



Intergenerational mobility measurement with latent transition matrices

Michele Bavaro¹ · Federico Tullio²

Received: 12 March 2020 / Accepted: 31 October 2022 / Published online: 31 December 2022
© The Author(s) 2022

Abstract

We propose a 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 · Transition matrix

1 Introduction

As income is often conceived as the main indicator of social status, economic studies on intergenerational mobility mostly focus on income or earnings mobility (Black and Devereux 2011; Björklund and Jäntti 2020). When dealing with the measurement of intergenerational mobility, a number of different methods are available to the researcher, even though there is no general agreement in the literature about the best way of measuring this phenomenon (Jäntti and Jenkins 2013). Intergenerational correlation and elasticity are the most popular measures, since the seminal papers by Solon (1992) and Zimmerman

The views expressed in this paper are those of the authors and do not involve the responsibility of the Bank of Italy. The usual disclaimers apply.

✉ Michele Bavaro
michele.bavaro@spi.ox.ac.uk

Federico Tullio
federico.tullio@bancaditalia.it

¹ University of Oxford, INET and Department of Social Policy and Intervention, Oxford, UK

² DG Economics, Statistics and Research, Bank of Italy, Rome, Italy

(1992). A limitation of these methods is that they are not informative about nonlinearities in mobility. Different methods have been adopted to overcome such limitation, including the Spearman (or rank) correlation (Chetty et al. 2014; Corak et al. 2014), transition matrices (Corak and Heisz 1999; O'Neill et al. 2007; Bhattacharya and Mazumder 2011), polynomial regressions (Bratsberg et al. 2007), and quantile regression approaches (Eide and Showalter 1999; Grawe 2004).

Transition matrices represent a useful tool to document the movement of individuals across different classes. The use of transition matrices offers a detailed depiction of intergenerational mobility and facilitates the interpretation of the phenomenon: mobility measured through transition matrices may be viewed as a reranking or positional phenomenon, in which individuals switch classes across generations. A key issue with this methodology consists in defining the thresholds between classes. Formby et al. (2004) distinguish between *quantile* and *size* transition matrices. In quantile transition matrices, class boundaries are set to have the total number of sample units equally divided across classes, whereas thresholds are exogenously set in size transition matrices. Transition matrices are also adopted in quantitative sociology, where they are known as class mobility tables and they are based on occupational classifications. An important contribution in this literature is provided by Erikson and Goldthorpe (1992), who developed a class schema and described the movement among classes through mobility tables. Similarly, Long and Ferrie (2013) use transition matrices to assess intergenerational occupational mobility in Britain and the US throughout the twentieth century.

We propose a model for the measurement of intergenerational class mobility in a multivariate framework, where individual social status is modeled as a discrete latent variable. Other recent works study intergenerational mobility adopting a latent variable approach in which multiple observable outcomes are considered *noisy* measures of an underlying latent variable, e.g. social status for Vosters and Nybom (2017), Vosters (2018) and Adermon et al. (2021), human capital for Fletcher and Han (2019), health status for Halliday et al. (2020). The use of multivariate statistical methods is becoming increasingly popular not only in mobility but more generally in studies of welfare economics.¹ Yet, to the best of our knowledge, the approaches developed in the literature of intergenerational mobility focus on the case of latent social status conceived as a continuous variable. While this choice is appropriate if the main interest lies in estimating intergenerational correlation coefficients, a discrete formulation of social status is more suited if the researcher's aim is to identify a finite number of classes in the population, and to study transition patterns among classes.

We fill this gap in the literature by proposing a model that adopts a latent class formulation, where individuals from the same family (e.g., father and adult son) are allowed to belong to different social classes in the two generations in which they are observed. The main feature of the model is represented by the matrix collecting transition probabilities, which explains the movement among classes across the two generations. This model deals effectively with two crucial issues in the methodological literature on intergenerational mobility: (i) accounting for multiple dimensions of individuals' social status, and (ii) establishing class boundaries when measuring mobility through transition matrices.

The main advantage of the model is that it allows data-driven clustering of subjects in different classes, thus avoiding any discretionary criterion in setting class boundaries. The

¹The use of latent variables models has also spread to multidimensional poverty analysis (Krishnakumar 2008; Dotto et al. 2018; Tullio and Bartolucci 2022) and to the *inequality of opportunity* literature (Li Donni et al. 2015).

