UNIVERSITA' CATTOLICA DEL SACRO CUORE

WORKING PAPER

DISCE

Dipartimenti e Istituti di Scienze Economiche

The Sampling Distribution of Directional Mobility Indices Applied to the Income of Italian Families

Camilla Ferretti

DISCE – September - 2014

UNIVERSITÀ CATTOLICA del Sacro Cuore

QUADERNI DEL DIPARTIMENTO DI SCIENZE
ECONOMICHE E SOCIALI

THE SAMPLING DISTRIBUTION OF DIRECTIONAL MOBILITY INDICES APPLIED TO THE INCOME OF ITALIAN FAMILIES

Camilla Ferretti

Serie Verde: Metodi quantitativi e Informatica – Quaderno N. 102 settembre 2014

UNIVERSITÀ CATTOLICA DEL SACRO CUORE
PIACENZA

The sampling distribution of directional mobility indices applied to the income of Italian families

Camilla Ferretti[*]

Abstract

In economics, transition matrices are often used to describe the dynamics of individuals among a discrete set of states, defined on the basis of an economically relevant variable. Typical examples are the analysis of flows of workers in the labor market or of householders among various income levels. We consider here the problem of comparing different matrices through a specific mobility index called *directional index*. Generally speaking, mobility indices are functions of a given transition matrix P. The directional index is a particular function able to indicate both the level of mobility and the prevailing direction (left/right) in the dynamics under study. Here we focus on the comparison between two different sampling matrices: consequently, in order to rigorously determine if the level of mobility has significantly changed, we provide the analysis of its asymptotic sampling distribution.

Empirical applications regard the analysis of sampling transition matrices about the income of Italian families. Such matrices cover four consecutive two-years periods, from 2004 to 2012. We make statistical inference to analyze changes of mobility with respect of time, and among families with different income levels. Starting from 2006, results show a prevailing negative mobility, that is the tendency of Italian families to move towards lower income classes. In particular, we observe negative peaks in the mobility values in the period 2010-2012, for middle/high income classes, indicating that Italian families are still suffering the 2008 downturn.

Keywords: Directional mobility index, income transition matrix

1 Introduction

In literature, transition matrices have been widely used to describe the dynamics of discrete state processes based on economic variables. In this sense relevant examples are, among others, the analysis of flows of individuals in the labor market (Fougere and Kamionka, 2003), of firms among size classes (Cipollini et al., 2012), or of citizens among income levels (Bourguignon and Morrison, 2002). As it is well-known, a given transition matrix $P = \{p_{ij}\}_{i,j=1,\dots,k}$ provides all the information on the movements of

[*]Department of Economic and Social Sciences, Università Cattolica del Sacro Cuore, Piacenza, Italy, e-mail: camilla.ferretti@unicatt.it

individuals among the states (labeled with $1,\dots,k$), in the sense that p_{ij} usually is the probability to be in j at time t, given the state i at time $t-1$. Assuming appropriate hypothesis (for example, P is ruling transitions in a Markov Chain), the knowledge of P permits to forecast the evolution of individuals among the k states.

In this work we focus on a specific issue related to the theory of discrete state processes: the comparison of the mobility of two different matrices. Indeed, any matrix P can be associated to a certain degree of mobility measured through a specific mobility index $I(P)$. The larger is $I(P)$, the higher is the degree of mobility in the group of individuals ruled by P. We shortly introduce the theory of mobility indices considering the set of the transition matrices with k states, with $k \in \mathbb{N}_+$:

$$\mathscr{P} = \left\{ P = (p_{ij}) \in \mathbb{R}^{k\times k} | p_{ij} \geq 0, \sum_{j=1}^{k} p_{ij} = 1, \forall i = 1,\dots,k \right\}.$$

Mobility indices for transition matrices are then functions $I(\cdot)$ mapping a given transition matrix $P \in \mathscr{P}$ in the value $I(P) \in \mathbb{R}$, and they make possible the comparison of two different matrices $P,Q \in \mathscr{P}$. Indeed, as specified in Ferretti (2012), mobility indices are able to introduce a *total order* in $\mathscr{P}$, that is a rule which permits to establish unequivocally if the mobility of P is lower/higher than the mobility of Q, for every couple $P,Q \in \mathscr{P}$. Basically, chosen the function $I(\cdot)$, we will say that $P \prec Q$ ("*P is less mobile than Q*") if and only if $I(P) < I(Q)$. It is worth noting that the order in $\mathscr{P}$ depends on the choice of $I(\cdot)$.

