)'PT \leq \pi(0)'T$ (that is, $\pi(0)'PT - \pi(0)'T = \pi(0)'D(\pi^e)^{-1}(T')^{-1}[T'D(\pi^e)(P-I)T] \leq 0$) it suffices to show that $T'D(\pi^e)(P-I)T \leq 0$.

Note that the last row and column of $T'D(\pi^e)(P-I)T$ consist of zeros; the (i,j)th element, when $i \leq j < n$, equals:

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

while, when $n > i > j$ we have:

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

[Necessity]: Suppose $\pi(0)'PT \leq \pi(0)'T$. Then $\pi(0)'D(\pi^e)^{-1}(T')^{-1}[T'D(\pi^e)(P-I)T] \leq 0$. For any matrix P with equilibrium vector π^e, it has just been shown that $T'D(\pi^e)(P-I)T \leq 0$. However, when P is tridiagonal, employing the steady-state equation $\pi^{e'} = \pi^{e'}P$ we get

$$T'D(\pi^e)PT = \begin{bmatrix} a & \pi_1^e & \pi_1^e & \cdots \\ \pi_1^e & b & \pi_1^e + \pi_2^e & \cdots \\ \pi_1^e & \pi_1^e + \pi_2^e & c & \vdots \\ \vdots & \cdots & \vdots & \ddots \end{bmatrix}.$$

Thus, $T'D(\pi^e)(P-I)T$ has nonpositive elements on the diagonal, and zeros elsewhere, which implies $\pi(0)'D(\pi^e)^{-1}(T')^{-1} \geq 0$, i.e. $\pi(0)'D(\pi^e)^{-1}$ is nonincreasing. □

4. Monotone comparative statics

Monotonicity considerations turn out to be important to give definite answers to problems regarding the effects of "perturbations" of the transition matrix P on the equilibrium income distribution. The comparative statics analysis of finite ergodic Markov chains is contained in an illuminating paper by Conlisk (1985), where he provides formulae for calculating the effects of infinitesimal changes in the transition matrix on the equilibrium vector and the mean first passage time matrix.

Consider a regular chain with transition matrix P and equilibrium vector π^e. The obvious elementary change in P would seem to be an increase in an element of P; however, to keep the corresponding row sum to one, at least another element of P must decrease. The simplest perturbation considered by Conlisk is composed of a gain of size ε in the gth element of row r, P_{rg}, at the expense of a fall of equal size in the element P_{rf}. P as a function of ε may be written as $P(\varepsilon) = P_0 + \varepsilon e_r'(e_g - e_f)$, where P_0 is the initial value of P (recall e_j is the jth $n \times 1$ unit vector). Let $\pi^e(\varepsilon)$ denote the equilibrium vector as function of ε. The interesting comparative statics result for our purposes is the sign of the derivative $\partial\pi^{e'}T/\partial\varepsilon$ evaluated at $\varepsilon = 0$, where $\pi^{e'}T$ is the equilibrium cumulative income distribution. When $\partial\pi^{e'}T/\partial\varepsilon$ is nonpositive, the perturbation causes a stochastically dominating shift in the equilibrium income distribution. Derivatives for a general perturbation can be built up from elementary perturbations by means of a perturbation matrix $\partial P/\partial\varepsilon = E$ which arrays elementary perturbations.

Suppose for example that we want to study the effect of an increase in the probability of transition to a given state g, at the expenses of a decrease in the probability of transition to state f for any origin state. This is an example of what Conlisk calls a "column gain" perturbation, where all gaining elements of P are in a single column and all falling elements are in another column. The perturbation matrix E has ones in column g, minus ones in column f and zeros elsewhere. Suppose for example that $g > f$, so that at each period it becomes easier to transit to an higher income class regardless of origin state. Intuitively, one would expect that eventually this society will settle in such a way that more people are allocated in richer income states (that is, the new equilibrium vector will dominate the old one). That this is not necessarily so is established by the following example:

$$P = \begin{bmatrix} 0.1 & 0.4 & 0.5 \\ 0.3 & 0.4 & 0.4 \\ 0.5 & 0.4 & 0.1 \end{bmatrix}; \qquad \pi^e = [0.3, 0.4, 0.3].$$

Suppose $f = 2$ and $g = 3$, so that the probability of transiting to the richer income class is increased for any origin class. Applying Conlisk's formula for $\partial\pi^e/\partial\varepsilon$ for the column gain of this sort we get $\partial\pi^{e'}T/\partial\varepsilon = [-0.143, 0.857, 0]$, which implies a *decrease* in the proportion of people in the richest class after the perturbation. This example is rather similar in spirit to the example considered in section three, and it is characterized by failure of monotonicity of P, as shown below.

