NBER WORKING PAPER SERIES

MOBILITY AS PROGRESSIVITY:
RANKING INCOME PROCESSES ACCORDING TO EQUALITY OF OPPORTUNITY

Roland Benabou
Efe A. Ok

Working Paper 8431
http://www.nber.org/papers/w8431

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2001

We are grateful to thank Andrea Ichino for providing us with the data on Italy and the United States. Roland Benabou gratefully acknowledges financial support from the National Science Foundation and the MacArthur Foundation. Efe A. Ok gratefully acknowledges financial support from the C.V. Starr Center. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

Mobility as Progressivity:
Ranking Income Processes According to Equality of Opportunity
Roland Benabou and Efe A. Ok
NBER Working Paper No. 8431
August 2001
JEL No. D31, D63, H20, J62

ABSTRACT

Interest in economic mobility stems largely from its perceived role as an equalizer of opportunities, though not necessarily of outcomes. In this paper we show that this view leads very naturally to a methodology for the measurement of social mobility which has strong parallels with the theory of progressive taxation. We characterize opportunity--equalizing mobility processes, and provide simple criteria to determine when one process is more equalizing than another. We then explain how this mobility ordering relates to social welfare analysis, and how it differs from existing ones. We also extend standard indices of tax progressivity to mobility processes, and illustrate our general methodology on intra- and intergenerational mobility data from the United States and Italy.

Roland Benabou
Department of Economics
Princeton University
Princeton, NJ 08544
and NBER
rbenabou@princeton.edu

Efe A. Ok
Department of Economics
New York University
269 Mercer Street
New York, NY 10003
efe.ok@nyu.edu

Introduction

Since the work of Kolm (1960), Atkinson (1970), Dasgupta, Sen and Starett (1973) and others, the measurement of income inequality is largely a settled issue. By contrast, the measurement of mobility remains in a state of flux, with the literature showing a somewhat bewildering array of approaches and concepts (see Fields and Ok (1999a) for a recent survey).[1] All emphasize different yet overlapping aspects of distributional dynamics, and without a prior view of *why* mobility is important, none stands out as uniquely compelling.

The point of this paper is that *equality of opportunity* provides a very natural approach to the evaluation of mobility processes, so natural that there is in fact no need for special concepts or indices to measure it. One cares about mobility not because income movements are intrinsically valuable, but primarily because of the view –or the hope– that it helps attenuate the effects of disparities in initial endowments, or social origins, on future income prospects (e.g., Stokey (1998)). From this view of mobility as an equalizer of opportunities (but not necessarily of outcomes), it follows quite naturally that one should measure it precisely by the extent to which it achieves such levelling. This, in turn, corresponds to a notion of progressivity quite similar to that used to assess tax functions in public finance. In other words, (desirable) mobility *is* progressivity, in the mapping between initial incomes and future opportunities. Measures of pure persistence and other existing mobility indices are of course not unrelated to this notion of opportunity–equalization, but none directly corresponds to it. In particular, movements in relative incomes may in general be equalizing or disequalizing, and yet many mobility criteria proposed in the literature fail to distinguish between the two.

The main insight underlying this paper is thus to view income mobility as just another form of *redistribution* –albeit a stochastic one.[2] Just like a tax scheme maps pre-tax incomes into post-tax incomes, a mobility process maps initial incomes into expected future incomes, or more generally into expected levels of intertemporal welfare.[3] The extent to which the terminal distribution is equalized compared to the initial one is then precisely measured by the degree to which the mapping is progressive, in the sense of having decreasing average "tax" rates. This simple idea allows us to build on the theory of progressivity measurement and Lorenz dominance (Jakobsson (1976), Fellman (1976)) and easily obtain a number of useful results for the study of income mobility. We characterize progressive or opportunity–equalizing mobility processes and their main properties, and provide simple criteria to determine when one process is more progressive than another. We also explain how this ordering relates to social welfare analysis, and discuss how it differs from

[1] The available panoply ranges from purely statistical measures of persistence based on the eigenvalues of transition matrices (see Conlisk (1989)), average numbers of changes in ranks (Bartholomew (1982)), or variations in relative incomes (Fields and Ok (1999b)), to ratios between single and multi–period measures of inequality (Shorrocks (19878a)), and finally to a number of more explicitly welfare–based criteria (Kanbur and Stiglitz (1986), Dardanoni (1993), (1995)).

[2] This view was first put forward in a political economy context in Bénabou and Ok (2001), which studies how mobility expectations affect the political support for redistributive policies.

[3] The relevant definition of "opportunity" naturally depends on agents' horizons and attitudes towards risk; our approach will be quite flexible in that respect.

some of those found in the literature. Since it is of course only a partial ordering, we also show how to extend standard indices of tax progressivity to dynamic processes, so as to obtain mobility indices consistent with the opportunity–equalization ordering. Finally, we illustrate the proposed methodology with intra– and inter–generational mobility from the United States and Italy.

As mentioned above, we are interested in mobility as an equalizer of ex–ante opportunities (or welfare), not of ex–post outcomes. Thus, unlike Kanbur and Stromberg (1988) and Dardanoni (1995), our concern is not whether future *realized* income distributions will be more or less equal than the current one. They could be much more unequal, but if this primarily reflects shocks which were unpredictable on the basis of initial conditions there is little disparity of opportunity. Moreover, in steady-state the income distributions in different periods must coincide in any case, making outcome–based comparisons inapplicable.

Our approach of assessing mobility from an "ex–ante" perspective is more closely related to the papers which take a welfare–based view, such as Kanbur and Stiglitz (1986) and, especially, Dardanoni (1993). Like them, we judge a mobility process to be more equal than another one if it results in a higher Lorenz curve for the distribution of individual levels of intertemporal welfare. Their approach, however, is restricted to mobility processes described by transition matrices between discrete income states. Most importantly, it only allows two processes to be compared if they have the same steady–state, and start in that steady–state. This may seem less of a problem for mobility analyses based on fractile matrices (which specify only rank transitions), since all bistochastic matrices have the same steady–state, namely the uniform distribution. As we explain later on, however, these matrices provide a very incomplete picture of income mobility because, in practice, interfractile income differences vary considerably from one fractile to the next, as well as across countries which have different income distributions. Furthermore, there is really no reason to limit the analysis to steady–states in the first place.[4]

The approach developed in this paper, by contrast, is applicable to continuous as well as discrete processes (or mixtures of the two), and to economies both in and out of steady–state. Moreover, the view of mobility as just another form of redistribution provides a natural "metric" to assess income processes, namely that of tax policy: one can compare the implied residual elasticities, progressivity indices, and marginal tax rates to those which are familiar from fiscal redistributions. Our approach also relates very naturally to the macroeconomic literature on "convergence," since progressivity in a mobility process correspond (in the simplest case) to the property that an agent's expected rate of income growth declines with her initial level of income.

The paper is organized as follows. Section 1 introduces basic concepts and notations. Section 2 presents our mobility ordering and a characterization result for general (Markovian) mobility processes. Section 3 focuses on the case where mobility is represented by transition matrices between discrete income states, offering in particular a very simple test of matrix dominance. Section 4 shows how summary indices of equalizing (or progressive) mobility follow naturally from

[4] This last point is also raised by Formby, Smith and Zheng (1995), who relax the assumption of a common steady–state vector. Their mobility ranking remains conditioned on a particular initial distribution, however, and still applies only within the context of transition matrices.

the general ordering. Section 5 applies the theory to mobility data from the United States and Italy. All proofs are gathered in the appendix.

1 Preliminaries

An *income distribution* is identified by a cumulative distribution function (cdf) $F : \mathbb{R}_+ \to [0,1]$, with finite mean μ_F.[5] Since one often needs to restrict the set of feasible incomes in an economy (the available data may also impose such a restriction), we shall denote as $\mathcal{F}(X)$ the class of all income distributions whose support is contained in a given subset $X \subseteq \mathbb{R}_+$.

The generalized inverse of a distribution $F \in \mathcal{F}(X)$ is defined as

$$F^{-1}(p) \equiv \inf\{y \in X : F(y) \geq p\}, \quad 0 \leq p \leq 1,$$

which corresponds to the income of the person whose rank in the distribution is p. The *Lorenz curve* associated with F can then be defined as the graph of the function:

$$L_F(p) \equiv \frac{1}{\mu_F} \int_0^p F^{-1}(q)\, dq, \quad 0 \leq p \leq 1.$$

Thus $L_F(p)$ is the proportion of the total resources owned by the poorest $100\,p$ percent of individuals. An income distribution F Lorenz–dominates another distribution G when

$$L_F(p) \geq L_G(p) \quad \text{for all } p \in [0,1],$$

which we denote as $F \succsim_L G$. The dominance is strict, and denoted $F \succ_L G$, when $F \succsim_L G$ and $L_F(p) > L_G(p)$ for some $p \in [0,1]$. It is a well-established tradition in welfare economics (especially since Atkinson, (1970)) to regard one income distribution as unambiguously more equal than another, whenever the former strictly Lorenz–dominates the latter.[6]

Let $X \subseteq \mathbb{R}_+$ stand for the set of all feasible income levels. A *mobility process* on X is a function $M : \mathbb{R}_+ \times X \to [0,1]$ such that $M(\cdot \mid y) \in \mathcal{F}(X)$ for all $y \in X$. Thus $M(x \,|\, y)$ is the probability that an individual with income y today will earn at most x tomorrow. Empirical plausibility requires that future income prospects increase smoothly with the current level, in the sense of *first order stochastic dominance*. We thus restrict attention to processes which are continuous and strictly *monotone:* for any $y_1, y_2 \in X$ with $y_2 > y_1$,

$$M(x \mid y_1) \geq M(x \mid y_2) \ \text{ for all } x \in \mathbb{R}_+,$$

with strict inequality for some x. The set of all such mobility processes on X is denoted $\mathcal{M}(X)$.

An *economy* will be defined as a triplet (X, F, M) consisting of set of feasible income levels $X \subseteq \mathbb{R}_+$, an initial income distribution $F \in \mathcal{F}(X)$, and a mobility process $M \in \mathcal{M}(X)$.

[5] By definition, F is increasing and right-continuous, with $F(0) = 0$, $F(\infty) = 1$ and $\mu_F \equiv \int_0^\infty y\, dF > 0$. Its support will be denoted $supp(F) \equiv \{x \geq 0 : F(x+\varepsilon) - F(x-\varepsilon) > 0 \text{ for all } \varepsilon > 0\}$.