Famous proposals for the function $I(\cdot)$ are, among others, in Shorrocks (1978), Sommers and Conlinsk (1979) and Bartholomew (1982). As an important example of $I(\cdot)$ we are going to use in the following, we recall here the *trace index* (Prais, 1955; Shorrocks, 1978):

$$I_{tr}(P) = \frac{k - \sum_{i=1}^{k} p_{ii}}{k-1} = \frac{\sum_{i=1}^{k}(1-p_{ii})}{k-1}. \tag{1}$$

The trace index furnishes an absolute measure of the mobility: in particular it measures the global tendency to leave from the current state ($1-p_{ii}$ is the probability to move away from the i-th state).

In Ferretti and Ganugi (2013) we proposed a new mobility index, the *directional index*:

$$I_{dir}(P) = \sum_i \omega_i \sum_j p_{ij} \cdot sign(j-i) \cdot v(|j-i|). \tag{2}$$

I_{dir} derives from the necessity to compare different evolutions when the set of states is ordered from the worst (state "1") to the best one (state "k") as in the case of incomes or firm size. $I_{dir}(P)$ is then a real value, equipped with the sign +/-, which is positive when the prevailing direction of individuals is towards the better states, and negative in the opposite case.

The directional index is also equipped with some parameters: 1) ω_i, $i = 1,\dots,k$ are weights to be assigned to individuals starting from the i-th state (we assume $\omega_i \geq 0$ for every i and $\sum_{i=1}^{k} \omega_i = 1$); 2) v is a function which measures the contribute to the whole mobility of individuals making jumps of magnitude $|j-i|$. In the end, the function *sign* is defined as follows:

$$sign(x) = \begin{cases} -1, & \text{if } x < 0; \\ +1, & \text{if } x > 0; \\ 0, & \text{if } x = 0; \end{cases}$$

and gives a positive (resp. negative) sign to jumps which lead to a better (resp. worse) state than the current one. For sake of shortness, we will indicate herein $sign(j-i) \cdot v(|j-i|)$ with v_{ij}.

In the aforementioned works we analyzed the general properties of the directional index, providing some results about the distribution of mobility in the whole set $\mathscr{P}$. Indeed, assuming that matrices are uniformly distributed in $\mathscr{P}$, we find that (Ferretti and Ganugi, 2013, p. 414) I_{dir} assumes values in $[m_1, m_2]$, where:

$$m_1 = \sum_{i=1}^{k} \omega_i v_{i1} \leq 0 \text{ and } m_2 = \sum_{i=1}^{k} \omega_i v_{ik} \geq 0, \tag{3}$$

and it has the following mean value and variance:

$$\mu_{\mathscr{P}}(I) = \frac{1}{k} \sum_i \omega_i \sum_j v_{ij}, \tag{4}$$

$$\sigma^2_{\mathscr{P}}(I) = \sum_j \omega_i^2 \sigma^2(I_i) = \frac{1}{k+1} \sum_i \omega_i^2 \left[\frac{1}{k} \sum_j v_{ij}^2 - \left(\frac{1}{k} \sum_j v_{ij} \right)^2 \right]. \tag{5}$$

Eqs. 4 and 5 describe the expected behavior of I_{dir} when it is evaluated on a given $P \in \mathscr{P}$. No assumptions are made on the theoretical nature of P, we only assume that different rows are independent one from each other. The distributional form of I_{dir} is not known, nevertheless previous results allow us to "place" the mobility of P with respect to the other matrices in $\mathscr{P}$, for example applying the Chebichev's theorem.

As a last result about the directional index we suggest to use the normalized version I^* of I_{dir}:

$$I^*(P) = \begin{cases} I_{dir}(P)/m_2 \text{ if } I_{dir}(P) \geq 0; \\ -I_{dir}(P)/m_1 \text{ if } I_{dir}(P) < 0; \end{cases} \tag{6}$$

which belongs to $[-1, +1]$ for every choice of ω and v and then simplifies the mobility comparison between two matrices.

Aim of this paper is to carry on the analysis of the directional index, facing the problem of its statistical properties. In the following we will assume the existence of an underlying theoretical model, observed through a sample of size n. In particular we suppose that the movements among the states are ruled by a Markov Chain with unknown transition matrix P. Following the ideas proposed in Schluter (1998) and Formby et al. (2004), we derive the asymptotic sampling distribution of $I_{dir}(\hat{P})$, where $\hat{P}$ is the estimate of P.

