

Welfare Properties of Transition Matrices

Michel Lubrano

January 2019

Contents

1	Introduction	2
2	Markov processes	3
2.1	Monotone mobility matrices	3
2.2	Other references	4
3	The welfarist approach for comparing transition matrices	4
3.1	Atkinson-Dardanoni	5
3.2	Benabou and Ok	6
4	Equality of expected opportunity	7
4.1	Mobility ordering	7
4.2	Application to discrete Markov processes	8
4.3	Discussion	10
4.4	An empirical illustration using the PSID	10
4.5	An empirical illustration on Italian data	11
5	Preferences for redistribution	12
5.1	Basic assumptions	13
5.2	A log normal example	14
5.3	The POUM effect and risk aversion	15
5.4	Stochastic dominance and regular income dynamics	17
5.5	A numerical example	18
5.6	Related empirical work	19
6	Conclusion	20

1 Introduction

A transition matrix models the transition probability to a future income state j when one starts from state i . For every starting state, one is faced to a separate lottery (which is considered as a multinomial process from a statistical point of view). Mobility indices summarize the degree of mobility or immobility that is contained in a transition matrix. However, we shall discover very soon that we cannot consider the class of all transition matrices (independent rows summing to 1). We have to restrict this class so that the indices could verify a set of axioms. But even within a restricted class of matrices, mobility was praised for itself. Perfect mobility is defined with reference to a matrix having all its lines equal to the equilibrium vector π . This correspond to a condition of independence with respect to the initial conditions. While perfect immobility at the other side of the spectrum is defined as the identity matrix. However, upward mobility is valued in the same way as downward mobility with no reference made to a social welfare function.

In this chapter, we want to go step further and measure the social consequences of income mobility. For this we have to restrict first our attention to a narrower class of transition matrices, those which are *monotone*. The notion of monotonicity is detailed in several papers such as Conlisk (1990) or Dardanoni (1995). Outside that class, it is difficult to say anything easily. Then we shall investigate how we can define *progressive* mobility with two different approaches in the literature: that of Atkinson (1983) and Dardanoni (1993) on one side and that of Benabou and Ok (2001a) on the other side. That second approach is used and applied in the QJE paper Benabou and Ok (2001b) to study the preferences for redistribution, the prospect of upward mobility.

The concerned literature can be divided in two parts. There are first surveys, mainly Fields and Ok (1999), or partial surveys, we shall use Formby et al. (2004), that can give an overview. The second part of the literature is mainly given by self contained papers, the most interesting one being the unpublished paper of Benabou and Ok (2001a) which is prelude to Benabou and Ok (2001b). Note also Atkinson (1981, 1983) and Dardanoni (1993).

2 Markov processes

When one thinks of P as specifying an inter-temporal income distribution sequence, for example, p_{ij} may stand for the probability that a parent in income state i will have an offspring in state j . Given this interpretation, therefore, we may interpret the probability vector (p_{i1}, \dots, p_{ik}) as the lottery that an agent with a parent from state i faces. Which type of mobility can we have and measure? From the sociological literature is made the distinction between structural versus exchange mobility. To quote the survey of Fields and Ok (1999), *A distinction can be made between*

1. *changes in mobility that can be attributed to the increased availability of positions in higher social classes and those*
2. *changes that can be attributed to an increased intergenerational movement among social classes, for a given distribution of positions among these classes.*

It seems that the increased availability of higher positions corresponds to income mobility matrices while the increase in intergenerational movements corresponds to rank mobility and quantile matrices. However this distinction does not make any difference of value judgement between these type of mobility. We shall now introduce a preliminary restriction that makes life easier, that of monotonicity.

2.1 Monotone mobility matrices

We have to restrict the class of mobility matrices, otherwise incoherences can arise. We have income states that are ordered by increasing order, $x_{i+1} > x_i$. We would like to impose that it is better to start from state $i + 1$ than from state i . Which means that agent in state $i + 1$ faces a better lottery than agent starting from state i . This relies first on the definition of classes. With quantiles or mean matrices, this is the case. However, note that the classification used in Prais (1955) is in the reverse order.

Definition 1 *A transition matrix is said to be monotone if for all $l = 1, \dots, k - 1$ we have*

$$\sum_{j=1}^l p_{i+1,j} \leq \sum_{j=1}^l p_{i,j}.$$

This notion was introduced in the literature by Keilson and Kester (1977). The restriction of monotonicity is very often verified empirically and leads

to many simplifications as noted in Conlisk (1990), Dardanoni (1995). It is worth recalling one example drawn from Dardanoni (1995) and quoted by Fields and Ok (1999).

Example 1 *Suppose that we have two income distributions, the evolution of each of them being governed by a Markov process of order one with transition matrices P and Q . We suppose that at time $t-1$ the first distribution Lorenz dominates the second distribution (it is more equal). Under which condition will this ordering be preserved at time t , that is the next period for the expected income. I suppose that if we are in the simple Galton-Markov model the answer can be simple (to be verified). But with a Markov chain model, the answer is not evident in the general case. We have to impose restrictions on both matrices P and Q . Dardanoni (1995) shows that if P is monotone, then one obtains a definitive answer very easily: the income distribution generated by P at time t will remain more equal than that generated by Q if, and only if, P stochastically dominates Q that is,*

$$\sum_{j=1}^l p_{ij} \leq \sum_{j=1}^l q_{ij} \quad \forall l = 1, \dots, k-1, \quad \forall i.$$

Monotonicity is a property that has to be tested. This is the object of Dardanoni and Forcina (1998), a quite complicated paper because of a quite complicated topic. See also Lee et al. (2009).