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 arbitrary thresholds. 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. Treating the individual status as a latent variable, we incorporate the popular concept that social 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 in economics and sociology.

The empirical application is based on US data from the National Longitudinal Survey of Youth (NLSY79). We further divide the sample in two subgroups based on race and ethnicity of adult sons, providing a comparative analysis of mobility among the two groups using absolute mobility measures and overall mobility indices.

The paper is organized as follows. In Section 2, we formulate the proposed specification. In Section 3, we present the empirical application. This section includes the data description, the discussion of the sample selection rules as well as the results. Concluding remarks are provided in Section 4.

2 A mixed-type data model for intergenerational mobility

To develop a model for intergenerational mobility, we rely on a latent class formulation (Lazarsfeld and Henry 1969; Goodman 1974) where individuals from the same family are allowed to belong to different social classes. Formally, our model belongs to the class of latent Markov models (LMMs; Bartolucci et al. 2013), also known as latent transition analysis models (Collins and Lanza 2010). In the following, we provide an overview on LMMs and we highlight the connection to related models adopted in the mobility literature. Then, we discuss the model assumption and the estimation procedure.

2.1 Overview on LMMs and relation to existing literature

LMMs represent a generalization of latent class models to longitudinal data, in which the units of observation are allowed to move between latent classes. They are designed to model class composition, as well as the incidence of transitions over time in latent class membership. In other words, a typical analysis of LMMs based on multivariate longitudinal data consists of finding clusters of units and studying the transition between these clusters. These models assume the existence of a latent process affecting the distribution of the observable (*manifest*) variables. The latent process follows a first-order Markov chain with a finite number of states, typically referred to as latent states. The manifest variables are assumed to be independent, conditional on the latent process: if we knew the latent state of an individual, then the realization of a manifest variable would not help in predicting the other manifest variables, as the latent process is the only explanatory factor of the observables. A thorough description of the assumptions is provided in the next section, in connection with the proposed formulation.

The use of LMMs to study intergenerational mobility presents distinct features with respect to standard applications of this type of models. 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 (studies that analyze more than two generations of individuals are rare; see for instance Adermon

et al. 2021 and Modalsli 2021). In typical applications of LMMs, the number of time periods is usually greater than or equal to three.² The second peculiarity is that the individuals are different in the two time occasions (e.g., fathers and adult sons). We propose a formulation with a grouping variable for each generation to better explicate this feature. Additionally, the proposed model allows the number of classes to change across the two generations, as in Anderson et al. (2019).

Similar frameworks combining several status proxies have been adopted in the mobility literature. Vosters and Nybom (2017) and Vosters (2018) adopt the methodology developed by Lubotsky and Wittenberg (2006) to provide estimates of intergenerational status persistence in Sweden and US. This method provides an estimate of intergenerational persistence by linearly combining the observable proxies, in a regression framework in which all the proxies are included simultaneously. In this framework, the latent factor is included as explanatory variable in a linear model where the dependent variable is son's log earnings. Differently from our approach, multidimensionality only enters into the *right-hand side* of the model. In general, the choice of the suitable latent variable model depends on the given context and application. Most of the times, latent status is modeled as a continuous latent variable.³ While this specification is often conveniently adopted (e.g., for income and self-reported health status, which are typically continuous and ordinal, respectively), it would still require setting exogenous thresholds on the latent scores to build a transition matrix. This choice is not particularly suited in the present context, as the main interest lies in partitioning the population into a finite number of groups (social classes) and studying transition patterns. In such case, a discrete variable with a finite number of states seems the most appropriate solution. Moreover, as far as the analysis is conducted assuming the number of latent classes as unknown, even if the underlying latent variable is actually continuous, a discrete approximation is known to work well. This is widely known in the literature on finite mixture models and related semi-parametric maximum likelihood estimation; see, for instance, Lindsay et al. (1991), Mroz (1999), Cameron and Heckman (2001) and Arcidiacono and Jones (2003).

2.2 Model set up

2.2.1 Main formulation

Let $T = 2$ be the number of time periods, and suppose that our units of observation are n parent-adult child pairs $c = 1, \dots, n$. For each generation t , we observe a realization \mathbf{y}_{ct} of the bivariate vector of manifest variables $\mathbf{Y}_{ct} = (Y_{1ct}, Y_{2ct})$. Assume that Y_{1ct} is a continuous variable (e.g., income), and Y_{2ct} is categorical (e.g., the type of occupation). Define $\tilde{\mathbf{Y}}_c$ as the vector of the manifest variables stacked along the time dimension and denote its realization as $\tilde{\mathbf{y}}_c$. In other words, the manifest variables of pair 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 social status, F_c and S_c for the first and second generation respectively, which are collected in the random vector $\mathbf{U}_c = (F_c, S_c)$. We allow the number

²For an example of an LMM with two time periods, see Collins and Lanza (2010, Chapter 7.3).

