Seminar Paper No. 339

INCOME TRANSITIONS AND
INCOME DISTRIBUTION DOMINANCE

by

S.M. Ravi Kanbur

and

Jan Olov Strömberg

INSTITUTE FOR INTERNATIONAL ECONOMIC STUDIES

University of Stockholm

ISSN 0347-8769

Seminar Paper No. 339

INCOME TRANSITIONS AND INCOME DISTRIBUTION DOMINANCE

by

S.M. Ravi Kanbur
Department of Economics
University of Essex, England
and
Woodrow Wilson School
Princeton University, U.S.A.

and

Jan Olov Strömberg
Institute of Mathematical and Physical Sciences
University of Tromsø, Norway

Abstract

This note analyses dominance relations between sequences of income distributions generated by income transition mechanisms. Necessary and sufficient conditions on these mechanisms are derived for a dominance relation to continue once established.

Seminar Papers are preliminary material circulated
to stimulate discussion and critical comment.

October, 1985

Institute for International Economic Studies
S-106 91 Stockholm
Sweden

1. Introduction

In a series of papers published in JET, a number of authors [Atkinson (1970), Dasgupta, Sen and Starrett (1973), Rothschild and Stiglitz (1973)] have addressed themselves to the problem of ranking income distributions. The object has been to find necessary and sufficient conditions for a pair of distributions to be ranked in the same way by each member of a given class of social welfare functions. When this happens one distribution is said to be "unambiguously more equal" than, or to "dominate", the other distribution. These are seminal results in the literature on inequality measurement, and their importance is not to be doubted. However, in recent years a number of authors have pointed out that these results refer to a "snapshot" comparison of two distributions - the next snapshot in time may well be different [see, for example, Hart, (1980)]. The rankings on the next snapshot may, for example, be reversed - it all depends on the transition mechanisms in play.

Intuitively one would expect the problem to be a serious one, in the sense that a ranking between two distributions in one snapshot would give little information on the rankings in the next snapshot. But can this intuition be made precise? We would expect the class of transition mechanisms which continued a dominance ranking into the next snapshot to be very restrictive. But how restrictive? What precisely is this class of transition mechanisms? The object of this note is to provide an answer to this question - to delineate precisely what one can and cannot say about dominance rankings in the future, given a dominance ranking in the present. Section 2 sets up the model and notation, while Sections 3 and 4 present the main results. Section 5 concludes the paper.

2. Model

We will analyse income transitions in a discrete time - continuous state space framework. Let x denote a variable representing income in period t , and let y represent income in the next period, $t + 1$. Both x and y are restricted to lie in the range $[\underline{z}, \bar{z}]$. Let $a(y|x)$ be the "income transition function" in a society representing the probability density of income in period $t + 1$ conditional on income in period t . The income transition function will be assumed to remain unchanged over time, and we will compare the sequences of income distributions produced by two transition functions $a(y|x)$ and $b(y|x)$. Let the initial income distributions to which these transition functions are applied be the densities $p_t(x)$ and $q_t(x)$, respectively. Then the income densities in the next period are of course given by

$$P_{t+1}(y) = \int_{\underline{z}}^{\bar{z}} a(y|x) p_t(x) dx \quad (1)$$

$$q_{t+1}(y) = \int_{\underline{z}}^{\bar{z}} b(y|x) q_t(x) dx \quad (2)$$

We are interested in the relation between $p_{t+1}(\cdot)$ and $q_{t+1}(\cdot)$, given a relation between $p_t(\cdot)$ and $q_t(\cdot)$.¹

The relation between distributions which is the focal point of our analysis is that of a mean preserving spread (see Rothschild and Stiglitz, 1970). For a density $p_t(\cdot)$, define

$$\hat{p}(x) = \int_{\underline{z}}^x p(\theta) d\theta \quad (3)$$

$$\hat{\hat{p}}(x) = \int_{\underline{z}}^x \hat{p}(\theta) d\theta \quad (4)$$

where $\hat{p}(x)$ is of course the cumulative distribution function while $\hat{\hat{p}}(x)$ is the area under this function up to x . We define a (weak) dominance relation D between income density functions as follows

$$[p D q] \Leftrightarrow [\hat{p}(x) \leq \hat{q}(x) \text{ for all } x; \text{ and } \hat{p}(\bar{z}) = \hat{q}(\bar{z})] \quad (5)$$

