IARIW 2021

Monday 23 – Friday 27 August

Statistisk sentralbyrå
Statistics Norway

Intergenerational Mobility Measurement with Latent Transition Matrices

Michele Bavaro

(University of Roma Tre)

Federico Tullio

(University of Rome “Tor Vergata")

‘

Paper prepared for the 36th IARIW Virtual General Conference
August 23-27, 2021
Theme 10: Inequalities in Opportunity for Income, Education and Health

Intergenerational mobility measurement with latent transition matrices

Michele Bavaro*

University of Roma Tre

and

Federico Tullio†

University of Rome “Tor Vergata”

Abstract

We propose a novel multivariate approach for the estimation of intergenerational transition matrices. Our methodology is grounded on the assumption that individuals' social status is unobservable and must be estimated. In this framework, parents and offspring are clustered on the basis of the observed levels of income and occupational categories, thus avoiding any discretionary rule in the definition of class boundaries. The resulting transition matrix is a function of the posterior probabilities of parents and young adults of belonging to each class. Estimation is carried out via maximum likelihood by means of an expectation-maximization algorithm. We illustrate the proposed method using National Longitudinal Survey Data from the United States in the period 1978-2006.

Keywords: Expectation-Maximization algorithm; Intergenerational mobility; Latent Markov models; Transition matrix

*Email: michele.bavaro@uniroma3.it

†Email: federico.tullio@alumni.uniroma2.eu

1 Introduction

The use of multivariate statistical methods in studies of welfare economics is becoming increasingly popular. Multidimensional poverty analysis is adopting latent variables models grounded on the fact that observed heterogeneity in the population is due to unobservable components (see, e.g., Moisio 2004, Whelan & Maitre 2006, Krishnakumar 2008, Dotto et al. 2018, Tullio & Bartolucci 2019). The use of latent variables models has also spread to the *inequality of opportunity* literature. For example, Li Donni et al. (2015) propose to identify social *types* with a latent class model.

However, applications of multivariate methods in the study of intergenerational mobility are not common. Economic studies on intergenerational mobility mostly focus on income mobility. For extended coverage of the measurement of income mobility, refer to the survey by Jäntti & Jenkins (2013). Still, as argued by Jäntti & Jenkins (2013), there is no general agreement in the mobility literature about the best way of measuring the phenomenon. The most popular measures are intergenerational elasticity and intergenerational correlation, useful summary measures that may hide interesting details about mobility at different points of the joint distribution of parental and adult child incomes. For instance, it could be the case that mobility is not uniform along the income distribution. Generally speaking, these standard ways of measuring mobility are not informative about nonlinearities across the income distribution. Several techniques have been adopted to deal with this issue, such as splines or locally weighted regressions, kernel density and quantile regression approaches (see, for example, Bratsberg et al. 2007, Eide & Showalter 1999, Grawe 2004, Mocetti 2007).

A common strategy for dealing with nonlinearities is to estimate transition matrices. A transition matrix documents the movement of individuals across different income classes or occupational categories. Transition matrices have been adopted as a tool for measuring mobility in economics and sociology. Among the several empirical papers in economics that adopt transition matrices, some of the most frequently cited are Corak & Heisz (1999), O'Neill et al. (2007) as well as Bhattacharya & Mazumder (2011). In quantitative sociology, transition matrices are usually known as class mobility tables; an important contribution in this literature is provided by Erikson & Goldthorpe (1992), who developed a class schema and then showed the movement among the classes through mobility tables. The use of transition matrices is also widespread in the axiomatic literature on intergenerational mobility that tries to identify the best mobility measure by focusing on the properties that the index should respect; see the original contribution by Shorrocks (1978) and the more recent works by Fields & Ok (1996) and Checchi & Dardanoni (2002).

The main advantage of the use of transition matrices is that they offer a more detailed depiction of intergenerational mobility. They also facilitate the interpretation of the phenomenon: Mobility measured through transition matrices may be viewed as a merely reranking or positional phenomenon, in which individuals switch income or occupational positions across generations. A key issue with this methodology consists of defining the thresholds between classes. Formby et al. (2004) distinguish between *quantile* and *size* transition matrices. In the case of quantile transition matrices, class boundaries are determined to have the total number of sample units equally divided across classes, in the distribution of parents and in the distribution of adult children. In size transition matrices, class thresholds are exogenously set. Different approaches lead to distinct interpretations of mobility as well as different sampling distributions.

Our main contribution is to propose an innovative model to appropriately deal with the measurement of intergenerational class mobility in a multivariate framework[1]. Formally, our model belongs to the class of latent Markov models (LMMs) (Bartolucci et al. 2013), also known as hidden Markov models (HMMs) for longitudinal data (Zucchini & MacDonald 2009) and latent transition analysis (LTA) models (Collins & Lanza 2010). This model deals effectively with two crucial issues in the methodological literature on intergenerational mobility outlined above. Treating the individual status as a latent variable, we incorporate the popular concept that the status is not observable but may be only proxied. As proxies for individuals' social status, we use income and occupation, in line with the previous literature on intergenerational mobility[2]. Moreover, the problem of class boundaries when using transition matrices is solved endogenously with the estimation algorithm, which classifies individuals based on the posterior probabilities of belonging to each class rather than setting thresholds. Therefore, by estimating the model, it is possible to assign each individual to a class, discussing the main features of the class itself and computing mobility measures based on the estimated transition matrix that encompasses the information coming from the income level and the occupation performed. Finally, the model allows the number of social classes to change across the two generations, following the approach proposed by Anderson et al.

[1]The model may be easily extended to study intragenerational mobility, that is, the transition in individuals' status positions over the life cycle.

[2]For a complete survey on mobility with various statuses' proxies see Black & Devereux (2011). For occupational mobility, we refer to the papers by Long & Ferrie (2013) and Mazumder & Acosta (2015).

(2019).

The paper is organized as follows. In Section 2, we formulate the proposed specification. In Section 3, we present the empirical application based on U.S. data from the National Longitudinal Survey of Youth (NLSY79). This section includes the data description, the discussion of the sample selection rules as well as the main empirical results. Finally, in Section 4, we provide some concluding remarks.

2 A mixed-type data model for intergenerational mobility

To develop a model for intergenerational mobility, we rely on the basic LMM formulation (Bartolucci et al. 2013). The main assumption is that the social status of an individual is unobservable and can be measured based on a set of observable variables, the *manifest* variables, which act as proxies for the individual latent characteristic.

The use of LMMs to study intergenerational mobility presents some peculiarities with respect to standard applications of this type of model. The first peculiarity is that the number of time occasions is limited. In studies of intergenerational mobility, the number of time occasions is usually equal to two, coinciding with parental and offspring generations (there have rarely been studies that analyze more than two generations of individuals; see for instance Adermon et al. 2019). In typical applications of LMMs, the number of time periods is usually greater than or equal to three. For an example of an LMM with only two time periods, see Collins & Lanza (2010, Chapter 7.3). The second peculiarity is that the individuals are different in the two time occasions (e.g., fathers and sons). Nonetheless, if we consider the family (or dynasty) as the unit of analysis, this could be considered constant over time, and therefore, respecting the main features of an LMM. However, we propose a formulation with one grouping variable for each generation, to better explicate this feature.