2.2 Other references

In his section 2, Conlisk (1990) review the main mathematical properties of monotonicity. Then he explores the properties of various immobility indices and concludes by a comparison between his results and those of Shorrocks (1978).

Dardanoni (1995) paper is more interesting for analysing the properties of monotone Markov processes when Conlisk (1990) sticks more to analysing the properties of indices.

3 The welfarist approach for comparing transition matrices

The welfarist approach aims at comparing mobility transition matrices in term of their welfare consequences. In other words, when we want to compare two income distributions in a static framework, we might seek a criterion which allows to compare these two distributions whatever the social

welfare function or the type of aversion to inequality. This is the famous paper of Atkinson (1970) about stochastic dominance at the order one, two, etc. Now we want to do the same in a dynamic context. Of course, things are much more complex and there is no unified approach as underlined in Fields and Ok (1999). In this literature, *mobility is seen in terms of its implications rather than from a direct consideration of what is meant by mobility*. Welfarist studies usually view income mobility from the angle of origin-independence, and in fact, sometimes identify it with the notion of equality of opportunity. Given the social welfare function, a variety of methods can be used to extract a particular mobility measure. We present two options: Atkinson (1981), Atkinson (1983, p. 61) and Dardanoni (1993) on one side and Benabou and Ok (2001a) on the other side.

3.1 Atkinson-Dardanoni

Atkinson (1981) develops dominance conditions under which we can rank mobility processes for a class of social welfare functions satisfying certain general properties. His approach is best demonstrated by transforming the original data reported in $x \rightarrow y$ into percentile classes and by confining attention to the mobility of individuals among these groupings. So this is well adapted to our context of mobility matrices. In particular it concerns quantile transition matrices.

Further on in the literature, this approach was called the Atkinson-Dardanoni approach by Formby et al. (2004). It may seem better to refer to this paper than to the survey of Fields and Ok (1999). In his seminal paper, Atkinson (1983, p. 61) proposed the first dominance approach for measuring income mobility. Atkinson's method relates mobility to the properties of a social welfare function defined over incomes at two different dates. Mobility per se is not directly measured. Instead, the welfare implications of mobility are explored and an indirect measure of mobility implied. Atkinson considers a utilitarian social welfare function defined as:

$$W(x, y) = \int_0^\infty \int_0^\infty U(x, y) dK(x, y),$$

where $U(x, y)$ is a utility function depending on x and y two income distributions at two points of time. The joint distribution is characterized by $K(x; y)$. We want to compare two different income trajectories defined by $K(x; y)$ and $\tilde{K}(x, y)$. For the comparison to be possible, a restriction is imposed which is that

the two distributions have identical marginal distributions or equal steady state.

More precisely as said in Benabou and Ok (2001a), *two processes can be compared in this approach only if they have the same steady state. This might not be a problem if using **quantile transition matrices** because these matrices have the same steady-state, namely the uniform distribution.*

The regime with $K(x, y)$ has greater social welfare than the regime with $\tilde{K}(x, y)$ according to all social welfare function $W(x, y)$ if and only if

$$K(x, y) \leq \tilde{K}(x, y), \quad \forall x, y,$$

which is a kind of stochastic dominance condition. When Atkinson's result is applied to transition matrices, the requirement of equal marginal distributions is reflected in the fact that the sums of rows and columns must be the same between the two matrices. For two transition matrices P and \tilde{P} , the dominance condition is expressed as:

$$\sum_{i=1}^l \sum_{j=1}^m \pi_j p_{ij} \leq \sum_{i=1}^l \sum_{j=1}^m \pi_j \tilde{p}_{ij}, \quad \forall l, m.$$

This dominance condition has also been characterized in Dardanoni (1993). In Formby et al. (2004), this is noted $P \succ_{AD} \tilde{P}$. We must also note a similar approach in Conlisk (1989) who compares two mobility matrices having the same equilibrium vector.

3.2 Benabou and Ok

Benabou and Ok (2001a) view mobility as a mechanism that equalizes income opportunities and derive a quite different dominance condition. They represent a person's opportunity as the expected income in the succeeding period. For example, given that (x_1, x_2, \dots, x_k) is the income vector associated with the k income classes in both regimes, a person initially in the i^{th} class will have an expected income of

$$e_i = \sum_{j=1}^k p_{ij} x_j, \quad \tilde{e}_i = \sum_{j=1}^k \tilde{p}_{ij} x_j.$$

Benabou and Ok show that the necessary and sufficient condition for a **size** transition matrix P to be more opportunity equalizing than \tilde{P} , denoted as $P \succeq_{BO} \tilde{P}$, for all possible initial income distributions of x is

$$\frac{\tilde{e}_1}{e_1} \geq \frac{\tilde{e}_2}{e_2} \geq \dots \geq \frac{\tilde{e}_k}{e_k}.$$

Note that, in contrast to the Atkinson-Dardanoni condition, Benabou and Ok's condition does not require the initial distributions to be equal.