Conlisk (1985; page 145) shows that the effect of a perturbation E on the cumulative distribution of π^e is given by $\partial\pi^{e'}T/\partial\varepsilon = \pi^{e'}EZT$, where Z is Kemeny's (1981) "fundamental matrix", given by $Z(d) = (I - P + 1d')^{-1}$ (d is any $n \times 1$ vector such that $d'1 \neq 0$). Choosing d equal to P's last row (denoted P_n) we get $\partial\pi^{e'}T/\partial\varepsilon = \pi^{e'}EZT = \pi^{e'}ET(I - T^{-1}PT + T^{-1}1P_nT)^{-1}T^{-1}T = \pi^{e'}ET(I - T^{-1}PT + T^{-1}1P_nT)^{-1}$. Under monotonicity, the matrix $(I - T^{-1}PT + T^{-1}1P_nT)^{-1}$ will be nonnegative; a perturbation E will cause a stochastically dominating shift in the equilibrium income distribution ($\partial\pi^{e'}T/\partial\varepsilon \leq 0$) whenever $\pi^{e'}ET$ is nonpositive. For the column gain considered above, the matrix ET has minus ones in the fth column and zeros elsewhere; it follows that under monotonicity a shift of probability mass towards a richer income class does imply a "richer" equilibrium income distribution in the long run.

Another interesting perturbation of the transition matrix P is obtained by considering a "diagonalising" shift of probability mass towards the main diagonal which represents a decrease in the mobility of the population, as analysed in Shorrocks (1978). When one of the diagonal element of P increases at the expense of an off-diagonal element in the same row, we may regard the new transition structure as indicating a lower level of mobility; a diagonalising perturbation along these lines may be obtained by assuming that in row i the ith element is the gaining and the $(i + 1)$th is the falling one, while in row $i + 1$ the element $i + 1$ is the gaining one at the expenses of the ith element. The effect of such an immobility shift can be signed, under monotonicity, by considering the sign of $\pi^{e'}ET$, which equals zero everywhere except the ith element which has the same sign as $(\pi_i^e - \pi_{i+1}^e)$. If income class i is more numerous in equilibrium than class $i+1$, the immobility increasing shift considered here will cause a dominated shift in the equilibrium income distribution.

These two examples show how comparative statics results may be obtained under monotonicity by simply considering the sign of the vector $\pi^{e'}ET$, given the nonnegativity of the matrix $(I - T^{-1}PT + T^{-1}1P_nT)^{-1}$ and the general formula for the effect of a perturbation E on the cumulative income distribution.

5. Lifetime inequality

For a given regular transition matrix P, we may derive the implied distribution of expected lifetime income for the individuals in a society whose mobility is governed by P and is in steady state. Consider a society consisting of identical individuals who are born simultaneously and live exactly for τ periods. Let Y^P denote an $(n \times 1)$ vector of expected discounted lifetime incomes, where the typical element Y_i^P denotes the expected lifetime discounted income of an individual who starts a life in income class i, and is given by the ith element of the vector $Y^P = y + \rho Py + \rho^2P^2y + \ldots + \rho^\tau P^\tau y$, where $0 \leq \rho < 1$ denotes the discount factor. Letting $\tau \to \infty$, $Y^P = [I - \rho P]^{-1}y$. The permanent income vector may be obtained by normalizing Y^P as $Y^P = (1-\rho)[I - \rho P]^{-1}y$. The steady state distribution of permanent incomes is given by $[\pi^e, Y^P]$. We will 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.