2.1 The model

Let $T = 2$ be the number of time periods, and suppose that our units of observation are n parent-adult child couples $c = 1, \dots, n$. For each generation t, we observe a realization $\boldsymbol{y}_{ct}$ of the bivariate vector of manifest variables $\boldsymbol{Y}_{ct} = (Y_{1ct}, Y_{2ct})$. Assume that Y_{1ct} is a continuous variable (e.g., income), and Y_{2ct} is categorical (e.g., the type of occupation). Define $\tilde{\boldsymbol{Y}}_c$ as the vector of the manifest variables stacked along the time dimension and denote its realization as $\tilde{\boldsymbol{y}}_c$. In other words, the manifest variables of couple c in period $t = 1$ refer to some measure of income and occupation of the parents' generation, while in $t = 2$ the same variables refer to adult children. Suppose the existence of two discrete latent grouping variables capturing the social status, one at $t = 1$ (first generation) and one at $t = 2$ (second generation), F_c and S_c, respectively, that are collected in the random vector: $\boldsymbol{U}_c = (F_c, S_c)$. We allow the number of latent classes to be different in the two generations, that is, to be time-specific: k_t, for $t = 1, 2$. As mentioned, the main interest of the model lies in the distribution of the latent process $\boldsymbol{U}_c$. We refer to the *initial* and *transition* probabilities as:

$$\begin{aligned} \pi_v &= P(F_c = v), && v = 1, \ldots, k_1, \\ \pi_{s|v} &= P(S_c = s | F_c = v), && v = 1, \ldots, k_1, s = 1, \ldots, k_2, \end{aligned}$$

for $c = 1, \ldots, n$, with $\pi_v \geq 0$, $\pi_{s|v} \geq 0$, $\sum_v \pi_v = \sum_s \pi_{s|v} = 1$, $\forall v, s$. The probability distribution of $\boldsymbol{U}_c$ is given by:

$$P(\boldsymbol{U}_c = \boldsymbol{u}) = \pi_f \times \pi_{s|v}, \tag{1}$$

where $\boldsymbol{u} = (v, s)$ denotes a realization of the random vector $\boldsymbol{U}_c$.

With respect to the measurement component, we make use of the *local independence* assumption, also known as the *contemporaneous independence* assumption. The manifest variables are assumed to be independent, conditional on the latent process: If we knew the latent state of an individual at time t, then the realization of a manifest variable would not help in predicting the other manifest, as the latent variable is the only explanatory factor of the observable variables. This assumption implies:

$$f_{\tilde{\boldsymbol{Y}}_c|\boldsymbol{U}_c}(\tilde{\boldsymbol{y}}_c|\boldsymbol{u}_c) = \prod_{r=1}^{2} f^{(1)}_{Y_{rc1}|F_c}(y_{rc1}|v) \prod_{r=1}^{2} f^{(2)}_{Y_{rc2}|S_c}(y_{rc2}|s), \tag{2}$$

where we use $f^{(t)}$ as a generic symbol for a density or probability function. Note that the conditional distribution of each variable Y_{rct} given the latent state is allowed to vary over time; that is, it is characterized by time-specific parameters, as indicated by the superscript (t).

We assume that the components of the continuous variable are Gaussian functions with statespecific means and variances:

$$\phi(y_{1c1}; \mu_v, \sigma_v^2) = f_{Y_{1c1}|F_c}(y_{1c1}|v),$$
$$\phi(y_{1c2}; \xi_s, \tau_s^2) = f_{Y_{1c2}|S_c}(y_{1c2}|s),$$

for $v = 1, \ldots, k_1, s = 1, \ldots, k_2$. Let j be the number of categories of Y_{2ct}, $t = 1, 2$. The number of categories, thus, is fixed across generations. Denote the conditional probabilities of Y_{2ct} as:

$$\eta^{(1)}_{y|v} = f_{Y_{2c1}|F_c}(y|v),$$
$$\eta^{(2)}_{y|s} = f_{Y_{2c2}|S_c}(y|s),$$

for $y = 1, \ldots, j$, and $v = 1, \ldots, k_1, s = 1, \ldots, k_2$, with $\sum_y \eta^{(t)}_{y|u} = 1$, $u = v, s$. Again, the measurement model's parameters refer to both generations. For instance, μ_{41} and σ_{41} represent the mean and standard deviation of the income density associated with the fourth class of the first generation, while $\eta^{(2)}_{3|1}$ is the probability of being in the third occupational category for a second-generation individual belonging to the first class.

Making use of Equations (1) and (2), we obtain the marginal distribution of the manifest variables (the *manifest* distribution) by marginalizing over the distribution of the latent process:

$$\begin{aligned} f_{\tilde{\boldsymbol{Y}}_c}(\tilde{\boldsymbol{y}}_c) &= \sum_{\boldsymbol{u}} P(\boldsymbol{U}_c = \boldsymbol{u}) \; f_{\tilde{\boldsymbol{Y}}_c|\boldsymbol{U}_c}(\tilde{\boldsymbol{y}}_c|\boldsymbol{u}) \\ &= \sum_{v=1}^{k_1}\sum_{s=1}^{k_2} \pi_v \times \pi_{s|v} \times \phi(y_{1c1}; \mu_v, \sigma_v^2) \times \phi(y_{1c2}; \xi_s, \tau_s^2) \times \eta^{(1)}_{y_{2c1}|v} \times \eta^{(2)}_{y_{2c2}|s}. \end{aligned} \tag{3}$$

Thus, the estimation of $f_{\tilde{\boldsymbol{Y}}_c}(\tilde{\boldsymbol{y}}_c)$ requires summing over all the possible $k_1 \times k_2$ configurations of the vector $\boldsymbol{u}$. This is done by resorting to the forward-backward recursions within the expectation-maximization (EM) algorithm (Baum et al. 1970, Welch 2003) developed in the HMM literature and implemented through suitable matrix notation (Bartolucci 2006, Zucchini & MacDonald 2009).

A straightforward interpretation of such a model arises. In particular, the following (and possibly rectangular) matrix:

$$\Pi_{k_1 \times k_2} = \begin{bmatrix} \pi_{1|1} & \pi_{2|1} & \dots & \pi_{k_2|1} \\ \pi_{1|2} & \pi_{2|2} & \dots & \pi_{k_2|2} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{1|k_1} & \pi_{2|k_1} & \dots & \pi_{k_2|k_1} \end{bmatrix},$$

that is, the matrix collecting the transition probabilities, may be interpreted as a matrix of intergenerational mobility, where social statuses in both time periods are measured based on the observed level of the manifest variables. Furthermore, the initial probabilities π_v, $v = 1, \dots, k_1$, represent the *sizes* of the k_1 social latent classes of the first generation, while the sizes of the second generation's k_2 classes may be retrieved as

$$\delta_s = P(S_c = s) = \sum_v P(S_c = s|F_c = v)P(F_c = v),$$

for $s = 1, \dots, k_2$.

The number of free parameters is equal to $k_1 - 1$ in the initial distribution (π_v), $k_2(k_1 - 1)$ in the transition distribution ($\pi_{s|v}$), $2(k_1 + k_2)$ for the Gaussian densities ($\mu_v, \xi_s, \sigma_v^2, \tau_s^2$) and $(k_1 + k_2)(c - 1)$ for the categorical responses ($\eta^{(1)}_{y|v}, \eta^{(2)}_{y|s}$).