This type of progressivity will be used in Benabou and Ok (2001b) where the main mechanism is the prospect of upward mobility and taxation. In the case of progressivity, households below the mean income vote against redistribution because they do not want their expected future income to be taxed and redistributed to those that will be under the mean. So Benabou and Ok (2001a) are interesting in the predicting mechanism for future income. This is the concept of *Equality of Expected Opportunity*.

4 Equality of expected opportunity

We present here more in detail NBER Benabou and Ok (2001a), a still unpublished paper which presents an alternative to the welfarist approach of Atkinson (1983) and Dardanoni (1993). The main point is that mobility *per se* is not the point of interest. The concept of equality of opportunity provides a natural way to gauge mobility. Desired mobility is progressive in the same sense as progressive taxation. Does mobility manages to compensate for unequal initial endowments in term of future opportunities? Mobility has to be seen as a stochastic redistribution. Is it an equalizing movement?

The difference with the Dardanoni (1993) approach is that the authors consider the ex-ante possibilities and not the realization of the future income distribution. They are concerned with what is predictable and not by the consequences of unpredictable shocks.

As a preliminary restriction, Benabou and Ok (2001a) imposes that the future income prospect increases smoothly with the current level of income in the sense of first order stochastic dominance. For any $y_2 > y_1$, we must have for the mobility process $M(\cdot)$, $M(x|y_2) > M(x|y_1)$ for all the possible values of x . If the mobility process is modelled using a Markov process, this means that the transition matrix has to be monotone in the sense detailed above and in Dardanoni (1995).

4.1 Mobility ordering

An economy is characterized by a triplet (X, F, M) , X a set of all the feasible income levels, F their initial distribution, and M the mobility process. Two questions are of main interest:

1. When would we say that mobility reduces inequality of opportunity, relative to social origins?

2. When would we say that mobility is more equalizing in an economy (X, F, M) than in an economy (X, G, P) (international comparisons).

Future opportunities are fully described by the conditional distribution $M(\cdot|y)$ which can be summarized by the strictly increasing function in y due to the monotonicity assumption:

$$e_M(y) = \int x dM(x|y).$$

The distribution of conditional expected incomes induced by (X, F, M) is given by $F(e_M^{-1}(x))$. Then the formal condition that mobility is opportunity equalizing is that we have the Lorenz ordering:

$$F(e_M^{-1}(x)) \succeq_L F(x).$$

This condition is relative to the initial income distribution. This is not so a complete ordering. The same mobility process can have opposed properties for a subgroup or in the future for a future realized income distribution. A Process M would be said equalizing if this ordering is valid for all the initial income distributions.

To answer the second question, we have to start from the same initial income distribution F for a local approach. We say that $M \succeq P$ for a global approach if:

$$F(e_M^{-1}(x)) \succeq_L F(e_P^{-1}(x)) \quad \forall F.$$

The previous condition is obtained as a special case if we consider $P = I$.

At last a characterization theorem is given by:

Theorem 1 *Let M and P be two mobility processes on X . The following statements are equivalent*

1. $M \succeq_{eq} P$
2. e_M/e_P is decreasing on X

The analysis can be generalized to take into account more than two periods and risk aversion.

4.2 Application to discrete Markov processes

We note $0 < y_1 < y_2 < \dots < y_k$ the income state vector y . An income distribution is the probability vector π .

A monotone transition matrix (facing a better lottery) is such that

$$\sum_{j=1}^l p_{i+1,j} \geq \sum_{j=1}^l p_{i,j} \quad \forall i, l \in \{1, \dots, k-1\}$$

The cumulative conditional distribution of y_l given the state y_i , or in other words the mobility process M_P , is defined by:

$$M_P(y_l|y_i) = \sum_{j=1}^l p_{i,j}.$$

Consequently monotone transition matrices can be ranked with \succeq_{eq} . Income opportunities are defined by:

$$e_P(y_i) = \sum_{j=1}^k p_{ij} y_j.$$

A matrix P will be said to be more equalizing for a given income distribution y than a matrix Q whenever $M_P \succeq_{eq}^y M_Q$. A matrix P will be said equalizing or progressive if $P \succeq_{eq}^y I$, where I is the identity matrix.

Theorem 2 *Let y be an ordered income state vector and let P and Q be two monotone transition matrices. The following statements are equivalent:*

1. $P \succeq_{eq}^y Q$,
2. $\frac{e_P(y_1)}{e_Q(y_1)} \geq \frac{e_P(y_2)}{e_Q(y_2)} \geq \dots \geq \frac{e_P(y_k)}{e_Q(y_k)}$.

Now let us consider the particular case the $Q = I$. We can state the condition to be tested empirically that a transition matrix is progressive. It corresponds to to the following ordering:

$$\frac{1}{y_1} \sum_{j=1}^k p_{1j} y_j \geq \frac{1}{y_2} \sum_{j=1}^k p_{2j} y_j \geq \dots \geq \frac{1}{y_k} \sum_{j=1}^k p_{kj} y_j.$$

The property of strong equalization corresponds to the case where the above condition is extended to all partial sums, replacing k by l and varying it between 1 and k .