[6] For exhaustive reviews of the literature on the Lorenz ordering and the theory of inequality measurement at large, we refer the reader to Foster (1985) and Sen (1997).

2 Opportunity–Equalizing Processes

2.1 A Mobility Ordering According to Equality of Opportunity

The following two questions summarize the main inquiry of this paper:

[Q1] When would we say that mobility over time reduces inequality of opportunity, relative to initial endowments or social origins, in an economy (X, F, M)?

[Q2] When would we say that mobility is a more powerful equalizer of opportunities in an economy (X, F, M) than in another economy (X, G, N)?

The simplest framework where opportunities can be distinguished from initial conditions and ex–post outcomes is a two–period stochastic framework. For an agent with income $y \in X$, future opportunities are then fully described by the conditional distribution $M(\cdot|y)$. To simplify further, we shall initially summarize these uncertain prospects by their conditional expectation,

$$e_M(y) \equiv \int_0^\infty x \; dM(x \mid y). \tag{1}$$

By identifying opportunities with expected incomes in a two–period context, we are clearly abstracting from agents' aversion to risk and intertemporal fluctuations, as well as their rate of time preference.[7] But, as explained in Section 2.3, this basic framework is readily extended to the more general case by redefining "opportunities" as permanent incomes or intertemporal utilities.

The monotonicity and continuity of $M(\cdot \mid y)$ imply that $e_M : X \to \mathbb{R}_+$ is a strictly increasing and continuous function, and therefore invertible on its range $e_M(X)$. *The distribution of conditional expected incomes* (or opportunities) induced by (X, F, M) is then given by the cdf

$$\Lambda_{F,M}(x) = F(e_M^{-1}(x)) \qquad \text{for all } x \in e_M(X), \tag{2}$$

with support $e_M(X)$.

We now consider the first question posed above. To build up intuition, let us start with an extreme case where everyone has the same expected future income, regardless of their current situation. It is natural to regard this situation as involving a perfect equality of opportunities: the *realized* income distribution next period may not be perfectly equal, or even be vastly unequal, but those differences would be due only to *unpredictable* shocks, as opposed to the persistence of initial disparities. Generalizing this intuition, one might say that mobility in an economy (X, F, M) equalizes opportunities, relative to social origins, whenever

$$\Lambda_{F,M} \succsim_L F. \tag{3}$$

This is, however, an inequality which is *conditional* on the current income distribution, and therefore provides only a local evaluation of mobility. To see why this is problematic, suppose that for some initial income distribution F we have $\Lambda_{F,M} \succ_L F$, but for the income distribution G which will prevail (with probability one) in the next period as a result of the mobility process, we have

[7] These aspects are the main topic of Gottschalk and Spolaore (2000), who show how to disentangle their effects on intertemporal social welfare.

$G \succ_L \Lambda_{G,M}$. In other words, M tends to equalize opportunities (relative to initial conditions or social origins) in period 1, but to disequalize them in period 2. Similarly, it could be that $\Lambda_{F,M} \succ_L F$, but that the ranking is reversed when we restrict attention to some subgroup of the population (e.g., the middle class, women, etc.), even though the evolution of their incomes is governed by the same dynamic process M as all other agents (they just have initial incomes distributed according to a different F'). Both of these examples show that declaring (X, F, M) as an intertemporally egalitarian economy just because (3) holds would not really be justified.

There is a sense in which the notion of equality of opportunity is inherently dynamic, and hence all the relevant information is contained in the *law of motion* that dictates the evolution of incomes, as opposed to the initial conditions. To provide such a global answer to [Q1] we need to rank *mobility processes,* as opposed to economies. Consequently, we shall declare a process $M \in \mathcal{M}(X)$ to be *equalizing* (or *progressive*) when

$$\Lambda_{F,M} \succsim_L F \quad \text{for } all \ F \in \mathcal{F}(X). \tag{4}$$

In words, a mobility process is equalizing when it leads to ex-ante income prospects that are more evenly distributed than initial incomes or endowments, regardless of this initial distribution.

Let us now turn to the second question of interest. Once again, [Q2] can be approached either locally or globally. For the same reasons as before, we deem the global approach to be preferable. We therefore declare a mobility process M on X *more equalizing (*or *more progressive)* than another $N \in \mathcal{M}(X)$, and write $M \succcurlyeq_{\text{eq}} N$, when

$$\Lambda_{F,M} \succsim_L \Lambda_{F,N} \quad \text{for } all \ F \in \mathcal{F}(X). \tag{5}$$

The expected incomes of agents starting from different social positions are then more equally distributed under M than under N, no matter what the profile of initial inequality looks like. Suppose for instance that some country B, whose economic structure and institutions (education system, labor market regulations, etc.) result in a mobility process M^B, were to adopt those of another country A, resulting in the process $M^A \succcurlyeq_{\text{eq}} M^B$. Such a reform would reduce inequalities of opportunity in B's population, even if its initial income distribution was very different from that of A; for instance, each country could initially be in its own steady–state. Conversely, if A were to adopt B's mobility structure, its opportunities would become more unequal. It is in this sense that $\succcurlyeq_{\text{eq}}$ qualifies society A as unambiguously more mobile than society B. As a special case, a mobility process M is equalizing, in the sense of (4), if and only if $M \succcurlyeq_{\text{eq}} I$, where $I(x \mid y) \equiv 1_{\{x \geq y\}}$ for all $(x, y) \in \mathbb{R}_+ \times X$, corresponds to the preservation of the status quo.

Finally, one should note that $\succcurlyeq$ is a (partial) ordering defined over $\mathcal{M}(X)$, and therefore depends on the set of feasible incomes X. The importance of this point will become more apparent when we focus on discrete processes modeled in terms of transition matrices.

2.2 A Characterization Theorem

While the ordering $\succcurlyeq_{\text{eq}}$ appears useful with regard to the measurement of inequality of opportunity, its definition makes it hard to determine when two random processes can actually be ranked in this way. The following theorem and corollaries, which draw on the literature on inequality measurement and progressive taxation (e.g., Jakobsson (1976), Fellman (1976)) provides several operational characterizations of $\succcurlyeq_{\text{eq}}$.

Theorem 1 *Let $X \subseteq \mathbb{R}_+$ and $M, N \in \mathcal{M}(X)$. The following statements are equivalent:*

(i) $M \succcurlyeq_{eq} N$;

(ii) e_M/e_N is decreasing on $X \cap \mathbb{R}_{++}$;

(iii) There exists a strictly increasing mapping $\xi : e_N(X) \to \mathbb{R}_+$ such that the mapping $x \mapsto \xi(x)/x$ is decreasing on $e_N(X) \cap \mathbb{R}_{++}$, and $e_M = \xi \circ e_N$.

These results are quite intuitive. Condition (ii) states that expected incomes ("opportunities") are equalized at a faster rate under the mobility process M than under the process N. This is equivalent, as stated in condition (iii), to the fact that expected income under M can be obtained from expected income under N via a *progressive* redistribution scheme. Theorem 1 thus reflects the main idea of this paper: a mobility processes is nothing else than a *redistribution scheme* –albeit a stochastic one. It is therefore equalizing, in terms of expected incomes, to the extent that it is progressive, in the formal sense of having an increasing *average tax rate,* $t(y) \equiv (y - e_M(y))/y$. It is disequalizing to the extent that it is regressive. The following corollaries provide simple characterizations in terms of residual elasticities and expected growth rates.

Corollary 1. *Let M and N be two mobility processes on X such that e_M and e_N are differentiable on $X \cap \mathbb{R}_{++}$. Then:*

$$M \succcurlyeq_{\text{eq}} N \quad \text{if and only if} \quad \eta_M(y) \equiv \frac{y e_M'(y)}{e_M(y)} \leq \frac{y e_N'(y)}{e_N(y)} \equiv \eta_N(y), \text{ for all } y \in X \cap \mathbb{R}_{++}.$$

Corollary 2. *A mobility process $M \in \mathcal{M}(X)$ is equalizing, $M \succcurlyeq_{\text{eq}} I$, if and only if poorer agents have higher expected income growth than richer ones: the mapping $y \to e_M(y)/y$ is decreasing on $X \cap \mathbb{R}_{++}$. When e_M is differentiable, this means that $\eta_M \leq 1$ everywhere.*

We can also exploit the connection with Lorenz dominance to provide a welfaristic support for the mobility ordering $\succcurlyeq_{\text{eq}}$. Indeed, it is easy to see from (5) that if M is more progressive than N and leads to higher average income next period, it yields higher social welfare in terms of any utilitarian social welfare function defined over individual's expected incomes (opportunities).

Corollary 3. *If $M \succcurlyeq_{\text{eq}} N$, then for any $F \in \mathcal{F}(X)$ such that $\int_0^\infty e_M \, dF \geq \int_0^\infty e_N \, dF$,*

$$\int_0^\infty u \, d\Lambda_{F,M} \geq \int_0^\infty u \, d\Lambda_{F,N}$$

for all concave and continuous utility functions u defined on $\mathbb{R}_+$.

2.3 Remarks and Extensions

2.3.1 Perfect Immobility and Perfect Mobility

In evaluating mobility orderings, it is often considered intuitive that the identity process ($I(x \mid y) \equiv \mathbf{1}_{\{x \geq y\}}$) should correspond to the smallest element, and be viewed –implicitly or explicitly – as the worst case scenario (e.g. Shorrocks (1978a), Dardanoni (1993)). More generally, according to what Kanbur and Stiglitz (1986) term "the diagonals view," *any* increase in relative income movement (in a transition matrix context, shifting probability weight from diagonal to off-diagonal elements) should imply a higher ranking in the mobility ordering.

Our ordering is quite different in that respect, because it recognizes that relative income movements can be disequalizing as well as equalizing, and it is only the latter type that will count positively as "mobility." Thus, in general, there does not exist a *smallest* element in $\mathcal{M}(X)$ with respect to $\succcurlyeq_{\text{eq}}$, just as there generally does not exist a most regressive tax scheme. In particular, the identity process (no movement) does *not* have his property. The next section provides a discrete example, but this most clearly seen for $X = \mathbb{R}_+$. Let M be such that future income is any convex function of current income, plus mean–zero noise (but ensuring that realizations remain positive with probability one). Clearly, the distribution of $e_M(y)$ is more unequal than that of y, under any initial conditions.