2.2 Maximum likelihood estimation

Under the formulation presented above, assuming independence of n sample units, the log-likelihood of the model may be expressed in the following way:

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{c=1}^{n} \log f_{\tilde{\boldsymbol{Y}}_c}(\tilde{\boldsymbol{y}}_c) \\ &= \sum_{c=1}^{n} \log \sum_{v=1}^{k_1} \sum_{s=1}^{k_2} (\pi_v \times \pi_{s|v}) \times \phi(y_{1c1}; \mu_v, \sigma_v^2) \times \phi(y_{1c2}; \xi_s, \tau_s^2) \times \eta^{(1)}_{y_{2c1}|v} \times \eta^{(2)}_{y_{2c2}|s}, \end{aligned} \tag{4}$$

where $\boldsymbol{\theta}$ is the set of all the model parameters. Equation (4) can be maximized by means of the EM algorithm (Dempster et al. 1977). The EM algorithm treats the individual latent states as missing data and finds the maximum likelihood estimates of the parameters in Equation (4) by maximizing the *complete data* log-likelihood (CDLL).

In the CDLL, we assume we know all the individual states. The *complete data* thus corresponds to the vectors $(\tilde{\boldsymbol{y}}_c, \boldsymbol{u}_c)$ for each couple c. Let $z_{cv} = I\{F_c = v\}$, $p_{cs} = I\{S_c = s\}$ and $z_{cs|v} = I\{F_c = v, S_c = s\}$, where $I\{\cdot\}$ is the indicator function taking value 1 if the argument is true. We can write the CDLL as:

$$\begin{aligned}
\ell(\boldsymbol{\theta}) = & \sum_{c=1}^{n}\sum_{v=1}^{k_1} z_{cv}\log\pi_v + \sum_{c=1}^{n}\sum_{v=1}^{k_1}\sum_{s=1}^{k_2} z_{cs|v}\log\pi_{s|v} \\
& + \sum_{c=1}^{n}\sum_{v=1}^{k_1} z_{cv}\log\phi(y_{1c1};\mu_v,\sigma_v^2) + \sum_{c=1}^{n}\sum_{s=1}^{k_2} p_{cs}\log\phi(y_{1c2};\xi_s,\tau_s^2) \\
& + \sum_{c=1}^{n}\sum_{v=1}^{k_1}\sum_{y=1}^{j} z_{cvy}\log\eta_{y|v}^{(1)} + \sum_{c=1}^{n}\sum_{s=1}^{k_2}\sum_{y=1}^{j} p_{csy}\log\eta_{y|s}^{(2)},
\end{aligned} \tag{5}$$

where the functions z_{cvy} and p_{csy} are equal to one for those individuals belonging to class v and s "responding" y; that is, $z_{cvy} = I\{F_c = v, Y_{2c1} = y\} = z_{cv} \times I\{Y_{2c1} = y\}$ and $p_{csy} = I\{S_c = s, Y_{2c2} = y\} = p_{cs} \times I\{Y_{2c2} = y\}$.

To maximize $\ell(\boldsymbol{\theta})$, the EM algorithm alternates two steps until convergence:

1. **E-Step:** Compute the expected value of the CDLL, given the data and the current value of the parameters. This boils down to computing the expected values of the indicator functions described above, that is, the following posterior probabilities:

$$\begin{aligned}
\hat{z}_{cv} &= P(F_c = v|\tilde{\boldsymbol{Y}}_c = \tilde{\boldsymbol{y}}_c), \\
\hat{p}_{cs} &= P(S_c = s|\tilde{\boldsymbol{Y}}_c = \tilde{\boldsymbol{y}}_c), \\
\hat{z}_{cs|v} &= P(S_c = s, F_c = v|\tilde{\boldsymbol{Y}}_c = \tilde{\boldsymbol{y}}_c),
\end{aligned}$$

   for all c, s and v. This step requires the forward-backward recursions.

2. **M-Step:** To update $\boldsymbol{\theta}$, maximize the expected value of the CDLL, replacing the probabilities $\hat{z}_{cv}$, $\hat{p}_{cs}$ and $\hat{z}_{cs|v}$ in Equation (5) and maximizing it with respect to the model's parameters. Note that the expression of the CDLL is made up of six different components, involving six separate maximizations. In a more general formulation with R manifest variables, the

maximization depends on the assumed state-dependent distributions, and numerical maximization is required if no explicit solutions are available. In our case, closed-form solutions are available in terms of latent and measurement models:

(a) For the initial and transition probabilities, the updates in the parameters are given by:

$$\hat{\pi}_v = \frac{\sum_c \hat{z}_{cv}}{n}, \tag{6}$$

$$\hat{\pi}_{s|v} = \frac{\sum_c \hat{z}_{cs|v}}{\sum_c \hat{z}_{cv}}, \tag{7}$$

for $v = 1, \ldots, k_1, s = 1, \ldots, k_2$.

(b) For the state-dependent Gaussian functions, we have:

$$\hat{\mu}_v = \frac{1}{\sum_c \hat{z}_{cv}} \sum_c \hat{z}_{cv}\, y_{1c1}, \tag{8}$$

$$\hat{\xi}_s = \frac{1}{\sum_c \hat{p}_{cs}} \sum_c \hat{p}_{cs}\, y_{2c2}, \tag{9}$$

$$\hat{\sigma}^2_v = \frac{1}{\sum_c \hat{z}_{cv}} \sum_c \hat{z}_{cv}\, (y_{1c1} - \hat{\mu}_v)^2, \tag{10}$$

$$\hat{\tau}^2_s = \frac{1}{\sum_c \hat{p}_{cs}} \sum_c \hat{p}_{cs}\, (y_{1c2} - \hat{\xi}_s)^2, \tag{11}$$

for $v = 1, \ldots, k_1, s = 1, \ldots, k_2$.

(c) Finally, the conditional response probabilities are given by:

$$\hat{\eta}^{(1)}_{y|v} = \frac{\sum_c \hat{z}_{cvy}}{\sum_c \hat{z}_{cv}}, \tag{12}$$

$$\hat{\eta}^{(2)}_{y|s} = \frac{\sum_c \hat{p}_{csy}}{\sum_c \hat{p}_{cs}}, \tag{13}$$

for $v = 1, \ldots, k_1, s = 1, \ldots, k_2$ and $y = 1, \ldots, c$.

The convergence of the algorithm is checked on the basis of the difference in the log-likelihood values of two consecutive steps. An important issue concerns the multimodality of the log-likelihood function. In practice, the convergence to a global maximum is not ensured; thus an appropriate initialization of the algorithm is required. We discuss our strategy in Section 3.2, with reference to the empirical application. Model identifiability is checked based on the numerical approximation of the observed information matrix.

3 Empirical application

3.1 Data

We use data from the National Longitudinal Survey of Youth 1979 cohort (NLSY79). The NLSY is a longitudinal survey that follows the lives of a sample of 12,686 young American men and women who were 14- to 22-year-old when first interviewed in 1979. Respondents were interviewed annually until 1994, and every other year since 1996. The project provides data available up to 2014 (Round 26), and the questions refer to the years before the interviews. During the first waves, for respondents living with their parents, a section of the survey (*Household Interview*) was addressed directly to the respondents' parents and includes information on family income at the parental level. In 1979, respondents were also asked to report parents' occupations. Given the long time period spanned by the survey, the NLSY dataset is adopted in several studies on intergenerational mobility (Jäntti et al. 2006, Bhattacharya & Mazumder 2011, Mazumder 2014). As standard in intergenerational mobility studies, the unit of observation is constituted by pairs of individuals linked across generations. Following Bhattacharya & Mazumder (2011), in our analysis we exclude daughters, to avoid labor force participation issues, and focus on father-son pairs[3]. The role of mothers as contributors to the social background is captured by the use of family income rather than fathers' earnings. Our final sample, thus, consists of 1,722 men (sons) living with their parents in 1979.