4.3 Discussion

1. A transition matrix is always defined with respect to a particular income state y . So mobility has to be discussed in that framework for a given income state, whatever the stationary distribution π .
2. The use of a quantile transition matrix does not seem recommendable, because it may correspond to unequal intervals, due to the skewness of the income distribution. There is no constant increase between the different classes when using quantiles.
3. Dardanoni (1993) propose another way to define progressivity, by reference to an equilibrium distribution π . Benabou and Ok (2001a) show that it is not possible to find an ordering that would be independent of both y and π . When $k \geq 3$, there is no transition matrix which is more equalizing than the identity matrix over all income supports.
4. In the ordering defined by Benabou and Ok (2001a) on monotone matrices, the identity matrix is not the smallest element, because it is not the worst from the point of view of inequality of opportunity. For instance

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

is a worse situation where the middle class falls down to the floor.

5. A new mobility index is not needed because here mobility is progressivity. We can use the existing progressivity indices such as that Reynolds-Smolensky (1977). We can take the difference between the Gini coefficients of the initial income distribution F and the predicted income distribution the next period

$$Gini(F) - Gini(F \circ e_M^{-1}).$$

I would add that we could use any other inequality index.

6. An aspect that can be interesting to develop is to investigate if mobility reduces poverty and how to test this assumption.

4.4 An empirical illustration using the PSID

Benabou and Ok report empirical results from Gottschalk (1997) based on interquintile transition matrices computed between 1974 and 1991 for individual male labour earnings. The mobility prospects over the complete

Table 1: Male earning mobility in the US

Initial and expected incomes	$y_{74} \rightarrow (M_{74}^{91}y_{74})$	$y_{74} \rightarrow (M_{74}^{91}y_{91})$
Gini	0.415 \rightarrow 0.226	0.415 \rightarrow 0.255
Δ Gini	0.137	0.160
Average Marginal tax rate	0.458	0.628
Dominance tests	$M_{74}^{91} \succeq_{eq} I$ for $y \in \{y_{74}, y_{91}\}$	

period reduce the Gini coefficient from 0.415 to 0.226 or 0.255 depending on whether the rise in inequality was unexpected (column 1) or fully anticipated (column 2). The Reynolds-Smolenski progressivity index is 0.189 or 0.160. The same index for the tax system was only of 0.031 in 1979. The average tax rates corresponding to this mobility process $t_i = 1 - e_M(y_i)/y_i$ are equal to

$$-282\%, -57.9\%, -16.9\%, +10\%, +41.4\%$$

for the five income classes of the study.

4.5 An empirical illustration on Italian data

Benabou and Ok (2001a) (NBER paper) use the data of Rustichini et al. (1999) to compare intergenerational income mobility between fathers and sons in the US and in Italy. The matrices consists of four occupational incomes with equi-proportional increases in income. The y_{US} and y_{IT} are roughly the same up to a constant of proportionality, but the equilibrium distributions π_{US} and π_{IT} are very different, so that the Atkinson-Dardanoni ordering would not be applicable.

Table 2: Intergenerational mobility in the US and in Italy

US Mobility		
Initial and expected incomes	$\pi_{US} \rightarrow \pi_{US}M_{US}$	$\pi_{IT} \rightarrow \pi_{IT}M_{US}$
Gini	0.200 \rightarrow 0.063	0.160 \rightarrow 0.044
Δ Gini	0.137	0.116
Average marginal tax rate	0.707	0.752
Italian Mobility		
Initial and expected incomes	$\pi_{US} \rightarrow \pi_{US}M_{IT}$	$\pi_{IT} \rightarrow \pi_{IT}M_{IT}$
Gini	0.200 \rightarrow 0.078	0.160 \rightarrow 0.056
Δ Gini	0.121	0.104
Average marginal tax rate	0.640	0.688

There is more inequality in the US (Gini = 0.200) than in Italy (Gini =

0.160). But after intergenerational mobility, the difference is not significant. The index of progressivity are more in favour of the US, but the way to go was more important. Look when the mobility matrix of one country is imported to the other country. The difference is reduced. Applying the ranking tests, we have

$$M_{US} \succeq_{eq} M_{IT} \succeq_{eq} I.$$

So according to these data, the US equalizing process is more efficient than the Italian one. The implicit tax rates are more important in the US than in Italy.

It is fair to note that the equalizing mobility process is very powerful, equivalent to marginal tax rates that the public would like to experience nowadays. Using another example based on individual male earnings collected from the PSID between 1974-1991, Benabou and Ok (2001a) note that the Reynolds-Smolenski progressivity index of the US fiscal system is much less important than the same index computed for income equalizing mobility.

5 Preferences for redistribution

The QJE paper of Benabou and Ok (2001b) builds a story around a simple idea. If individuals anticipate that natural income mobility is progressive, as we have seen in the previous section, individuals could decide that there is no need for further taxation-redistribution, even for some of them who are below the mean income. This is not said exactly in these terms, but this is the essential message of the POUM hypothesis, prospect of upward mobility.

The classical model of Meltzer and Richard (1981) determines an optimal taxation rate leading for redistribution just by inspecting the gap between the median and the mean incomes. This is a static model based on the theory of the median voter of Romer (1975). Individuals who have an income below the mean will vote for redistribution because they expect to receive more than their tax contribution. This model predicts that increased inequality will induce increased redistribution. Even if it has received a recent renewal of interest with Karabarbounis (2011), this model is too simple, just because it is a static model that cannot capture income dynamics. If the same voters anticipate that their future income will be greater than the mean, their future status will be changed from being tax earners to being tax payers, and they will vote against redistribution. Benabou and Ok (2001b) formalised this idea with the POUM hypothesis.