³The most relevant examples of these models are represented by factor analysis models, MIMIC (multiple indicators and multiple causes) and structural equation models (SEM) including LISREL-type models. See Jöreskog and Goldberger (1975) and Bollen (1989) for a complete coverage of the topic and Krishnakumar and Nagar (2008) for a comparison between these different methodologies.

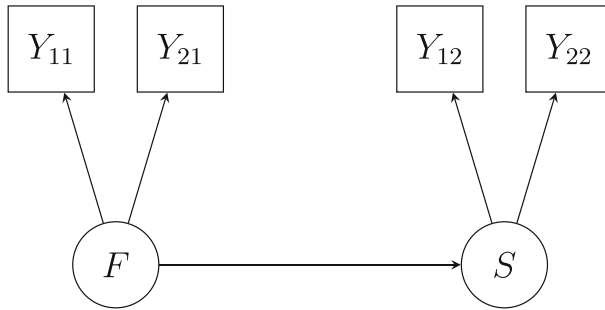


Fig. 1 Path diagram of the model

of states (also named 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 denoted by U_c . We refer to the *initial* and *transition* probabilities as:

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

for $c = 1, \dots, n$, 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 U_c is given by:

$$P(U_c = \mathbf{u}) = \pi_v \times \pi_{s|v}, \tag{1}$$

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

With respect to the measurement component, we make use of the *local independence* assumption, also known as the *contemporaneous independence* assumption. This assumption implies:

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

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). A path diagram of the proposed model is illustrated in Fig. 1, where unobservable random variables are indicated by circles and observables are indicated by squares.

We assume that the components of the continuous variable are Gaussian functions with state-specific means and variances:⁴

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

⁴Given their tractability, taking the component densities to be normal is a common assumption in mixture modeling (McLachlan and Peel 2000). Examples of recent works using income level as manifest variable and assuming normality for the component densities include Anderson et al. (2016) and Anderson et al. (2016). An alternative parametrization employs log-normal densities (Flachaire and Nunñez 2007), which is equivalent to log transforming the income variable Y_{1ct} and assuming a mixture of normal densities for the distribution of $\log Y_{1ct}$, and it requires the support of Y_{1ct} to be strictly positive.

for $v = 1, \dots, k_1, s = 1, \dots, k_2$. 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:

$$f_{Y_{2c1}|F_c}^{(1)}(y|v) = \eta_{y|v}^{(1)},$$

$$f_{Y_{2c2}|S_c}^{(2)}(y|s) = \eta_{y|s}^{(2)},$$

for $y = 1, \dots, j$, and $v = 1, \dots, k_1, s = 1, \dots, k_2$, with $\sum_y \eta_{y|u}^{(t)} = 1, u = v, s$. Again, the measurement model's parameters are generation-specific. For instance, μ_4 and σ_4 represent the mean and standard deviation of the income density associated with the fourth class of the first generation, while $\eta_{3|1}^{(2)}$ is the probability of being in the third occupational category for a second-generation individual belonging to the first class.

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

$$f_{\tilde{Y}_c}(\tilde{y}_c) = \sum_{\mathbf{u}} P(U_c = \mathbf{u}) f_{\tilde{Y}_c|U_c}(\tilde{y}_c|\mathbf{u})$$

$$= \sum_{v=1}^{k_1} \sum_{s=1}^{k_2} \pi_v \times \pi_{s|v} \times \phi(y_{1c1}; \mu_v, \sigma_v^2) \times \phi(y_{1c2}; \xi_s, \tau_s^2) \times \eta_{y_{2c1}|v}^{(1)} \times \eta_{y_{2c2}|s}^{(2)}. \quad (3)$$