The paper is organized as follows: Sect. 2 recalls some important results about the asymptotic sampling distribution of $\hat{P}$ and proves that $I_{dir}(\hat{P})$ is asymptotically normally distributed; Sect. 3 reports the results of a numerical example which support the asymptotic Normality of the sampling directional index; Sect. 4 shows the application of statistical inference applied to the directional mobility of four samples of Italian families provided by Banca d'Italia for the years 2004-2012. The last section contains comments and possible further developments.

2 Inference about Mobility Indices

Usually measures of mobility are evaluated on samples of size n, drawn from a larger population. In this case the function $I(\cdot)$ is applied to the estimated transition matrix $\hat{P}$ instead of the theoretical matrix P. In literature there exist some results about the sampling distribution of the elements $\hat{p}_{ij}$, and consequently about the sampling value $I(\hat{P})$. In particular we will refer to the results obtained by Anderson and Goodman (1957).

2.1 One observed matrix

We suppose to have at disposal observations about the individuals in two consecutive instants of time t_0 and t_1. On such basis we build the $k \times k$ empirical matrix X such that x_{ij} is the number of individuals being in state i at time t_0 and in state j at time t_1. It holds

$$\sum_{j=1}^{k} x_{ij} = n_i = \text{ nr. of individuals starting from } i$$

and

$$\sum_{i,j=1}^{k} x_{ij} = \sum_{i=1}^{k} n_i = n = \text{ size of the sample.}$$

In this case the MLE estimator $\hat{P}$ of P has elements

$$\hat{p}_{ij} = \frac{x_{ij}}{n_i} \tag{7}$$

(Anderson and Goodman, 1957, p. 92).

Let $x_i = (x_{i1}, \ldots, x_{ik})$ be the observed vector of individuals moving from i. It is proved that

$$x_i \sim \text{Multinomial}(p_i, n_i),$$

where $p_i = (p_{i1}, \ldots, p_{ik})$. Consequently (Anderson and Goodman, 1957, p. 95) $\hat{p}_i = (\hat{p}_{i1}, \ldots, \hat{p}_{ik})$ is asymptotically distributed as a multivariate Normal $\mathcal{N}(p_i, \Sigma)$, with

$$\Sigma_{jl} = \text{cov}(\hat{p}_{ij}, \hat{p}_{il}) = \begin{cases} \frac{p_{ij}(1-p_{ij})}{n_i} \text{ if } j = l; \\ -\frac{p_{ij}p_{il}}{n_i} \text{ if } j \neq l. \end{cases} \tag{8}$$

(Rows of P are still supposed to be independent one from the each others, and n_i is supposed to be known for every $i = 1, \ldots, k$).

On such basis we are able to derive the asymptotic sampling distribution of many mobility indices. We recall the results about the trace index (Schluter, 1998, p. 159):

$$I_{tr}(\hat{P}) \overset{asym.}{\sim} \mathcal{N}\left(\frac{k-\sum_{i=1}^{k} p_{ii}}{k-1}, \frac{1}{(k-1)^2}\sum_{i=1}^{k}\frac{p_{ii}(1-p_{ii})}{n_i}\right). \tag{9}$$

Analogously the directional index is asymptotically distributed as a Normal with mean $\mu(I_{dir}(\hat{P}))$ and variance $\sigma^2(I_{dir}(\hat{P}))$. By linearity we have that

$$\mu(I_{dir}(\hat{P})) = I_{dir}(P). \tag{10}$$

The formula of $\sigma^2(I_{dir}(\hat{P}))$ needs some additional computation. We firstly compute the variance of $\hat{I}_i = I_i(\hat{P})$, where

$$I_i(P) = \sum_j p_{ij} \cdot v_{ij}.$$

$$\sigma^2(\hat{I}_i) = \sigma^2\left(\sum_{j=1}^{k} \hat{p}_{ij} v_{ij}\right) = \sum_{j=1}^{k}\left[\sigma^2(\hat{p}_{ij} v_{ij}) + 2\sum_{l=j+1}^{k} Cov(\hat{p}_{ij} v_{ij}, \hat{p}_{il} v_{il})\right] =$$

$$= \sum_{j=1}^{k}\left[\frac{p_{ij}(1-p_{ij})}{n_i} \cdot v_{ij}^2 - 2\sum_{l=j+1}^{k}\frac{p_{ij} p_{il}}{n_i} \cdot v_{ij} \cdot v_{il}\right].$$