When the income support is bounded above and below, on the other hand, one can show that I is a *minimal* element in $\mathcal{M}(X)$ with respect to $\succcurlyeq_{\text{eq}}$, i.e. there is no $M \in \mathcal{M}(X)$ such that $I \succ_{\text{eq}} M$. Indeed, by Theorem 1 this would imply $e_M(y_2)/y_2 \geq e_M(y_1)/y_1$ for all pairs $y_1, y_2 \in X$ with $y_2 \geq y_2 > 0$. But we must also have $e_M(\min X) \geq \min X$ and $e_M(\max X) \leq \max X$. Both conditions are compatible only if $e_M(y)/y = 1$ for all y, which contradicts $I \succ_{\text{eq}} M$. Note that this minimal property of I is entirely due to the assumption of fixed upper and lower bounds on feasible incomes, which seems rather arbitrary. Moreover, I is still not a smallest element of $\preccurlyeq_{\text{eq}}$, since one can again find mobility processes which are more regressive than immobility on some subset of X. The next section will provide specific examples in the context of transition matrices.

Is there a *greatest* element in $\mathcal{M}(X)$ with respect to our mobility ordering? Because of the requirement of strict monotonicity, there does not exist one either but –in contrast to the absence of a smallest element– this is only a minor technical wrinkle. If we extend the ordering $\succcurlyeq_{\text{eq}}$ to the set $\mathcal{M}_*(X)$ which includes processes with $M(x \mid \cdot)$ that are only weakly increasing, we immediately observe that any process M such that $M(\cdot \mid y)$ is independent of y is a greatest element in $\mathcal{M}_*(X)$ with respect to $\succcurlyeq_{\text{eq}}$. This is because under such a process, all agents have the same conditional distribution of future incomes, and in particular the same expected income.

2.3.2 Permanent Income and Intertemporal Utility

When introducing our ordering as one based on the equalization of opportunities, we defined the latter as conditional expected incomes. In terms of welfare, this corresponds to agents who live for two periods, care only about the second one (e.g., their children's expected lifetime income), and are risk–neutral. It is, however, straightforward to extend the analysis to a multiperiod setting (say, infinite horizon) where agents care about permanent incomes, or more generally about some additively separable intertemporal utility. Given the one–period transition M, let $M^{(t)} \in \mathcal{M}(X)$ denote the t–period–ahead mobility process, which is defined recursively by $M^{(1)} \equiv M$ and

$$M^{(t+1)}(x|y) \equiv \int_X M(x|z) dM^{(t)}(z|y) \tag{6}$$

for all $t \in \mathbb{N}$. When agents are risk–neutral and have discount factor ρ, we need only replace $e_M(y)$ in Theorem 1 and its corollaries by

$$e_M(y;\rho) \equiv (1-\rho)\sum_{t=0}^{\infty} \rho^t e_{M^{(t)}}(y), \tag{7}$$

provided course the series converges. All the results hold unchanged, for what is now a mobility ranking relating to the reduction of *lifetime inequality.* Similarly, when agents' instantaneous utility function is $u(\cdot)$, $e_M(y)$ is replaced by

$$e_M(y; u, \rho) \equiv (1-\rho)\sum_{t=0}^{\infty} \rho^t \int_X u(x) dM^{(t)}(x|y), \tag{8}$$

yielding parallel results for a mobility ordering based on the equalization of the *lifetime utilities.*[8] Of course, computing $e(y;\rho)$ or $e_M(y;u,\rho)$ from M may be easy or difficult. In the case of discrete mobility processes represented by transition matrices (see Section 3), it will be extremely simple.

2.3.3 Strongly Equalizing Mobility Processes

In the previous subsection we showed how to rank mobility processes according to equalization of permanent incomes, for a given discount factor ρ. It may also be of interest to know whether a mobility process equalize opportunities, relative to initial conditions, *over any horizon t* (say, for grandchildren as well as children), or *for any discount factor ρ*. Formally, the question is whether M being equalizing guarantees that $M^{(t)}$ is equalizing for all $t \in \mathbb{N}$. Unfortunately, the answer is negative, as one can show by means of simple examples. We can, however, identify a stronger condition on M which ensures that this appealing "horizon–independence" property holds.

Let $M \in \mathcal{M}(X)$, and define

$$E_M(y,\theta) \equiv \int_0^{\theta} x \; dM(x \mid y), \quad y \in X, \; \theta \in (0,\infty].$$

Clearly, E_M is a nonnegative-valued mapping on $X \times (0,\infty]$. We say that M is *strongly equalizing* if, *for each* $\theta \in (0,\infty]$, the mapping

$$y \mapsto \frac{1}{y} E_M(y,\theta) \text{ is decreasing in } y \text{ on } X \cap \mathbb{R}_{++}.$$

[8] An informal but perceptive antecedent to this idea can be found in Loury (1981), who wrote (about his model): *"the graph of the ... indirect utility function will be "flatter" when society is more mobile.... This makes the cross–sectional distribution of welfare... less unequal than would be the case with little or no mobility".* Our results formalize and generalize this intuition, showing in particular that what matters is the elasticity, rather than the slope, of the mapping between initial conditions and lifetime utilities.

Note that we have $E_M(y,\infty) = e_M(y)$ for all y, so by Theorem 1 a strongly equalizing process is equalizing. More interesting is the following key property.

Proposition 1 *Let X be a closed subset of $\mathbb{R}_+$. If $M \in \mathcal{M}(X)$ is strongly equalizing, so is $M^{(t)}$ for any $t \in \mathbb{N}$.*

If mobility in an economy is governed by the same strongly equalizing process over t-periods, then expected incomes are equalized, relative to initial conditions, over all relevant horizons. Clearly, expected present values of incomes then have the same property, for any discount factor.[9] Unfortunately, it does not seem possible to extend this result so as to define a "more strongly equalizing" ordering between arbitrary processes M and N which would be similarly preserved through iteration.

3 Discrete Markov Processes

3.1 Equalizing Transition Matrices: Properties and Characterization

In empirical applications, it is common to focus on discrete income distributions and represent mobility by transition matrices between n income states. In this section we elaborate on the properties of our mobility ordering $\succcurlyeq_{\text{eq}}$ in this case.

A transition matrix $P \equiv [p_{ij}]$ is any $n \times n$ stochastic matrix (i.e., $p_{ij} \geq 0$ and $\sum_j p_{ij} = 1$ for all i and j). To interpret it as modeling the evolution of individuals' *incomes*, we attach an income level y_i to income state i. We assume that $0 < y_1 < \cdots < y_n$ and denote the *income state vector* as $\mathbf{y} \equiv (y_1, ..., y_n)'$, where a prime signifies transposition. An income distribution in this setting corresponds to a probability row vector $\boldsymbol{\pi} = (\pi_1, ..., \pi_n)$, where π_i is interpreted as the proportion of individuals who are in income state i. Or, consistently with our general definition, one may think of an income distribution as any cdf with support contained in $X = \{y_1, ..., y_n\}$.

A transition matrix P is said to be *monotone* if an individual in income state $i+1$ faces a better lottery over her future income than an individual in income state i :

$$\sum_{j=1}^{k} p_{i+1,j} \geq \sum_{j=1}^{k} p_{ij} \quad \text{for all } i,\, k \text{ in } \{1, ..., n-1\}, \tag{9}$$

with strict inequality for some k.[10] A transition matrix P induces a mobility process M_P on $\{y_1, ..., y_n\}$ in a natural way:[11]

$$M_P(y_k \mid y_i) \equiv \sum_{j=1}^{k} p_{ij}, \quad k = 1, ..., n-1. \tag{10}$$

[9] As an example, the familiar loglinear–lognormal process, where $\ln y_{t+1} = \alpha \ln y_t + \varepsilon_{t+1}$, and the ε_t's are i.i.d. and normal, is strongly equalizing for all $\alpha < 1$.

[10] Monotone transition matrices were introduced to the mathematical literature by Keilson and Ketser (1977), and are now widely used in modeling income mobility; see Conlisk (1989), (1990) and Dardanoni (1993), (1995).

[11] $M_P(\cdot \mid y_i)$ is easily extended to a cdf on $\mathbb{R}_+$ by making it a step function that is constant on any $[y_j, y_{j+1})$.

Consequently, monotone transition matrices will be ranked by ordering with respect to $\succcurlyeq_{\text{eq}}$ the mobility processes induced over the income states in $\mathbf{y}$. We shall thus write $P \succcurlyeq_{\text{eq}}^{\mathbf{y}} Q$, and say that P is more *equalizing* (or *progressive*) than Q over the state space X, whenever $M_P \succcurlyeq_{\text{eq}} M_Q$.[12] In words, $P \succcurlyeq_{\text{eq}}^{\mathbf{y}} Q$ means that the distribution of conditional expected incomes (opportunities) induced by P, which for a fraction π_i of the population are equal to

$$e_P(y_i) \equiv \sum_{j=1}^{n} p_{ij}\, y_j, \tag{11}$$

is more equal than that similarly induced by Q, *for all* initial income distributions $\boldsymbol{\pi}$ defined on $\{y_1, ..., y_n\}$. Similarly, given an income state vector $\mathbf{y}$, we say that P is equalizing (or progressive) if $P \succcurlyeq_{\text{eq}}^{\mathbf{y}} I$, where I is the $n \times n$ identity matrix.

The following theorem provides easy-to-apply methods for checking whether or not two transition matrices can be ranked on the basis of $\succcurlyeq_{\text{eq}}^{\mathbf{y}}$, given an income state vector $\mathbf{y}$. In its statement, we denote by $D^*[A]$ the first superdiagonal of any square matrix A, and by $D_*[A]$ its first subdiagonal.

Theorem 2 *Let $\mathbf{y} \in \mathbb{R}^n_{++}$ (with $y_1 < \cdots < y_n$) and let P and Q be two $n \times n$ monotone transition matrices. The following statements are equivalent.*

(i) $P \succcurlyeq_{eq}^{\mathbf{y}} Q$;

(ii) $\dfrac{e_P(y_1)}{e_Q(y_1)} \geq \cdots \geq \dfrac{e_P(y_n)}{e_Q(y_n)}$;

(iii) $(D^* - D_*)[P\mathbf{y}\mathbf{y}'Q'] \geq 0$.