As a measure of income, we use total net family income, that is, the sum of a number of income values for household members related to the survey respondent by blood or marriage. A limitation of the income variable is the top coding of the upper tail of each year's income distribution[4]. Given that the true income levels are not always observed, the standard EM algorithm may deliver undesirable parameter estimates, and it may sometimes fail to converge (Atkinson 1992). Accounting

[3]In the second generation, the percentage of unemployed women is twice that of unemployed men. In the supplementary material (Section 2), we provide the results based on the full sample with sons and daughters.

[4]In particular, from 1979 to 1984, every income level above USD 75,000 was set equal to USD 75,001. However, starting from 1996, a different algorithm has been implemented. This algorithm takes the top 2% of respondents with valid entries and averages their value. The averaged value then replaces the income levels of the top 2% of the distribution.

for censored observations would require some modifications to the model log-likelihood, and thus, to the estimation algorithm (Atkinson 1992, Lee & Scott 2012, McLachlan & Jones 1988). However, we resort to an alternative strategy. In particular, taking time averages as a measure of individual income reduces the problem, as we use information from several years for each individual. The advantages of using time averages as proxies of permanent incomes have been extensively discussed by Solon (1992) and Zimmerman (1992); consequently, it is a common practice in empirical studies on intergenerational mobility measurement.

Thus, as a measure of parental income, we take the first three waves of the survey, 1979-1981, and we average the total net family income. To measure sons' incomes, we average the same variable over five consecutive waves (1998-2006), when respondents were 37 to 45 years old. For both generations, we use any available year of data, and we include in the sample only pairs in which sons have at least two valid income records. We convert all the income variables into constant 2000 U.S. dollars using the Organisation for Economic Co-operation and Development (OECD) consumer price index.

To have comparable occupational statuses across generations, we map the original variable labels (Census 1970 occupational codes) into a new variable that splits the occupations into three broad skill-based categories, and we refer to these categories as high-, medium- and low-skilled occupations[5]. We exclude unemployed individuals from the sample.

As stated above, data on parental employment is available only for 1979. We use father's occupation whenever reported; otherwise, we drop the father-son pair from the sample. We follow the same procedure for the respondents' generation in each of the five consecutive waves from 1998 to 2006. The occupational codes (Census 1970 to 2000 and Census 2000 from 2002 on) refer to the respondents' main job in the year before the interview. Regarding the long-term occupational status, Mazumder & Acosta (2015) point out that it is a more salient issue today than in the past, in particular due to a higher degree of occupational switching during the life course. We selected the most recurrent category among the nine ISCO-88 categories. If an individual had the same

[5]This simplified categorization is obtained from the International Standard Classification of Occupations (ISCO-88). Low-skilled occupations comprise i) plant and machine operators and ii) assemblers and elementary occupations; medium-skilled occupations comprise iii) clerks, iv) service workers and shop and market sales workers, v) skilled agricultural and fishery workers and vi) craft and related trades workers; and high-skilled occupations comprise vii) legislators, senior officials and managers, professionals and viii) technicians and associate professionals. As a robustness check, we estimate the model using the nine major ISCO categories. Given the sparseness of these categories, the information matrix is not invertible due to some estimated probabilities lying at the boundary of the parameter space. Still, results in terms of mobility remain substantially unchanged.

number of observations for two different categories, we selected the most recent one. Then, the ISCO categories were reduced to our skill-based three-group categorization.

Table 1 shows the descriptive statistics arising from the selection procedure. The gap in the

Table 1: Summary statistics ($n = 1722$)

	Average	Std. Dev.
Adult sons' income (1998-2006)	64,065	50,897
Fathers' income (1978-1980)	50,439	29,881
Sons' age in 2002	39.8	2.1
Fathers' age in 1979	46.4	7.2
Ethnic Group (%)		
Hispanic (H)	17.6	
Black (B)	22.2	
Non H, Non B	60.2	
Occupation (%)	Fathers	Adult sons
Low-skilled	31.8	21.9
Medium-skilled	42.6	38.8
High-skilled	25.6	39.3

Notes: The top panel of the table reports the income means and standard deviations of adult sons and fathers, as well as their age, in the final sample. The bottom panel reports the distribution of the occupational variable in the two generations.

average family income in the two generations is due to the structural growth of the U.S. economy in the period of analysis as well as to the role of the two different top-coding algorithms adopted in the two time periods. The lower average age of sons, usually identified as the cause of life-cycle bias in the estimates of intergenerational elasticities, is in line with the previous studies of mobility using NLSY79 data. Moreover, the almost 40-years-old average sample of sons fulfills the age requirements to reduce the bias at the minimum, given that the seminal work by Haider & Solon (2006) sets this optimal age between 35 and 45 years old for the United States. The evolution of the occupational distributions in the two generations reflects changes that occurred in the labor market structure in the years between the 1980s and the 2000s with the increase in high-skilled jobs to the detriment of low-skilled jobs.

3.2 Model fitting

We now present the main results of the empirical application of the model described in Section 2. We first discuss the selection of the number of latent classes. Then, we show in detail the latent class composition in terms of income and occupation. Finally, we display the estimated transition matrix that captures the degree of intergenerational mobility.

To select the number of latent states, we run the model for each possible combination of (k_1, k_2), with $k_1 = \{2, 3, 4\}$ and $k_2 = \{2, 3, 4, 5\}$. A suitable number of latent states is $k_1 = 3$ and $k_2 = 4$, according to the Bayesian's information criterion (BIC), as shown in Table 2.

Table 2: Model selection

k_1/k_2	2	3	4	5
	40625.48	40279.23	40251.52	40261.20
2	[-20241.95]	[-20046.47]	[-20010.27]	[-19992.75]
	(19)	(25)	(31)	(37)
	40540.31	40149.96	**40118.04**	40135.52
3	[-20177.01]	[-19955.76]	[**-19913.72**]	[-19896.38]
	(25)	(32)	(**39**)	(46)
	40544.04	40151.66	40124.48	40150.90
4	[-20156.53]	[-19930.53]	[-19887.14]	[-19870.54]
	(31)	(39)	(47)	(55)

Notes: The table reports, for each combination of (k_1, k_2), the number of free parameters of the model (brackets), the maximum of the log-likelihood function (square brackets) and the value of the Bayesian information criterion.

Due to the usual problem of possible multimodality of the likelihood function, we estimate the model 15 times by randomly selecting the starting values of the algorithm. We are quite confident that the solution, occurring all the 15 times up to negligible differences in the Gaussian densities' means and variances, is the global maximum of the likelihood function. Standard errors for the estimates are computed via non-parametric bootstrap.

Table 3: Class sizes and Gaussian densities' parameter estimates

Panel A: Fathers

v	1	2	3	-
π_v	0.411	0.462	0.126	
μ_v	28.943	56.767	97.234	
σ_v	13.793	18.138	37.399	

Panel B: Adult sons

s	1	2	3	4
δ_s	0.325	0.419	0.208	0.047
ξ_s	28.113	57.851	96.340	223.637
τ_s	12.641	17.603	31.536	83.381

Notes: The table shows the estimated class size, the mean (shown in US dollar/1000) and the standard deviation of the Gaussian density of each component, for fathers' (Panel A) and adult sons' (Panel B).