From the independence between I_i and I_j, the final variance of $I(\hat{P})$ is given by

$$\sigma^2(I_{dir}(\hat{P})) = \sigma^2(\sum_i \omega_i I_i(\hat{P})) = \sum_i \omega_i^2 \sigma^2(\hat{I}_i). \tag{11}$$

2.2 Two or more observed matrices

In many cases, relevant economic variables are observed in more than two consecutive instants of time: $t_0, \ldots, t_T$. Usually we consider equi-spaced instants of time, indicating them simply with $0, 1, ..., T$. Coherently with the previous section, let $x_{ij}(t)$ be the number of individuals being in the i-th state at time $t-1$ and in the j-th state at time t, for every $t = 1, \ldots, T$. Entries of P are in this case estimated with:

$$\hat{p}_{ij} = \frac{\sum_{t=1}^{T} x_{ij}(t)}{\sum_{t=1}^{T}\sum_{j=1}^{k} x_{ij}(t)}. \tag{12}$$

Note that the time-homogeneity of Markov Chains is here needed.

As before, we consider the number of individuals being in the i-th state at time t, for $t = 0, \ldots, T$:

$$n_i(t) = \sum_{j=1}^{k} x_{ij}(t),$$

and the number of individuals passing from i before time T:

$$n_i^* = \sum_{t=0}^{T-1} n_i(t).$$

Then it is possible to prove that

$$\hat{p}_i \overset{asym.}{\sim} \mathcal{N}(p_i, \Sigma^*),$$

with

$$\Sigma_{jl} = \text{cov}(\hat{p}_{ij}, \hat{p}_{il}) = \begin{cases} \frac{p_{ij}(1-p_{ij})}{n_i^*} \text{ if } j = l; \\ -\frac{p_{ij}p_{il}}{n_i^*} \text{ if } j \neq l. \end{cases} \tag{13}$$

Consequently, $I_{dir}(\hat{P})$ is asymptotically distributed as a Normal with expected values equal to $I_{dir}(P)$ and variance obtained by Eq. 11 simply substituting n_i with n_i^*.

3 Simulation

The purpose of the following numerical example is to support the Normality of the index asymptotic distribution. With this aim we consider the randomly chosen theoretical matrix

$$P = \begin{bmatrix} .2 & .3 & .1 & .4 \\ .3 & .5 & .1 & .1 \\ .2 & .4 & .4 & 0 \\ .1 & .2 & .6 & .1 \end{bmatrix}.$$

Our data consist in 5000 simulated realizations of the corresponding Markov chain. For any simulation we set $n = 500$ (the number of individuals), with the given starting frequencies $p_0 = (.2, .1, .3, .4)$. In the first case we consider $T = 1$ (that is individuals are observed in two consecutive instants of time, as in Sect. 2.1). In the second case we consider $T = 5$ (as in Sect. 2.2). Consequently $\hat{P}$ is estimated as in Eqs. 7 and 12.

We evaluate $I_{dir}(\hat{P})$ setting $v(|j-i|) = |j-i|$, which gives a linearly increasing weight to jumps from i to j, as specified in the following section. We also analyze two different cases for the weights ω_i: on one hand we suppose that the starting state gives no information about the mobility, that is $\omega = (.25, .25, .25, .25)$; on the other hand we choose $\omega = (.1, .2, .3, .4)$, which means that individuals in the 4-th state carry more weight than others in the global mobility measure[1]. We obtain

$$I_{dir}(P) = -0.1,$$

when $\omega = (.25, .25, .25, .25)$, and

$$I_{dir}(P) = -0.59,$$

[1] As explained in Ferretti and Ganugi (2013), we may need to make distinction among different states. For example, in the empirical application it is useful to give more weight to states with more individuals, by setting ω equal to the starting frequency distribution $n_1(0)/n, \ldots, n_k(0)/n$.

when $\omega = (.1, .2, .3, .4)$. In both the cases we find that the mobility related to P is negative, that is individuals mainly tend to move towards left (i.e. to make their condition worse, if they were, for example, families moving among income classes, as in the following section).

