Intragenerational Income Mobility: a Nonparametric Estimate

Irene Brunetti - Davide Fiaschi *

February 27, 2013

Abstract

This paper analyses the level of intragenerational income mobility in Italy. We provide two novel methodologies to estimate income mobility based on nonparametric methods, and we apply it to the analysis of mobility of a sample of Italian individuals (between 16 and 65 years old) from the Survey on Household Income and Wealth (SHIW) by the Bank of Italy in the period 1987-2010.

First, a linear specification of the Markovian model is estimated removing the assumption of no serial correlation in the error term suggesting a low level of income mobility; second, a nonlinear specification of Markovian model is estimated providing both "local" and global measures of income mobility.

Income mobility appears to be low; in particular it reaches a minimum in the middle of income distribution and maximum values at the extreme bounds, with an income elasticity ranging from 0.4 to 0.8 in the relevant range of income (0.5-2). Moreover, from 1987-1998 to 2000-2010 income mobility has increased over time, in particular in the middle of distribution.

Keywords: Relative Income Mobility, Mobility Indexes, Markov Chain, Nonparametric Estimate.

JEL Classification Numbers: C14; J60; J62

---

*Irene Brunetti, Department of Economics and Management, University of Pisa, e-mail: ireneb@ec.unipi.it; Davide Fiaschi, Department of Economics and Management, University of Pisa, e-mail: dfiaschi@ec.unipi.it. We thank the Phd Whorkshop' participants at Scuola Superiore Sant'Anna in Pisa and the PhD committee for useful comments. Usual disclaimers apply.

1 Introduction

Intragenerational mobility deals with the individual's changes in social status (measured by income, earnings or occupation) over the lifetime or the work carrier. This paper focuses on income mobility which, as Fields (2006) discusses, has various features and different implications in terms of social welfare.

In literature there is not consensus on a precise definition/concept of income mobility but the relationship between actual and future income is an essential ingredient of its measurement. Indeed several ways of summarizing this relationship have been proposed.

Atkinson et al. (1992), Fields (2006), Fields and Ok (1996) contain a review of income mobility concepts and of their measures. It can be distinguished between mobility as: i) positional change, ii) income growth, iii) reduction of long-term inequality and iv) income risk.

In this paper, we refer to the concept of mobility as positional change (relative mobility). The idea is that mobility depends on the relative variations of individuals, that is, the definition of actual and future social conditions of an individual should consider the positions of everyone else in the society (Jenkins (2011)). Here, therefore, mobility depends not on whether individual income has increased or decreased over time, but on how his/her social condition has changed with respect to the average of (income) distribution. Thus, any equi-proportionate income negative variations of individual has not impact on mobility as positional change but can has a negative impact in terms of income growth, a positive impact in terms of reduction of long-term inequality, and a negative impact on mobility as income risk.

Taken mobility as positional change, *Perfect Mobility* occurs when the future income of each individual is independent of his/her actual income. In according to this definition, there will be infinitely perfectly mobile society and, in particular, there will be a *Perfectly Mobile Society* with *ex-post Minimum Inequality*, as we will discuss in Section 5, and a *Feasible Perfectly Mobile Society* where the stochastic process reflects the equilibrium (ergodic) distribution (see Prais (1955)).

We focus in particular on the quantitative measurement of the intragenerational income mobility in Italy.

We provide two novel methodologies to estimate income mobility based on non-parametric methods, and we apply it to the analysis of mobility of a sample of Italian individuals (between 16 and 65 years old) from the Survey on Household Income and Wealth (SHIW) by the Bank of Italy in the period 1987-2010. To our scope, individual disposable earnings (wage plus self-employment and business income) with respect to the sample average appear the most appropriate measure of relative income. In literature, on the assumption that the log of in-

come follows a linear Markovian model, the estimate of constant elasticity of (relative) income between different periods is the usual measure of income mobility considered in literature (see Atkinson et al. (1992)). However, we will show how is severely biased both by the presence of serial correlation in error term, and overall by the presence of nonlinearities. First, a linear specification of the Markovian model is estimated removing the assumption of no serial correlation suggesting a low level of income mobility; second, a nonlinear specification of Markovian model is estimated, providing both a local and synthetic measures of income mobility.
The local measure of income mobility consists in the estimate of the elasticity of (relative) income at period t conditioned to the level of (relative) income at period $t-1$, (LIE), by estimating the stochastic kernel of income dynamics in a continuous state space and the related conditioned mean (Quah (1997)). Synthetic measures of income mobility consist in indexes based on the estimate of stochastic kernel and related ergodic distribution largely inspired by Shorrocks (1978a) and Bartholomew (1973); at the same time they also provide a complementary estimate of the "local" income mobility (LIMI), i.e. income mobility for different ranges of income.
Income mobility reaches a minimum in the middle of income distribution and maximum values at the extreme bounds, with an income elasticity ranging from 0.4 to 0.8 in the relevant range of income (0.5-2). The estimate of local component of the synthetic mobility indexes confirms these results. Overall income mobility in Italy appears to be low with respect other developed countries (e.g. in U.S. is estimated equal to about 0.4 in terms of income elasticity, see Altonji and Dunn (1991)). We also analyse the different dynamics of income mobility into two sub-periods: 1987-1998 and 2000-2010. Income mobility has increased over time, in particular in the middle of distribution.
The paper is outlined as follows. In Section 2 we explains the standard methodology used to estimate mobility measures and its drawbacks. Given these limits, in Section 3 we introduce an alternative methodology to the study of mobility. Section 5 discusses the concept of *Perfect Mobility* and its welfare implications, while the empirical application is presented in Section 6. Finally, Section 7 contains some concluding remarks.

2 The Methodology

This section discusses the methodological issues in the measurement of income mobility. Firstly we critically review the standard approach, and then we propose two new methodologies based on non-parametric methods.

2.1 Standard Approach to the Measurement of Income Mobility

In literature the standard (Markovian) model describing income dynamics of individual i at period t is given by:

$$w_{i,t} = \beta w_{i,t-1} + \eta_{i,t}, \tag{1}$$

where w_{it} and w_{it-1} are the (logarithm of) relative income y_{it} (normalized with respect to sample average of period) (see Atkinson et al. (1992)). The following assumptions on the stochastic term η_{it} guarantee an unbiased estimate of the coefficient β:

$$\eta_{it} \sim \mathcal{N}\left(-\frac{\sigma_\eta^2}{2\left(1+\beta\right)}, \sigma_\eta^2\right); \tag{2}$$

$$cov\left(\eta_{it}, \eta_{it+s}\right) = 0 \;\; with \;\; s \neq 0; \tag{3}$$

$$cov\left(\eta_{it}, \eta_{jt}\right) = 0 \;\; with \;\; j \neq i; \;\; \text{and} \tag{4}$$

$$cov\left(w_{it-1}, \eta_{it}\right) = 0. \tag{5}$$

Assumption (2) implies an exogenous variability and independent of the income level. The negative expected mean of the stochastic term derives from the constraints that $E\left[e^{w_{it}}\right] = E\left[y_{it}\right] = 1$.