5.1 Basic assumptions

In a static world, individuals have a preference for redistribution if their income is below the mean and if the taxation-redistribution scheme relating their disposable income x_i to their gross wage w_i , to the mean wage \bar{w} and to the taxation rate τ has the linear form:

$$x_i = (1 - \tau)w_i + \tau\bar{w},$$

a scheme where the government budget is in equilibrium and that cannot be changed in the next future. A consequence of this scheme is that when $w_i < \bar{w}$, we have $x_i > w_i$. This is the classical model of Meltzer and Richard (1981) based on the theory of the median voter of Romer (1975). In practice however, we often find that some individuals, despite being below the mean wage are against redistribution. This is the POUM (prospect of upward mobility) effect of Benabou and Ok (2001b). If individuals take into account their future income, they might anticipate that income mobility will make them better off so that their future income might be greater than the future mean income of the distribution. Benabou and Ok show that this effect does exist, provided we impose a mild restriction on the income mobility process. Consequently, studying preferences for redistribution becomes just equivalent to analysing the properties of the income mobility process. See Alesina and Giuliano (2009) for a review.

Three main assumptions are made in Benabou and Ok that are:

1. *Individual incomes x_{it} are drawn from a common skewed distribution.*
2. *Income grows according to a continuous function f with a well defined expectation in x .*
3. *The function f is a concave non-affine function.*

As a consequence, there exists a current value of x , $x^* < E_t(x)$ such that the individuals belonging to the income interval $[x^*, E_t(x)]$ have a future income which is greater than $E_{t+1}(x)$. With this simple consequence, as soon as individuals integrate their future income in their utility function, all those having an income greater than x^* will vote against redistribution, and not only those with an income greater than $E_t(x)$, provided of course that they are not too much risk adverse.

There is a side assumption that is made in Benabou and Ok (2001b), but that does not seem to be used in their proof:

4. *Future income increases with current income in the sense of first order stochastic dominance.*

This last condition would be the equivalent of a Pareto assumption (everybody is better off with the future income distribution). This assumption is not innocent as it imposes a further restriction on the dynamics of the income mobility process, a regularity condition which entails the existence of a steady state.

In this section, we shall develop a small model around the lognormal distribution. We shall illustrate the importance of each of the above assumptions when deriving the result of Benabou and Ok and show, when relaxing slightly some of them, notably the stochastic dominance assumption, the fragility of their result. As a consequence, studying the income mobility process is certainly interesting per se, but will not exhaust the question of explaining individual preferences for redistribution. A complete econometric model has to be build, explaining status mobility, income mobility and the entailed preferences for redistribution for each group.

5.2 A log normal example

Let us consider a population of n individuals that have an income which is log-normally distributed at time t with parameters μ_t and σ_t^2 . That means:

$$x_{it} \sim \Lambda(\mu_t, \sigma_t^2) \quad \log x_{it} \sim N(\mu_t, \sigma_t^2).$$

In order to exploit the properties of the lognormal process, we suppose that individual income grows according to an autoregressive process. This is a variant of the Galton-Markov model extensively used for instance in Atkinson et al. (1992) or Hart (1976):

$$\log x_{it} = \log a + b \log x_{i,t-1} + \epsilon_{it}, \tag{1}$$

where ϵ_{it} a Gaussian white noise of zero mean and variance ω^2 .¹ The function f is thus defined as being:

$$f(x) = ax^b \exp(\epsilon).$$

This function is concave as soon as $b < 1$. Under these conditions, the income distribution in the next period will be also log normal, but with parameters $\log(a) + b\mu_t$ and $b^2\sigma_t^2 + \omega^2$ so that:

$$x_{i,t+1} \sim \Lambda(\log(a) + b\mu_t, b^2\sigma_t^2 + \omega^2).$$

We shall show that $b < 1$ is a necessary but not a sufficient condition.

¹See also Benabou and Ok (2001b) p. 475 for their income distribution and transition exercise.

5.3 The POUM effect and risk aversion

Let us now suppose that the utility of income has the form of a Constant Relative Risk Aversion function:

$$U(x_i) = \frac{x_i^{1-\alpha}}{1-\alpha}, \quad 0 < \alpha \leq 1.$$

The POUM effect requires that we can find individuals having an income at time t that is lower than the mean, but with a future expected income that is greater than the mean of the next period income distribution. Then they will vote against redistribution, which means against the taxation of their future expected income if the expected utility of their future income is greater than the utility of the mean of the future income distribution. The expected utility of the predicted future income is computed as a conditional expectation:

$$E_\epsilon U(x_{i,t+1}|x_{it}) = \int_\epsilon \frac{1}{1-\alpha} [ax_{it}^b \exp(\epsilon)]^{1-\alpha} dF_\epsilon.$$

Factorising all the elements which are not a function of ϵ out of the integral, we get

$$E_\epsilon U(x_{i,t+1}|x_{it}) = \frac{1}{1-\alpha} a^{1-\alpha} x_{it}^{(1-\alpha)b} \int_\epsilon [\exp((1-\alpha)\epsilon)] dF_\epsilon.$$