The interpretation of conditions (i) and (ii) in terms of progressivity is similar to that of Theorem 1. Condition (iii) is the new, "operational" one, being immediate to check.[13] Turning next to social welfare, we note again that, provided P does not yield a lower mean income than Q, the statement $P \succcurlyeq_{\text{eq}}^{\mathbf{y}} Q$ has a fairly strong utilitarian implications.

Corollary 1 *Let $\mathbf{y} \in \mathbb{R}^n_{++}$ (with $0 < y_1 < \cdots < y_n$) and let P and Q be two $n \times n$ monotone transition matrices. If $P \succcurlyeq_{eq}^{\mathbf{y}} Q$, then for all probability vectors $\boldsymbol{\pi}$ such that $\boldsymbol{\pi}P\mathbf{y} \geq \boldsymbol{\pi}Q\mathbf{y}$, we have*

$$\sum_{i=1}^{n} \pi_i u(e_P(y_i)) \geq \sum_{i=1}^{n} \pi_i u(e_Q(y_i))$$

for all concave and increasing utility functions u defined on $\mathbb{R}_+$.

Finally, the defining properties of an equalizing (progressive) transition matrix P naturally correspond to the particular case where $Q = I$ in the above results.

[12] The reason why we make explicit in our notation the dependence of this ordering on the income state space $\mathbf{y}$ will be discussed below.

[13] Note that, for any square matrix A, $(D^* - D_*)[A] = D^*[A - A']$.

3.2 Permanent Income and Intertemporal Utility

We now extend the analysis to multiperiod settings where agents care about their permanent income (lifetime or dynastic), or more generally about some intertemporal utility, as in (8). Following Dardanoni (1993), let us denote $P(\rho) \equiv (1-\rho)(I-\rho P)^{-1}$, for any transition matrix P and discount factor $\rho \in (0,1)$; given that P is monotone, so is $P(\rho)$. Next, for any increasing utility function u on $\mathbb{R}_+$, let (by a slight abuse of notation), $u(\mathbf{y}) \equiv (u(y_1), ...u(y_n))'$ be the utility state vector. Since $P(\rho)u(\mathbf{y})$ is the vector of conditional lifetime utilities, we have by Theorem 2:

Proposition 2 *Let $\mathbf{y} \in \mathbb{R}^n_{++}$ (with $y_1 < \cdots < y_n$) and let P and Q be any two $n \times n$ monotone transition matrices. The following statements are equivalent:*

(i) $P(\rho) \succcurlyeq^{u(\mathbf{y})}_{eq} Q(\rho)$: starting from any initial distribution $\boldsymbol{\pi}$, the Lorenz curve for agents' intertemporal utilities which obtains under the mobility process P is everywhere below that which obtains under Q.

(ii) $(D^* - D_*)[P(\rho)u(\mathbf{y})u(\mathbf{y})'Q(\rho)'] \geq 0$.

Note that this ranking (and the associated test) are conditional on the value chosen for the discount factor ρ –as in Kanbur and Stiglitz (1986), and Formby, Smith and Zheng (1995). By contrast, Dardanoni (1993) provides sufficient conditions which ensure that, if P and Q can be ranked according to his ordering (to start with, they must have the same steady–state), then $P(\rho)$ and $Q(\rho)$ have the same ranking for all values of ρ. It does not seem that any ρ–independent criterion can be provided for our ordering, except in special cases.

One such case which is of interest relates to the question of whether a mobility process is equalizing, relative to initial conditions, not just for next period's expected incomes, but also for lifetime incomes or welfare levels. For this, we shall make use again of strongly equalizing mobility processes, for which progressivity is actually a ρ–independent property. Consistent with the general definition in Section 2.3.3, a monotone matrix P is said to be *strongly equalizing* over the income space $\mathbf{y}$ if, for each $i = 1, ..., n-1$,

$$\frac{1}{y_i}\sum_{j=1}^{k} p_{ij}\, y_j \geq \frac{1}{y_{i+1}}\sum_{j=1}^{k} p_{i+1,j}\, y_j, \quad k = 1, ..., n.$$

Proposition 3 *Let P be an $n \times n$ monotone transition matrix. Given a utility state vector $u(\mathbf{y}) \in \mathbb{R}^n$ (with $u(y_1) < \cdots < u(y_n)$), if P is strongly equalizing over $u(\mathbf{y})$, then so is P^t for all t. Furthermore, $P(\rho)$ is then equalizing for all $\rho \in (0,1)$.*

3.3 Discussion and Relation to Other Orderings

Remark 1. A discrete mobility process was defined by a transition matrix P *and* an income state vector $\mathbf{y}$ over which it operates. In contrast to the view implicitly taken in much of the mobility literature, *income* mobility cannot, we believe, be adequately studied or even defined independently of the values taken by income, and this for several reasons.

The first issue goes back to the empirical meaning of a transition matrix. When we represent a country's mobility process by a transition matrix P, these coefficients correspond to the frequencies of transitions which were observed to occur between certain well-defined income levels –more precisely, income intervals– such as (say): 0–25,000, 25,000–50,000, 50,000–75,000, 75,000–100,000, 100,000–∞. *Nothing allows us to pretend* that the same transition probabilities obtain between any arbitrary five income levels or intervals, such as 0-1,000, 1,000–60,000, 60,000–61,000, 61,000-1,000,000, 1,000,000–∞. In fact, the data would surely contradict this notion.[14]

One may hope to avoid this problem by using interfractile transition matrices, where the ith state always corresponds to the ith fractile of the income distribution. This route leads to another difficulty, however, because it does not take into account the magnitude of the (absolute or relative) income changes associated to movements between fractiles. Clearly, the same transition probabilities P between some income states (y_1, y_2, y_3) on the one hand, and between $(y_1, y_1 + \varepsilon, y_2 \times 10^6)$ on the other (where ε is small), represent very different mobility processes in any economically meaningful sense of the term. Both the prospects faced by individuals (expected incomes, risks) and the implied magnitude and persistence of inequality are radically different. This problem arises very concretely with intefractile matrices, due in particular to the skewedness of empirical income distributions. For instance, in the 1979 US income distribution (estimated from PSID data, see Hungerford (1993)), a move from the 1^{st} to the 2^{nd} decile would correspond on average to a near doubling of family income, while rising from the 5^{th} to the 6^{th} or the 6^{th} to the 7^{th} would bring an average gain of only about 15%. The increase would become again much more significant for a move from the 9^{th} to the 10^{th}, which raises average income by 55%. What is more, such ratios typically differ from one country to another, thereby rendering intercountry mobility comparisons based solely on interfractile transition matrices conceptually problematic.

To summarize, an ordering or index purely based on changes in ranks (independently of their income implications), or more generally on the properties of P alone, as most of those found in the literature are, can in general not fully account for economic mobility, whether in the sense of equalization of opportunities or in terms of intertemporal welfare consequences. This leads us to the view that the measurement of income mobility should be based on $(P, \mathbf{y})$, as opposed to P alone.

Remark 2. It is interesting to note the kind of "duality" which exist between our ordering and those of Dardanoni (1993), or Formby, Smith and Zheng (1995). In both cases one compares the distributions of expected future incomes (or their present values) under P and Q. In these two papers this comparison is conditional on a particular value of $\boldsymbol{\pi}$ (common steady–state vector in the first case; arbitrary fixed $\boldsymbol{\pi}$ in the second), but required to hold for all $\mathbf{y}$. In our ordering it is conditional on $\mathbf{y}$, but required to hold for all $\boldsymbol{\pi}$.

Could one insist on an ordering which was independent of the income state vector $\mathbf{y}$, as well as of the initial distribution $\boldsymbol{\pi}$? A simple example will make clear why no such "global–global"

[14] This is of course a problem with any analysis based on transition matrices. The fundamental difficulty is that a discrete transition process provides only a very partial representation of the actual law of motion for incomes, which in reality operates on a large subset of $\mathbb{R}_+$.

ordering exists, as soon as $n \geq 3$. Suppose that one could find some 3×3 monotone transition matrix $P = [p_{ij}]$ which implied a progressive or equalizing mobility process over *all* possible income supports, i.e. $P \succ_{\text{eq}}^{\mathbf{y}} I$, for all $y_1 < y_2 < y_3$. This would mean that

$$\frac{p_{11}y_1 + p_{12}y_2 + p_{13}y_3}{y_1} > \frac{p_{21}y_1 + p_{22}y_2 + p_{23}y_3}{y_2} > \frac{p_{31}y_1 + p_{32}y_2 + p_{33}y_3}{y_3},$$

For the first inequality to remain valid as y_2 tends to y_1 it must be that $p_{23} \leq p_{13}$, which can be rewritten as $p_{21} + p_{22} \geq p_{11} + p_{12}$. Similarly, by letting y_2 tend to y_3 in the second inequality we get $p_{21} \leq p_{31}$. But both of these conditions imply that future income is (weakly) stochastically *de*creasing in current income. The intuition is simple: as y_2 and y_1 become very close, for instance, transition probabilities between these two states become almost irrelevant. Thus, starting from a current income of y_1 versus y_2, having higher expected income growth $e_M(y)/y$ becomes equivalent to having a higher probability of rising to y_3. Progressivity is thus incompatible with strict monotonicity, and consistent with weak monotonicity only when $p_{13} = p_{23}$ and $p_{21} = p_{31}$. But then, when y_3 becomes very large compared to y_2 the second inequality clearly becomes violated, even in its weak form. In summary, for $n \geq 3$ there is no transition matrix which is more equalizing (progressive) than the identity over all income supports.