Moreover, under Assumptions (2)-(5) and $\beta \in (-1, 1)$ Central Limit Theorem applies, i.e.:

$$w_{it} \sim \mathcal{N}\left(-\frac{\sigma_\eta^2}{2\left(1-\beta^2\right)}, \frac{\sigma_\eta^2}{1-\beta^2}\right), \tag{6}$$

and therefore

$$y_{it} \sim \ln\mathcal{N}\left(-\frac{\sigma_\eta^2}{2\left(1-\beta^2\right)}, \frac{\sigma_\eta^2}{1-\beta^2}\right). \tag{7}$$

Assumptions (3) and (4) implies that the stochastic term is $i.i.d$ over time and across individuals. Assumption (5) implies that there is no any omitted variable.

In order to have a meaningful model of income dynamics β should be lower than 1[1].

From the estimate of the Markovian Model (1) the literature proposes two measures of income mobility (see Boeri and Brandolini (2005) and Pisano and Tedeschi (2008)):

[1] Indeed, from Eq. (1) $w_{i,t} = \beta^t w_{i,0} + \sum_{j=0}^{t} \beta^j \eta_{i,t-j}$, hence, $\lim_{t\to\infty} \beta^t w_{i,0} = 0$ for $\beta \in (-1, 1)$, and given a sequence of random independently distributed variables, as $t \to \infty$ proves that the Central Limit Theorem is applicable to $w_{i,t}$.

- $\hat{\beta}$: a high value of the estimated elasticity of current income to past income, i.e. $\hat{\beta}$, implies a low level of income mobility;

- $\hat{\rho}_{t,t-1} = \hat{\beta}\hat{\sigma}_{w(t-1)}/\hat{\sigma}_{w(t)}$: a high value of the estimated serial correlation of w_i implies a low value of income mobility.

It is worth nothing that the correlation coefficient, $\hat{\rho}$, is proportional to $\hat{\beta}$ but it is inversely related to the income variance.
The standard deviation of income, $\sigma_{w(t)}$ is a measure of income inequality; this suggests an inverse relationship between mobility and inequality: a higher value of $\sigma_{w(t)}$ (given $\sigma_{w(t-1)}$) means an increase of inequality but, also a decrease in $\hat{\rho}$, i.e. an increase in mobility.

2.2 Serial Correlation in the Error Term

The Markovian model in Eq. (1) is crucially based on the assumption that the stochastic term is uncorrelated over time. However, individuals are able to move through the income distribution in a quite systematic way, or incomes improvements may depend crucially on previous success. To evaluate the bias in the estimates due to the presence of serial correlation assume that the stochastic term $\eta_{i,t}$ in Eq. (1) follows the first-order auto-regressive process:

$$\eta_{it} = \phi\eta_{it-1} + \varepsilon_{it}, \tag{8}$$

where $\phi \in (-1, 1)$ is assumed to be the same for all individuals, and ε_{it} is $i.i.d$ with variance σ_{ε}^2. $\phi > 0$ means that success breeds success, $\phi < 0$ means that success in one period tends to be followed by a reverse in the next.
If ϕ are serially correlated, the estimate of β in Eq. (1) by OLS is not consistent.
Given Eq. (8), the Cochrane-Orcutt procedure (see Creedy (1974)) allows to adjust the estimate of linear model for the serial correlation; in particular the model (1) is transformed into a model where OLS leads to unbiased estimate of β. The first step is to take Eq. (1) at the period $t-1$:

$$w_{it-1} = \beta w_{it-2} + \eta_{it-1}, \tag{9}$$

Multiplying Eq. (9) for ϕ and subtracting from Eq. (1), we obtain:

$$w_{it} - \phi w_{it-1} = \beta w_{it-1} - \phi\beta w_{it-2} + \varepsilon_{it}, \tag{10}$$

where $\eta_{it} - \phi\eta_{it-1} = \varepsilon_{it}$ (see Eq. (8)). Eq. (10) can be written as:

$$w_{it} = a w_{it-1} - b w_{it-2} + \varepsilon_{it}, \tag{11}$$

where a is equal to $\phi + \beta$ and b is equal to $\phi\beta$.
Eq. (11) can be consistently estimated by OLS. Given an estimate of a and b, $\hat{a} = \phi + \beta$ and $\hat{b} = \phi\beta$, Creedy (1974) shows that β and ϕ are the positive roots of the following equation:

$$x^2 - ax + b = 0 \tag{12}$$

Applying this procedure to a sample 12,999 observations (the sample is limited by the necessity to have three wave transitions, i.e. 4-years lag):

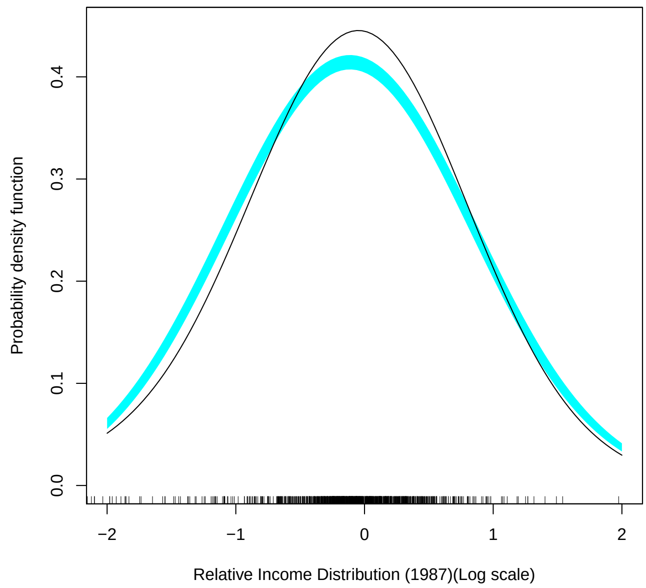
$$w_{it} = \underset{(0.008)}{0.458} w_{it-1} - \underset{(0.007)}{0.235} w_{it-2} \tag{13}$$

Given these estimates, β and ϕ are respectively equal to 0.76 and 0.30. The estimate of Eq. (1) provided a value of $\hat{\beta}$ equals to 0.57; the resulting bias is therefore equal to of 0.19.

2.2.1 The Drawbacks of the Standard Methodology

The standard approach to measure income mobility presents two drawbacks.
First, the model involves that the log of relative incomes are normally distributed, but this not hold in the Italian data. Figure (1) shows that the income data of our sample in 1987 are not normally distributed. The blue curve represents the confidence bands calculated by bootstraps under the null hypothesis of normally distributed observations, while the black curve is the estimated distribution of the log relative income in 1987. The estimate is largely outside the confidence bands for wide ranges of income and the Jarque-Bera test of the hypothesis of normal distribution is rejected at 5% significance level.

