INTERGENERATIONAL MOBILITY AND DYNASTIC INEQUALITY

S. M. Ravi Kanbur
University of Essex
Princeton University

Joseph E. Stiglitz
Princeton University

Econometric Research Program
Research Memorandum No. 324

April 1986

ECONOMETRIC RESEARCH PROGRAM
Princeton University
207 Dickinson Hall
Princeton, NJ 08544

Abstract

This paper is a contribution to the comparison of mobility in a social welfare framework. In particular, we emphasize the consequences of alternative mobility structures for dynastic inequality. This means that we consider social welfare functions which are not simply additive across dynastic welfare. We derive a necessary and sufficient characterization of social welfare dominance in this framework, and relate the implied partial ordering of mobility structures to several conventional views on what constitutes a more mobile society.

Contents

1. Introduction
2. Conventional views and a framework of analysis
3. A benchmark proposition and a research agenda
4. Mobility and dynastic inequality: A characterization result
5. Interpretations
6. Further research

1. Introduction

In a survey of the statics and dynamics of income distribution, Hart (1980) contrasts a "snapshot" income distribution at a point in time with a "movie" of incomes showing mobility over time. The former provided the focus for much of the normative analysis of the economics of inequality in the 1970's, including the papers by Atkinson (1970), Dasgupta, Sen and Starrett (1973), Rothschild and Stiglitz (1973), and others. In the late 1970's, Atkinson (1979) noted that "Economists have typically paid little attention to questions of mobility, leaving the field largely to sociologists." The situation has been redressed somewhat since then, with papers by Shorrocks (1978a, 1978b), Atkinson (1980a, 1980b), Markandya (1982, 1984), Kanbur and Stiglitz (1982), King (1983), Chakravarty, Dutta and Weymark (1985), Kanbur and Stromberg (1985), and others. The present paper is a contribution to this strand of the literature, emphasizing in particular the <u>dynastic inequality</u> implied by different mobility structures.

The spirit of our analysis is close to that of Atkinson's (1970) original paper on the measurement of static inequality. Just as he set out to rank income distributions with respect to a class of static social welfare functions, we wish to rank mobility mechanisms with respect to a class of dynamic social welfare functions that emphasize dynastic inequality. Analogously to his use of the Lorenz curve ranking as an operational method of checking for static social welfare dominance, we wish to develop an operational rule for comparing mobility mechanisms that will tell us whether dynamic social welfare dominance holds or not. In the same way that he held "the size of the cake" constant across distributions being compared, we wish to hold the size of the "intertemporal cake" constant. In Atkinson (1980b), the results from Atkinson and Bourgignon (1982) on ranking bivariate distributions were used to rank mobility mechanisms, with a two period social welfare function which emphasized non-separability of welfare over time. In the present paper an alternative approach is taken which emphasizes non-separability across dynasties i.e. dynastic inequality.

The plan of the paper is as follows. In the next section we discuss some conventional views on "greater mobility", introduce some notation and discuss the framework of analysis. Section 3 presents a benchmark result which serves to highlight alternative routes within the social welfare function approach to the evaluation of mobility. Section 4 follows one of these routes, which emphasizes lifetime or dynastic inequality, and provides a necessary and sufficient characterization of social welfare ranking on the space of mobility matrices. Section 5 interprets this characterization in terms of the conventional views discussed in Section 2. Section 6 concludes the paper with suggestions for further research.

2. Conventional Views and a Framework of Analysis

To say that interest in income mobility is recent in the formal economic literature is not of course to deny that it has been of considerable interest to the informal, policy oriented literature. In his well known polemic, Friedman (1962) argues as follows:

"Consider two societies that have the same annual distribution of income. In one there is great mobility and change so that the positions of particular families in the income hierarchy varies widely from year to year. In the other there is great rigidity so that each family stays in the same position year after year. The one kind of inequality is a sign of dynamic change, social mobility, equality of opportunity; the other, of a status society. The confusion between the two kinds of inequality is particularly important precisely because competitive free enterprise capitalism tends to substitute the one for the other."

Whatever the truth of Friedman's characterization of the capitalist system, there remains the question of formalizing and making precise the many different conceptions of mobility present in this passage, and in the writings of other authors.

To aid precision, let us specify the n x n transition matrix:

$$A = ||a_{ij}||; \; a_{ij} \geq 0, \; \sum_{j=1}^{n} a_{ij} = 1 \qquad (2.1)$$

whose typical element a_{ij} is the probability of movement from income category i to income category j. The transition can be thought of as either inter-generational or intra-generational; both aspects are present in informal writings. Let there be n categories of incomes with income levels

$$y_1 \leq y_2 \leq y_3 \leq \ldots . \leq y_n \qquad (2.2)$$

For simplicity, the transition matrix A will be assumed to be constant over time and, where appropriate, the probability that a particular income category is occupied will be identified with the proportion of people who actually end up in that category.