Remark 3. Because it makes mobility a "signed" concept, the ordering we define on monotone matrices also differs from previous ones in the sense that the identity mapping is not the smallest element –because it is not the worst one from the point of view of inequality of opportunities. This was shown earlier with continuous processes defined on all of $\mathbb{R}_+$ but, as a simpler example, consider the transitions

$$J \equiv \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

defined over $y_1 < y_2 < y_3$. In this scenario, which could be made stochastic, the middle class "falls through the floor" and joins the ranks of the poor –an obvious oversimplification of a real policy concern. Clearly, for any initial distribution such that $\pi_1 = 0$ (this could describe the whole population, or only some subgroup), the distribution of conditional expected incomes associated to J is more unequal than the one associated to I. (Moreover, the restriction $\pi_1 = 0$ is entirely due to the fact that a fixed lower bound $y_1 > 0$ is imposed on all incomes). The converse holds when $\pi_3 = 0$, so neither $I \succ_{eq} J$ nor $J \succ_{eq} I$ holds.[15]

That "immobility" is not the worst–case scenario is a natural property of *any* mobility criterion which cares about relative income movement not as an end in itself, but because of its impact on individual and aggregate welfare. Dardanoni (1993) also argued that mobility processes should be evaluated on the basis of their (ex–ante) equalizing properties, but restricted his ordering to mobility matrices which share the same steady–state, and his Lorenz comparisons to initial situations where

[15] This example shows that $\succcurlyeq_{\text{eq}}^{\mathbf{y}}$ satisfies neither the *monotonicity* assumption of Shorrocks (1978b) nor the axiom of *diagonalizing switches* of Atkinson (1983). Of course, this is not surprising since these properties really concern the measurement of relative income movement (progressive or regressive), as opposed to equality of opportunity.

the economy finds itself in this steady–state. In terms of our example, this means comparing I and J's welfare implications only when $\pi_2 = 0$, i.e. when no one is located in the range where the "tax scheme" imposed by J is regressive. In that case, and only then, $\Lambda_{\pi,J,} = \Lambda_{\pi,I}$. It is therefore the local nature of the inequality comparisons underlying Dardanoni's ordering (justified by a focus on "exchange mobility" within the invariant distribution) which prevents it from registering the fact that some forms of mobility may aggravate existing inequalities, hence result in a Lorenz curve for expected or permanent incomes which is strictly outside that which prevails under "immobility".[16]

A related point can be made in relation to non–welfare based approaches, which make no distinction between equalizing and disequalizing income movements. We shall illustrate it in relation to Shorrocks' (1978a) immobility criterion, but it applies equally to the "diagonals view" of mobility matrices (e.g., Atkinson (1983)), including eigenvalue–based criteria. For simplicity, let there be two periods, and let inequality be measured as the variance of relative incomes. Shorrock's index of immobility is then the ratio of the variance in two–period (undiscounted) relative incomes to the sum of the one-period variances. This ratio, call it R, is shown to be always less than one, and equal to one only in the case of perfect immobility. In discussing its economic interpretation, Shorrocks states that: *"mobility is regarded as the degree to which equalization occurs as the period is extended. This view captures the prime importance of mobility for economists".* We agree with these statements (taken in an ex–ante sense), but not with the claim that R necessarily captures this notion of mobility as an equalizing force: what R records is just relative movement, *whether equalizing or disequalizing.*[17] Conversely, if mobility is defined according to R, it is not the case that "mobility is unambiguously good," as stated by Shorrocks, even staying within the context of risk–neutral agents.

Consider, for instance, three individuals with period–zero incomes 2, 3 and 4, and period one incomes $2+\varepsilon$, 3 and $4-\varepsilon$, respectively. The two-period average income vector is thus $(4+\varepsilon, 6, 8-\varepsilon)$. Since this mobility process, call it $M(\varepsilon)$, simply transfers ε dollars from rich to poor, it is regressive for $\varepsilon < 0$, and progressive for $\varepsilon > 0$. Yet we know that for *any* value of ε, $M(\varepsilon)$ ranks lower according to Shorrock's (1978a) criterion than $M(0) = I$, even though when $\varepsilon < 0$ all that "mobility" does is to aggravate existing inequalities. Moreover, it is easily observed that: $R(2/3) = R(-2) = .80$, even though the first process increases the poor's income by a third at the expense of the rich, while the latter takes away all the poor's income to give it to the rich! Similarly, $R(.30) > R(-.5) > R(-1) > R(-2)$: as we move from the first process, which reduces initial disparities, to the next three, which increasingly accentuate them, the index records rising mobility.[18] **Figure 1**, which

[16] Formby, Smith and Zheng (1995) relax Dardanoni's (1993) joint–steady state requirement, and compare mobility processes starting from any *given* distribution π. This is still a local comparison, although no longer restricted to a common steady-state. They do not examine the issue of a minimal or smallest element.

[17] Indeed, Shorrocks proves that $R \leq 1$ but not that inequality in m–period incomes must necessarily decline with m (as pictured on his graph), nor that such declines are directly related to decreases in R. The example he provides does have these properties, however, as the process is not only progressive but even non-monotonic (incomes have a serial correlation of -1).

[18] Note that $M(\varepsilon)$ is deterministic, for simplicity. One could obviously introduce some uncertainty without changing any of what follows, or interpret the table's entries as conditional expected incomes rather than probability–one

plots $R(\varepsilon)$, makes clear more generally that the kind of mobility measured by Shorrock's index need not always be of the type such that *"equalization is more pronounced in a very mobile society".* By contrast, our ordering clearly conclude that $M(\varepsilon) \succ_{\text{eq}} M(\delta)$ whenever $\varepsilon > \delta$.[19]

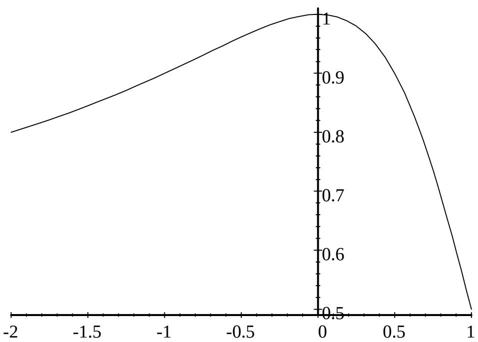

Figure 1: Shorrocks's immobility criterion R, versus the degree of equalization ε.

4 Summary Indices of Equalizing Mobility

Since our mobility ordering is the direct translation of inequality and progressivity measurement to a dynamic context, the same practical issues arise as in those literatures. When $P \succ_{\text{eq}} Q$ the first process is unambiguously better from the point of view of equalizing opportunities, but one would still like to quantify this difference. Even more importantly, when P and Q are not rankable according to $\succ_{\text{eq}}$, one would still like to compare them according to some unidimensional criterion, consistent with this ordering.

Fortunately, we do not need to devise and defend a new mobility index. According to our view, *mobility is progressivity,* so one should simply use existing and familiar measures of the latter –more specifically, of residual progressivity. Thus, given an economy (X, FM), one can simply:

(1) Compare inequality of initial incomes and inequality of conditional expected incomes (more generally, permanent incomes or intertemporal utilities), for example by taking the difference in the corresponding Gini coefficients:

$$\rho_M^{RS} \equiv Gini(F) - Gini(F \circ e_M^{-1}). \tag{12}$$

This gap, equal to the area between the two Lorenz curves, increases as M rises in the mobility ordering $\succcurlyeq_{\text{eq}}$. And indeed, ρ_M^{RS} is nothing but the familiar Reynolds–Smolensky (1977) *index of residual progressivity*, applied to the "redistributive scheme" e_M.[20]

realizations. In any case, we impose $\varepsilon \leq 1$ to maintain monotonicity.

[19] Of course, a mobility process which is globally regressive is very unrealistic. But, more generally, there might be equalization of opportunities over a certain range and disequalization over another, so it is important for a mobility index to be sensitive to the difference between the two. This is especially true when comparing a process M not to the identity but to some other process N, since what matters then is *relative progressivity,* which could well be decreasing even when both processes are progressive.

[20] For an extensive discussion of progressivity indices, see Lambert (1993).

(2) Alternatively, compute the average *residual elasticity* if the process is differentiable,

$$\bar{\eta} \equiv \int_X \frac{y e'_M(y)}{e_M(y)}\, dF(y), \tag{13}$$

or a discrete analogue (also consistent with $\succcurlyeq_{\text{eq}}$), if it is not:

$$\bar{\eta} \equiv \sum_{i=1}^{n-1} \tilde{\pi}_i \left(\frac{\ln(e_M(y_{i+1})/e_M(y_i))}{\ln(y_{i+1}/y_i)} \right), \tag{14}$$

where the weights are for instance set to $\tilde{\pi}_i \equiv \pi_i \Big/ \sum_{i=1}^{n-1} \pi_j$, $i = 1, ...n-1$.[21,22]

Given the parallel with taxation, it may also be interesting to characterize the mobility process in terms of implicit *average and marginal tax rates.* This metric is both familiar and intuitive, and readily allows comparisons between the reshuffling of incomes due to the workings of the economy and those due to public policy. Normalizing both the initial and the expected distribution by their means, the expected average tax rate at any income level $y \in X$ is thus $t(y) \equiv 1 - (e_M(y)/y) \,/ \left(\mu_{\Lambda_{F,M}}/\mu_F\right)$; that is, it equals *one minus the expected growth rate* of relative income. The average marginal tax rate over the population is then $\bar{\tau} \equiv \int_X t'(y)\, dF(y)$ in the differentiable case.[23] The drawback is that this aggregate index is not always consistent with true progressivity.

5 Empirical Applications

5.1 Medium Term Earnings and Income Mobility in the US

We shall first illustrate our general methodology using PSID data from two sources, which correspond to different horizons and economic units. The first one is Gottschalk (1997), who provides the interquintile transition matrix M_{74}^{91} for individual male labor earnings over the 17 year period between 1974 and 1991. To (re)construct the income state vector, $\mathbf{y}$, we assign to each quintile its mean income level, as observed either in 1974 ($\mathbf{y}_{74}$) or in 1991 ($\mathbf{y}_{91}$). The distribution over $\mathbf{y}$ is $\boldsymbol{\pi} = (.2, .2, .2, .2, .2)$.

Table 1 here

[21] Alternatively, $\tilde{\pi}_i \equiv \pi_{i+1} \big/ \sum_{i=1}^{n-1} \pi_{j+1}$, or some combination of the two values. In our empirical applications we use the midpoint, but this choice makes almost no difference. Finally, one can similarly define the *income–weighted* elasticity, by replacing $dF(y)$ with $(y/\mu_F)dF(y)$ in (13), or the π_j's by $\pi_j\, y_j$'s in the definition of the weights $\tilde{\pi}_i$ entering (14).

[22] In the discrete case the $e_M(y_i)$'s are the components of Py, making $\bar{\eta}$ easy to compute. In the continuous case it is clear that when the mobility process is loglinear, $\ln x = \alpha + \beta \ln y + \varepsilon$, $\bar{\eta}$ is just equal to β, so the standard regression on individual data estimates the "right" measure of mobility.