Figure 1: **The Distribution of the Log of Relative Income of Individuals in 1987. Source: *SHIW*.**

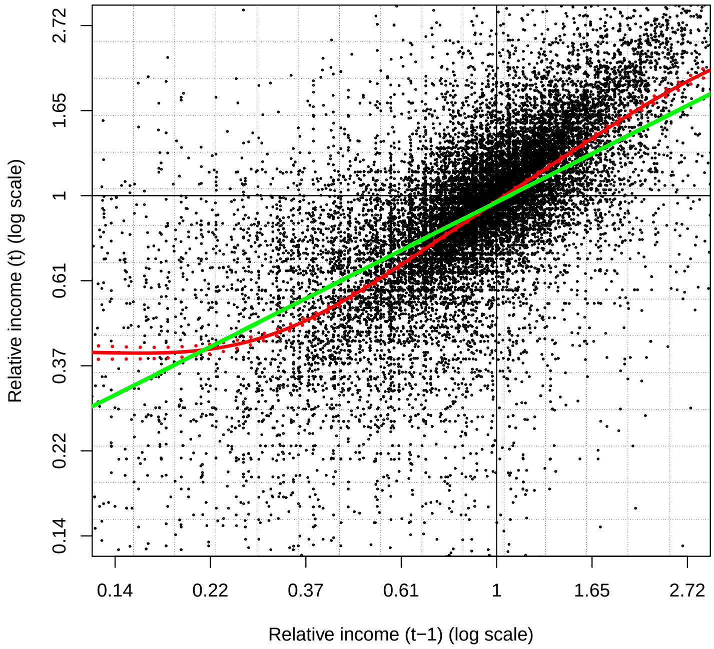


The second drawback is the implicit assumption that mobility is independent of the level of (relative) income. Figure (2) reports the estimates both of the linear model against the estimate of the nonlinear Markov model reported in Eq. (14) below[2].

Figure (2) highlights statistically significant difference between the estimate of linear and non-linear Markov model (see Section 6.2 for more details).

[2]Nonlinear model is estimated by the (*mgcv*) routine (see Wood (2011) for more details).

Figure 2: **The Estimate of Markov Models for the period 1987-2010. Confidence bands at 5% significance levels for non-parametrics estimate are reported by red dotted lines. Source: *SHIW*.**



3 A New Methodology based on the Stochastic Kernel

Given the drawbacks of the standard methodology, we consider the alternative nonlinear Markov model:

$$w_{i,t} = \beta\left(w_{i,t-1}\right) w_{i,t-1} + \eta_{i,t} \tag{14}$$

where β is assumed to be a function of the income level at period $t-1$. The estimate of $\beta\left(w_{i,t-1}\right)$ should be still in the range [0,1) $\forall t$ in order to maintain a meaningful model of income dynamics. In fact, from Eq. (14) $w_{i,t} = \prod_{j=0}^{t} \beta\left(w_{i,j}\right) w_{i,0} + \sum_{j=0}^{t} \prod_{q=0}^{j} \beta\left(w_{i,q}\right) \eta_{i,j}$. Therefore $\lim_{t\to\infty} \prod_{j=0}^{t} \beta\left(w_{i,j}\right) w_{i,0} = 0$ for $\beta\left(w_{i,j}\right) \in (-1, 1)$. Given the sequence of random dependently distributed variables, the Bernstein's conditions guarantees the applicability of the law of large numbers (see Gnedenko (1978)), i.e. the first two moments of the distribution are finite. However the Central Limit Theorem cannot be applied, i.e. the limiting distribution of w_i is not normal in general. Therefore Eq. (14) allows for not normal equilibrium distribution of income.

The estimate of Eq. (14) is obtained by estimating the stochastic kernel in a continuous state space and the related conditioned mean. The stochastic kernel is the conditional distribution of w_t given w_{t-1}[3]:

$$g\left(w_t|w_{t-1}\right) = \frac{f_{w_t,w_{t-1}}\left(w_t, w_{t-1}\right)}{f_{w_{t-1}}\left(w_{t-1}\right)} \tag{15}$$

The estimate of the conditioned mean leads to a nonparametric estimated of the Markovian Model (14):

$$E\left[\hat{g}\left(w_t|w_{t-1}\right)\right] = \hat{\beta}\left(w_{it-1}\right) w_{it-1}. \tag{16}$$

In the nonparametric estimation of the Markovian Model, the problem of serial correlation can be easily settled. Bowman and Azzalini (1997) explain that the additive nature of the kernel estimator makes the correlation between w_i, w_j irrelevant[4]. Therefore, the expectation of the kernel estimator is exactly the same as for independent data. This result is common to other estimators, for instance when the sample mean is used to estimate the population mean for dependent data.
Our methodology provides two classes of measures of income mobility:

- *Local Measures* and
- *Measures of Mobility by Synthetic Indexes.*

4 Mobility Measures

4.1 Local Indexes of Income Mobility

Model (14) admits that income mobility may change with the level of income; the presence of non linearities suggests to use a *local index* of income mobility defined as:

$$LIE = \frac{dw_{it}}{dw_{it-1}} = \hat{\beta}'\left(w_{it-1}\right) w_{it} + \hat{\beta}\left(w_{it-1}\right), \tag{17}$$

i.e. to use a measure of Local Income Elasticity (LIE).
The relationship between LIE and income crucially depends on the behaviour of β'. If $\hat{\beta}''$ is positive there will be always a positive relationship between LIE and w_{it}; if $\hat{\beta}''$ is negative is instead a *necessary* condition to observe a negative relationship.

[3] To estimate the stochastic kernel we follow the methodology proposed by Silverman (1986) known as *adaptive kernel*. In appendix A there is a brief description of this procedure.

[4] The general form of the kernel estimator is: $\hat{f}(y) = \frac{1}{n}\sum_{i=1}^{n}\omega(y - y_i; h)$ where ω is itself a probability density, called kernel function, whose variance is controlled by the parameter h.

4.2 Synthetic Mobility Indexes

The second class of mobility indexes is represented by three indexes generally used in literature to measure mobility. A higher value of these indexes means higher income mobility.
The first index was inspired by Shorrocks (1978b). In general the Shorrocks index with continuous state space appears as:

$$I_S = 1 - \int_{\underline{w}}^{\bar{w}} \omega(q)\, g(q|q)\, dq; \tag{18}$$

where $\omega(q)$ represents a weighting function that can assume different specifications[5].
The closest counterpart to the original Shorrocks index in discrete state space is:

$$I_S^U = 1 - \int_{\underline{w}}^{\bar{w}} U(q)\, g(q|q)\, dq; \tag{19}$$

where $U(q)$ is the uniform distribution. Using this distribution it implicitly assumes that there are no differences between classes.
Alternatively $\omega(q)$ can be represented by the equilibrium distribution $\pi(q)$, that is:

$$I_S^E = 1 - \int_{\underline{w}}^{\bar{w}} \pi_w(q)\, g(q|q)\, dq; \tag{20}$$

According to this specification, transition probabilities are measured in the long-run. In our analysis we assume that the weighting function is equal to the marginal density of the actual distribution. The Shorrocks index is the following:

$$I_S^E = 1 - \int_{\underline{w}}^{\bar{w}} f_w(q)\, g(q|q)\, dq; \tag{21}$$

I_S is in the range $[0, 1]$ and it measures the level of persistence since it considers only the elements on the main diagonal (represented by $g(q|q)$).
Bartholomew (1973) proposed another index that takes into account the transition outside of the main diagonal, known as the Bartholomew index. It can be computed as follow:

$$I_B^\alpha = \int_{\underline{w}}^{\bar{w}} \pi_w(q) \int_{\underline{w}}^{\bar{w}} g(s|q)\, \omega(s, q, \underline{w}, \bar{w}, \alpha)\, ds dq, \tag{22}$$

where $\pi_w(q)$ is the ergodic distribution of w and $\omega(s, q, \underline{w}, \bar{w}, \alpha)$ is a weighting function. I_B^α is in $[0, 1]$.