The one period Lorenz curve for $[\pi^e, y]$ has horizontal coordinates given by $(\pi_1^e, \pi_1^e + \pi_2^e, \ldots, \pi_1^e + \pi_2^e + \ldots + \pi_i^e, \cdots, 1)$; while the vertical coordinates are given by $[\pi_1^e y_1, \pi_1^e y_1 + \pi_2^e y_2, \ldots, \pi_1^e y_1 + \pi_2^e y_2 + \ldots + \pi_i^e y_i, \ldots, \pi^{e'}y]/\pi^{e'}y$. Note that when P is monotone also $P(\rho)$ is monotone; it follows that the permanent income vector Y^P is nondecreasing for any nondecreasing income vector y. When Y^P is nondecreasing, the Lorenz curve for $[\pi^e, Y^P]$ has horizontal coordinates given by $(\pi_1^e, \pi_1^e + \pi_2^e, \ldots, \pi_1^e + \pi_2^e + \ldots + \pi_i^e, \ldots, 1)$ and vertical coordinates given by $[\pi_1^e Y_1^P, \pi_1^e Y_1^P + \pi_2^e Y_2^P, \ldots, \pi_1^e Y_1^P + \pi_2^e Y_2^P + \ldots + \pi_i^e Y_i^P, \ldots, \pi^e Y^P]/\pi^{e'}Y^P$.

Given income distributions $[\pi^e, x]$ and $[\pi^e, y]$ for two income state vectors x and y with equal average income (i.e. satisfying $\pi^{e'}x = \pi^{e'}y$), the Lorenz ordering $[\pi^e, x] \succeq_L [\pi^e, y]$ is given by the condition that the Lorenz curve for $[\pi^e, x]$ lies nowhere below that for $[\pi^e, y]$. Using the summation matrix T, we may write $[\pi^e, x] \succeq_L [\pi^e, y]$ compactly as $T'D(\pi^e)x \geq T'D(\pi^e)y$; and given two permanent income distributions $[\pi^e, X^P]$ and $[\pi^e, Y^P]$, the Lorenz ordering $[\pi^e, X^P] \succeq_L$ $[\pi^e, Y^P]$ is written compactly as $T'D(\pi^e)X^P \geq T'D(\pi^e)Y^P$ Lambert (1993) contains a thorough review of the uses and properties of the Lorenz curve for the analysis of the distribution and redistribution of income.

Before proceeding, we need the following definition: the *reverse Markov chain* for a regular Markov chain with transition matrix P and equilibrium vector π^e is a Markov chain with transition matrix given by $D(\pi^e)^{-1}P'D(\pi^e)$. The typical element of a transition matrix for the reverse Markov chain gives the probability that an individual who is in class i at time t came from class j in the previous period. If the backward and forward transition matrices are equal, i.e. if $P = D(\pi^e)^{-1}P'D(\pi^e)$, the process in equilibrium will appear the same looked at backwards as forwards: in this case we call the chain time reversible. Reverse

Markov chains and time reversibility are considered in detail in Kemeny and Snell (1976).

Dardanoni (1992) considers the conditions under which Lorenz dominance of the single period income distributions is preserved in the permanent income distributions when the chain is in steady state:

Proposition 4. *Let P be a monotone regular transition matrix with equilibrium vector π^e, and let x and y be two different income state vectors, with $X^P = P(\rho)x$ and $Y^P = P(\rho)y$ being their respective permanent income vectors. Then the following conditions are equivalent:*

(1) $[\pi^e, x] \succeq_L [\pi^e, y]$ implies $[\pi^e, X^P] \succeq_L [\pi^e, Y^P]$ for all income state vectors x and y with equal average income;

(2) The lifetime transition matrix for the reverse Markov chain, $D(\pi^e)^{-1} P(\rho)' D(\pi^e)$, is monotone.

Proof. [(2) implies (1)]. By Lemma 2, monotonicity of P implies that X^P and Y^P are nondecreasing. Condition (1) may then be written as $T'D(\pi^e)(\rho)y \geq T'D(\pi^e)P(\rho)x$ for any y and x such that $T'D(\pi^e)y \geq T'D(\pi^e)x$ and $\pi^{e\prime}y = \pi^{e\prime}x$. Condition (2) may be written as $T^{-1}D(\pi^e)^{-1}P'(\rho)D(\pi^e)T \geq 0$, which can be written as $T'D(\pi^e)P(\rho)D(\pi^e)^{-1}(T^{-1})' \geq 0$. But this implies $T'D(\pi^e)P(\rho)[y - x] = T'D(\pi^e)P(\rho)D(\pi^e)^{-1}(T^{-1})'T'D(\pi^e)[y - x] \geq 0$ when $T'D(\pi^e)[y - x] \geq 0$.