An oft encountered view on what constitutes greater mobility may be characterized as the "diagonals view". Thus a transition matrix $A=||a_{ij}||$ is said to exhibit greater mobility than a transition matrix $B=||b_{ij}||$ if

$a_{ii} \leq b_{ii}$ for all i with strict inequality for some i, in other words each diagonal element of A is no larger than the corresponding diagonal element of B. A "strong diagonals view" can be characterized as saying that $a_{ij} \geq b_{ij}$ for all $i \neq j$ with strict inequality for some $i \neq j$. Shorrocks (1978a) called this property of a mobility measure "monotonicity". Bartholomew (1973) presents some other measures along these lines.

The extreme of immobility in the "diagonals view" is of course the case of the Identity matrix:

$$A \equiv I;\ a_{ij} = \delta_{ij} = \begin{cases} 1 \text{ for } i=j \\ 0 \text{ for } i \neq j \end{cases} \qquad (2.3)$$

The other extreme is sometimes argued to be the matrix which has ones along the non-leading diagonal and zero's elsewhere:

$$a_{ij} = k_{ij} = \begin{cases} 1 \text{ for } j=n+1-i \\ 0 \text{ for } j \neq n+1-i \end{cases} \qquad (2.4)$$

The 4x4 example is the following:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Kanbur and Stiglitz (1982) refer to the view which holds this matrix to be the best as the "negative correlation" view. Atkinson (1980a) calls this the case of "complete reversal" and contrasts this with another view of perfect mobility, namely that of "equality of opportunity" as indicated by independence of future prospects from the current situation. This is given by a matrix whose rows are identical

$$E = \begin{bmatrix} \underset{\sim}{a}' \\ \underset{\sim}{a}' \\ \underset{\sim}{a}' \end{bmatrix} \tag{2.5}$$

where $\underset{\sim}{a}'$ represents the vector of transition probabilities, identical for every initial state. Shorrocks (1978a) also uses (2.5) to define his property PM ("perfect mobility"). As Shorrocks goes on to show, one cannot both hold the "diagonals view" and the "equality of opportunity view" simultaneously. For example, with

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

it should be clear that A is better than B on the diagonals view while the reverse is true on the equality of opportunity view.

The problem is, of course, that we have not specified a clear and explicit social welfare function with respect to which we can rank transition matrices. For example, one argument could be that it is the income distribution consequences of a transition matrix that are important - the matrix which gives rise to the most equal distribution of income is the one which should be preferred. This is still an intermediate step because more generally we would require an intertemporal social welfare function with respect to which the sample paths of the stochastic processes represented by different matrices could be aggregated and evaluated. The next section will in fact present a social welfare function which rationalizes a focus on the snapshot distribution of income, but the nature of the task facing us should now be clear. We have to first of all specify an explicit intertemporal social welfare function. Secondly, we have to derive the (possibly partial) ordering that a class of social welfare functions induces on the set of transition matrices. In other words, we wish to find a necessary and sufficient characterization of the relation between two transition matrices such that the social welfare from the stochastic process generated by one is greater than that from the other. Finally, given this characterization our task is to assess the precise implications for the conventional views discussed here - which family of social welfare functions rationalizes which view, if any?

There is, of course, some choice regarding the nature of social welfare. In this paper we view social welfare as being evaluated in two steps. First, individuals (or dynasties) evaluate their own lifetime prospects. Second, these valuations become arguments of society's valuation function, defined over "dynastic" welfares. This view of social welfare could be termed Utilitarian, although some might object to this usage. Thus our object is to make sense of conventional views in such a social welfare framework. It must be conceded that if we fail, or even if we succeed, there is still the possibility that these views are rooted in values that fall outside of our formulation of social welfare.

3. A Benchmark Proposition and A Research Agenda

Consider a two period world with dynasty i currently having income y_i. Thus income in the next period is y_j with probability a_{ij}, from (2.1). If $W(y_i,y_j)$ represents the welfare from incomes (y_i,y_j), then the expected welfare of dynasty i is

$$V_i = \sum_{j=1}^{n} a_{ij} W(y_i,y_j) \qquad (3.1)$$

If p_i^o is the proportion of total units in category i, then social welfare will be some function of the V_is and the p_i^os. If we suppose that social welfare S is additive in dynastic welfares, then

$$S = \sum_{i=1}^{n} p_i^o \sum_{j=1}^{n} a_{ij} W(y_i, y_j) \tag{3.2}$$

This is, in fact, the social welfare function considered by Atkinson (1980a,b).