In particular the weighting function is :

$$\omega(s, q, \underline{w}, \bar{w}, \alpha) = |q - s|^\alpha A(\underline{w}, \bar{w}, \alpha, q) \tag{23}$$

[5] See Schluter and Van de gaer (2003).

where $A(\underline{w}, \bar{w}, \alpha, q)$ is a constant such that $\int_{\underline{w}}^{\bar{w}} |q-s|^{\alpha} A(\underline{w}, \bar{w}, \alpha, q)\, ds = 1$. Therefore, $A(\underline{w}, \bar{w}, \alpha, q) = \frac{1}{\int_{\underline{w}}^{\bar{w}} |q-s|^{\alpha}}$. If α, a parameter higher than zero, is equal to 2, Bartholomew index weights the transition probabilities more than proportionally with respect to the length of jumps between income levels.

Finally, we present a modified version of the Bartholomew index, known as Fiaschi-Lavezzi index:

$$I_{FL} = \int_{\underline{w}}^{\bar{w}} f_w(q) \int_{\underline{w}}^{\bar{w}} g(s|q)\, w(s, q, \underline{w}, \bar{w}, \alpha)\, dsdq \tag{24}$$

The use of the marginal distribution of w instead of the ergodic distribution responds to the fact that in income contexts the ergodic distribution could not provide a faithful picture of the ultimate consequences of the current income distribution because intra-distribution patterns do not remain unchanged (Maza et al. (2010)). Also this index varies between 0 and 1.

The last two indexes contain a *Local Income Mobility Index* (LIMI), i.e.:

$$LIMI = \int_{\underline{w}}^{\bar{w}} g(s|q)\, \omega(s, q, \underline{w}, \bar{w}, \alpha)\, ds \tag{25}$$

This local mobility index should not be confused with the other local indexes since, in this case, the transition probabilities are weighted.

5 Perfect Mobility and Welfare Implications

In Section 1 we have introduced the concept of *Perfect Mobility*. In general, it occurs if future income of each individual is independent of his/her actual income. According to this definition, there will be infinitely many possible mobility processes, and so infinitely perfectly mobile society, but, following the approach of Prais (1955), we choose the mobility process corresponding to the equilibrium distribution (*Feasible Perfect Mobile Society*).

Perfect Mobility doesn't imply *ex-post Minimum Inequality*, but it is one of the possibility. In this case, if incomes are measured as ratio with respect to the sample average, a *Perfect Mobile Society* shows a global convergence to the sample average (*Social Optimum Perfect Mobility*). According to this definition, mobility is important not because income movements are intrinsically valuable, but because it can help to attenuate the effects of disparities in initial endowments on future income prospects (Benabou and Ok (2001)).

From this view, mobility is considered as an equalizer of ex-ante opportunities (but not necessarily of outcomes). Future realised income distributions can be more unequal than the current one since, if this is due to shocks unpredictable on the basis of initial conditions, there is little

disparity of opportunity, that is, society appears fair.
The *Perfect Mobility* can be represented by the following transition matrix:

$$\mathbf{P_E} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{26}$$

where each row is equal to each other and the mobility process permits to reach the average of the distribution.
In this context, measures of pure persistence and other mobility indexes (e.g. I_S, I_B and I_{FL}) are related to the notion of equalizer of opportunities but don't directly correspond to it. In particular, movements in relative incomes may be equalizing or disequalizing, and mobility indexes, generally proposed in literature, fail to distinguish between the two. In evaluating mobility, in fact, it often considered that the identity process (or identity matrix) should correspond to the smallest element and be viewed as the worst scenario (see Shorrocks (1978b)). More generally, according to this "diagonals view", any increase in relative income movement (any shifts from diagonal to off-diagonal elements) should imply a higher level of mobility and a higher ranking the mobility ordering. Unfortunately, relative income movements can be disequalizing as well equalizing, and only the latter type count positively as mobility.
In this paper we propose a method to overcome this drawback. To understand if the existing mobility process is also equalizing we apply to mobility indexes, and in particular to I_B and I_{FL}, a structure of weights which gives higher weights to mobility towards sample average. The structure can be expressed as follow:

- if the actual income (q) is lower than the sample average (poor people) and $(q-s)<0$ $\Rightarrow \omega_{ij} = |q-s|^{\alpha}$[6];
- if the actual income (q) is lower than the sample average (poor people) and $(q-s)>0$ $\Rightarrow \omega_{ij} = -|q-s|^{\alpha}$;
- if the actual income (q) is higher than the sample average (reach people) and $(q-s)<0$ $\Rightarrow \omega_{ij} = -|q-s|^{\alpha}$;
- if the actual income (q) is higher than the sample average (reach people) and $(q-s)>0$ $\Rightarrow \omega_{ij} = |q-s|^{\alpha}$;

The Bartholomew index (I_B) shows another limit. Supposing that the mobility process is described by matrix 26, in this case I_B, weighting transition probabilities with the ergodic distribution, is equal to 0 suggesting that there is no mobility because all the mass is concentrated

[6] s represents the future income state.

on the central class. However, following the notion of mobility as equalizer of opportunities, the matrix 26 represents a society with *Perfect Mobility* and *ex-post Minimum Inequality*. Thus I_B leads to a wrong conclusion. This drawback can be solved using the uniform or the actual distribution to weights the transition probabilities instead of the ergodic one.

6 The Empirical Application

6.1 The Data

Data used in the analysis are drawn from the historical database of the Bank of Italy: "Survey on Household Income and Wealth" (SHIW). We study the changes in the individual relative income in the period 1987-2010. In this period we have 12 waves (1987, 1989, 1991, 1993, 1995, 1998, 2000, 2002, 2004, 2006, 2008, 2010). We consider all individuals that remain in the sample at least for two consecutive waves, male and females, aged from 16 up to 65 with positive income, obtaining a sample of 13,090 individuals.

The variable used in the analysis is the logarithm of relative income of each individual, defined as the ratio between the individual income and the sample average of the distribution. In particular we consider the net income including income both from wages and self-employment/business[7]. The transitions are defined as the movements between two consecutive waves, i.e. we consider 2-years lag. Given this lag we have 25,858 transitions[8].

Table 1 reports descriptive statistics about the number of observations, the mean, the median of the individual income and the Gini index, both for our individual sample and for the total sample provided by the Bank of Italy (values in brackets). There are no relevant differences between the two samples.