[23] In the discrete case, the implicit marginal tax rate between y_i and y_{i+1} is given by

$$1 - \tau_i \equiv \left(\frac{e_M(y_{i+1}) - e_M(y_i)}{y_{i+1} - y_i} \right) \Big/ \left(\frac{\Sigma_{i=1}^n \pi_j e_M(y_j)}{\Sigma_{i=1}^n \pi_j y_j} \right),$$

for all $i = 1, ...n-1$. The average marginal tax rate is then $\bar{\tau} \equiv \sum_{i=1}^{n-1} \tilde{\pi}_i \tau_i$,with the weights $\tilde{\pi}_i$'s defined as before, or as in footnote 21.

As shown in **Table 1,** mobility prospects over the 17–year period reduce the Gini coefficient from .415 for initial incomes in 1974 to .226 or .255 for 1991 expected incomes, depending on whether it is assumed that the general rise in inequality which occurred in the late 70's and 80's was initially unexpected (column 1), or fully anticipated (column 2).[24] The Reynolds–Smolensky (1977) index of progressivity is thus . 189 or .160, respectively. This represents a substantial degree of equalization: by comparison, the same index for the US income tax system was only .031 in 1979 and .025 in 1988 (see Bishop, Chow and Formby (1997)).[25] In fact, the mobility process was globally *progressive* ($M_{74}^{91} \succ_{\text{eq}} I$), with average tax rates $t_i \equiv 1 - e_M(y_i)/y_i$ equal to $-282\%, -57.9\%, -16.9\%, +10.0\%$ and $+41.4\%$ (column 1). This means that the average person in the bottom quintile in 1974 could expect their earnings to grow 3.82 times as fast as the population mean over the following 17 years. Conversely, someone starting in the top quintile had expected relative losses of 10%. Note also that while inequality of outcomes had risen to .466 by 1991, at most $.255/.466 = .55\%$ of this total reflected ex–ante unequal opportunities as of 1974. In other words, only 55% could have been predicted on the basis of initial earnings disparities.[26] Finally, column (3) describes mobility prospects starting in 1991, assuming that the transition matrix and quintile shares remain unchanged. The numbers obtained are comparable to those of the other columns.

Our second source is Hungerford (1993), who describes shorter transitions over the 7 year intervals $1969 - 1976$ and $1979 - 1986$, at the decile level. This allows for a finer description of states and transitions, but at the cost of higher sampling errors. There are also differences in the nature of the data; for instance Gottschalk's transitions relate to individual male earnings, whereas Hungerford's is for family incomes.[27] The income state vector is constructed as explained above, either from $\mathbf{y}_{69}$ or from $\mathbf{y}_{76}$ (used in place of $\mathbf{y}_{79}$, which is not reported by Hungerford); $\boldsymbol{\pi}$ is now ten–dimensional, with all entries equal to .1.

Table 2 here

The results are presented in **Table 2.** Mobility reduces the Gini coefficient from . 363 for 1969 initial conditions to between .217 and .235 for 1976 income prospects, depending on the terminal income state vector which is used (columns 1 and 2, respectively). The Reynolds–Smolensky index of progressivity is thus .146 or .128, which is still large compared to fiscal redistributions, but clearly less than in the earlier data. As intuition suggests, longer horizons allow more equalization

[24]Formally, the difference is whether conditional probabilities over quintiles in 1991 (computed using M_{74}^{91}) are translated into expected incomes using the initial distribution of relative incomes $(\boldsymbol{\pi}, \mathbf{y}_{74})$, or the ex–post realized one, $(\boldsymbol{\pi}, \mathbf{y}_{91})$. The fist assumption seems more realistic, since no one in the early 70's foretold the tide of rising inequality. Clearly, it makes little difference whether column 1 or column 2 is used.

[25]These numbers reflect only the progressivity of income taxes. If the incidence of public transfers and in–kind benefits was taken into account, the overall degree of fiscal progressivity would undoubtedly be higher.

[26]The Lorenz curve for 1991 income realizations is everywhere below the curve for 1994 incomes, which itself is below the curve for expected 1991 incomes (conditional on 1974 levels). This makes clear the fact that mobility equalizes (ex–ante) *opportunities,* but not (ex–post) *outcomes.*

[27]We refer the reader to these two sources for a full description of their data.

of opportunities –provided the mobility process is indeed progressive (rather than regressive, as in the last example of Section 3.3).

The numbers for the 1979–1986 period, given in the bottom panel of the table, are virtually identical to those of the earlier one.[28] This is consistent with Hungerford's conclusion, based on traditional tests, that there was no measurable change in mobility between the two periods. Indeed, the *"superdiagonals test"* of Proposition 3 shows that neither M_{69}^{76} nor M_{89}^{86} dominates the other in the sense of $\succcurlyeq_{\text{eq}}$. On the other hand, progressivity is satisfied by both matrices at almost every decile: the sequence $\mathbf{t} \equiv (t_i)_{i=1}^{10}$ of average tax rates for the ten deciles is decreasing, except at one or two points showing a slight increase, probably due to measurement error.[29] For instance, the 1979–1986 transition (as measured in column 3) yields

$$\mathbf{t} = (-136.4\%, -61.2\%, -41.0\%, -15.7\%, -16.8\%, -5.8\%, -4.2\%, -7.9\%, 15.0\%, 33.9\%).$$

Note that all but those families who start in the top two deciles have expected gains in their relative incomes.

5.2 Intergenerational Mobility: the U.S. versus Italy

Finally, we turn to intergenerational mobility, which is probably where equality of opportunity matters most. For this purpose we use the data of Rustichini, Ichino and Checchi (1999) for the United States and Italy. This consists of father–to–son transition probabilities between four "occupational income" classes, whose boundaries correspond to equiproportional increases in income.[30] This presentation of the data is particularly well suited to our purpose, as the transition matrices M_{US} and M_{IT} operate on (nearly) the same income state vector $\mathbf{y}$, up to a constant of proportionality (see the discussion in Section 3.3). By contrast, the distribution of incomes π_{US} and π_{IT} are very different, so that orderings which require a common (steady–state) distribution would not be applicable.

Based on standard indicators, Rustichini et al. find greater social mobility in the United States than in Italy. In addition to revisiting the issue with indices of progressive mobility, we shall apply our more stringent test of dominance (according to $\succ_{\text{eq}}$), asking which process better equalizes children's opportunities, for any arbitrary distribution of parental backgrounds. The results are summarized in **Table 3.**[31]

[28] The most relevant comparison is between numbers in the bottom and top panels of the *same column,* which compare how M_{69}^{76} and M_{79}^{86} operate on the same income distribution.

[29] These points often coincide with those where the transition matrices reported by Hungerford (1993) show slight non–monotonicities

[30] Fathers and sons are described by their occupations, and to each occupation is assigned its median income, as an indicator of long–term economic status. Finally, these "occupational incomes" are grouped into four intervals, whose boundaries differ by the same growth factor: denoting as $\underline{y}$ and $\bar{y}$ the minimum and maximum levels of income, class $k \in \{1, ...4\}$ corresponds $y \in [\underline{y}(1+g)^{k-1}, \underline{y}(1+g)^k]$, where $1 + g \equiv (\bar{y}/\underline{y})^{1/4}$. In the data, $g_{US} = 1.476$ and $g_{IT} = 1.467$. Since these numbers are quite close to each other, we define the (normalized) income state vector for both countries as $\mathbf{y} = (1, 1+\bar{g}, (1+\bar{g})^2, (1+\bar{g})^3)'$, where $\bar{g} \equiv \sqrt{g_{US}g_{IT}} = 1.472$.

[31] In particular, each country's mobility process, M_{US} or M_{IT}, is evaluated on both the US and the Italian income distributions, that is, on π_{US} (first column) as well as on π_{IT} (second column).

Table 3 here

The first observation is that there is much more cross-sectional inequality in the US than Italy. As illustrated on **Figure 2**, the US Lorenz curve for fathers' occupational incomes is everywhere below its Italian counterpart, with respective Gini coefficients of $G_{US} = .200$ and $G_{IT} = .160$. When we look at the extent to which these differences in social origins determine the next generation's opportunities, however, the picture is very different. The two Lorenz curves for sons' conditional expected incomes are virtually indistinguishable, with Ginis of $\hat{G}_{US} = .063$ and $\hat{G}_{IT} = .056$. The corresponding indices of progressivity are $\rho_{US}^{RS} = .137$ and $\rho_{IT}^{RS} = .104$. This comparison is not really fair, however, because there is less to equalize in Italy in the first place. As explained in the earlier sections, the appropriate comparison is between the effects of the two mobility processes on a *common* initial income distribution. This means comparing the entries in the top and bottom panels of **Table 3** within the *same column,* as opposed to across columns. For instance, had the Italian mobility process M_{IT} operated on the US distribution of fathers' incomes, the Gini in sons' opportunities would have been reduced by $\rho^{RS} = .121$; this is still less than the US number of $\rho_{US}^{RS} = .137$, but the gap is much smaller than before. Conversely, had Italy "imported" the US mobility process, inequality of opportunities would have fallen by only $\rho^{RS} = .116$.

Note that whether looking at Ginis, elasticities or average marginal tax rates, the ranking is the same for both initial distributions. And indeed, the last row of **Table 4** reveals that the US mobility process is in fact *unambiguously more egalitarian* than the Italian one: the superdiagonals test of Theorem 2 yields

$$M_{US} \succ_{\text{eq}} M_{IT} \succ_{\text{eq}} I.$$

This is rather remarkable, given the stringency of the requirements, even if the magnitude of the differences is fairly moderate. Thus, according to this data, the American intergenerational mobility process is a greater equalizer of *opportunities* than the Italian one. This contrasts sharply with *outcomes*, which are not at all equalized: in both countries, sons' ex–post income realizations exhibit the same degree of inequality as fathers' incomes. Finally, it is interesting to compute the actual profile of Italy–to–US relative net–of tax rates, or *relative growth rates in relative incomes,* $(1 - t_{IT}(y_i)) \, / \, (1 - t_{US}(y_i)) \; = \; e_{M_{IT}}(y_i)/e_{M_{US}}(y_i)$. The values are .90, . 91, . 95, . and .97, which shows that most of the mobility difference between the two countries occurs when the father rises above the second income class. The advantage thereby conferred to the son is markedly stronger in Italy than in the United States.

Appendix

Proof of Theorem 1. The equivalence of (ii) and (iii) is easy to establish, so we focus here on the equivalence of (i) and (ii).[32] For this purpose, we need the following two claims.