If $W(y_i, y_j)$ was additively separable, i.e.,

$$W(y_i, y_j) = U^o(y_i) + U^1(y_j) \tag{3.3}$$

then (3.2) becomes

$$S = \sum_{i=1}^{n} p_i^o U^o(y_i) + \sum_{j=1}^{n} p_j^1 U^1(y_j) \tag{3.4}$$

where

$$p_j^1 = \sum_{i=1}^{n} p_i^o a_{ij} \quad ; \quad j = 1,2,\ldots, n \tag{3.5}$$

is the spot distribution of income in the next period. As can be seen from (3.4), with additive separability in intertemporal valuations, and additivity in social valuation of dynastic welfares, the problem reduces to one of evaluating the static distributions of income p^o and p^1, with respect to utility functions U^o and U^1. This implication of additive separability over time and across individuals is noted by Atkinson

(1980b) and Markandya (1982). If U^o and U^1 are concave, then we know from Atkinson (1970) that the Lorenz curve ranking can be applied period by period when comparing two societies with different mobility structures. One question of interest is the conditions on mobility structures under which a Lorenz curve ranking in the first period is preserved in the next period. Kanbur and Stromberg (1985) derive a necessary and sufficient characterization of this property.

Let us now impose the condition that income transitions are in a steady state, so that the snapshot distribution of income remains unchanged over time. Then

$$p^o = p^1 = p^*$$

and

$$S = \sum_{i=1}^{n} p_i^*[U^o(y_i) + U^1(y_i)] \tag{3.6}$$

The problem is now one of ranking the steady state distributions of Markov income transition matrices. So far as we know this problem has not received a great deal of attention in the literature, although Kanbur and Stromberg (1985) derive sufficient conditions under which the steady state income distribution of one transition matrix dominates the steady state income distribution of another transition matrix.

From (3.6) it is clear that, under the assumptions made so far, if two transition matrices have identical steady state distributions of income, then they will be ranked indifferent according to the social welfare function. In fact, the class of bistochastic transition matrices i.e. those for which

$$a_{ij} \geq 0 \quad ; \quad \sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n} a_{ij} = 1 \tag{3.7}$$

all have a steady state distribution of the form

$$p^{*} = (\frac{1}{n},\frac{1}{n},\frac{1}{n}, \ldots, \frac{1}{n}) \tag{3.8}$$

i.e. equal numbers in each category. The bistochastic assumption is made by Atkinson (1980a,b) and by Rothschild and Yaari (1979).

We have thus proved the following proposition:

If (a) the social welfare function is additive across expected dynastic welfare, (b) dynastic welfare is additively separable in incomes over time, (c) the income distribution is in a steady state and (d) the transition matrix is bistochastic, then all societies are ranked indifferent according to the utilitarian social welfare function.

With these assumptions, therefore, the social welfare function approach has no cutting power whatsoever. This is not surprising, of course, since one by one the assumptions remove any influence that income mobility may have on intertemporal social welfare.

The proposition is of course fairly straightforward. Its usefulness lies in the fact that it provides a systematic framework and an agenda for research. Relaxing these assumptions will reintroduce different aspects of mobility into the analysis. Thus relaxing (c) while maintaining the other assumptions focuses attention on the time path of income distribution snapshots. This is the route followed in Kanbur and

Stromberg (1985). Relaxing (d) on its own allows us to concentrate on the effect of mobility on the steady state income distribution - Kanbur and Stromberg (1985) have some results on this.

Relaxation of (b) while maintaining (a), (c) and (d) has been the focus of the work of Atkinson (1980a,b); see also Kanbur and Stiglitz (1982). Intertemporal non-separability is invoked to distinguish between the social welfare consequences of different transition mechanisms. By analogy with multivariate stochastic dominance, a condition is derived for one transition matrix to give a higher social welfare than another for a class of social welfare functions. If

$$A = ||a_{ij}|| \text{ and } B = ||b_{ij}||$$

are the two transition matrices being compared and social welfare is given

by (3.2) with $p_i^o = p_i^* = \frac{1}{n}$ for all i, then it can be shown that

$$S^A \geq S^B \text{ for all } W(\cdot,\cdot) \text{ such that } \frac{\partial^2 W}{\partial y_i \partial y_j} \leq 0$$

if and only if

$$M = \sum_{i=1}^{k} \sum_{j=1}^{m} (a_{ij} - b_{ij}) \leq 0 \text{ for all } k,m. \qquad (3.9)$$

We will take up the interpretation of this result later in the paper, but we note that the condition on transition matrices is essentially equivalent to the cumulative of the bivariate distribution of (y_i, y_j) being nowhere greater in society A than in society B.

The assumption of social welfare being additive in dynastic expected welfare rules out the possibility of showing concern for inequality among dynasties along this dimension. The next section relaxes assumption (a) in order to extend the analysis in this direction, but maintains the other assumptions for reasons of simplicity and clarity.