Fig. 1 shows the results of our simulation for $T = 1$ and $T = 5$. We compare the simulated density distribution of the index (built up using 50 class intervals) and the Normal density distribution with expected value and variance calculated as in Eqs. 10 and 11. The graphical comparison supports the Normality of the sample index. Tab. 1 instead contains the theoretical and simulated values for the mean and the variance of $I_{dir}(\hat{P})$. We display also the p-value obtained with the Shapiro-Wilk test of normality, which still confirms our results.

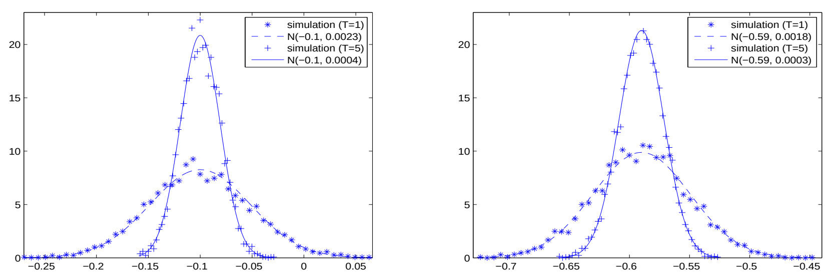

Figure 1: Simulated versus theoretical density distribution of $I^*(\hat{P})$, with $\omega = (.25, .25, .25, .25)$ (left) and $\omega = (.1, .2, .3, .4)$ (right).

From a descriptive point of view a last remark regards the comparison between different indices. Since m_1 and m_2 depend on ω and v, it is preferable to compare the normalized indices I^*, which belongs to $[-1, +1]$ for every choice of ω and v. In the previous example we find respectively -0.067 (-6.7%) and -0.295 (-29.5%). Then mobility is toward a worsening in both the cases and in addition we can say that it is higher in the second case.

Table 1: Comparison between simulated and theoretical mean and variance, and Shapiro-Wilk test's p-value for Normality.

	$\mu(\hat{I}_{dir})$	simulated mean	$\sigma^2(\hat{I}_{dir})$	simulated variance	p-value
$\omega = (.25, .25, .25, .25)$					
$T = 1$	-0.1	-0.0998	0.0024	0.0023	0.1787
$T = 5$	-0.1	-0.0997	0.0004	0.0004	0.5074
$\omega = (.1, .2, .3, .4)$					
$T = 1$	-0.59	-0.5909	0.0018	0.0016	0.6756
$T = 5$	-0.59	-0.5901	0.0004	0.0003	0.5027

4 An empirical illustration

As an empirical illustration of the previous results, we propose now the analysis of the mobility of Italian families subdivided according with their income. We aim to establish if mobility among income classes before and after the crisis is negative, positive, or equal to zero, and if it has significantly changed in the time.

For our proposals we use the data provided every two years by Banca d'Italia, regarding samples of Italian families which have never changed their composition. In particular we consider the surveys about the two-years periods 2004-2006, 2006-2008, 2008-2010 ad 2010-2012 (Banca d'Italia , 2008, 2010, 2012, 2014). In the same surveys sampling transition matrices are already proposed, but they are based on quintile classes, whose extremes vary with the time. Because of that we prefer to reconstruct the transition matrices using five fixed classes $q_1,\ldots,q_5$, delimited by the 2004 quintiles and displayed in Tab. 2.

Table 2: Income classes for Italian families (Euros).

q_1	q_2	q_3	q_4	q_5
$[0;17097)$	$[17097;24453)$	$[24453;34036)$	$[34036;48762)$	$[48762;+\infty)$

To reconstruct the matrices we consider for every couple of years families belonging to both the samples, and to avoid the bias caused by the inflation we convert the observed incomes into the 2013 value using the ISTAT revalorization's coefficients (http://www.istat.it/it/archivio/ 30440). The estimated matrices are obtained as in Eq. 7 and they are shown in Tab. 3. The same table contains also the sample size and the starting distribution[2] for every matrix (that is the distribution of Italian families among income classes in 2004, 2006, 2008 and 2010).