[(1) implies (2)]: We argue by contradiction: Assume the (i,j)th element of $T^{-1}D(\pi^e)^{-1}P'(\rho)D(\pi^e)T$ is negative; taking the transpose, this implies that the (j,i)th element of $T'D(\pi^e)P(\rho)D(\pi^e)^{-1}(T^{-1})'$ is negative. Then choose all the elements of x (except the elements i and $i+1$) equal to y, and let $x_i = y_i - k/\pi_i^e$ and $x_{i+1} = y_{i+1} + k/\pi_{i+1}^e$, with k being a positive constant which does not rearrange the x vector. Then $T'D(\pi^e)[y - x]$ equals a vector with all elements except the ith equal to zero, while the ith element equals $k > 0$. But this implies that the jth element of $T'D(\pi^e)P(\rho)[y - x] = T'D(\pi^e)P(\rho)D(\pi^e)^{-1}(T^{-1})'T'D(\pi^e)[y - x]$ is negative, a contradiction. □

A necessary and sufficient condition for ensuring that a "static" reduction in inequality due to a change in the income state vector will imply a reduction in lifetime inequality is that the lifetime transition matrix for the reverse chain is monotone. Intuitively, this means that the "backward lifetime lottery" that a person in class i has faced is better, in terms of stochastic dominance, than the lottery that has been faced by an individual in class $i - 1$; this seems intuitively plausible. Note that monotonicity of the transition matrix of the reverse chain is sufficient to ensure monotonicity of the lifetime reverse matrix for any discount factor of interest. Lifetime inequality reduction is guaranteed, given a reduction in inequality of the "spot" income distribution, by monotonicity of the reverse chain. Finally, note that if the chain is monotone and time reversible, the lifetime reverse matrix will trivially be monotone.

6. Mobility orderings

Consider two societies which follow respectively transition matrices P and Q, and assume that they have a common steady state distribution vector π^e, so that $\pi^{e\prime}P = \pi^{e\prime}Q = \pi^{e\prime}$. Dardanoni (1993) proposes a mobility partial ordering of monotone transition matrices with equal steady state vectors. The equal steady state assumption plays the same role as the equal mean assumption in the literature on static inequality measurement, as pionereed by Kolm (1969) and Atkinson (1970). Dardanoni argues that the proposed partial order is the 'natural' extension of Lorenz curve ordering employed to rank "one shot" income distributions. Letting Y^P and Y^Q denote the vectors of permanent incomes in the steady state, he shows that the mobility ordering is equivalent to the Lorenz ordering on the permanent income vectors. In other words, denoting by $\succeq_M$ the mobility partial ordering, Dardanoni shows, among other things, that under monotonicity $P(\rho) \succeq_M Q(\rho)$ if and only if $[\pi^e, Y^P] \succeq_L [\pi^e, Y^Q]$.

Consider two monotone regular transition matrices P and Q which share a common steady state vector π^e. Assume that both societies also share the initial income distribution vector $\pi(0)$, which satisfies the sufficient condition of Proposition 3, namely that $\pi(0)_i/\pi_i^e$ is nonincreasing in $i = 1, 2, \ldots, n$, so that the sequences of income distributions stemming from P and Q are stochastically increasing over time and will eventually converge to the common steady state distribution π^e.

Intuitively, one might expect that if P is a more mobile society than Q, it will be more effective in changing the initial income distribution, and approach the steady state distribution more rapidly. Thus, if $P \succeq_M Q$, even though both societies share by assumption the initial income distribution $\pi(0)$ and the steady state distribution π^e, one might expect that at all times $t = 1, 2, \ldots$ the income distribution under P, $\pi(0)P^t$, dominates that under Q, $\pi(0)Q^t$. This turns out to be so:

Proposition 5. *Let P and Q be regular monotone transition matrices with monotone reverse chains and equal steady state vector π^e. The following two conditions are equivalent:*

(1) $P^t \succeq_M Q^t$ for $t = 1, 2, \ldots$;

(2) $\pi(0)'P^t \succeq \pi(0)'Q^t$ for $t = 1, 2, \ldots$ for any initial income distribution vector $\pi(0)$ satisfying the condition that $\pi(0)_i/\pi_i^e$ is nonincreasing in $i = 1, 2, \ldots, n$.