Thus, the estimation of $f_{\tilde{Y}_c}(\tilde{y}_c)$ requires summing over all the possible $k_1 \times k_2$ configurations of the vector \mathbf{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 Hidden Markov models literature and implemented through suitable matrix notation (Bartolucci 2006; Zucchini and 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. The proposed model formulation allows to separately analyze the two main mobility concepts, namely, *structural* and *exchange* mobility. These two notions, though inherited from sociology, are frequently adopted in the economic literature (Markandya 1982; Schluter and Van de Gaer 2011). Structural mobility is captured by changes in the classes' expectations of the manifest variables, and by the change in the number of classes over generations. Therefore, the transition matrix depicts exchange mobility entailing the reranking mechanism at work from one generation to another. Furthermore, the initial probabilities $\pi_v, v = 1, \dots, k_1$, 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

$$\delta_s = P(S_c = s) = \sum_{v=1}^{k_1} \pi_v \times \pi_{s|v},$$

for $s = 1, \dots, k_2$. Interestingly, comparisons of exchange mobility between two distinct societies may be obtained by the differences in their transition matrices. If the number of classes were to differ between countries, testing the difference in each single transition probability would be unfeasible. However, mobility indices summarizing overall transition patterns into a scalar are available in the literature (e.g., Anderson 2018). Some caution would be required if state-dependent distributions were to differ significantly. Nonetheless, the study of class definitions and compositions is relevant itself, in a comparative perspective, to highlight differences in social classes among countries. The same considerations hold for a given country observed over time.

2.2.2 Including categorical covariates in the latent model

The model allows in a simple way to encompass the role of explanatory variables in the formation of social classes and the dynamics among them. Suppose that for each pair c we also observe a categorical variable, G_c , which identifies L subgroups in the sample of parents-child pairs, that is, $G_c = 1, \dots, L$. In the United States, for instance, one of the main covariates that must be taken into account is given by the individuals' ethnicity and race. We can model the initial and transition probabilities conditional on belonging to group g , namely,

$$\begin{aligned} \pi_v^g &= P(F_c = v | G_c = g), \\ \pi_{s|v}^g &= P(S_c = s | F_c = v, G_c = g). \end{aligned}$$

The δ_s^g parameters may be retrieved accordingly. The same set up stems from the inclusion of two or more categorical covariates in the model, where the L groups are identified by the interaction of all the covariates.

2.3 Maximum likelihood estimation

Under the main formulation presented in Section 2.2.1, assuming independence of n sample units, the log-likelihood of the model may be expressed in the following way:

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

where θ is the set of all the model parameters. The number of free parameters is equal to $k_1 - 1$ in the initial distribution, $k_2(k_1 - 1)$ in the transition distribution,⁵ $2(k_1 + k_2)$ for the Gaussian densities, and $(k_1 + k_2)(c - 1)$ for the categorical responses. The function in Eq. 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 Eq. 4 by maximizing the *complete data* log-likelihood (CDLL).

⁵These numbers increase to $L(k_1 - 1)$ and $Lk_2(k_1 - 1)$ if we include G_c in the model, and parameters $(\pi_v^g, \pi_{s|v}^g)$ replace $(\pi_v, \pi_{s|v})$ in the log-likelihood function.

A complete description of the estimation algorithm is provided in Appendix A in supplementary material. In Appendix B in supplementary material we provide the modifications to the EM algorithm required for the extended setting outlined in Section 2.2.2.

A crucial issue is the choice of the number of latent classes. A common approach to determine the number of states is based on the Bayesian's Information Criterion (Schwarz 1978, BIC;), which is expressed as $BIC = -2\hat{\ell} + p \log(n)$, where p is the number of nonredundant parameters. In words, the BIC is a penalized version of the maximum log-likelihood where the penalty term depends on the model complexity (i.e., the number of parameters), and it is known to perform well in the context of LMMs (e.g., Bartolucci and Farcomeni 2009). The model to be selected is the one with the smallest BIC.

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 further 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 and Mazumder 2011; Mazumder 2014). As standard in intergenerational mobility studies, the units of observation are pairs of individuals linked across generations. Following Bhattacharya and Mazumder (2011), in our analysis we exclude daughters, to avoid labor force participation issues, and focus on father-son pairs.⁶ 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 1722 men (sons) who were 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.⁷ 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 for censored observations would require some modifications to the model log-likelihood, and thus, to the estimation algorithm (McLachlan and Jones 1988; Atkinson 1992; Lee and Scott 2012). However, we resort to an alternative strategy. In particular, taking time averages as a measure of individual income reduces the

⁶In 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.

⁷In 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.

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), and 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 US 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.⁸ 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;⁹ 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 and 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 number of observations for two different categories, we selected the most recent one. Then, the ISCO categories were reduced to the afore-mentioned skill-based three-group categorization.