The integral then represents the expectation of a lognormal with parameters 0 and $(1-\alpha)^2\omega^2$ so that:

$$E_\epsilon U(x_{i,t+1}|x_{it}) = \frac{1}{1-\alpha} a^{1-\alpha} x_{it}^{(1-\alpha)b} \exp\left(\frac{(1-\alpha)^2\omega^2}{2}\right).$$

This expected utility has to be greater than the utility of the mean of the future income distribution, namely $U(E_\epsilon(x_{i,t+1}))$ which is equal to:

$$\begin{aligned} U(E(x_{i,t+1})) &= U\left(\exp\left(\log a + b\mu + \frac{b^2\sigma^2 + \omega^2}{2}\right)\right) \\ &= \frac{1}{1-\alpha} a^{1-\alpha} \exp\left((1-\alpha)b\mu + \frac{(1-\alpha)b^2\sigma^2 + (1-\alpha)\omega^2}{2}\right). \end{aligned} \quad (2)$$

Equating these two expectations, we find the current value of income, x_{it}^* above which an individual will vote against redistribution:

$$x_{it}^* = \exp\left(\mu_t + \frac{b^2\sigma_t^2 + \alpha\omega^2}{2b}\right).$$

A POUM effect exist if in the current income distribution we have $x_{it}^* < E_t(X)$, which means:

$$\exp\left(\mu + \frac{b^2\sigma^2 + \alpha\omega^2}{2b}\right) < \exp\left(\mu + \frac{\sigma^2}{2}\right),$$

the greater the distance $E(X) - x_{it}^*$, the greater the POUM effect.

We have two interesting cases, depending on risk aversion:

1. When $\alpha = 0$ (risk neutrality), the POUM effect requires simply that $b < 1$. This is a simple concavity restriction on the transition function f which corresponds to the core assumption in Benabou and Ok.
2. In the case of risk aversion ($\alpha > 0$), the condition is more complicated:

$$b(b-1) + \frac{\omega^2}{\sigma^2}\alpha < 0. \quad (3)$$

It includes a quadratic function of b and depends on the relative noise ratio ω^2/σ^2 . When the noise in the income mobility tends to zero, we are back to the previous condition. For a strictly positive noise, more risk aversion implies more concavity and thus a lower b . This is true till a certain point because α has to be lower than a given number:

$$\alpha < \sigma^2/(4\omega^2). \quad (4)$$

Otherwise (3) has no solution. For a high degree of risk aversion, it is not possible to find a feasible value for b . So a large risk aversion can kill any possibility of a POUM effect.

The lower bound can be easily reached as can be seen from a rough calibration. A value of $\sigma^2 = 0.30$ corresponds to a Gini equal to 0.30, a most common value for gross income in Europe.² If $\omega^2 < 0.075$, then α has just to be lower than 1, which corresponds anyway to its upper bound. This value of ω^2 means a residual variance of 7.5% in regression (1) describing the mobility process. When $\omega^2 = \sigma^2$ meaning a much higher variance in the mobility process, then α cannot be greater than 0.25 to allow for a POUM effect. The parameter ω^2 represents uncertainty in the mobility process. For a small value, society evolve at a regular pace. A larger value corresponds to a higher risk of falling down in the social ladder due for instance to a greater risk of unemployment, or getting up. This is illustrated in Table 3. We note that with those calibrated values and $b = 0.75$, the percentage of individuals with an income below the mean still voting against redistribution is rather small. In an empirical application, risk aversion is going to be individual dependant, introducing thus heterogeneity.

²In the lognormal process, the Gini index is equal to $G = 2\Phi(\sigma/\sqrt{2}) - 1$. So that for a given value of G , we have $\sigma^2 = 2[\Phi^{-1}(\frac{G+1}{2})]^2$.

Table 3: Percentage of a POUM effect
with risk aversion

α	0.00	0.20	0.40	0.60	0.80	1.00
ω^2						
0.300	3.68	0.00	0.00	0.00	0.00	0.00
0.150	3.68	1.74	0.00	0.00	0.00	0.00
0.075	3.68	2.71	1.74	0.75	0.00	0.00
0.037	3.68	3.20	2.71	2.23	1.74	1.24

5.4 Stochastic dominance and regular income dynamics

Stochastic dominance for the lognormal process was first analysed in Levy (1973). More precisely, his theorem 4 states:

Theorem 3 *Let F_1 and F_2 be two lognormal distributions with parameters μ_j and σ_j . F_2 dominates F_1 at the order one if $\mu_2 > \mu_1$ and $\sigma_1 = \sigma_2$.*

Stochastic dominance at the order one requires that the two processes have the same log variance. In our case, this condition implies

$$\sigma_t^2 = b^2 \sigma_t^2 + \omega^2 \Rightarrow b^2 = 1 - \omega^2 / \sigma_t^2.$$

The second condition is a kind of growth condition. In order that $F_{t+1} > F_t$, we must have $\mu_{t+1} > \mu_t$, which translated in our example requires

$$\log a > (1 - b)\mu_t.$$

The growth parameter a has first to be greater than 1 and secondly has to be an inverse function of b : the lower b , the greater a .