[7] SHIW income code: YL (wage) and YM (self-employment/business income).

[8] Each sample unit is assigned a weight to take into account the probability of inclusion in the sample and, only for the panel section of the survey, the correction for the attrition. The variable used is "PESOFL2", obtained by multiplying PESOFL. Weights obtained by raking for alignment with the distributions derived from socio-demographic and labor force statistics from ISTAT by a constant (different for each survey) providing the estimate of the totals for the universe (Italian resident population).
In the analysis we use the historical database that includes sampling weights slightly different from those of the annual waves.

Table 1: **Descriptive Statistics**

Year	*N.Obs*	*Mean*	*Median*	*Gini*
1987	1,010 (9,034)	17,068 (18,180)	15,406 (16,160)	0.27 (0.30)
1989	2,503 (9,249)	18,278 (18,282)	17,332 (16,408)	0.23 (0.26)
1991	3,309 (6,670)	17,490 (16,851)	17,056 (16,203)	0.20 (0.21)
1993	3,818 (7,985)	17,565 (16,946)	15,994 (15,528)	0.30 (0.31)
1995	3,555 (8,085)	16,666 (15,844)	15,600 (14,181)	0.32 (0.32)
1998	3,911 (7,276)	17,221 (16,998)	15,815 (15,816)	0.31 (0.32)
2000	3,918 (7,845)	17,740 (17,390)	15,813 (15,813)	0.29 (0.30)
2002	3,704 (7,397)	18,087 (17,944)	16,303 (15,604)	0.31 (0.32)
2004	3,763 (7,265)	19,379 (19,145)	16,716 (16,270)	0.32 (0.34)
2006	3,879 (7,077)	20,237 (19,875)	17,188 (16,651)	0.32 (0.33)
2008	4,149 (7,048)	18,770 (18,304)	16,574 (16,369)	0.28 (0.29)
2010	2,834 (6,847)	18,215 (18,085)	16,380 (16,500)	0.28 (0.30)

6.2 Income Mobility during 1987-2010

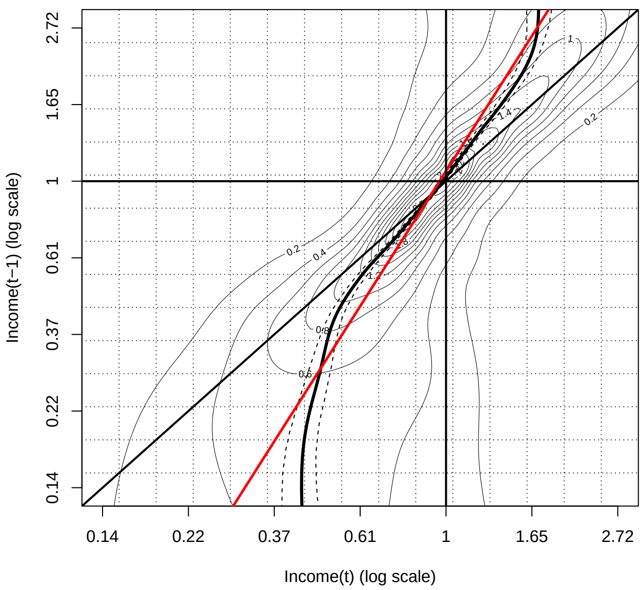
Figure 3 reports the estimate of the stochastic kernel for the relative incomes of Italian individuals during the period 1987-2010.

The red line represents the unbiased estimate of linear model, while the black curve is the conditional mean, i.e. the expected income at wave t conditional to income of wave $t-1$[9]. This stochastic process governing the income distribution appears to be strongly non linear but with just one equilibrium in 1. Indeed the black curve crosses the bisector from below in one point, around 1, leading to an actual and equilibrium (ergodic) distribution with one peak. The fact that the average, both for high and low level of income at time $t-1$, is far above the bisector

[9] The estimate of confidence bands for the conditional mean is made by the bootstrap procedure (1,000 replications) with 5,000 transitions randomly drawn from the original sample (this is for the huge computational burden) (see Efron and Tibshirani (1993)).

suggests that there is convergence towards the mean value of income in the considered period. The vertical line represents the situation with *Perfect Mobility and ex-post Minimum Inequality*. The horizontal distance between this line and the estimate of the stochastic kernel shows that we are far from a *Perfect Mobility* situation, in particular at the extremes of the distribution.

Figure 3: **Estimated Stochastic Kernel of Relative Income for the period 1987-2010. Confidence bands at 5% significance levels for non-parametrics estimate are reported by black dashed lines.**



To compare our analysis with others studies we estimate the linear model. The estimate of the elasticity reported in Table 2 shows that there is a low level of income mobility. This finding is confirmed by $\hat{\rho}$.

Pisano and Tedeschi (2008) corroborate this result. They measure the level of earnings mobility using $\hat{\rho}$ for two periods finding that, for the first period (1995-1998), $\hat{\rho}$ is equal to 0.47, whereas, for the second one (2004-2006), it is equal to 0.60 suggesting a decrease of income mobility.

Table 2: **Estimates of the Elasticity and Correlation Coefficient for the period 1987-2010.**

Indexes	1987 – 2010
$\hat{\beta}_{biased}$	0.567 (0.005)
$\hat{\rho}_{biased}$	0.597 (0.006)
$\hat{\beta}_{unbiased}$	0.766 (0.008)
$\hat{\rho}_{unbiased}$	0.626 (0.007)

Standard errors are reported in parenthesis

Moreover, Table 2 highlights that the assumption of no serial correlation in error term leads to an overestimate of mobility of mobility. In fact, removing it, $\hat{\beta}$ and $\hat{\rho}$ increase showing a lower level of income mobility.
The Figure 4 reports the biased and unbiased estimate of the linear model and the estimate of the Local Income Elasticity for the period 1987-2010. The Figure 4 shows that observations in the range [0.61-1] displays a lower level of mobility than those in the tails of the distribution.

Figure 4: **Local Income Elasticity for the period 1987-2010. Confidence bands at 5% significance levels for the estimates are reported by red dashed lines**

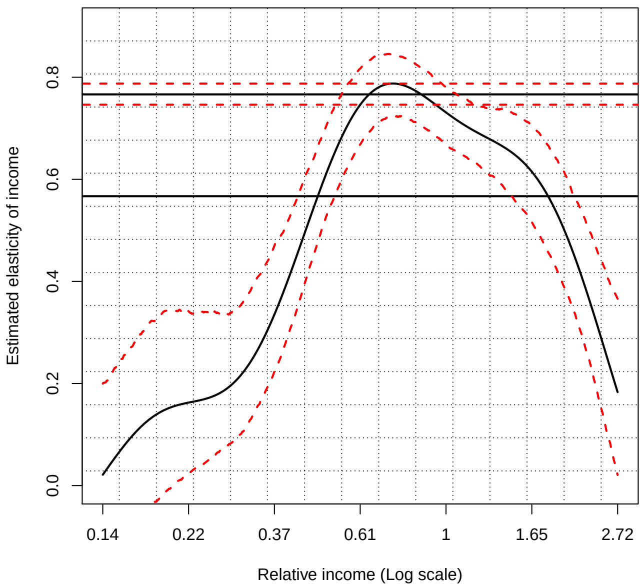


Table 3 reports the value of synthetic mobility indexes described in Section 4. I_S is close to 1 while the other indexes are lower. However, the low value of the last three indexes doesn't means that the society is far to be *Perfectly Mobile*, in the sense described by Prais (1955). To establish whether the society shows *Perfect Mobility* first, we construct the perfect transition matrix where the probability of entering a particular class is independent of the class of the previous period and where its elements are equal to those of the ergodic distribution and then, we apply the mobility indexes to it.