As a first descriptive step we evaluate the trace index shown in Eq. 1. We recall that I_{tr} gives a value in $[0,1]$, that is it is an absolute measure of the mobility. Furthermore, as previously noted, it provides a measure of the turbulence of families among income classes. From Tab. 4 we see that the global tendency to move away from the starting income class tends to decrease until 2010, that is Italian families undergo a decline of their capacity to move from one class to the others. In 2010-2012 we observe instead an increase in the mobility. It is a relevant result, nevertheless I_{tr} is not able to provide information about the kind of movements of families. Indeed, the same value of I_{tr} may correspond both to an increase and a decrease of the mean income level. In this light, the importance of a directional index is supported.

The second step of our analysis consists in using the statistical inference to rigorously determine if Italian families have experienced a worsening or an improvement of their conditions, and if the mobility has significantly changed after the crisis. We will use $I_{dir}(\hat{P})$ to apply statistical hypothesis tests about the theoretical value $I_{dir}(P)$ and about the difference between two indices. Tests will be applied on $\hat{I}_{dir}$ instead of $\hat{I}^*$ because, being the latter a piecewise linear function of the former, we may run into some

[2]The 2004 distribution is not equal to (0.2, 0.2, 0.2, 0.2, 0.2) because quintiles in Tab. 2 are evaluated on the whole 2004's sample (8012 families) , and matrices are estimated on the 3957 families which appear in both the 2004's and the 2006's samples.

Table 3: Estimated income transition matrices for the years 2004-2006, 2006-2008, 2008-2010, 2010-2012 (values are expressed as percentages).

	Income in the final year					
Income in the initial year	q_1	q_2	q_3	q_4	q_5	p_0
2004-2006 ($n = 3957$)						
q_1	64.10	24.35	7.57	3.44	0.55	18.37
q_2	17.10	43.21	27.42	9.27	3.00	19.36
q_3	6.35	17.79	41.68	26.18	8.01	19.89
q_4	2.42	6.40	18.48	49.15	23.55	20.92
q_5	0.71	2.47	6.01	20.26	70.55	21.46
2006-2008 ($n = 4345$)						
q_1	69.39	22.46	5.21	2.14	0.80	17.22
q_2	18.87	52.33	20.38	6.79	1.64	18.30
q_3	7.05	21.27	48.90	18.38	4.39	19.91
q_4	1.46	5.61	20.27	55.51	17.15	22.14
q_5	0.31	1.44	5.74	19.69	72.82	22.44
2008-2010 ($n = 4621$)						
q_1	73.52	19.09	4.56	1.60	1.23	17.57
q_2	19.27	53.82	20.64	5.59	0.68	18.98
q_3	5.71	19.21	54.00	17.76	3.32	20.84
q_4	1.78	4.95	18.69	56.28	18.30	21.88
q_5	0.63	1.36	4.38	19.42	74.22	20.73
2010-2012 ($n = 4611$)						
q_1	78.75	16.09	3.24	1.68	0.24	18.07
q_2	27.21	53.10	15.99	3.10	0.60	18.17
q_3	10.44	26.52	45.26	15.34	2.45	20.36
q_4	3.05	8.66	25.05	47.86	15.38	21.30
q_5	0.88	2.75	6.67	23.75	65.95	22.10

Source: processing of microdata from Banca d'Italia, *Indagine sui bilanci delle famiglie italiane*.

Table 4: Sampling trace index for the estimated income matrices.

Years	2004-2006	2006-2008	2008-2010	2010-2012
$\hat{I}_{tr}$	0.5783	0.5026	0.4704	0.5227

problems in the variance calculus, especially if $I^*(P) = 0$. Nevertheless tests results on $\hat{I}_{dir}$ can be extended to $\hat{I}^*$ exploiting the fact that $I^*(P) = 0$ if and only if $I_{dir}(P) = 0$. The normalized directional index will be shown to make easier the descriptive comparison between different samples.

To evaluate I_{dir} we need to choose the parameters ω and v. The vector of weight ω is here chosen to be equal to the starting distribution of every two-years span of time (the last column of Tab. 3). The choice of v is quite more tricky: indeed, in the index formula, the term $v(|j-i|)$ is introduced to give more weight to larger jumps. As for example, let $P \in \mathscr{P}$ be such that $p_{12} = p_{13} = 0.5$. If $v \equiv 1$, individuals moving from state 1 to state 2 and individuals moving from 1 to 3 have the same role in the whole mobility measure. If $v(|j-i|) = |j-i|$, then jumps from 1 to 3 are worth twice than jumps from 1 to 2. Other choices for v may be for example polynomial or exponential functions of $|j-i|$. In this study we see from Tab. 2 that income intervals based on the 2004 quintiles are approximately linearly increasing in their size. In this light we are allowed to set $v(|j-i|) = |j-i|$.