We are now in a position to interpret this condition of stochastic dominance. The first condition which says that $\sigma_t^2 = \sigma_{t+1}^2$ guaranties a stability of the process. In particular, if condition (4) is verified at the initial state of the system, it will be verified all the time. The proportion of individuals having an income between x^* and $E_t(x)$ will remain constant and the upper bound on α will also remain constant. The absence of stochastic dominance at the first order can create a situation where for instance σ_t^2 decreases over time. That is a condition for Lorenz ordering. But at the same time, it becomes harder to meet the requirement on α when times elapses. So we can start from a situation where there is a POUM effect and that the POUM effect disappears after a certain time. We shall illustrate that situation on a small numerical example. Stochastic dominance at the order one eliminates

irregular dynamic situations and thus might be an oversimplification when confronted to real data. It excludes for instance situation where inequality is decreasing if the log normal assumption is verified.

5.5 A numerical example

Let us propose the following calibration for our dynamic system: $\sigma^2 = 0.30$ justified by a Gini index of 0.30, $\omega^2 = 0.15$ a somewhat not so erratic mobility process, $\alpha = 0.34$ an intermediate risk aversion, $\mu_1 = 0.5$ (a scalar for the graph), $b = \sqrt{(1 - \omega^2/\sigma^2)} - 0.2$ meaning a small departure from stochastic dominance implying a decreasing σ_t^2 and $a = \exp((1 - b) * \mu) + 0.1$ implying a small growth rate of the mean income. We have the following results in Table 4. From periods 1 to 3, the POUM condition (3) is verified. But starting

Table 4: The POUM effect with transitory dynamics

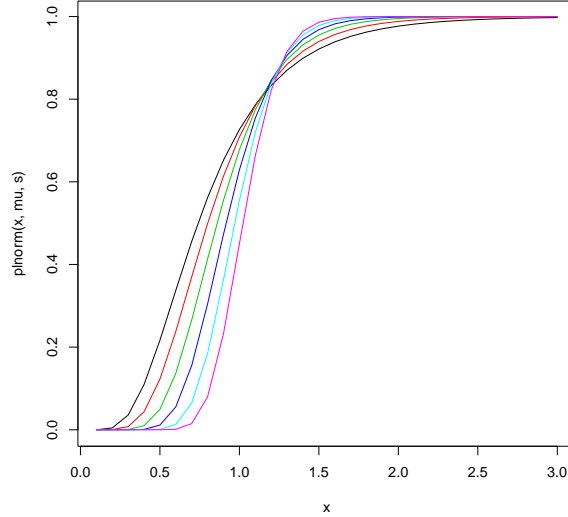
Period	μ_t	σ_t^2	b	x^*	E(x)	% Poum	α lower
1	0.500	0.300	0.507	1.871	1.916	1.668	0.500
2	0.575	0.227	0.507	1.980	1.991	0.464	0.379
3	0.613	0.208	0.507	2.047	2.050	0.092	0.347
4	0.633	0.204	0.507	2.085	2.085	0.000	0.339
5	0.643	0.202	0.507	2.105	2.104	0.000	0.337
6	0.648	0.202	0.507	2.115	2.114	0.000	0.337

with period 4, the upper bound on α is violated because σ_t^2 has decreased. The POUM effect disappears because of risk aversion. We see in Figure 1 that we are too far from stochastic dominance, the dominance curves are intersecting in their upper part. This is an example of an irregular dynamics that can reflect a transitional situation.

Remark:

Feri (2012) in a similar lognormal model considers the case where the population is partitioned in two groups, a large group of unskilled individuals in proportion p with a low μ_1 experiencing a slow change of their income and a small group of skilled individuals in proportion $1-p$ with a higher μ_2 experiencing a quick mobility. Consequently x_1^* will be greater than the mean income of the whole population $\mu = p\mu_1 + (1-p)\mu_2$ while $x_2^* < \mu$. The total effect will depend on p and the mobility differences. In an empirical illustration on Italian data, Feri shows that the POUM effect is present when income mobility is treated as a whole and disappears when allowing for heterogeneity in income mobility. His result is

Figure 1: Dominance curves and transitory dynamics



obtained supposing a steady state and would be much more complex to derive outside this framework. This is thus another example where the POUM effect disappears.

5.6 Related empirical work

The POUM effect is easy to interpret. The income distribution has to evolve in such a way that it dynamically implements directly a redistributive scheme, so that no extra taxation-redistribution has to be implemented in order to reduce inequality. Provided that individuals anticipate correctly this mechanism and that they integrate future income in their utility function, they will vote against an extra redistributive scheme even if they are under the mean. However, we have shown that any departure from the four assumptions made in Benabou and Ok leads to situations where the POUM effect can disappear. Income mobility, heterogenous risk aversion are complex processes that have to be detailed in order to relate them correctly to the preference for redistribution. We have thus to focus on the individual level. From the previous section, we have seen that individual preferences can be highly complex and non-linear even if the impact of each separate element is trackable. We have to study how different factors enter into the individual utility function.