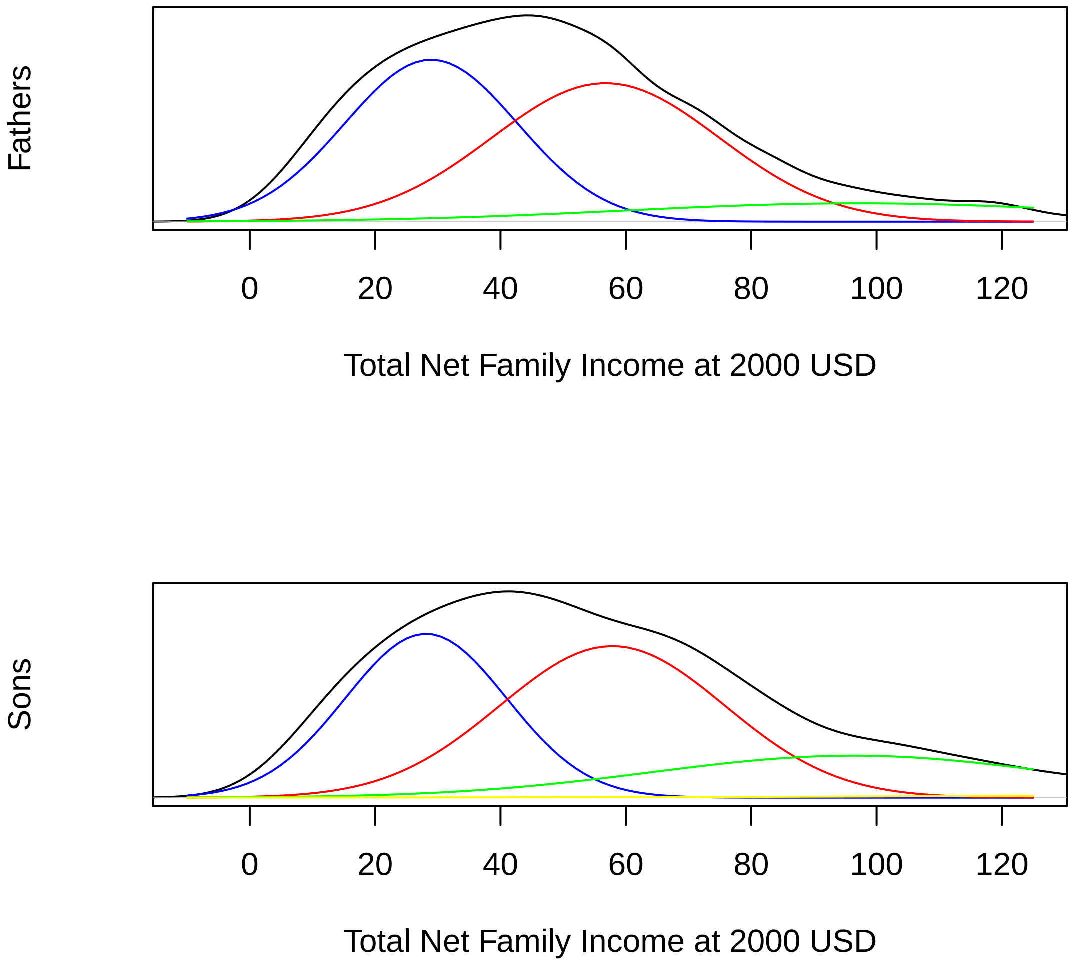
Table 3 reports the estimated class sizes along with the means and variances of the income components. Figure 1 plots the kernel density estimates of the income distributions along with the scaled components' densities. We recall that LMMs are, in general, globally identified up to a switching of the latent states. Therefore, we identify the classes based on the ordering of the Gaussian means at each time period.

We emphasize that the latent states have different interpretations in the two time periods. In this way, we are able to account for structural changes in the economy, that is, changes in the income distribution and in the labor market. This is particularly true when the number of classes varies over generations, as in this case[6]. In the second generation, the top class seems to be composed of "super rich" individuals, as the average of the income component is well above the overall average, and the high variance captures the positive skewness of the observed distribution.

Interestingly, the proposed model formulation allows to analyze separately the two main mo-

[6]The selection of the model with $(k_1, k_2) = (3, 4)$ classes may be a consequence of the top-coding rule that impacts the mean and variance at the top of the fathers' income distribution. In the supplementary material (Section 1), we provide the main results of the model with $k_1 = k_2 = 4$, entailing a more standard square transition matrix and providing the same qualitative results, in terms of mobility, of the main analysis.

Figure 1: State-dependent distributions: Income



Notes: The figure shows kernel density estimates (black) and the components' estimated scaled Gaussian densities of the income variable (shown in US dollar/1000) in both generations. Right tails truncated at USD 125,000. First class (blue), second class (red), third class (green), and fourth class (yellow).

bility concepts, that is, *structural* and *exchange* mobility. These two definitions, though inherited from sociology, are frequently adopted in the economic literature (Markandya 1982, Schluter & Van de Gaer 2011). Structural mobility is captured by the changes in the classes' expectations of the manifest variables, and by the change in the number of classes over generations. Therefore, the transition matrix summarizes exchange mobility entailing the reranking mechanism at work from one generation to another.

In Figure 2, we plot the estimated probabilities of belonging to each occupational category (on the horizontal axis) given the latent state, for each generation. The bottom class is associated with a high probability of being employed in an low-skilled occupation. In contrast, the top class is associated with a high probability of being employed in a high-skilled occupation. In the second generation, the third and fourth classes' behavior in terms of occupation is similar, as both classes are mainly composed of high-skilled employees. As noted, the fourth class in the second generation, which is characterized by higher family income, emerges as an "extension" of the third class.

The estimated transition matrix is summarized in Table 4. As stated in Section 1, there are