Now $\hat{p}(x) \leq \hat{q}(x)$ for all x is a statement that $\hat{p}(\cdot)$ second order dominates $\hat{q}(\cdot)$, while $\hat{p}(\bar{z}) = \hat{q}(\bar{z})$ is a statement that the means of the two distributions are the same (see Hadar and Russell, 1969, and Rothschild and Stiglitz, 1970). In other words, we say that p dominates q if q is a means

¹ There are two interpretations that can be put on this transition mechanism. One is that $a(y|x)$ is a frequency density, representing the fraction of those with income x today who end up getting income y tomorrow. For any individual this could be consistent with the transition itself being stochastic or non-stochastic - it is the overall change in the distribution of individuals across income levels which is important from our point of view. If we view the process as being stochastic for any given individual, then (1) and (2) are of course the expected income distributions in the next period, but we can appeal to large numbers to argue that these should indeed be the focus of our attention.

preserving spread of p . Not surprisingly, we will refer to $\hat{p}(x)$ as an "indicator of spread of $p(\cdot)$ at x ". The significance of mean preserving spreads in terms of social welfare rankings at a point in time is well known [see Atkinson (1970), Dasgupta, Sen and Starrett (1973)] and need not be elaborated upon here. Our task, however, is to consider the relationship between a ranking in one period and the rankings as time rolls forward. Since we give no account of incentive effects in the model, it behooves us to concentrate on situations where mean income is constant, focusing purely on aspects of inequality.

Let us now make precise the question to be asked. Suppose we know that the income distribution in one society is a mean preserving spread of the income distribution in another at a point in time, without necessarily knowing the income distributions in detail. Under what conditions will these income distributions stay in the same relation as time goes on? Intuitively we might expect the class of transition mechanisms which lead to this result to be fairly restricted. But what exactly is this class of transition mechanisms? In other words, what are the necessary and sufficient conditions on $a(y|x)$ and $b(y|x)$ such that

$$p_t(x) D q_t(x) \Rightarrow p_{t+1}(y) D q_{t+1}(y)$$

for all $p_t(x)$ and $q_t(x)$ satisfying the dominance and range condition?

3. Results

In order to answer the question posed at the end of the last section, we have to consider the indicators of spread of $p_{t+1}(\cdot)$ and $q_{t+1}(\cdot)$: From (1) and (2), the linearity of the integral operator gives us the following

$$\hat{p}_{t+1}(y) = \int_{\underline{z}}^{\bar{z}} \hat{a}(y|x) p_t(x) dx \quad (6)$$

$$\hat{q}_{t+1}(y) = \int_{\underline{z}}^{\bar{z}} \hat{b}(y|x) q_t(x) dx \quad (7)$$

where

$$\begin{aligned} \hat{a}(y|x) &= \int_{\underline{y}}^y \int_{\underline{z}}^z a(\theta|x) d\theta dy \\ \hat{b}(y|x) &= \int_{\underline{y}}^y \int_{\underline{z}}^z b(\theta|x) d\theta dy \end{aligned} \quad (8)$$

Expressions (6) and (7) are basic to our analysis. They establish that the indicator of spread of the period $t+1$ distribution is a weighted sum of the indicators of spread of conditional prospects, the weights being given by the actual distribution in period t .