Table 5: Sampling directional index for the estimated income matrices.

Years	2004-2006	2006-2008	2008-2010	2010-2012
$\hat{I}_{dir}$	0.0695	-0.0304	-0.0249	-0.1945
$\hat{se}$	0.0144	0.0124	0.0117	0.0123
p-value	1.30E-06	0.0143	0.0329	0.0000
$\hat{I}^*$	3.61%	-1.42%	-1.19%	-9.2%

Tab. 5 shows the values of $I_{dir}(\hat{P})$. In the same table we show also the estimated standard error and the normalized values $I^*(\hat{P})$. From a descriptive point of view we see that the mobility is positive in 2004-2006 and assumes negative values in the years over the crisis. Still more relevant is the negative sign joined with an increased absolute value of the index in the last period 2010-2012 (-9.2%). Italian families have then apparently experienced a slowing down in their mobility over the crisis (confirmed by the trace index, too), which has led to a negative-near-to-zero index value, followed by a stronger worsening of the mean conditions in the last two years. This result is supported by a one-sample test about the mean value $I_{dir}(P)$: the null hypothesis H_0: $I_{dir}(P) = 0$ is rejected for the first and the last span of time, and accepted only at 1% level for the other two periods (p-values are in the third row of Tab. 5). As an additional result, we apply a difference-of-mean test based on two samples to test H_0: "the mobility has not significantly changed in the time". H_0 is always rejected, except for the couple 2006-2008 and 2008-2010. In this case the corresponding p-value is equal to 0.747. We conclude that from 2006 to 2010 families have undergone a sort of stagnation in their mobility.

The last step of this analysis regards the measurement of the mobility related to every single income class. Indeed the presence of the weights ω_i in the directional index allows us to measure the mobility for families starting from the i-th class, simply by setting $\omega = e_i$ (the i-th canonical vector in $\mathbb{R}^k$).

Table 6: $\hat{I}_{dir}$ values evaluated on families moving from different income classes.

Starting class	2004-2006	2006-2008	2008-2010	2010-2012
q_1	0.5199	0.4251	0.3793	0.2857
$\hat{se}$	0.0304	0.0277	0.0265	0.0219
q_2	0.3786	0.2000	0.1460	-0.0322
$\hat{se}$	0.0351	0.0311	0.0275	0.0270
q_3	0.1169	-0.0821	-0.0623	-0.2716
$\hat{se}$	0.0357	0.0313	0.0275	0.0303
q_4	-0.1498	-0.1871	-0.1563	-0.3615
$\hat{se}$	0.0324	0.0269	0.0264	0.0302
q_5	-0.4252	-0.3672	-0.3476	-0.4887
$\hat{se}$	0.0265	0.0220	0.0222	0.0253

Tab. 6 contains the sampling directional index $\hat{I}_{dir}$ restricted to every class, and the corresponding standard error. From such results we derive the following issues:

1. Mobility is always significantly different from zero (p-values are not displayed for shortness, there is only one exception that will be stressed in the following). The restricted directional indices confirm the fact that mobility has not significantly changed between 2006-2008 and 2008-2010 (difference-of-mean's p-values are displayed in Tab. 7).

Table 7: Difference-of-means test's p-values on 2006-2008 versus 2008-2010.

Class	q_1	q_2	q_3	q_4	q_5
p-value	0.2318	0.1933	0.6350	0.4130	0.5304

2. Families starting from the first class have necessarily a positive mobility (they are allowed to move only towards right), which changes significantly in the time (comparing 2004-2006 and 2010-2012) and shows a decreasing trend.

3. Families in q_2 exhibit a positive mobility except for the 2010-2012 period. Nevertheless we can not reject the hypothesis H_0: "mobility in 2010-2012 is equal to zero" in favor of H_1: "mobility in 2010-2012 is negative", since the corresponding one-tailed test about the mean has p-value equal to 0.116. Then families in the second class show a tendency to improve their income level. The same tendency is slowing down in the time.

4. Families in q_3 and q_4 have the worst results. Mobility is always negative except for the third class in the first period. In particular we note the negative minimum reached in both the classes in the last two years, indicating that families are still suffering the economic crisis.