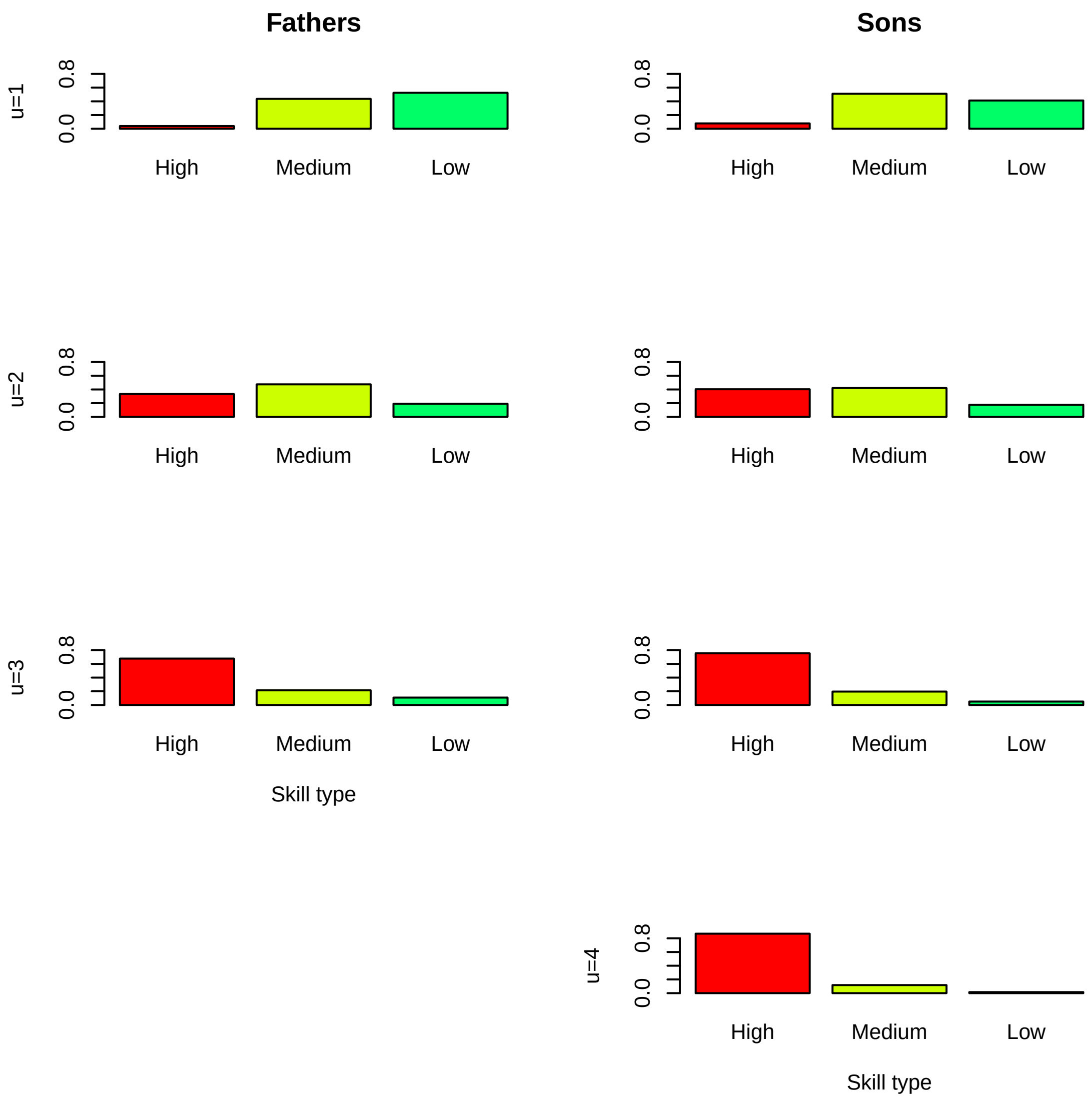
Table 4: Estimated transition matrix ($\hat{\Pi}_{3\times4}$)

	Adult sons			
Fathers	First	Second	Third	Fourth
First	0.653	0.240	0.094	0.013
	(0.122)	(0.104)	(0.030)	(0.014)
Second	0.112	0.605	0.260	0.023
	(0.064)	(0.115)	(0.143)	(0.019)
Third	0.036	0.324	0.391	0.250
	(0.031)	(0.136)	(0.130)	(0.093)

Notes: The table shows the matrix collecting the estimated transition probabilities. Standard errors (in parentheses) are computed via non-parametric bootstrap with 100 replications.

several ways through which it is possible to summarize the information arising from the transition

Figure 2: State-dependent distributions: Occupation



Notes: The figure shows the probabilities of belonging to each occupational category (1: high-skilled, 2: medium-skilled, 3: low-skilled), conditional on the latent state, for both generations. Latent states are ordered by row.

matrix, Π. Initially, we focus on *absolute* mobility measures[7]. We look at upward mobility as the percentage of sons born to fathers in the lowest class who reach the top class, a measure that is of great normative interest for equality of opportunity reasons. This value corresponds to the value at the top-right corner of the transition matrix, and in our case, it is a probability of 0.013 (i.e., 1.3 percentage points). On the other hand, the level of downward mobility is the percentage of sons born to fathers in the top class who end up in the bottom one. This value corresponds to the value at the lower-left corner of the transition matrix. In our case the value is a probability of 0.036 (i.e., 3.6 percentage points). These results provide evidence of a low degree of mobility, when looking at a latent status composed of the income and occupation dimensions. According to the estimates, those who have fathers in the first class are unlikely to reach the higher classes. The opposite happens for those whose fathers are in the top class. The degree of persistence at the top and at the bottom (i.e. the probability, for those who have fathers in the first and bottom class, to remain in the same class), is an additional significant information provided by the transition matrix. In our application, the persistence at the bottom is considerably higher than the persistence at the top, showing that it may emerge a serious issue with class poverty traps[8].

3.3 Race and ethnicity: mobility comparison

An advantage of the model described in Section 2 is that it allows in a simple way to encompass the role of other determinants in the formation of the latent transition matrix. In the United States, one of the main covariates that must be taken into account is given by the respondents' ethnicity and race.

The ethnic and racial differences in the income distribution in the U.S. have well-known foundations that originate in the last centuries; these inequalities may seriously hamper the process of development if they are persistent over generations, which is why estimates of intergenerational mobility by ethnic group or race can provide insights into whether racial differences in the United States are likely to be eliminated, and, if so, how long it might take.

[7]As Chetty et al. (2014) argue, there is a broad distinction between relative and absolute mobility measures based on the different questions to which they apply. Relative mobility measures identify the outcomes of adult children from low-income families relative to those of adult children from high-income families. Absolute mobility measures identify the outcomes of adult children from families of a given level in absolute terms. For instance one may be interested in measuring the outcomes of adult children who grew up in low-income families.

[8]This result may be induced by the different class structures in the two generations. However, as shown in the supplementary material (Section 1), the persistence at the bottom is higher than the persistence at the top also when looking to the (4×4) transition matrix.

Other studies have analyzed the role that racial or ethnic origins play in terms of intergenerational mobility; see for example Mazumder (2014) and Bhattacharya & Mazumder (2011) who use the same dataset, resorting to different methodologies[9]. The authors conclude that in recent decades black individuals have experienced substantially less upward intergenerational mobility and substantially more downward intergenerational mobility than white individuals.

In our case, we focus on the distinction between two groups into which the sample is divided, African American and Hispanic individuals (who account for about the 46% of the sample;see Table 1) and the rest of the sample who identify as white individuals. In the following, we refer to the two groups as white and non-white individuals. The choice to consider the two minority groups together is driven by estimation issues related to the restricted sample size.

The empirical model allows to study class composition and transition paths for the two groups. The class definition procedure does not change with respect to the previous section. All the classes are defined over the same parameter estimates.

We retrieve the initial and transition probabilities as defined by (6) and (7) conditional on the group of the sons (G_c), that is:

$$\hat{\pi}_v^g = \hat{P}(F_c = v | G_c = g) = \frac{\sum_c \hat{z}_{cv} \times I\{G_c = g\}}{n}, \tag{14}$$

$$\hat{\pi}_{s|v}^g = \hat{P}(S_c = s | F_c = v, G_c = g) = \frac{\sum_c \hat{z}_{cs|v} \times I\{G_c = g\}}{\sum_c \hat{z}_{cv} \times I\{G_c = g\}}, \tag{15}$$

for $v = 1 \dots k_1, s = 1 \dots k_2$ and $g = white, non - white$. Thus, now π_v^g may be interpreted as the probability of belonging to class v for a father whose son belongs to group g. Similarly, $\pi_{s|v}^g$ is now the probability that a son from group g belongs to class s conditional on his father's class being v. Finally, we have the marginal probabilities:

$$\hat{\delta}_s^g = \hat{P}(S_c = s | G_c = g) = \sum_v \hat{\pi}_{s|v}^g \hat{\pi}_v^g, \tag{16}$$

for $s = 1 \dots k_2$ and $g = white, non - white$.