Table 3: **Synthetic Mobility Indexes for the period 1987-2010.**

Index\Period	$1987-2010$
I_S	0.958 (0.0009)
$I_B(\alpha=1)$	0.267 (0.0012)
$I_B(\alpha=2)$	0.248 (0.0014)
I_{FL}	0.266 (0.011)
Num.Obs	25858

Standard errors are reported in parenthesis

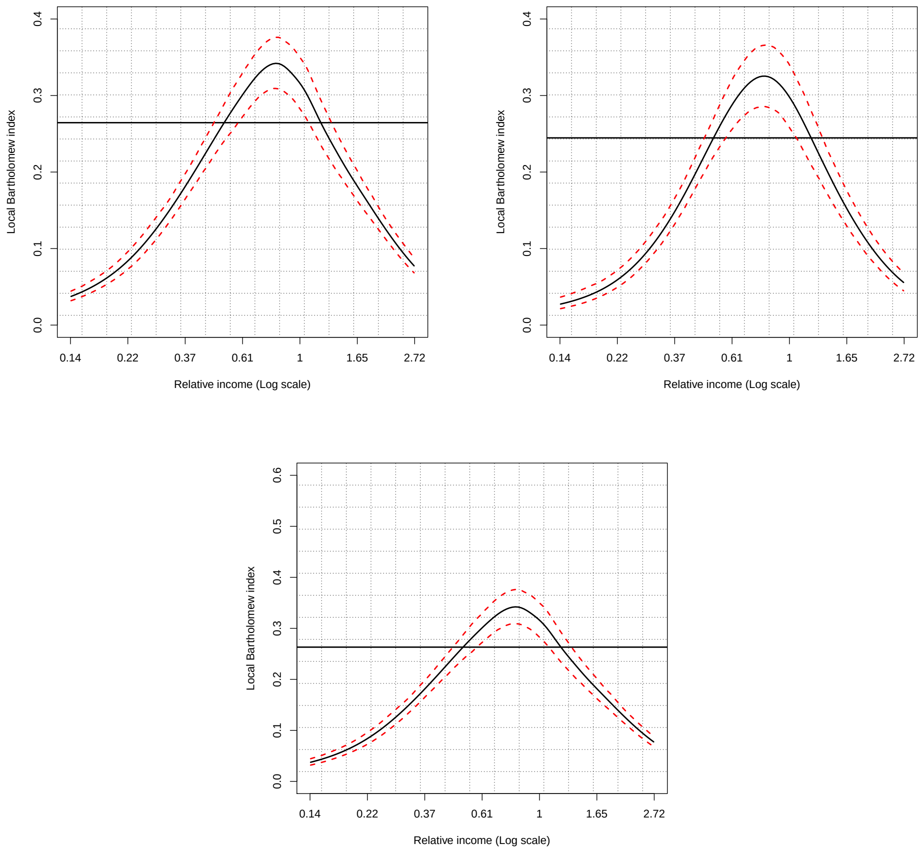
First we compute the mobility indexes on the perfect mobility transition matrix and then we compare them with mobility indexes calculated on the actual transition matrix. Table 4 reports the ratio between the two types of mobility indexes. With the exception of I_S, the level of income mobility measured is about 0.40% of a *Perfect Mobile Society*.

Table 4: **"Distance" from a *Perfect Mobile Society* for the period 1987-2010.**

Index\Period	$1987-2010$
I_S	0.96
I_B	0.38
I_{BM}	0.38
I_{FL}	0.35

The magnitude of the last three indexes is important, but we are interesting also into evaluate the level of mobility along the distribution using their local component. Figure 5 shows that, at low income level, mobility is low, then it starts to increase as income increases and, after a threshold, mobility comes back to decrease. Looking at the dynamics of the indexes's local component, those individuals that are in the middle part of the distribution display a higher level of mobility than those that are at the extremes of the distribution.

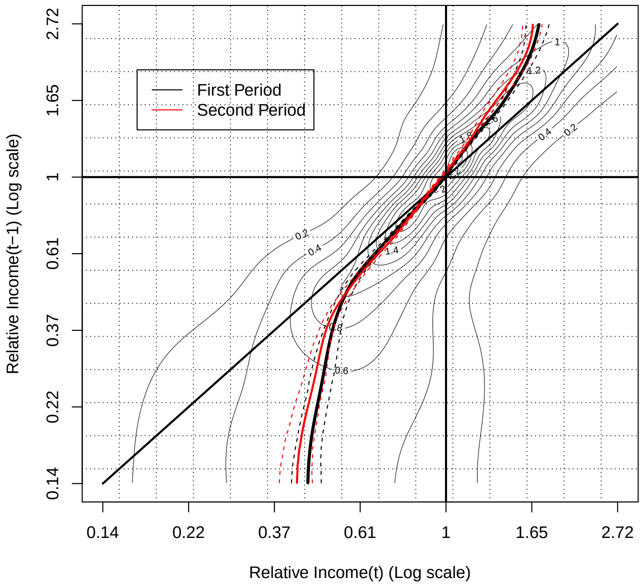
Figure 5: **Local Income Mobility Index for the period 1987-2010. Confidence bands at 5% significance levels for the estimates are reported by red dashed lines.**



6.3 Income Mobility during 1987-1998 and 2000-2010.

To control the dynamics both of the stochastic kernel and of the mobility indexes we divide the whole period into two sub-periods: 1987-1998 and 2000-2010. Figure 6 shows the estimate of the stochastic kernel in the sub-periods. In the second period the richest part of the distribution shows a higher level of mobility than in the first, they seem to be more close to the perfect mobility situation. The opposite occurs for the poorest part of the distribution. However, these changes are just slight. For the middle class, the level of mobility doesn't change.

Figure 6: **Estimated Stochastic Kernel of Relative Income for the two periods. Confidence bands at 5% significance levels for nonparametrics estimates are reported by black and red dashed lines.**



Looking at the estimates of the local income elasticities for both periods, the Figure 7 shows a shift downwards and to the left.

Therefore, from the first to the second period, mobility decreases for the the poorest individuals (with relative income in the range [0.14-0.61]), increases for the middle class (with relative income in the range [0.61-1.65]), and doesn't change for the richest individuals (with relative income higher than 1.65)[10].

[10]The statistical significance is tested by the bootstrap procedure.