Table 1 shows the descriptive statistics for the variables of the analysis.

The gap in the average family income in the two generations is due to the structural growth of the US economy in the period of analysis as well as 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 and Solon (2006) sets the optimal age between 35 and 45 years old for the United States. The evolution of the occupational

⁸This 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.

⁹We use family income and fathers' occupation to have comparable variables with the sons' generation. In a previous version of this paper, for those respondents whose father's occupation is not available, we used mother's occupation. Results remained essentially unchanged.

Table 1 Summary statistics ($n = 1, 722$)

	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. Income figures are converted into constant 2000 US dollars using the OECD consumer price index. The middle panel reports the sample racial/ethnic composition, in percentage points. The three categories are mutually exclusive, namely, each individual belongs to a single ethnic group. The bottom panel reports the distribution of the occupational variable in the two generations, in percentage points

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

In this section we present the results of the empirical application. We recall that the assumptions of our model can be summarized as follows:

$$\begin{cases} P(\mathbf{U}_c = \mathbf{u}) = \pi_v \times \pi_{s|v} \\ f_{\tilde{\mathbf{Y}}_c|\mathbf{U}_c}(\tilde{\mathbf{y}}_c|\mathbf{u}) = \phi(y_{1c1}; \mu_v, \sigma_v^2) \times \phi(y_{1c2}; \xi_s, \tau_s^2) \times \eta_{y_{2c1}|v}^{(1)} \times \eta_{y_{2c2}|s}^{(2)} \end{cases}$$

where $\mathbf{u} = (v, s)$. We discuss the selection of the number of latent classes and we comment on class composition, state-dependent distributions and transition patterns. Finally, we provide an analysis of mobility comparison among racial and ethnic groups, where the latent process is assumed to be group-specific with probability distribution $P(\mathbf{U}_c = \mathbf{u}|G_c = g) = \pi_v^g \times \pi_{s|v}^g$.

3.2.1 Main formulation

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 $k_2 = \{2, 3, 4, 5\}$. According to the BIC, a suitable number of latent states is $k_1 = 3$ and $k_2 = 4$, as shown in Table 2. Due to the problem of possible multimodality of the likelihood function, we estimate the model 15 times by randomly

Table 2 Model selection

k_1/k_2	2	3	4	5
2	40,625.48	40,279.23	40,251.52	40,261.20
	[-20,241.95]	[-20,046.47]	[-20,010.27]	[-19,992.75]
	(19)	(25)	(31)	(37)
3	40,540.31	40,149.96	40,118.04	40,135.52
	[-20,177.01]	[-19,955.76]	[-19,913.72]	[-19,896.38]
	(25)	(32)	(39)	(46)
4	40,544.04	40,151.66	40,124.48	40,150.90
	[-20,156.53]	[-19,930.53]	[-19,887.14]	[-19,870.54]
	(31)	(39)	(47)	(55)

Notes: The table reports for each combination of (k_1, k_2) , listed by row and column respectively, the value of the Bayesian information criterion, the maximum of the log-likelihood function (square brackets), and the number of free parameters of the model (brackets). The values of the selected model are shown in bold

selecting the starting values of the algorithm. We are confident that the solution is the global maximum of the likelihood function, as it occurs 14 out of 15 times up to negligible differences in the Gaussian densities' means and variances. Standard errors for the estimates are computed with a non-parametric bootstrap.

Table 3 reports the estimated class sizes (π_v and δ_s , $v = 1, \dots, 3$, $s = 1, \dots, 4$) along with the means and standard deviations of the income distributions (μ_v , δ_v , ξ_s and τ_s), whereas Fig. 2 plots the kernel density estimates of the income distributions 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. This is particularly true when the number of classes varies over generations as in our application.¹⁰ 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. In the second generation, the top class is composed of "super rich" individuals as the average of their component is well above the overall average, and the high variance of the same component captures the positive skewness of the observed distribution.

In Fig. 3, we plot for each generation the estimated probabilities of belonging to each occupational category given each latent state ($\eta_{y|v}^{(1)}$ and $\eta_{y|s}^{(2)}$).

Bottom classes are associated with a high probability of being employed in a low-skilled occupation. In contrast, top classes are associated with a high probability of being employed in a high-skilled occupation. In the second generation, the third and fourth classes' occupational composition is similar, as both classes are mainly composed of high-skilled workers.