A substantial difference exists in terms of the class compositions and transition patterns. In particular, most of the white individuals belong to the upper classes, whereas the non-white indi-

[9]They adopt transition probabilities of relative income status and measures of *directional rank mobility*.

viduals belong, on average, to the bottom classes, as shown in Table 5. Tables 6 and 7 show the

Table 5: Average posterior probabilities

Panel A: Fathers

v	1	2	3	-
White	0.276	0.552	0.172	
Non-white	0.616	0.326	0.057	

Panel B: Adult sons

s	1	2	3	4
White	0.221	0.47	0.247	0.062
Non-white	0.482	0.343	0.149	0.026

Notes: The table shows the average posterior probability of belonging to each latent class, conditional on race or ethnicity.

transition probabilities for the two groups. The upward mobility of white individuals is almost twice the upward mobility of non-white individuals, although this figure is very low. In contrast, downward mobility rate of non-white individuals is more than twice that of white individuals. These results are in line with previous empirical studies, and provide evidence of the presence of an ethnic gap in terms of upward and downward mobility. The degree of persistence in the top classes, which corresponds to the transition from the third to the fourth class, is instead similar: For white individuals, the degree equals 0.252, and for non-white individuals it equals 0.239.

However, the overall transition patterns as described by the two matrices do not seem very different. To provide formal evidence of this, we introduce a set of summary measures of mobility, and then we test the differences in these measures computed for the two groups.

The overall mobility described in transition matrices may be summarized by indices that are built up aggregating the movement among classes occurring in each father-son couple. More specifically, a mobility measure is a function $M(\Pi)$ that maps Π into a scalar (see, e.g., Formby et al. 2004). Several mobility indices are available to researchers (Checchi & Dardanoni 2002, Fields & Ok 1996). In our framework, the rectangular matrix constitutes a peculiarity, because perfect immobility cannot be characterized by an identity matrix. This implies that not all the standard

Table 6: Estimated transition matrix for white individuals

	White adult sons			
Fathers	First	Second	Third	Fourth
First	0.595	0.281	0.106	0.018
	(0.131)	(0.111)	(0.035)	(0.017)
Second	0.094	0.609	0.273	0.024
	(0.055)	(0.121)	(0.146)	(0.020)
Third	0.029	0.326	0.393	0.252
	(0.026)	(0.139)	(0.133)	(0.090)

Notes: The table shows the matrix collecting the estimated transition probabilities for white individuals. Standard errors (in parentheses) are computed via non-parametric bootstrap with 100 replications.

Table 7: Estimated transition matrix for non-white individuals

	Non-white adult sons			
Fathers	First	Second	Third	Fourth
First	0.693	0.213	0.085	0.009
	(0.116)	(0.099)	(0.029)	(0.012)
Second	0.157	0.595	0.229	0.020
	(0.083)	(0.103)	(0.134)	(0.019)
Third	0.068	0.312	0.381	0.239
	(0.056)	(0.129)	(0.131)	(0.117)

Notes: The table shows the matrix collecting the estimated transition probabilities for non white individuals. Standard errors (in parentheses) are computed via non-parametric bootstrap with 100 replications.

mobility indices may be applied. We use the index developed by Anderson (2018), defined as:

$$A = 1 - \frac{\sum_{s=1}^{k_2}(max(\pi_{s|\cdot}) - min(\pi_{s|\cdot}))}{k_1},$$

where $max(\pi_{s|\cdot})$ and $min(\pi_{s|\cdot})$ are the operator returning the maximum and minimum values, respectively, of the sth column of Π. The A index satisfies the normalization, immobility and perfect mobility axioms proposed by Shorrocks (1978).

We further adopt two modified versions of the Bartholomew index (Bartholomew 1982), B_1 and B_2. The idea behind the original index is that each class transition is weighted by the number of crossed boundaries. Our two modified versions are distinguished depending on the weights attached to each transition, given that in the context of different numbers of classes over generations the interpretation of the class transition is not straightforward. In B_1, the weights are given by the number of crossed boundaries. In B_2, the weights w_{sv} are set equal to 1 for transitions to subsequent classes, equal to 2 for transitions to classes 1-step more distant than the subsequent classes, and so on. The indices are defined as follows:

$$B_1 = \frac{1}{k_2 - 1} \sum_{v=1}^{k_1} \sum_{s=1}^{k_2} \pi_v \pi_{s|v} |s - v|,$$

$$B_2 = \frac{1}{k_2 - 1} \sum_{v=1}^{k_1} \sum_{s=1}^{k_2} \pi_v \pi_{s|v} w_{sv}.$$

Similar to the standard Bartholomew index, these modified versions satisfy the normalization and the immobility axioms, but not the perfect mobility axiom. Moreover they will not be comparable across studies based on transition matrices of different orders. Nonetheless, their interpretation is straightforward. They are bounded between 0 and 1, and the higher their values, the greater the level of mobility in the transition matrix.

In Table 8, the mobility indices extracted from the transition matrix $\Pi_{k_1 \times k_2}$ and computed on the whole sample are compared with those extracted from the transition matrices of the two sample subgroups. The point estimates suggest that the overall degree of mobility is higher for non-white individuals, and this is true for all the considered indices. However, this differences are not statistically different from zero.

In contrast, when testing the significance of the difference in terms of the initial probabilities

Table 8: Mobility comparison among racial or ethnic groups

	Whole sample	White	Non-white	Group diff
A	0.474	0.492	0.476	0.198
	(0.055)	(0.053)	(0.061)	
B_1	0.155	0.163	0.144	0.352
	(0.038)	(0.041)	(0.035)	
B_2	0.387	0.390	0.382	0.395
	(0.013)	(0.012)	(0.016)	

Notes: The table shows the Anderson (A) and modified Bartholomew (B_1, B_2) indices computed on the model's estimated transition matrices for the whole sample and the two racial or ethnic groups ($\Pi^g_{k_1 \times k_2}$, for $g = white, non - white$ and $(k_1, k_2) = (3, 4)$). The last column (Group diff) reports the test statistics of a two-tailed test for the equality of the indices computed in the two groups. Standard errors (in parentheses) are computed via non-parametric bootstrap with 100 replications.

among the two groups, we reject the null hypothesis of the difference being zero in all cases. This results provide evidence of the existence of a racial or ethnic gap in terms of class composition, rather than transition patterns.

4 Conclusions

This paper aims at proposing a new approach based on a multivariate framework to the study of intergenerational mobility. The model adopted may be cast in the class of latent Markov models, where an individual status is treated as a discrete latent variable with a finite number of states. Individuals are aggregated in different classes and a latent transition matrix between the two generations' classes is estimated. The peculiarity of this model is that it allows to analyze a multidimensional status and to avoid the issue of fixing class boundaries. Individual assignment is based on the posterior probabilities of belonging to each class. Another key feature of the model is that it allows the number of classes to differ in the two time periods, potentially delivering a rectangular transition matrix.

The results from the empirical application based on the National Longitudinal Survey Data from the U.S. show that the level of upward and downward mobility of the latent transition matrix is low, thus implying a strong class persistence at the extremes of the distribution. We further look to the more comprehensive mobility indices that encompass all the movements between the selected number of classes. If we disentangle the classes' composition and the transition patterns based on the racial or ethnic origin of the sons, we find that overall mobility is not statistically different among the two groups, although a key discrepancy is found in terms of class belonging, in particular for the fathers' generation.