Figure 7: **Local Income Elasticity for the two periods. Confidence bands at 5% significance levels for the estimates are reported by black and red dashed lines**

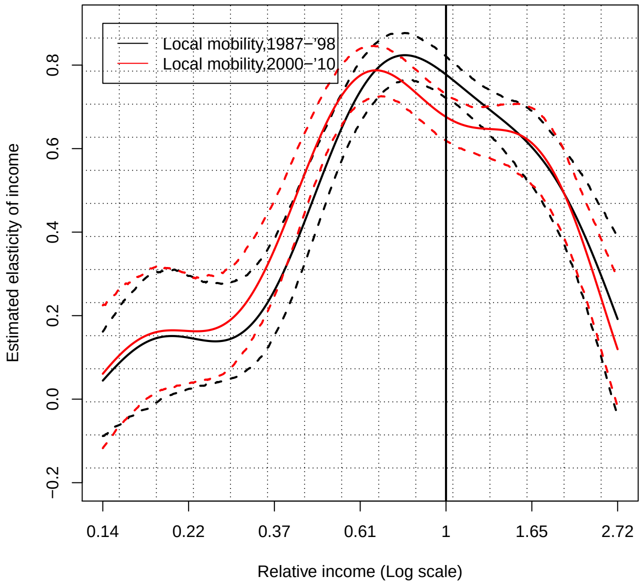


Table 5 reports the estimate of two classes of mobility indexes for both periods. The Table displays that the unbiased estimate of β and ρ decreases over time suggesting an increase in income mobility[11]. The estimate of the synthetic indexes proves the same result. Indeed I_B and I_{BM} slightly increase showing a rise in the income mobility, while I_S and I_{FL} doesn't change. From a statistical point of view, I_B and I_{BM} are statistically different[12].
The Table 6 highlights that, from the first period to the second one, the "distance" from a *Perfect Mobile Society* decreases.

[11]We can reject the null hypothesis of equality for the two indexes at the usual confidence level of 5%.
[12]We can reject the null hypothesis of equality at the usual confidence level of 5%.

Table 5: **Synthetic Mobility Indexes for the two periods.**

Index\Periods	$1987-1998$	$2000-2010$
$\hat{\beta}_{unbiased}$	0.81 (0.008)	**0.75** * (**0.006**)
$\hat{\rho}_{unbiased}$	0.77 (0.009)	**0.71** * (**0.007**)
I_S	0.96 (0.007)	0.96 (0.0006)
$I_B(\alpha = 1)$	0.25 (0.009)	**0.26** * (**0.010**)
$I_B(\alpha = 2)$	0.23 (0.001)	**0.24** * (**0.012**)
I_{FL}	0.26 (0.009)	0.26 (0.009)
Num.Obs	9,624	13,472

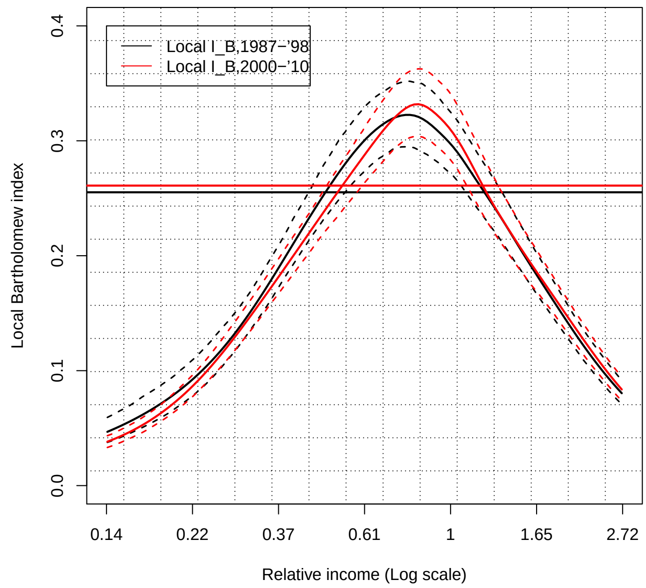
Standard errors are reported in parenthesis

Table 6: **"Distance" from a *Perfect Mobile Society* for the two periods.**

Index\Period	$1987-1998$	$2000-2010$
I_S	0.97	0.97
I_B	0.36	**0.38***
I_{BM}	0.33	**0.34***
I_{FL}	0.37	0.37

Finally, Figure 8 reports the local component of I_B computed for the two periods. Individuals with a relative income in the range [0.35-0.65] show a decrease of mobility, whereas those individuals in the range [0.65-1.5] display an increase of mobility, as suggested by the Figure 7.

Figure 8: **Local Income Mobility Index for the two periods. Confidence bands at 5% significance levels for the estimates are reported by black and red dashed lines.**



7 Concluding Remarks

In this paper we have studied intragenerational income mobility of a sample of italians individuals during the period 1987-2010. First of all, we have found that, making the strong assumption of no serial correlation in error term, the linear Markovian model provides a biased estimate of the level of mobility. In particular, it leads to an overestimate.

Secondly, we have proposed two new methodologies to study income mobility because of the presence of strong nonlinearities in the estimates. We have introduced two different types of mobility measures based on the estimate of the nonlinear Markovian model: a local measure given by the estimate of *Local Income Elasticity* (LIE) and measures given by the estimate of synthetic mobility indexes.

The local mobility measure provides an estimate of the income mobility for each level of income. It is obtained by estimating the stochastic kernel in a continuous state space and the related conditioned mean.

The second type of mobility measures consist in three synthetic indexes (I_S, I_B and I_{FL}) that are computed starting from the estimate of the stochastic kernel and that allow to measure income mobility at aggregate level. The last two synthetic indexes contain a local measure of income mobility (LIMI).

The analysis shows that Italian individuals are characterized by a low level of intragenerational mobility with respect to other countries (for example in U.S.A. $\hat{\beta}$ is around 0.4), and that the income mobility is very high at the extreme bounds of distribution, but low in the middle. Moreover, income mobility is increased over time only for the middle part of the distribution.

A Appendix A

A.1 Adaptive Kernel Estimates

The kernel density estimator can be considered as a viewing window that slides over the data and the estimate of the density depends on the number of observations that fall into the window (Pittau and Zelli (2002)).
When observations are scattered over the support of the distribution, we ca not use a fixed bandwidth in density estimation since we want to estimate long-tailed or multi-modal density function. The fixed bandwidth approach may result in under-smoothing in areas with only sparse observations while over-smoothing in others. The adaptive kernel estimation is a two-stage procedure which mitigates this problem (see Silverman (1986), p. 101).