4. Mobility and Dynastic Inequality: A Characterization Result

We assume then, a bistochastic transition matrix with the income distribution in steady state. The bistochastic assumption is restrictive but (i) as Atkinson (1980a) argues, if the income categories are quantiles then the transition matrix is by definition bistochastic and (ii) as Rothschild and Yaari (1979) argue, given the steady state assumption any Markov process can be approximated arbitrarily closely by a bistochastic matrix of appropriate dimension. Notice that since the steady state of a bistochastic transition matrix has equal numbers in each category, we can normalize to the situation where there is one person in each category.

We will further simplify the dynasty's intertemporal utility function to be the expected discounted value of a given instantaneous utility function with a constant discount rate. Then for a two period world dynasty i's lifetime expected welfare is given by

$$V_i = U(y_i) + \gamma \sum_{j=1}^{n} a_{ij} U(y_j) \; ; \; i = 1,2,\ldots, n \qquad (4.1)$$

where γ is the discount factor (lying between 0 and 1) and $U(\cdot)$ is the instantaneous utility function. If we define

$$\underset{\sim}{U} = (U(y_1), U(y_2), \ldots, U(y_n))' \tag{4.2}$$

as the column vector of utilities of income, then we can stack the expressions in (4.1) to give

$$\underset{\sim}{V} = \underset{\sim}{U} + \gamma A \underset{\sim}{U} \tag{4.3}$$

where $\underset{\sim}{V}$ is the vector of dynastic expected welfares. For a T period world (4.3) becomes

$$\underset{\sim}{V} = [I + \gamma A + \gamma^2 A^2 + \ldots + \gamma^T A^T] \underset{\sim}{U} \tag{4.4}$$

If we let $T \to \infty$ then under standard conditions we get

$$\underset{\sim}{V} = [I - \gamma A]^{-1} \underset{\sim}{U} \tag{4.5}$$

Now it can be shown that

$$\underset{\sim}{e}' \underset{\sim}{V} = \frac{1}{1-\gamma} \underset{\sim}{e}' \underset{\sim}{U} \text{ where } \underset{\sim}{e}' = (1,1,\ldots,1) \tag{4.6}$$

so that the sum of lifetime utilities is the same for all transition mechanisms in this class. This is of course the proposition discussed in the last section. But in this section, instead of simply taking the sum of V_1, V_2, $\ldots V_n$, we will let social welfare be

$$S = S(V_1, V_2, \ldots, V_n) \tag{4.7}$$

where $S(\cdot,\cdot,\ldots \cdot,\cdot)$ is a symmetric, quasi-concave function, assumptions which are meant to capture, respectively, anonymity and egalitarianism with respect to dynastic prospects.

Our object is to rank transition matrices according to the social welfare function (4.7). Denote $\underset{\sim}{V}^A$ and $\underset{\sim}{V}^B$ as the vector of lifetime utilities under two transition matrices A and B:

$$\begin{aligned} \underset{\sim}{V}^A &= [I - \gamma A]^{-1} \underset{\sim}{U} \\ \underset{\sim}{V}^B &= [I - \gamma B]^{-1} \underset{\sim}{U} \end{aligned} \tag{4.8}$$

and let

$$S^A = S(\underset{\sim}{V}^A); \; S^B = S(\underset{\sim}{V}^B) \tag{4.9}$$

we can now state the basic result of this section:

<u>Theorem</u>: $S^A \geq S^B$ for all symmetric, quasi-concave $S(\cdot)$ and all $U(\cdot)$ which are unique up to a positive linear transformation, if and only if there exists a bistochastic matrix Q such that

$$B = \frac{1}{\gamma} [I - Q] + AQ$$

<u>Proof of Sufficiency</u>

$$B = \frac{1}{\gamma} [I - Q] + AQ$$

$$=> [I - \gamma B] = [I - \gamma A]Q$$

$$=> [I - \gamma A]^{-1} = Q[I - \gamma B]$$

$$=> [I - \gamma A]^{-1}\underset{\sim}{U} = Q[I - \gamma B]^{-1}\underset{\sim}{U} \quad \text{for all } U(\cdot)$$

$$=> \underset{\sim}{V}^A = Q\, \underset{\sim}{V}^B$$

Now if it is not the case that $V_1^A \leq V_2^A \leq \ldots \leq V_n^A$, simply permute them to give

$$\hat{\underset{\sim}{V}}^A = P_A\, \underset{\sim}{V}_A$$

where the vector $\hat{\underset{\sim}{V}}^A$ is ordered so that $\hat{V}_1^A \leq \hat{V}_2^A \ldots \leq \hat{V}_n^A$, and P_A is the appropriate permutation matrix. Similarly, use an appropriate permutation matrix P_B to reorder $\underset{\sim}{V}^B$ to $\hat{\underset{\sim}{V}}^B$. The argument now continues as follows:

$$\underset{\sim}{V}^A = Q\, \underset{\sim}{V}^B$$

$$=> P_A\, \underset{\sim}{V}^A = P_A Q\, P_B^{-1}\, P_B\, \underset{\sim}{V}^B$$

$$=> \hat{\underset{\sim}{V}}^A = [P_A Q\, P_B^{-1}]\, \hat{\underset{\sim}{V}}^B$$

$$=> \hat{\underset{\sim}{V}}^A = \hat{Q}\, \hat{\underset{\sim}{V}}^B$$

It is easy to check that the inverse of a permutation matrix is itself a permutation matrix and hence $\hat{Q} = P_A\, Q\, P_B^{-1}$ is also a bistochastic matrix. But from Dasgupta, Sen and Starrett (1973, Theorem 1),

$$\hat{\underset{\sim}{V}}^A = \hat{Q}\, \hat{\underset{\sim}{V}}^B \quad \text{where } \hat{Q} \text{ bistochastic}$$

$$\Rightarrow S(\hat{\underset{\sim}{V}}^A) \geq S(\hat{\underset{\sim}{V}}^B)$$

for all symmetric quasi-concave functions $S(\cdot)$. Moreover, by symmetry of the $S(\cdot)$ function,

$$S(\hat{\underset{\sim}{V}}^A) = S(\underset{\sim}{V}^A) \; ; \; S(\hat{\underset{\sim}{V}}^V) = S(\underset{\sim}{V}^B) \; ; \text{ so that}$$

$$S(\underset{\sim}{V}^A) \geq S(\underset{\sim}{V}^B) \text{ and sufficiency is proved.}$$

<u>Proof of Necessity</u>

$$S(\underset{\sim}{V}^A) \geq S(\underset{\sim}{V}^B) \quad \text{for all symmetric quasi-concave functions } S(\cdot) \text{ and all } U(\cdot) \text{ unique up to a positive linear transformation}$$

$$\Rightarrow S(\hat{\underset{\sim}{V}}^A) \geq S(\hat{\underset{\sim}{V}}^B) \quad \text{by symmetry of } S(\cdot)$$

$\Rightarrow$ there exists a bistochastic matrix $\hat{Q}$ such that

$$\hat{\underset{\sim}{V}}^A = \hat{Q}\, \hat{V}^B \quad \text{from Dasgupta, Sen and Starrett (1973), Theorem 1}$$

$$\Rightarrow P_A^{-1}\, \hat{\underset{\sim}{V}}^A = P_A^{-1}\, \hat{Q}\, P_B\, P_B^{-1}\, \hat{V}^B$$

$$\Rightarrow \underset{\sim}{V}^A = [P_A^{-1} \hat{Q} P_B] \underset{\sim}{V}^B$$

$$\Rightarrow \underset{\sim}{V}^A = Q \underset{\sim}{V}^B$$

$$\Rightarrow [I - \gamma A]^{-1} \underset{\sim}{U} = Q[I - \gamma B]^{-1} \underset{\sim}{U}$$

$$\Rightarrow [I - \gamma A]^{-1} = Q[I - \gamma B]^{-1}$$

Since the above equality is true for all $U(\cdot)$ unique up to a positive linear transformation

$$\Rightarrow B = \frac{1}{\gamma} [I - Q] + AQ$$

and proof of necessity is completed

5. Interpretations

It should be noted that the theorem of the previous section provides us with a necessary and sufficient characterization result. Within the class of bistochastic transition matrices and restricting ourselves to steady states; if we are not willing to specify the social welfare function in any greater detail than that it by symmetric and quasi-concave in dynastic expected welfares; and if we accept that lifetime utility is the present discounted sum of non-Neumann Morgenstern expected utility over an infinite horizon; then to say that social welfare is higher under transition matrix A than under transition matrix B, is equivalent to saying that A and B stand in the relation

$$B = \frac{1}{\gamma} [I - Q] + AQ \tag{5.1}$$

where Q is a bistochastic matrix. Of course, this is a very strong requirement. But under the conditions stated this is all that can be said. One route for further research may be to relax some of these conditions, but here we will restrict ourselves to relating (5.1) to some conventional views on "greater mobility".

Given any two bistochastic transition matrices A and B, from (5.1) we simply have to check whether or not

$$Q = [I - \gamma A]^{-1} [I - \gamma B] \tag{5.2}$$

is a bistochastic matrix. Then, and only then, will A's welfare ranking be higher than that of B for the class of social welfare functions being considered. Now using

$$[I - \gamma A]^{-1} = \sum_{t=o}^{\infty} \gamma^t A^t$$

It can be seen that

$$\underset{\sim}{e}' Q = \underset{\sim}{e}' \; ; \; Q \underset{\sim}{e} = \underset{\sim}{e} \tag{5.3}$$

Thus all that remains to be checked is whether each element of the matrix in (5.2) is non-negative ie whether $q_{ij} \geq 0$.