5. Finally, families in the fifth class can have only negative mobility. As previously remarked, mobility is always statistically different from zero. It is worth noting again the negative peak in the last two years.

Lastly, we show the normalized index for families starting from every income class (Tab. 8). Normalized values of mobility still confirm the previous results. It is interesting to note that, in 2010-2012, families starting from q_3, q_4 and q_5 reach similar levels of mobility (around -13%).

Table 8: Normalized directional index for income classes.

	2004-2006	2006-2008	2008-2010	2010-2012
q_1	13%	10.63%	9.48%	7.14%
q_2	12.62%	6.67%	4.87%	-3.22%
q_3	5.84%	-4.10%	-3.12%	-13.58%
q_4	-4.99%	-6.24%	-5.21%	-12.05%
q_5	-10.63%	-9.18%	-8.69%	-12.22%

5 Conclusions and further research

In this paper we develop an analysis of the sampling distribution for directional mobility index proposed in Ferretti and Ganugi (2013). The index results to be asymptotically Normally distributed, with explicit expression for the expected value and the variance.

The asymptotic distribution of the index allows us to rigorously analyze movements of Italian families among income classes, exploiting the sampling microdata provided by Banca d'Italia in the years 2004-2006, 2006-2008, 2008-2010 and 2010-2012. Families are subdivided according with the 2004 income quintiles.

Results show a prevailing tendency to move towards the lowest income classes. Only families starting from the second class (having income comprised between 17097 and 24453 Euros) show a positive behavior, also if their mobility is decreasing in the time. A relevant result regards the 2010-2012 mobility: in this span of time families in the last three classes (that is having income greater that 24435 Euros) show a negative minimum in their mobility value (around -13%).

Further researches will regard the update of the analysis using more recent data. From a theoretical point of view we aim to extend the theory of directional indices, and the related statistical inference, to more refined models, such as mixtures of Markov Chains and continuous-in-time models.

References

Anderson T.W. and Goodman L.A. (1957) Statistical Inference about Markov Chains. Ann Math Statist 28(1): 89-110

Bartholomew D.J. (1982) Stochastic Models for Social Processes. 3nd ed. Wiley, London

Bourguignon F. and Morrison C. (2002) Inequality among World Citizens. Am Econ Rev 92(4):727-744

Cipollini F., Ferretti C. and Ganugi P. (2012) Firm size dynamics in an Industrial District: the Mover-Stayer model in action. In: Di Ciaccio A. et al. (eds) Advanced Statistical Methods for Large Datasets, pp 443-452

Ferretti C. (2012) Mobility as Prevailing Direction for Transition Matrices. Stat Appl 10(1):67-78

Ferretti, C. and Ganugi, P. (2013) A new Mobility Index for Transition Matrices. Stat Meth and App 22(3): 403-425

Formby J. P., Smith, W. J. and Zheng, B. (2004) Mobility measurement, transition matrices and statistical inference. J of Econometrics 120:181-205

Fougere D. and Kamionka T. (2003) Bayesian inference for the Mover - Stayer model of continuous time. J Appl Econ 18:697723

Geweke J., Robert C.M. and Gary A.Z. (1986) Mobility Indices in Continuous Time Markov Chains. Econometrica 54(6):1407-1423

Banca d'Italia (2008) I bilanci delle famiglie italiane nel 2006, in: Supplementi al bollettino statistico http://www.bancaditalia.it/statistiche/

Banca d'Italia (2010) I bilanci delle famiglie italiane nel 2008, in: Supplementi al bollettino statistico http://www.bancaditalia.it/statistiche/

Banca d'Italia (2012) I bilanci delle famiglie italiane nel 2010, in: Supplementi al bollettino statistico http://www.bancaditalia.it/statistiche/

Banca d'Italia (2014) I bilanci delle famiglie italiane nel 2012, in: Supplementi al bollettino statistico http://www.bancaditalia.it/statistiche/

Kendall M.G., Stuart A. (1977) The Advanced Theory of Statistics, vol I, Griffin, London

Prais S.J. (1955) Measuring social mobility, J Royal Stat Soc Series A, part I, 188:56-66

Schluter C. (1998) Statistical inference with mobility indices. Econ Lett 59:157-162

Shorrocks A.F. (1978) The measurement of Mobility. Econometrica 46(5):1013-1024

Sommers P. and Conlinsk J. (1979) Eigenvalue Immobility Measure for Markov Chains. J Math Soc 6:253-276