We start the analysis with the case where $a(y|x) \equiv b(y|x)$ i.e. the transition mechanisms in the two societies are the same. If the common transition function is ergodic we know that in the long run the two societies will have the same distribution of income. But what will the paths to the steady state look like? In particular, when will a dominance ranking be preserved over time? Denoting $a(y|x)$ as the common transition function, the following proposition gives the answer:

Proposition I

$$[p_t(x) D q_t(x) \Rightarrow p_{t+1}(y) D q_{t+1}(y)]$$

if and only if

$$[\hat{a}(y|x) \text{ is convex in } x \text{ for all } y; \text{ and } \hat{a}(\bar{z}|x) \text{ is linear in } x] \quad (9)$$

We will first of all prove this proposition and then proceed to a discussion and interpretation of the condition. From (6) and (7), setting $b \equiv a$,

$$\hat{p}_{t+1}(y) - \hat{q}_{t+1}(y) = \int_{\underline{z}}^{\bar{z}} \hat{a}(y|x) [p_t(x) - q_t(x)] dx \quad (10)$$

The sufficiency part of proposition (9) follows straightforwardly by noting that since $q_t(x)$ is a mean preserving spread of $p_t(x)$, if $\hat{a}(y|x)$ is convex in x then the expression in (10) must be non-positive (this is analogous to comparisons of expected utility for a convex utility function - see Rothschild and Stiglitz, 1970). The equality of $\hat{p}_{t+1}(\bar{z})$ and $\hat{q}_{t+1}(\bar{z})$ also follows straightforwardly from the linearity of $\hat{a}(\bar{z}|x)$ in x , and from the fact that $p_t(x)$ and $q_t(x)$ have the same mean.

For proof of necessity, first of all choose $p_t(x)$ to be the degenerate distribution concentrated at x_0 and let $q_t(x)$ have probability mass π at $x_1 < x_0$ and $(1-\pi)$ at $x_2 > x_0$ such that $x_0 = \pi x_1 + (1-\pi)x_2$. Then it is clear that

$$\begin{aligned} p_{t+1}(y) - q_{t+1}(y) &= \hat{a}(y|x_0) - [\pi \hat{a}(y|x_1) + (1-\pi) \hat{a}(y|x_2)] \\ &\leq 0 \Rightarrow \hat{a}(y|x) \text{ is convex in } x. \end{aligned}$$

Also,

$$\begin{aligned} p_{t+1}(\bar{z}) - q_{t+1}(\bar{z}) &= \hat{a}(\bar{z}|x_0) - [\pi \hat{a}(\bar{z}|x_1) + (1-\pi) \hat{a}(\bar{z}|x_2)] \\ &= 0 \Rightarrow \hat{a}(\bar{z}|x) \text{ is linear in } x. \end{aligned}$$

Proposition (I) provides a complete characterisation of the common transition mechanism which will maintain a dominance ranking once established. Not surprisingly, the class of transition mechanisms which permit this is very restricted. But the proposition does tell us precisely when the snapshot at a point in time characterises all the snapshots in the future. Let us now look more closely at the conditions in proposition (9) and interpret them. The condition that $\hat{a}(\bar{z}|x)$ be linear in x is easily interpreted. Denoting $\mu(x)$ as the mean of income in period $t+1$ conditional on the income in period t , it is easy to establish (by integration by parts) that

$$\mu(x) = \int_{\underline{z}}^{\bar{z}} y a(y|x) dy = \bar{z} - \hat{a}(\bar{z}|x)$$

Hence we require that the conditional mean of tomorrow's income be a linear function of today's income. If a transition mechanism satisfies this requirement then equality of means of p_t and q_t implies equality of means of p_{t+1} . [Notice that this does not say anything about the means of p_t and p_{t+1} being equal, or about the means of q_t and q_{t+1} being equal].

The condition that $\hat{a}(y|x)$ be convex in x for all y is somewhat more difficult to interpret. For given x , $\hat{a}(y|x)$ is an indicator of spread of the probability distribution facing an individual with current income x .

The values of $\hat{a}(y|x)$ at all values of y present an entire picture of the spread of riskiness of the transition facing an individual with current income x . Our condition can be interpreted as saying that going from the lowest x to the highest x , the change in riskiness accelerates. The riskiness of the transition is a convex function of current income. The three possibilities are depicted in Figure 1. As can be seen, our condition is consistent with riskiness increasing or decreasing with current income - what is important is its convexity. The convexity of riskiness as a function of current income means that the "average degree of riskiness" is greater when the current distribution is more spread out. Hence our result that with such transition mechanisms the society which starts off more unequal remains more unequal.