Claim 1. For any integrable functions $f, g : [0,1] \to \mathbb{R}_{++}$ such that f/g is decreasing, we have

$$\left(\int_0^p f(q)\ dq\right) \Big/ \left(\int_0^1 f(r)\ dr\right) \geq \left(\int_0^p g(q)\ dq\right) \Big/ \left(\int_0^1 g(r)\ dr\right) \text{ for all } p \in [0,1]. \tag{A.1}$$

Proof of Claim 1. Fix any $p \in (0,1]$, and note that

$$\begin{aligned} & \left(\int_0^p f(q)\ dq\right)\left(\int_0^1 g(r)\ dr\right) - \left(\int_0^p g(q)\ dq\right)\left(\int_0^1 f(r)\ dr\right) \\ = & \int_0^p \int_p^1 \left(f(q)g(r) - g(q)f(r)\right)\ drdq = \int_0^p \int_p^1 g(r)g(q)\left(\frac{f(q)}{g(q)} - \frac{f(r)}{g(r)}\right)\ drdq. \end{aligned}$$

That the last expression is nonnegative follows immediately from the fact that f/g is decreasing. ‖

Claim 2. For all $p \in (0,1]$ and $F \in \mathcal{F}(X)$, $\Lambda^{-1}_{F,M}(p) = e_M(F^{-1}(p))$ and $\Lambda^{-1}_{F,N}(p) = e_N(F^{-1}(p))$.

Proof of Claim 2. Take any $p \in (0,1]$, and note that

$$\begin{aligned} \Lambda^{-1}_{F,M}(p) &= \inf\{x \in supp(\Lambda_{F,M}) : \Lambda_{F,M}(x) \geq p\} = \inf\{x \in e_M(X) : F(e_M^{-1}(x)) \geq p\} \\ &= \inf\{e_M(y) \in X : F(y) \geq p\} = e_M\left(\inf\{y \in X : F(y) \geq p\}\right) = e_M(F^{-1}(p)), \end{aligned}$$

where the fourth equality follows from the continuity and strict monotonicity of e_M on X. ‖

Now, to show that (ii) implies (i), first extend e_M and e_N to $X \cup \{0\}$ by setting $e_M(0) = \inf e_M(X)$ and $e_N(0) = \inf e_N(X)$. Next, let $f \equiv e_M \circ F^{-1}$ and $g \equiv e_N \circ F^{-1}$. Since $F^{-1}(0) = 0$ and $F^{-1}(0,1] \subseteq supp(F) \subseteq X$, the functions f and g are well-defined. It is easy to see that they are also monotonic, and satisfy all the assumptions of Claim 1. Therefore, by Claims 1 and 2:

$$L_{\Lambda_{F,M}}(p) = \int_0^p f(q)\ dq \Big/ \left(\int_0^1 f(r)\ dr\right) \geq \int_0^p g(q)\ dq \Big/ \left(\int_0^1 g(r)\ dr\right) = L_{\Lambda_{F,N}}(p)$$

for all $p \in (0,1]$; that is, $M \succcurlyeq_{\text{eq}} N$.

Conversely, suppose that there exist $a, b \in X$ with $b > a > 0$ such that $e_M(b)/e_N(b) > e_M(a)/e_N(a)$. We need to show that $\Lambda_{F,M} \succsim_L \Lambda_{F,N}$ does not hold for some $F \in \mathcal{F}(X)$. Indeed, let $F(q) = \frac{1}{2}\mathbf{1}_{[a,b)} + \frac{1}{2}\mathbf{1}_{[b,\infty)}$, and observe that

$$L_{\Lambda_{F,M}}(1/2) = \frac{e_M(a)}{e_M(a) + e_M(b)} < \frac{e_N(a)}{e_N(a) + e_N(b)} = L_{\Lambda_{F,N}}(1/2),$$

which completes the proof. ■

[32] Various versions of this result in the *tax context* have been proved, for instance, in Jakobsson (1976), Fellman (1976), Lambert (1993) or Le Breton et al. (1996). We were, however, unable to find in the literature the more general version needed for Theorem 1, which covers both the continuous and discrete cases (as well as mixtures of the two). At any rate, our proof is short and self-contained.

Proof of Proposition 1. Fix any $t \geq 1$ and any $\theta \in (0, \infty]$, and observe first that:

$$\begin{aligned} E_{M^{(t+1)}}(y, \theta) &= \int_0^\theta x\, dM^{(t+1)}(x \mid y) = \int_0^\theta x\, d\left(\int_X M(x \mid z)\, dM^{(t)}(z \mid y)\right) \\ &= \int_0^\theta \int_X x\, dM(x \mid z)\, dM^{(t)}(z \mid y) = \int_X E_M(z, \theta)\, dM^{(t)}(z \mid y). \end{aligned}$$

Now, define $\Psi(z) \equiv E_M(z, \theta)/z$, for all $z \in X \cap \mathbb{R}_{++}$. Since X is closed, $\mathbb{R}_{++} \backslash X$ is open (in $\mathbb{R}_{++}$) and hence there exists a countable collection of disjoint open intervals $\{(a_n, b_n)\}_{n=1}^\infty$ in $\mathbb{R}_{++}$ such that $\mathbb{R}_{++} \backslash X = \cup(a_n, b_n)$. By linear interpolation we can then extend Ψ to a continuous, decreasing function on all of $\mathbb{R}_{++}$; we denote this extension again by Ψ for simplicity. We then have:

$$\frac{1}{y} E_{M^{(t+1)}}(y, \theta) = \int_X \Psi(z) \frac{z}{y}\, dM^{(t)}(z \mid y) = \int_0^\infty \Psi(z) \frac{z}{y}\, dM^{(t)}(z \mid y). \tag{A.2}$$

Integrating by parts with respect to z, then, we find

$$\begin{aligned} \frac{1}{y} E_{M^{(t+1)}}(y, \theta) &= \left[\Psi(z) \int_0^z \frac{u}{y}\, dM^{(t)}(u \mid y)\right]_0^\infty - \int_0^\infty \left(\int_0^z \frac{u}{y}\, dM^{(t)}(u \mid y)\right) d\Psi(z) \\ &= \left[\Psi(z) \frac{E_{M^{(t)}}(y, z)}{y}\right]_0^\infty - \int_0^\infty \frac{E_{M^{(t)}}(y, z)}{y}\, d\Psi(z) \\ &= \lim_{z \to \infty} \Psi(z) \left(\frac{e_{M^{(t)}}(y)}{y}\right) - \int_0^\infty \frac{E_{M^{(t)}}(y, z)}{y} d\Psi(z). \end{aligned} \tag{A.3}$$

Since Ψ is positive and decreasing in z, it is clear from this expression that $E_{M^{(t+1)}}(y, \theta)/y$ is decreasing in y whenever $M^{(t)}$ is strongly equalizing. The claimed result follows by induction. ■

Proof of Theorem 2. The equivalence of (i) and (ii) follows from Theorem 1. To show the equivalence with (iii), note that (i) holds if and only if

$$\frac{\mathbf{e}_i' P \mathbf{y}}{\mathbf{e}_i' Q \mathbf{y}} \geq \frac{\mathbf{e}_{i+1}' P \mathbf{y}}{\mathbf{e}_{i+1}' Q \mathbf{y}}, \quad i = 1, ..., n-1, \tag{A.4}$$

where $\mathbf{e}_i$ is the ith unit vector. But his can be rewritten as $\mathbf{e}_i' P \mathbf{y}\mathbf{y}' Q' \mathbf{e}_{i+1} \geq \mathbf{e}_i' Q \mathbf{y}\mathbf{y}' P' \mathbf{e}_{i+1}$, or

$$\mathbf{e}_i' P \mathbf{y}\mathbf{y}' Q' \mathbf{e}_{i+1} - \mathbf{e}_{i+1}' P \mathbf{y}\mathbf{y}' Q' \mathbf{e}_i \geq 0, \quad i = 1, ..., n-1,$$

which is equivalent to (iii). ■

Proof of Proposition 3 The first statement follows directly from Proposition 1. For the second one, take any $\rho \in (0, 1)$ and observe that since $P(\rho) = (I - \rho P)^{-1} = \sum_{t=0}^\infty \rho^t P^t$, we have:

$$\begin{aligned} (D^* - D_*)[P(\rho) u(\mathbf{y}) u(\mathbf{y})' I)] &= (1 - \rho)(D^* - D_*) \left[\sum_{t=0}^\infty \rho^t \left(P^t u(\mathbf{y}) u(\mathbf{y})'\right)\right] \\ &= (1 - \rho) \sum_{t=0}^\infty \rho^t (D^* - D_*)[P^t u(\mathbf{y}) u(\mathbf{y})']. \end{aligned}$$

Since P is strongly equalizing (given $u(\mathbf{y})$) we have $P^t \succcurlyeq_{\text{eq}}^{u(\mathbf{y})} I$ for each $t \in \mathbb{N}$, hence $(D^* - D_*)[P^t u(\mathbf{y}) u(\mathbf{y})'] \geq 0$ by Theorem 2. Thus $(D^* - D_*)[P(\rho) u(\mathbf{y}) u(\mathbf{y})'] \geq 0$, implying $P(\rho) \succcurlyeq_{\text{eq}}^{u(\mathbf{y})} I$ by Theorem 2. ■

References

[1] Atkinson, A. B. (1970). "On the Measurement of Inequality," *Journal of Economic Theory* 2, 244-263.

[2] Atkinson, A. B. (1983). "The Measurement of Economic Mobility," in *Social Justice and Public Policy*, ed. by A. B. Atkinson, Cambridge: MIT Press.

[3] Bartholomew, D.J (1982) "Stochastic Models for Social Processes," 3^d ed., Wiley: New York.

[4] Bénabou, R. and Ok, E. A. (2001). "Social Mobility and the Demand for Redistribution: the POUM Hypothesis," *Quarterly Journal of Economics,* 116(2), 447–487.

[5] Bishop. J., Chow, V. and Formby J. (1995). "The Redistributive Effect of Direct Taxes: A Comparison of Six Luxembourg Income Study-Countries," *Journal of Income Distribution*, 5, 65-90.

[6] Conlisk, J. (1989) "Ranking Mobility Matrices," *Economics Letters,* 29, 231-235.

[7] Conlisk, J. (1990). "Monotone Mobility Matrices." *Journal of Mathematical Sociology,* 15, 173-191.