The general strategy used is the following: given a multivariate data set $X = X_1, ..., X_n$ and a vector of sample weights $W = \omega_1, ..., \omega_n$, where X_i is a vector of dimension d and $\sum_{i=1}^{n} \omega_i = 1$, first we have to run the pilot estimate:

$$\tilde{f}(x) = \frac{1}{n \det(H)} \sum_{i=1}^{n} \omega_i k\{H^{-1}(x - X_i)\} \tag{27}$$

where $k(u) = (2\pi)^{-1} exp(-1/2u)$ is Gaussian kernel. The estimate of the density function at each point is determined directly from the sample data, without assuming any functional form a priori. The restriction on the kernel function K(-) is to be nonnegative and integrated to 1 over its support. Any probability density function satisfies this condition and, as a general rule, the kernel estimates do not depend much on the kernel chosen. For large samples, any kernel function will be close to an optimal one, thus the choice of kernel is a minor issue BW (1986).

The bandwidth matrix H is a diagonal matrix $(d \times d)$ with diagonal elements $(h_1, ..., h_d)$ given by the optimal normal bandwidths, i.e. $h_i = [4/(d+2)]^{1/(d+4)} \hat{\sigma}_i n^{1/(d+4)}$; $\hat{\sigma}$ is the estimated standard error of the distribution of X_i. The use of a diagonal bandwidth matrix instead of a full covariance matrix follows the suggestions in Wand and Jones (1993). In the case of $d = 2$ we have $H = \det(H) = (1)^{1/6} n{-}1/6\hat{\sigma}$. In the mobility estimate $W = \{p_i, ..., p_n\}$, where p_i is the weight associated to each individual i. We then define local bandwidth factors λ_i by:

$$\lambda_i = [\tilde{f}(X_i)/g]^{\alpha} \tag{28}$$

where $\log(g) = \sum_{i=1}^{n} \omega_i \log(\tilde{f}(X_i))$ and $\alpha \in [0, 1]$ is a sensitivity parameter. We set $\alpha = 1/2$ as suggested by Silverman (1986), 103. Finally the adaptive kernel estimate $\hat{f}(x)$ is defined as:

$$\hat{f}(x) = \frac{1}{n\det(H)} \sum_{i=1}^{n} \lambda_i^{-d} \omega_i k\{\lambda_i^{-1} H^{-1}(x - X_i)\} \tag{29}$$

The Gaussian kernel guarantees that the number of modes is a decreasing function of the bandwidth; such a property is at the basis of the test for unimodality (see Silverman (1986)).

A.2 Derivation of the Linear Model

To derive Eq.(1) start with the following equation:

$$\log(\tilde{y}_{it}) = \beta \log(\tilde{y}_{it-1}) + \eta_t \tag{30}$$

where $\tilde{y}_{it}$ is equal to $y_{it}/\bar{y}_t$ and $\tilde{y}_{it-1}$ is equal to $y_{it-1}/\bar{y}_{t-1}$. Replacing into Eq.(30):

$$\log(y_{it}) - \log(\bar{y}_t) = \beta(\log(y_{it-1}) - \log(\bar{y}_{t-1})) + \eta_t = \beta \log(y_{it-1}) + \log(\bar{y}_t/\bar{y}_{t-1}^{\beta}) + \eta_t \tag{31}$$

Assuming that $\bar{y}_t = (1+g)\bar{y}_{t-1}$, where g is a growth rate, we obtain:

$$\log(y_{it}) = [\beta \log(y_{it-1})] + \log[(1+g)\bar{y}_{t-1}] - [\beta \log((1+g)\bar{y}_{t-1})] + \eta_t, \tag{32}$$

Rewriting:

$$\log(y_{it}) = \beta \log(y_{it-1}) + \log(1+g) + [(1-\beta)\log(\bar{y}_{t-1})] + \eta_t, \tag{33}$$

We assume that $\alpha = \log(1+g) + [(1-\beta)\log(\bar{y}_{t-1})]$ where the first term represents the deterministic trend which increases earnings at time t thanks to common growth whereas, the second term represents an omitted variable.
Finally we obtain the following equation:

$$\log(y_{it}) = \alpha + \beta \log(y_{it-1}) + \eta_t \tag{34}$$

To verify the stationarity of this model, we estimate it and we find a $\hat{\beta} < 1$ suggesting that the model converges toward the steady state.

References

Altonji, J. and T. Dunn (1991). "Relationship among the Family Incomes and Labor Market Outcomes of Relatives". *Research in Labor Economics 2*.

Atkinson, A. B., F. Bourguignon, and C. Morrison (1992). *Empirical Studies of Earnings Mobility.* Harwood Academic Publishers.

Bartholomew, D. (1973). *Stochastic Models for Social Processes.* New York: John Wiley & Son.

Benabou, R. and E. Ok (2001). "Mobility as Progressivity: Ranking Income Processes According to Equality of Oppurtunity". *NBER Working Paper Series*.

Boeri, T. and A. Brandolini (2005). "The Age of Discontent: Italian Household at the Beginning of the Decade". *IZA Discussion Paper* (1530).

Bowman, A. and A. Azzalini (1997). *Applied Smoothing Tecniques for Data Analysis.* Clarendon Press, Oxford.

BW, S. (1986). *Density Estimation for Statistics and Data Analysis.* Chapman and Hall, London.

Creedy, J. (1974). "Income Changes over the Life Cycle". *Oxford Economic Papers 26*(3), 405–423.

Efron, B. and R. Tibshirani (1993). *Introduction to the Bootstrap.* Chapman & Hall/CRC.

Fields, G. (2006). *The Many Facets of Economic Mobility.* Houndsmills: Palgrave Macmillan: in M. McGillivrary (ed), Inequality, Poverty, and Well-Being.

Fields, G. and E. Ok (1996). "The Measurement of Income Mobility: An Introduction to the Literature". *Working Papers*.

Gnedenko, B. (1978). *The Theory of Probability.* Moscow: Mir Publishers.

Jenkins, S. (2011). *Changing Fortunes. Income Mobility and Poverty Dynamics in Britain.* Oxford University Press.

Maza, A., M. Hierro, and J. Villaverde (2010). "Measuring Intra-distribution Income Dynamics: An Application to the European Regions ". *The Annals of Regional Science 45*(2).

Pisano, E. and S. Tedeschi (2008). "Income Inequality, Poverty and Mobility in Italy 1977-2006: Some Evidence". *CRISS Working Papers* (38).

Pittau, M. and R. Zelli (2002). "Income Distribution in Italy: a Nonparametric Analysis". *Statistical Method and Application 10*, 175–190.

Prais, S. (1955). "Measuring Social Mobility ". *Journal of the Royal Statistical Society 118* (1), 56–66.

Quah, D. (1997). "Empirics for Growth and Distribution: Stratification Polarization and Convergence and Clubs". *Journal of Economic Growth 2*, 27–59.

Schluter, C. and D. Van de gaer (2003). "Mobility as Distributional Difference". *Working Paper*.

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

Shorrocks, A. (1978b). "The Measurement of Mobility". *Econometrica 46*, 1013–1024.

Silverman, B. (1986). *Density Estimation for Statistics Data Analysis*. Chapman and Hall.

Wand, M. and M. Jones (1993). "Comparison of Smoothing Parameterizations in Bivariate Kernel Density Estimation". *Journal of the American Statistical Association 88*, 520–528.

Wood, S. (2011). "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models". *Journal of the Royal Statistical Society 73* (1).