For any two matrices, calculation of (5.2) and checking if the resultant matrix is non-negative is the operational method of identifying social welfare dominance in our setting. It is the equivalent of the

Lorenz curve comparison in the static case, and the Atkinson criterion (3.9) in the case of measuring mobility when dynastic inequality is ignored but intertemporal non-separability is introduced. Let us illustrate (5.2) with some examples. Start with an intuitively obvious case. How does the identify matrix compare to other bistochastic matrices? Let B = I in (5.2) and let A be any other bistochastic matrix, noting that from the many steady state distributions for I we choose

$p^* = (\frac{1}{n},\frac{1}{n},\frac{1}{n},\ldots, \frac{1}{n})$ to maintain comparability. Then from (5.2)

$$Q = (1 - \gamma)\ [I - \gamma A]^{-1} \tag{5.4}$$

all of whose elements are non-negative. Thus <u>any</u> matrix in the bistochastic class is ranked better than the identity matrix. Moreover, letting A = I in (5.2) we get

$$Q = \frac{1}{1-\gamma}\ [I - \gamma B] \tag{5.5}$$

some of whose off-diagonal elements will always be negative if B is not itself the identity matrix. Thus I is not ranked better than any other bistochastic matrix. It is in this sense that I is the unique globally "worst" transition matrix in the bistochastic class.

Consider now the "equality of opportunity view", which would regard the matrix

$$E = \frac{1}{n}\begin{bmatrix} \underset{\sim}{e}' \\ \underset{\sim}{e}' \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \underset{\approx}{e}' \end{bmatrix} = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \tag{5.6}$$

as the best in the class of bistochastic transition matrices. Let us compare this with the matrix K defined in (2.4), where correlation between one period's income and the next period's income is the most negative it can be. First of all, let A = E and B = K in (5.2). Noting that (see Kanbur and Mukerji, 1980)

$$[I - \gamma E]^{-1} = [I + \frac{\gamma}{1-\gamma} E] \tag{5.7}$$

we have

$$Q = [I + \frac{\gamma}{1-\gamma} E] [I - \gamma K]$$

$$= I - \gamma K + \gamma E \tag{5.8}$$

In deriving the final form of (30) we have used the fact that KE=E. Now from (5.8) and the definition of K in (2.4),

$$i \neq j;\ q_{ij} = \gamma(\frac{1}{n} - k_{ij}) = \begin{cases} \gamma(\frac{1}{n} - 1) & \text{for } j=n-i+1\neq i \\ \frac{\gamma}{n} & \text{for } j\neq n-i+1,\ j\neq i \end{cases}$$

$$i = j;\ q_{ij} = 1 - \gamma k_{ii} + \gamma \frac{1}{n} = \begin{cases} 1 - \gamma + \gamma \frac{1}{n} \text{ for } i = \frac{n+1}{2} \text{ (n odd)} \\ 1 + \gamma \frac{1}{n} \text{ for other } i \end{cases} \tag{5.9}$$

Thus at least some elements of Q (those for which $j=n-i+1\neq i$) will be negative and E <u>cannot</u> be established as unambiguously better than K.

Can K be established an unambiguously better than E? Letting A = K, B = E in (5.2), and noting that

$$[I - \gamma K]^{-1} = \frac{1}{1-\gamma^2} [I + \gamma K] \tag{5.10}$$

we get

$$Q = [I - \gamma K]^{-1} [I - \gamma E]$$

$$= \frac{1}{1-\gamma^2} I + \frac{\gamma}{1-\gamma^2} K - \frac{\gamma}{1-\gamma} E \tag{5.11}$$

Hence

$$i\neq j;\ q_{ij} = \frac{\gamma}{1-\gamma^2} k_{ij} - \frac{\gamma}{1-\gamma} \cdot \frac{1}{n} = \begin{cases} \frac{\gamma}{1-\gamma^2} - \frac{\gamma}{1-\gamma} \cdot \frac{1}{n} \text{ for } j=n-i+1\neq i \\ - \frac{\gamma}{1-\gamma} \cdot \frac{1}{n} \text{ for } j\neq n-i+1,\ j\neq i \end{cases}$$

$$i=j;\ q_{ii} = \frac{1}{1-\gamma^2} + \frac{\gamma}{1-\gamma^2} k_{ii} - \frac{\gamma}{1-\gamma} \cdot \frac{1}{n} = \begin{cases} \frac{n-\gamma(1+\gamma)}{n(1-\gamma^2)} + \frac{\gamma}{1-\gamma^2} \text{ for } i = \frac{n+1}{2} \text{ (n odd)} \end{cases}$$

$$\frac{n-\gamma(1+\gamma)}{n(1-\gamma^2)} \quad \text{for other } i$$

(5.12)

The only dimension for which the $j \neq n-i+1$ and $j \neq i$ case is not a possibility is $n=2$. For $n=3$, the element q_{23}, for example, satisfies $j \neq n-i+1$ and $j \neq i$.