There exists an important empirical literature containing evidences about the relation between preferences for redistribution and income mobility, see e.g. Alesina and La Ferrara (2005) or Clark and D'Angelo (2008) and the references quoted there. However, Clark and D'Angelo (2008) focus on between generation mobility. Alesina and La Ferrara (2005) illustrates the importance of future income expectations which can dominate the current income effect. But they do not take into account heterogeneity and risk aversion. They measure income mobility by mean of an homogeneous Markov transition matrix and show that their result (the importance of upper income expectations on preferences for redistribution) is robust to individual heterogeneity. In this paper, we point out that individual heterogeneity is of prime importance. Ravallion and Lokshin (2000), using Russian panel data insisted on heterogenous individual effect and on dynamics. But their panel survey contains information on preference for redistribution only for 1996. They concluded that differences in speed for predicted income mobility was a determinant factor.

6 Conclusion

There are interesting readings for this topic. We have stressed the interest of Formby et al. (2004) which is more than just an econometric paper. There is a complementary bibliography with:

Conlisk, J. 1974. "Can Equalization of Opportunity Reduce Social Mobility?" *American Economic Review* 64:80-90.

Conlisk, J. 1985. "Comparative Statics for Markov Chains." *Journal of Economic Dynamics and Control* 9:139-151.

Conlisk, J. 1989. "Ranking Mobility Matrices" *Economics Letters* 29: 231-235.

Conlisk, J. 1990. "Monotone Mobility Matrices" *Journal of Mathematical Sociology* 15: 173-191.

There is the survey of Fields and Ok (1999): The main interest for us are now sections 4 and 5 of this paper, concerned, for section 4, by *Welfarist Approaches to Income Mobility Measurement*, especially the paragraph on Atkinson (1981)'s paper and for section 5 *The Markovian Approach to Mobility Measurement*.

References

- Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89(5-6):897–931.
- Alesina, A. F. and Giuliano, P. (2009). Preferences for redistribution. Technical report, National Bureau of Economic Research.
- Atkinson, A. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2:244–263.
- Atkinson, A. (1983). The measurement of economic mobility. In Atkinson, A., editor, *Social Justice and Public Policy*. MIT Press, Cambridge, MA.
- Atkinson, A. B. (1981). The measurement of economic mobility. In Atkinson, A. B., editor, *Essays in Honor of Jan Pen*. reprinted in *Social Justice and Public Policy* (Brighton: Wheatsheaf 1983, Chapter 3).
- Atkinson, A. B., Bourguignon, F., and Morrisson, C. (1992). *Empirical Studies of Earnings Mobility*, volume 52. Harwood Academic Publishers.
- Benabou, R. and Ok, E. A. (2001a). Mobility as progressivity: Ranking income processes according to equality of opportunity. Working Paper 8431, National Bureau of Economic Research, Cambridge MA.
- Benabou, R. and Ok, E. A. (2001b). Social mobility and the demand for redistribution: The POUM hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487.
- Clark, A. and D’Angelo, E. (2008). Upward social mobility, wellbeing and political preferences: Evidence from the BHPS. Working paper, Paris School of Economics, PSE, Paris.
- Conlisk, J. (1989). Ranking mobility matrices. *Economics Letters*, 29:231–235.
- Conlisk, J. (1990). Monotone mobility matrices. *The Journal of Mathematical Sociology*, 15(3-4):173–191.
- Dardanoni, V. (1993). On measuring social mobility. *Journal of Economic Theory*, 61:372–394.
- Dardanoni, V. (1995). Income distribution dynamics: monotone Markov chains make light work. *Social Choice and Welfare*, 12:181–192.

- Dardanoni, V. and Forcina, A. (1998). A unified approach to likelihood inference on stochastic orderings in a nonparametric context. *Journal of the American Statistical Association*, 93(443):1112–1123.
- Feri, F. (2012). A note on the POUM effect with heterogeneous social mobility. *Economics Letters*, 115:258–262.
- Fields, G. S. and Ok, E. A. (1999). The measurement of income mobility: An introduction to the literature. In Silber, J., editor, *Handbook on income inequality measurement*, pages 557–596. Kluwer Academic Publishers, Norwell, MA.
- Formby, J. P., Smith, W. J., and Zheng, B. (2004). Mobility measurement, transition matrices and statistical inference. *Journal of Econometrics*, 120(1):181–205.
- Hart, P. E. (1976). The comparative statics and dynamics of income distributions. *Journal of the Royal Statistical Society. Series A (General)*, 139(1):108–125.
- Karabarbounis, L. (2011). One dollar, one vote. *The Economic Journal*, 121(553):621–651.
- Keilson, J. and Kester, A. (1977). Monotone matrices and monotone Markov chains. *Stochastic Processes and Their Applications*, 5:231–241.
- Lee, S., Linton, O., and Whang, Y.-J. (2009). Testing for stochastic monotonicity. *Econometrica*, 77(2):585–602.
- Levy, H. (1973). Stochastic dominance among log-normal prospects. *International Economic Review*, 14(3):601–614.
- Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *The Journal of Political Economy*, 89(5):914–27.
- Prais, S. (1955). Measuring social mobility. *Journal of the Royal Statistical Society. Series A*, 118:56–66.
- Ravallion, M. and Lokshin, M. (2000). Who wants to redistribute?: The tunnel effect in 1990s Russia. *Journal of Public Economics*, 76(1):87–104.
- Romer, T. (1975). Individual welfare, majority voting and the properties of a linear income tax. *Journal of Public Economics*, 7:163–188.
- Shorrocks, A. F. (1978). The measurement of mobility. *Econometrica: Journal of the Econometric Society*, pages 1013–1024.