[8] Dardanoni, V., (1993). "Measuring Social Mobility." *Journal of Economic Theory*, 61, 372-394.

[9] Dardanoni, V. (1995). "Income Distribution Dynamics: Monotone Markov Chains Make Light Work," *Social Choice and Welfare* 12, 181-192.

[10] Dasgupta, P., Sen, A. and Starett, D. (1973) "Notes on the Measurement of Inequality," *Journal of Economic Theory,* 6(2), 180–187.

[11] Fellman, J. (1976). "The Effect of Transformation on Lorenz Curves," Econometrica, 44, 823-4.

[12] Fields, G. and Ok, E. A. (1999a). "The Measurement of Income Mobility: An Introduction to the Literature," in *Handbook of Inequality Measurement,* ed. by J. Silber, Dordrecht, Germany: Kluwer Academic Publishers, 557–596.

[13] Fields, G. and Ok, E. A. (1999b). "Measuring Movement of Incomes," *Economica*, 66, 455–471.

[14] Formby, J., Smith, J. and Zheng B. (1995). "Economic Growth, Welfare and the Measurement of Social Mobility". University of Alabama–Tuscaloosa, mimeo, November.

[15] Foster, J. E. (1985). "Inequality Measurement," in *Fair Allocation*, ed. by H. P. Young. Providence: American Mathematical Society Proceedings in Applied Mathematics, Vol. 33.

[16] Gottschalk, P. (1997) "Inequality, Income Growth and Mobility: The Basic Facts," *Journal of Economic Perspectives,* 11(2), 21–40.

[17] Gottschalk, P. (1997) and Danziger, S. "Family Income Mobility: How Much is There and Has It Changed?", Boston College mimeo.

[18] Gottschalk, P. and Spolaore E. (2000) "On the Evaluation of Economic Mobility," Boston College mimeo, March.

[19] Jakobsson, U. (1976). "On the Measurement of the Degree of Progressivity," *Journal of Public Economics,* 5, 161–168.

[20] Hungerford, T.L., (1993). "U.S. Income Mobility in the Seventies and Eighties." *Review of Income and Wealth*, 31(4), 403–417.

[21] Kanbur, R. and Stromberg, J-O. (1988). "Income Transitions and Income Distribution Dominance," *Journal of Economic Theory* 45, 408-416.

[22] Kanbur, R. and Stiglitz, J. (1986). "Intergenerational Mobility and Dynastic Inequality," Princeton University mimeo.

[23] Keilson, J. and A. Ketser, (1977). "Monotone Matrices and Monotone Markov Chains." *Stochastic Processes and their Applications,* 5, 231-241.

[24] Kolm, S.C. (1960) "The Optimal Production of Justice," in *Public Economics*, J. Margolis and H. Guitton eds., MacMillan: London.

[25] Lambert, P. (1993) "The Distribution and Redistribution of Income: A Mathematical Analysis." 2nd. ed. Manchester: Manchester University Press.

[26] Le Breton, M., Moyes, P. and Trannoy A. (1996). "Inequality Reducing Properties of Composite Taxation," *Journal of Economic Theory* 69, 71-103.

[27] Loury, G. (1981) "Intergenerational Transfers and the Distribution of Earnings," *Econometrica,* 49, 843–867.

[28] Reynolds, M. and Smolensky, E.E. (1977) *"Public Expenditures, Taxes and the Distribution of Income: The United States 1950, 1961, 1970".* New York: Academic Press.

[29] Rustichini, A., Ichino, A. and Checchi, D. (1999) "More Equal But Less Mobile? Education Financing and Intergenerational Mobility in Italy and the US," *Journal of Public Economics,* December, 70(3), pp. 399.424.

[30] Sen, A. K. (1997). *On Economic Inequality.* Expanded Edition, Oxford: Clarendon Press.

[31] Shorrocks, A. (1978a). "Income Inequality and Income Mobility," *Journal of Economic Theory,* 19, 376–393.

[32] Shorrocks, A. (1978b). The Measurement of Mobility," *Econometrica* 45, 1013-1024.

[33] Stokey, N. (1998), "Shirtsleeves to Shirtsleeves: The Economics of Social Mobility," in *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures 1983-1997*, N. L. Schwartz, D. Jacobs and E. Kalai eds., Cambridge: Econometric Society Monographs No. 23.

[34] United States Bureau of the Census, *"Current Population Reports"* and data available at http://www.census.gov/ftp/pub/hhes/www/income.html. U.S. Department of Commerce, Income Statistics Branch, HHES Division, Washington, D.C. 20233-8500.

Initial and Expected Incomes	$(\pi, y_{74}) \rightarrow (\pi, M^{91}_{74} \cdot y_{74})$	$(\pi, y_{74}) \rightarrow (\pi, M^{91}_{74} \cdot y_{91})$	$(\pi, y_{91}) \rightarrow (\pi, M^{91}_{74} \cdot y_{91})$
Gini	.415 $\rightarrow$ .226	.415$\rightarrow$. 255	.466 $\rightarrow$.255
ΔGini = ρ^{RS}	.137	.160	.211
1-Average Residual Elasticity	.442	.391	.450
Average Marginal Tax Rate	.458	.628	.463
Dominance Tests: $M^{91}_{74} \succ^{\mathbf{y}}_{\text{eq}} I$ and $M^{91}_{74} \succ^{\mathbf{y}}_{\text{eq}} I$, for $y \in \{y_{74}, y_{91}\}$			

Table 1: Male earnings mobility in the US

Transitions: M^{76}_{69}			
Initial and Expected Incomes	$(\pi, y_{69}) \rightarrow (\pi, M^{76}_{69} \cdot y_{69})$	$(\pi, y_{69}) \rightarrow (\pi, M^{76}_{69} \cdot y_{79})$	$(\pi, y_{79}) \rightarrow (\pi, M^{76}_{69} \cdot y_{79})$
Gini	.362 $\rightarrow$ .217	.362 $\rightarrow$. 235	.393 $\rightarrow$. 235
ΔGini = ρ^{RS}	.137	.127	.157
1-Average Residual Elasticity	.432	.378	.466
Average Marginal Tax Rate	.349	.304	.340
Transitions: M^{86}_{79}			
Initial and Expected Incomes	$(\pi, y_{69}) \rightarrow (\pi, M^{86}_{79} \cdot y_{79})$	$(\pi, y_{69}) \rightarrow (\pi, M^{86}_{79.} \cdot y_{79})$	$(\pi, y_{79}) \rightarrow (\pi, M^{86}_{79} \cdot y_{79})$
Gini	.362 $\rightarrow$. 219	.362 $\rightarrow$. .238	.393 $\rightarrow$. .238
ΔGini = ρ^{RS}	.144	. 15 9	.155
1-Average Residual Elasticity	.444	.392	.430
Average Marginal Tax Rate	.353	.382	.318
Dominance Tests: see the text.			

Table 2: Famility income mobility in the US

United States Mobility		
Initial and Expected Incomes	$(\pi_{US}, y) \rightarrow (\pi_{US}, M_{US} \cdot y)$	$(\pi_{IT}, y) \rightarrow (\pi_{IT}, M_{US} \cdot y)$
Gini	.200 → .063	.160→.044
ΔGini = ρ^{RS}	.137	.116
1-Average Residual Elasticity	.727	.733
Average Marginal Tax Rate	.707	.752
Italian Mobility		
Initial and Expected Incomes	$(\pi_{US}, y) \rightarrow (\pi_{US}, M_{IT} \cdot y)$	$(\pi_{IT}, y) \rightarrow (\pi_{IT}, M_{IT} \cdot y)$
Gini	.200→.078	.160 → . 056
ΔGini = ρ^{RS}	.121	.104
1-Average Residual Elasticity	..640	.669
Average Marginal Tax Rate	.640	.688
Dominance Tests: $M_{US} \succ^{\mathbf{y}}_{\text{eq}} M_{IT} \succ^{\mathbf{y}}_{\text{eq}} I$		

Table 3: Intergenerational mobility in the US and Italy

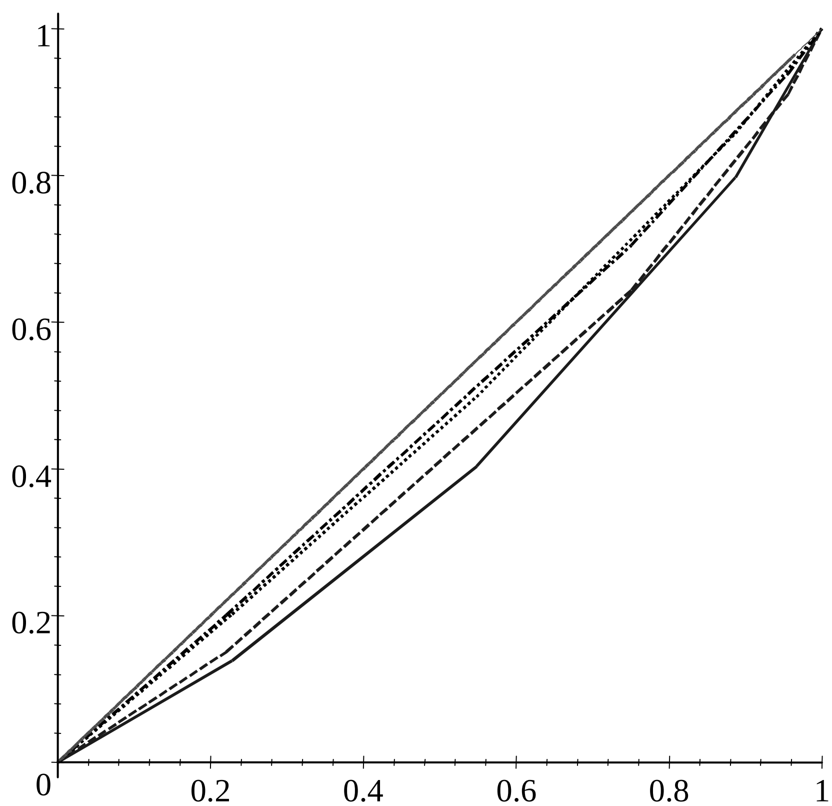

Figure 2: Lorenz curves for fathers' incomes and son's predicted incomes in the US and Italy

——— : L_{US}, – – – – : L_{IT}, ••••••• : $\hat{L}_{US}$, –•–•–•– : $\hat{L}_{IT}$