4. Generalisations

~~So far we have restricted attention to the case where the two societies~~
have identical transition mechanisms. In this section we will consider the general case of two societies with different transition mechanisms. Before doing this, however, we note here a curiosum. Following an argument similar to that used in the proof of proposition I, the following proposition can also be proved:

$$[p_t(x) \text{ D } q_t(x) \Rightarrow q_{t+1}(y) \text{ D } p_{t+1}(y)]$$

if and only if

$$[\hat{a}(y|x) \text{ is concave in } x \text{ for all } y; \text{ and } \hat{a}(\bar{z}|x) \text{ is linear in } x]$$

(11)

In this case the dominance relation switches every period. Since transition risk is now a concave function of initial income, the "average transition risk" is now lower for the more unequal distribution - hence the dominance switches. Of course if the transition function is ergodic then both societies will tend in to the same distribution in the long run, but the path will be such that the income distribution in a given society will be the better of the two in one period, worse in the next, and so on.

Let us now consider the case where the transition mechanisms in the two societies differ but the initial distributions are the same i.e.

$p_t(x) \equiv q_t(x)$ but $a(y|x) \neq b(y|x)$. In this case

$$\hat{p}_{t+1}(y) - \hat{q}_{t+1}(y) = \int_{\underline{z}}^{\bar{z}} [\hat{a}(y|x) - \hat{b}(y|x)] p_t(x) dx \quad (12)$$

so that we can prove the fairly obvious result:

Proposition II

$$[p_{t+1}(y) D q_{t+1}(y) \text{ for all } p_t(x) \equiv q_t(x)]$$

$$\Leftrightarrow [a(y|x) D b(y|x) \text{ for all } x] \quad (13)$$

In other words, if given the same initial distribution, every conditional transition in one society dominates the corresponding conditional transition in another society, then the resulting income distribution in the former dominates the resulting distribution in the latter. The proof of sufficiency in (13) follows directly from the definition of the dominance relations D , as specified in (5), and necessity follows by choosing degenerate initial distributions at each value of x .

Finally, let us consider the general case $p_t(x) \neq q_t(x)$ and $a(y|x) \neq b(y|x)$. In other words, the two societies start from non-identical distributions and have different transition mechanisms. As we might expect, the class of transition mechanisms which preserve dominance is very restricted. The following proposition specifies exactly how restricted it is:

Proposition III

$$[p_t(x) \text{ D } q_t(x) \Rightarrow p_{t+1}(y) \text{ D } q_{t+1}(y)]$$

if and only if

[There exists a function $C(x,y)$, convex in x for all y , such that

$$\hat{a}(y|x) \leq C(x,y) \leq \hat{b}(y|x) \text{ for all } x,y; \text{ and } \hat{a}(\bar{z}|x) = \hat{b}(\bar{z}|x) = \alpha + \beta x] \quad (14)$$

The proof of sufficiency goes as follows:

$$\begin{aligned} \hat{p}_{t+1}(y) - \hat{q}_{t+1}(y) &= \int_{\underline{z}}^{\bar{z}} [\hat{a}(y|x)p_t(x) - \hat{b}(y|x)q_t(x)]dx \\ &\leq \int_{\underline{z}}^{\bar{z}} C(x,y)[p_t(x) - q_t(x)]dx \leq 0 \quad (15) \end{aligned}$$

since $C(x,y)$ is convex in x and $q_t(x)$ is a mean preserving spread of $p_t(x)$ [see Rothschild and Stiglitz (1970)]. Also

$$\hat{p}_{t+1}(\bar{z}) - \hat{q}_{t+1}(\bar{z})$$

is easily seen to be zero when $\hat{a}(\bar{z}|x) = \hat{b}(\bar{z}|x) = \alpha + \beta x$, i.e. when both $\hat{a}(\bar{z}|x)$ and $\hat{b}(\bar{z}|x)$ are identical linear functions in x .