Given the characteristics of the model, future work might aim to study intragenerational mobility and the transition between social classes over the individual life-cycle, or to conduct a cross-country comparative analysis to verify the differences in terms of class composition in different countries.

References

Adermon, A., Lindahl, M. & Palme, M. (2019), 'Dynastic human capital, inequality and intergenerational mobility', *CESifo Working Paper* .

Anderson, G. (2018), 'Measuring aspects of mobility, polarization and convergence in the absence of cardinality: indices based upon transitional typology', *Social Indicators Research* **139**(3), 887–907.

Anderson, G., Farcomeni, A., Pittau, M. G. & Zelli, R. (2019), 'Rectangular latent Markov models for time-specific clustering, with an analysis of the well being of nations', *Journal of the Royal Statistical Society (Series C)* **68**, 603–621.

Atkinson, S. E. (1992), 'The performance of standard and hybrid em algorithms for ml estimates of the normal mixture model with censoring', *Journal of Statistical Computation and Simulation* **44**, 105–115.

Bartholomew, D. J. (1982), *Stochastic Models for Social Processes*, 3rd edn, Wiley, London, UK.

Bartolucci, F. (2006), 'Likelihood inference for a class of latent Markov models under linear hypotheses on the transition probabilities', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**, 155–178.

Bartolucci, F., Farcomeni, A. & Pennoni, F. (2013), *Latent Markov Models for Longitudinal Data*, Chapman and Hall/CRC press, Boca Raton, FL.

Baum, L., Petrie, T., Soules, G. & Weiss, N. (1970), 'A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains', *Annals of Mathematical Statistics* **41**, 164–171.

Bhattacharya, D. & Mazumder, B. (2011), 'A nonparametric analysis of black–white differences in intergenerational income mobility in the United States', *Quantitative Economics* **3**, 335–379.

Black, S. & Devereux, P. (2011), Recent Developments in Intergenerational Mobility, first edn, Vol. 4B, Elsevier, chapter 16, pp. 1487–1541.

Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T. & Österbacka, E. (2007), 'Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons', *Economic Journal* **117**, C72–C92.

Checchi, D. & Dardanoni, V. (2002), 'Mobility comparisons: does using different measures matter?', *Research on Economic Inequality* **9**.

Chetty, R., Hendren, N., Kline, P. & Saez, E. (2014), 'Where is the land of Opportunity? the Geography of Intergenerational Mobility in the United States', *The Quarterly Journal of Economics* **129**, 1553–1623.

Collins, L. M. & Lanza, S. T. (2010), *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*, John Wiley and Sons Inc., Hoboken, NJ.

Corak, M. & Heisz, A. (1999), 'The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data', *The Journal of Human Resources* **34**, 504–556.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via the EM algorithm', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **39**, 1–38.

Dotto, F., Farcomeni, A., Pittau, M. G. & Zelli, R. (2018), 'A dynamic inhomogeneous latent state model for measuring material deprivation', *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **182**, 495–516.

Eide, E. & Showalter, M. (1999), 'Factors affecting the transmission of earnings across generations: A quantile regression approach', *Journal of Human Resources* **34**(2), 253–267.

Erikson, R. & Goldthorpe, J. (1992), *The Constant Flux: A Study of Class Mobility in Industrial Countries.*, New York: Oxford University Press.

Fields, G. & Ok, E. (1996), 'The Meaning and Measurement of Income Mobility', *Journal of Economic Theory* **71**(2), 349–377.

Formby, J. P., Smith, W. J. & Zheng, B. (2004), 'Mobility measurement,transition matrices and statistical inference', *Journal of Econometrics* **120**, 181–205.

Grawe, N. (2004), 'Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints', *The Journal of Human Resources* **39**, 813–827.

Haider, S. & Solon, G. (2006), 'Life-Cycle Variation in the Association between Current and Lifetime Earnings', *American Economic Review* **96**, 1308–1320.

Jäntti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Österbacka, E., Bjorklund, A. & Erikson, T. (2006), 'American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States', *IZA Discussion Paper 1938* .

Jäntti, M. & Jenkins, S. (2013), 'Income Mobility', *IZA Discussion Papers 7730* .

Krishnakumar, J. (2008), Multidimensional Measures of Poverty and Well-being Based on Latent Variable Models, *in* N. Kakwani & J. Silver, eds, 'Quantitative Approaches to Multidimensional Poverty Measurement', Palgrave Macmillan, London, UK, chapter 7, pp. 118–134.

Lee, G. & Scott, C. (2012), 'EM algorithms for multivariate Gaussian mixture models with truncated and censored data', *Computational Statistics and Data Analysis* **56**, 2816–2829.

Li Donni, P., Rodirguez, J. & Rosa Dias, P. (2015), 'Empirical definition of social types in the analysis of inequality of opportunity: a latent classes approach', *Social Choice and Welfare* **44**, 673–701.

Long, J. & Ferrie, J. (2013), 'Intergenerational Occupational Mobility in Great Britain and the United States since 1850', *American Economic Review* **103**, 1109–1137.

Markandya, A. (1982), 'Intergenerational exchange mobility and economic welfare', *European Economic Review* **17**(3), 307–324.

Mazumder, B. (2014), 'Black–white differences in intergenerational economic mobility in the United States', *Economic Perspectives* **38**.

Mazumder, B. & Acosta, M. (2015), 'Using Occupation to Measure Intergenerational Mobility.', *The ANNALS of the American Academy of Political and Social Science* **657**, 174–193.

McLachlan, G. J. & Jones, P. N. (1988), 'Fitting Mixture Models to Grouped and Truncated Data via the EM Algorithm', *Biometrics* **44**, 571–578.

Mocetti, S. (2007), 'Intergenerational earnings mobility in Italy', *The BE Journal of Economic Analysis & Policy* **7**.

Moisio, P. (2004), 'A Latent Class Application to the Multidimensional Measurement of Poverty', *Quality & Quantity* **38**, 703–717.

O'Neill, D., Sweetman, O. & Van de Gaer, D. (2007), 'The effects of measurement error and omitted variables when using transition matrices to measure intergenerational mobility.', *Journal of Economic Inequality* **5**(2), 159–178.

Schluter, C. & Van de Gaer, D. (2011), 'Upward structural mobility, exchange mobility, and subgroup consistent mobility measurement: Us–german mobility rankings revisited', *Review of Income and Wealth* **57**(1), 1–22.

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

Solon, G. (1992), 'Intergenerationai Income Mobility in the United States', *American Economic Review* **82**, 393–408.

Tullio, F. & Bartolucci, F. (2019), 'Evaluating time-varying treatment effects in latent Markov models: An application to the effect of remittances on poverty dynamics', *MPRA Paper 91459* .

Welch, L. R. (2003), 'Hidden Markov models and the Baum-Welch algorithm', *IEEE Information Theory Society Newsletter* **53**, 1–13.

Whelan, C. T. & Maitre, B. (2006), 'Comparing poverty and deprivation dynamics: Issues of reliability and validity', *The Journal of Economic Inequality* **4**, 303–323.

Zimmerman, D. (1992), 'Regression toward Mediocrity in Economic Stature', *American Economic Review* **82**, 409–29.

Zucchini, W. & MacDonald, I. L. (2009), *Hidden Markov models for time series: An introduction using R*, Chapman and Hall/CRC press, Boca Raton, FL.