Thus in the 2x2 case K is unambiguously ranked better than E. However, more generally the two cannot be ranked without specifying the social welfare function further. In fact in the 2x2 case if we let

$$A = \begin{bmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{bmatrix}; \quad B = \begin{bmatrix} \beta & 1-\beta \\ 1-\beta & \beta \end{bmatrix} \tag{5.13}$$

then

$$Q = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}; \quad q = \frac{1+\frac{1}{\gamma}-(\alpha+\beta)}{1+\frac{1}{\gamma}-2\alpha} \tag{5.14}$$

Clearly in this case q is always non-negative and

$$\beta \geq \gamma \iff q \leq 1 \tag{5.15}$$

So that if B is worse than A on the "diagonals view" then, and only then, is it unambiguously worse from a social welfare point of view. Of course, it should be noted that in the 2x2 case the "diagonals view" and the "strong diagonals view" or "monotonicity", as Shorrocks (1978a) called it coincide exactly. When presenting the incompatibility of the equality of opportunity view with monotonicity in the context of the 2x2 example, Shorrocks (1978a) argued that "on balance monotonicity seems to be the less artificial restriction." The result here makes this intuition precise in the social welfare framework.

Of course no comparable general validation of the "diagonals view" is available for $n \geq 3$. A counter example is easily presented using the result comparing K and E for n=4.

$$K = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} ; \quad E = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \tag{5.16}$$

On the diagonals view (though not on the strong diagonals view) K is superior to E for n=4. But we already know from (5.12) that an unambiguous ranking on social welfare grounds is not available. More generally, for any two bistochastic matrices A and B, let

$$D = B - A \tag{5.17}$$

If B and A stand in the relation (5.1), then

$$d_{ij} = \frac{1}{\gamma}[\delta_{ij} - q_{ij}] + \sum_{\ell=1}^{n} a_{i\ell} q_{\ell j} - a_{ij} \tag{5.18}$$

In particular,

$$d_{ii} = \sum_{\ell \neq 1} a_{i\ell} q_{\ell i} + \left(\frac{1}{\gamma} - a_{ii}\right)(1 - q_{ii}) \tag{5.19}$$

$$\geq 0$$

Thus any two matrices which stand in the relation (5.1) satisfy the diagonals view. But the converse is not true: any two matrices which satisfy the diagonals view need not necessarily stand in the relation (5.1) to each other - the counter example in (5.16) establishes that.

A general interpretation of (5.1) can be provided by writing it as

$$B = [\frac{1}{\gamma} I] [I-Q] + AQ.$$

B can thus be seen as a matrix weighted sum of $\frac{1}{\gamma}$ I and A, the weights being Q and [I-Q]. If there exists a Q such that B can be written as this weighted sum, in a general sense B is closer to the Identity matrix than A. If interpreted in this wide sense, the "diagonals view" has a rationale in a social welfare function that is egalitarian with respect to dynastic prospects.

For a specific application, consider the matrices A, B and C given below, which are taken from Atkinson (1980a).

$$A = \begin{bmatrix} 0.48 & 0.42 & 0.10 & 0.0 \\ 0.25 & 0.34 & 0.27 & 0.14 \\ 0.19 & 0.17 & 0.37 & 0.27 \\ 0.08 & 0.07 & 0.26 & 0.59 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.35 & 0.24 & 0.27 & 0.14 \\ 0.26 & 0.30 & 0.27 & 0.17 \\ 0.19 & 0.17 & 0.37 & 0.27 \\ 0.20 & 0.29 & 0.09 & 0.42 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.44 & 0.23 & 0.19 & 0.14 \\ 0.32 & 0.26 & 0.25 & 0.17 \\ 0.18 & 0.36 & 0.27 & 0.19 \\ 0.06 & 0.15 & 0.29 & 0.50 \end{bmatrix}$$

Matrix C is a father to son, quartile to quartile, transition matrix for hourly earnings in Britain, derived from the work of Atkinson, Maynard and Trinder (1983). Matrices A and B are alternative specifications of father to son occupational transition matrices which are got by adjusting the Goldthorpe (1980) data from its seven classes into four and enforcing bistochasticity. Atkinson (1980a) provides a fuller discussion of the procedure, but essentially matrix A has been derived by shifting weight towards the diagonal and matrix B by shifting weight away from the diagonal. Atkinson compares C with A and C with B using his criterion (3.9), and finds that C and A cannot be so ranked, but that "the earnings transition matrix is welfare inferior to case B of the occupational status matrix."