For proof of necessity, choose $p(x)$ to be the degenerate distribution concentrated at x_0 and let $q(x)$ have probability π at $x_1 < x_0$ and $1 - \pi$ at $x_2 > x_0$ such that $x_0 = \pi x_1 + (1-\pi)x_2$. Then it is clear that

$$\hat{p}_{t+1}(y) \leq \hat{q}_{t+1}(y) \Rightarrow \hat{a}(y|x_0) \leq \pi \hat{b}(y|x_1) + (1-\pi) \hat{b}(y|x_2) \quad (16)$$

By choosing π , x_0 , x_1 and x_2 appropriately, we can derive from (16) the implication that for given y , $\hat{a}(y|x)$ must lie everywhere below the convex hull of $\hat{b}(y|x)$, as illustrated in Figure 2. We then have the convex function $C(x,y)$ needed in the statement $\hat{a}(y|x) \leq C(x,y) \leq b(y|x)$.

Proof of necessity is completed by considering

$$\int_{\underline{z}}^{\bar{z}} [\hat{a}(\bar{z}|x)p(x) - \hat{b}(\bar{z}|x)q(x)]dx = 0$$

Choosing $p(x) \equiv q(x)$ to be degenerate at x shows that $\hat{a}(\bar{z}|x) \equiv \hat{b}(\bar{z}|x)$.

Given this identity, now choose $q(x)$ to be a mean preserving spread of $p(x)$; it follows that the above expression is zero only if $\hat{a}(\bar{z}|x) \equiv \hat{b}(\bar{z}|x) = \alpha + \beta x$ [see Rothschild and Stiglitz (1970)].

Proposition III is a generalisation of proposition I, as is seen by noting that (14) implies (9) when $a(y|x) = b(y|x)$. When the two societies have non-identical transition mechanisms, their ergodic distributions will also be different. However, notice that the condition in (14) is a sufficient condition for the ergodic distribution of $a(y|x)$ to dominate the ergodic distribution of $b(y|x)$ since if the dominance relation is maintained at each step of the sequence it must be maintained in the limit as well. The condition is not necessary, of course, since the ergodic distributions are the limits of other sequences also. The question of the necessary and sufficient set of conditions for the ergodic distribution of one transition matrix to dominate that of another is an open area for research.

How can we interpret the condition in proposition (14)? Intuitively speaking, if the average transition was riskier for the society in which income distribution was already more unequal, then we would expect this society to stay the more unequal one. The condition in the proposition makes this intuition precise. In what sense can we say that one transition mechanism is riskier than another? One obvious answer is that every transition in one be riskier than the corresponding transition in the other. As Figure 2 shows, this requirement is not strong enough. Something more is needed to take care of "averaging out" of riskiness - the function $C(x,y)$ specifies exactly what is required.

5. Conclusion

Suppose we were only interested in dominance rankings between income distributions. Under what conditions is it then the case that a snapshot contains all the information we require - in the sense that a dominance ranking in that snapshot guarantees the same ranking at all future points in

time? Intuitively, we would expect the class of transition mechanisms that ~~permits this to be extremely restrictive. But what precisely is that class~~ of mechanisms? This note provides the answer by means of necessary and sufficient conditions on transition mechanisms which carry forward an established dominance relation into the future.

References

Atkinson, A. B. (1970). "On the Measurement of Inequality", Journal of Economic Theory.

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

Hadar, J., and Russell, W. R. (1969). "Rules for Ordering Uncertain Prospects", American Economic Review.

Hart, P. E. (1980). "The Statics and Dynamics of Income Distributions: A Survey", in W. A. Klevmarken and J. A. Lybeck (eds.). The Statics and Dynamics of Income, Tieto.

Rothschild, M. and Stiglitz, J. E. (1970), "Increasing Risk I: A Definition", Journal of Economic Theory.

Rothschild, M. and Stiglitz, J. E. (1973), "Some Further Results on the Measurement of Inequality", Journal of Economic Theory.

Figure 1

Indicators
of spread
for given
 y

Figure 2