Proof. [(1) implies (2)]: Rewrite $\pi(0)'P^t \succeq \pi(0)'Q^t$ as $\pi(0)'P^tT \leq \pi(0)'Q^tT$, which can written as $\pi(0)'D(\pi^e)^{-1}(T')^{-1}T'D(\pi^e)(P^t - Q^t)T \leq 0$ using by now familiar manipulation. From Dardanoni's (1993) condition (1) can be written as $T'D(\pi^e)(P^t - Q^t)T \leq 0$. Rewriting the condition that $\pi(0)_i/\pi_i^e$ is nonincreasing in $i = 1, 2, \ldots, n$ as $\pi(0)'D(\pi^e)^{-1}(T')^{-1} \geq 0$ the result follows.

[(2) implies (1)]: We argue by contradiction. Rewrite $\pi(0)'P^t \succeq \pi(0)'Q^t$ as $\pi(0)'D(\pi^e)^{-1}(T')^{-1}T'D(\pi^e)(P^t - Q^t)T \leq 0$ and assume that, contrary to condition (1), the (j, l)th element of $T'D(\pi^e)(P^t - Q^t)T$ is strictly positive. Now choose $\pi(0) = (\pi_1^e/\sum_{i=1}^j \pi_i^e, \pi_2^e/\sum_{i=1}^j \pi_i^e, \ldots, \pi_j^e/\sum_{i=1}^j \pi_i^e, 0, \ldots, 0)$. It follows that the vector $\pi(0)'D(\pi^e)^{-1}(T')^{-1}$ has zeros everywhere except for the jth element, which is strictly positive, and we obtain the desired contradiction. □

The result confirms the simplifying nature of the monotonicity assumption in obtaining insights into the dynamics of the income distribution over time in a simple and unified fashion.

7. Conclusions

The transition matrix of a Markov chain of income mobility is monotone when richer individuals face a better income prospect than poorer individuals. Monotonicity is a strong mathematical assumption, and yet is intuitively plausible and empirically valid. The mathematical structure implied by the monotonicity assumption makes many interesting results very easy to demonstrate. It is hoped that this illustrative paper will help clarify the use of the monotonicity assumption in various income distribution and redistribution problems.

References

Atkinson AB (1970) On the measurement of inequality. J Econ Theory 2: 244–63

Atkinson AB, Bourguignon F, Morrisson C (1992) Empirical Studies of Earnings Mobility. Harwood, Chur

Bartholomew DJ (1982) Stochastic Models for Social Processes. 3rd Ed., Wiley and Sons, New York

Billingsley P (1961) Statistical Inference for Markov processes. University of Chicago Press, Chicago

Conlisk J (1985) Comparative statics for Markov chains. J Econ Dynamics Control 9: 139–51

Conlisk J (1990) Monotone mobility matrices. J Math Sociol 15: 173–191

Dardanoni V (1992) The Fiscal System and the Dynamic Redistribution of Income. In: Design and Reform of Taxation Policy. Galeotti G, Marrelli M (Eds), Kluwer, Dordrecht

Dardanoni V (1993) Measuring social mobility. J Econ Theory 61: 372–394

Dardanoni V, Forcina A (1993) You'd better have a rich father: Testing monotone mobility matrices. Mimeo

Lambert PJ (1993) The distribution and redistribution of income: A mathematical analysis. 2nd Ed., Manchester University Press, Manchester

Kanbur SMR, Stromberg J-O (1988) Income transitions and income distribution dominance. J Econ Theory 45: 408–416

Keilson J, Kester A (1977) Monotone matrices and montone Markov chains. Stochastic Process Appl 5: 231–241

Kemeny JG (1981) Generalization of a fundamental matrix. Linear Algebra Appl 38: 193–206

Kemeny JG, Snell JL (1976) Finite Markov Chains. Springer, New York

Kolm S-Ch (1969) The optimal production of social justice. In Public Economics. J Margolis, Guitton H (Eds), MacMillan, London

Shorrocks A (1976) Income mobility and the Markov assumption. Econ J 86: 556–578

Shorrocks A (1978) The measurement of mobility. Econometrica 45: 1013–1024