Can the same be said of the matrices using our criterion? For γ = 0.5 (an intergenerational interest rate of 100%), we get the following Q matrices:

$$Q_{CA} = [I-\gamma C]^{-1}[I-\gamma A]$$

$$= \begin{bmatrix} 0.978 & -0.111 & 0.051 & 0.082 \\ 0.035 & 0.953 & -0.009 & 0.021 \\ -0.002 & 0.095 & 0.947 & -0.040 \\ -0.011 & 0.063 & 0.011 & 0.938 \end{bmatrix}$$

$$Q_{BC} = [I-\gamma B]^{-1}[I-\gamma C]$$

$$= \begin{bmatrix} 0.948 & -0.001 & 0.049 & 0.003 \\ -0.034 & 1.016 & 0.016 & 0.002 \\ 0.009 & -0.101 & 1.050 & 0.042 \\ 0.076 & 0.086 & -0.115 & 0.953 \end{bmatrix}$$

As in the case of Atkinson's criterion, C cannot be ranked as better than A. However, what is interesting is that B cannot be ranked as better than C, in contrast to the Atkinson finding. Thus while it should be clear that there are a number of empirical problems in comparing earnings transition matrices with occupational ones, our calculations nevertheless indicate that a stress on dynastic inequality could alter rankings which emphasize intertemporal non-separability.

6. Further Research

Much remains to be done. While we have shown that our criterion for comparing transition matrices can lead to different rankings than the Atkinson criterion, the exact nature of the difference between them will bear further investigation. More generally, the benchmark proposition in Section 3 provides an ordered agenda for research. Relaxation of any one of the four assumptions listed there allows us to drive a wedge between alternative mobility structures in terms of their social welfare implications. In this paper we have concentrated on relaxing the assumption that social welfare is simply additive in dynastic prospects, and in this first cut we have restricted attention to the case where dynastic prospects are intertemporally separable in their utility consequences. An analysis of the intertemporally non-separable <u>and</u> interdynastically non-separable social welfare function is the logical next step. Also waiting in the wings are the analysis of non-bistochastic transition matrices and of non-steady time paths.

References

Atkinson, A.B. (1970): "On the measurement of inequality," Journal of Economic Theory.

Atkinson, A.B. (1979): "Intergenerational income mobility," IHS Journal.

Atkinson, A.B. (1980a): "On intergenerational income mobility in Britain," Journal of Post Keynesian Economics.

Atkinson, A.B. (1980b): "The measurement of economic mobility," in Essays in Honour of Jan Pen.

Atkinson, A.B. and F. Bourgignon (1982): "The comparison of multi-dimensional distributions of economic status, "Review of Economic Studies.

Atkinson, A.B., A.K. Maynard and C.G. Trinder (1983): Parents and children: incomes in two generations, London: Heinemann Educational Books.

Bartholomew, D.J. (1973): Stochastic models for social processes, 2nd edition, New York: John Wiley & Sons.

Chakravarty, S.R., B. Dutta and J.A. Weymark (1985): "Ethical indices of income mobility," Social Choice and Welfare.

Dasgupta, P., A.K. Sen and D. Starrett (1973): "Notes on the measurement of inequality," Journal of Economic Theory.

Friedman, M. (1962): Capitalism and freedom, University of Chicago Press.

Goldthorpe, J.H. (1980): Social mobility and class structure in modern Britain. London: Oxford University Press.

Hart, P. (1980): "The statics and dynamics of income distribution: A survey," in N.A. Klevmarken and J.A. Lybeck (eds.) The Statics and Dynamics of Income, Tieto.

Kanbur, S.M.R. and V. Mukerji (1980): "Probabilistic equity and equality of opportunity in a stochastic framework," Arthaniti.

Kanbur, S.M.R. and J.E. Stiglitz (1982): "Mobility and inequality: a Utilitarian analysis," Economic Theory Discussion Paper No. 57, Department of Applied Economics, University of Cambridge.

Kanbur, S.M.R. and J.O. Stromberg (1985): "Income Transitions and Income Distribution Dominance," forthcoming in Journal of Economic Theory.

Markandya, A. (1982): "Intergenerational exchange mobility and economic welfare," European Economic Review.

Markandya, A. (1984): "The welfare measurement of changes in economic mobility," Economica.

Rothschild, M. and J.E. Stiglitz (1973): "Some further results on the measurement of inequality," Journal of Economic Theory.

Rothschild, M. and M. Yaari (1979): "Comparing discounted streams of income," Mimeographed.

Shorrocks, A.F. (1978a): "The measurement of mobility," Econometrica.

Shorrocks, A.F. (1978b): "Income inequality and income mobility," Journal of Economic Theory.