¹⁰The selection of the model with $(k_1, k_2) = (3, 4)$ classes may be a consequence of the top-coding rule affecting 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.

Table 3 Class sizes and Gaussian densities' parameter estimates

Panel A: Fathers				
v	1	2	3	–
π_v	0.411 (0.059)	0.462 (0.042)	0.126 (0.043)	
μ_v	28.943 (2.486)	56.767 (2.972)	97.234 (11.447)	
σ_v	13.793 (1.311)	18.138 (1.607)	37.399 (2.706)	
Panel B: Adult sons				
s	1	2	3	4
δ_s	0.325 (0.060)	0.419 (0.036)	0.208 (0.060)	0.047 (0.017)
ξ_s	28.113 (2.944)	57.851 (4.738)	96.340 (12.173)	223.637 (42.617)
τ_s	12.880 (1.502)	17.976 (1.743)	32.216 (5.932)	82.579 (17.773)

Notes: Panel A shows the estimated size (π_v), the mean (μ_v) and the standard deviation (σ_v) of the Gaussian density of each class v in thousand dollars, for fathers. Panel B shows the estimated size (δ_s), the mean (ξ_s) and the standard deviation (τ_s) of the Gaussian density of each class s in thousand dollars, for adult sons. Standard errors (in parentheses) are computed with a non-parametric bootstrap with 100 replications

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 presented in Table 4.

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

Fathers	Adult sons			
	First	Second	Third	Fourth
First	0.653 (0.100)	0.240 (0.086)	0.094 (0.028)	0.013 (0.012)
Second	0.112 (0.058)	0.605 (0.088)	0.260 (0.104)	0.023 (0.019)
Third	0.036 (0.033)	0.324 (0.124)	0.391 (0.137)	0.250 (0.099)

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

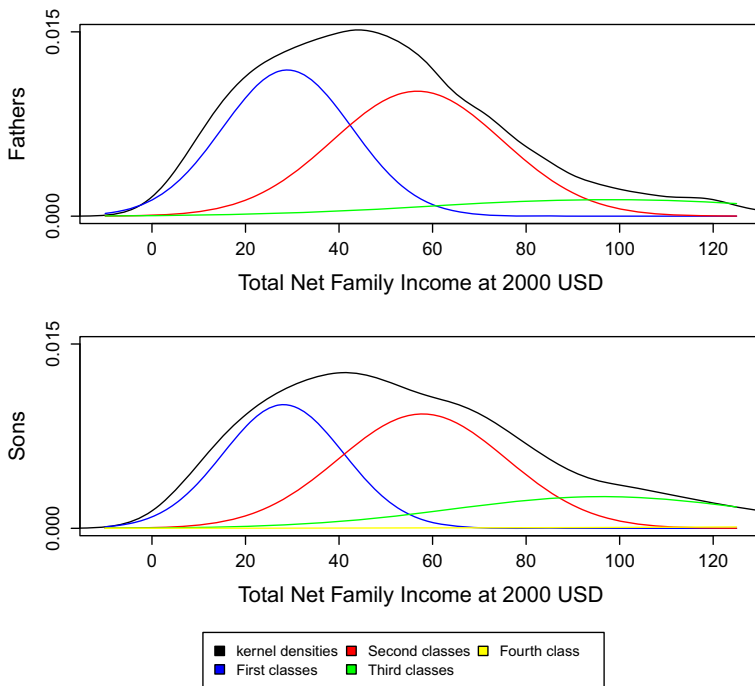


Fig. 2 State-dependent distributions: Income. *Notes:* The figure shows the kernel density estimates and the components' estimated scaled Gaussian densities of the income variable (in thousand US dollars) in both generations. Right tails truncated at USD 125,000

There are several ways to analyze transition patterns in Π , starting from *absolute* mobility measures.¹¹ We look at upward mobility as the probability of reaching the top class for sons born to fathers in the lowest class, which is a measure 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 equal to 0.013 (i.e., 1.3 percentage points). On the other hand, downward mobility is represented by the probability of moving to the bottom class for sons born to fathers in the top class. This value corresponds to the value at the lower-left corner of the transition matrix, and it is equal to 3.6 percentage points. These results provide evidence of a low level of mobility. 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) provides additional significant information: The persistence at the

¹¹ 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-status families relative to those of adult children from high-status families. Absolute mobility measures identify the outcomes of adult children from families of a given status in absolute terms.

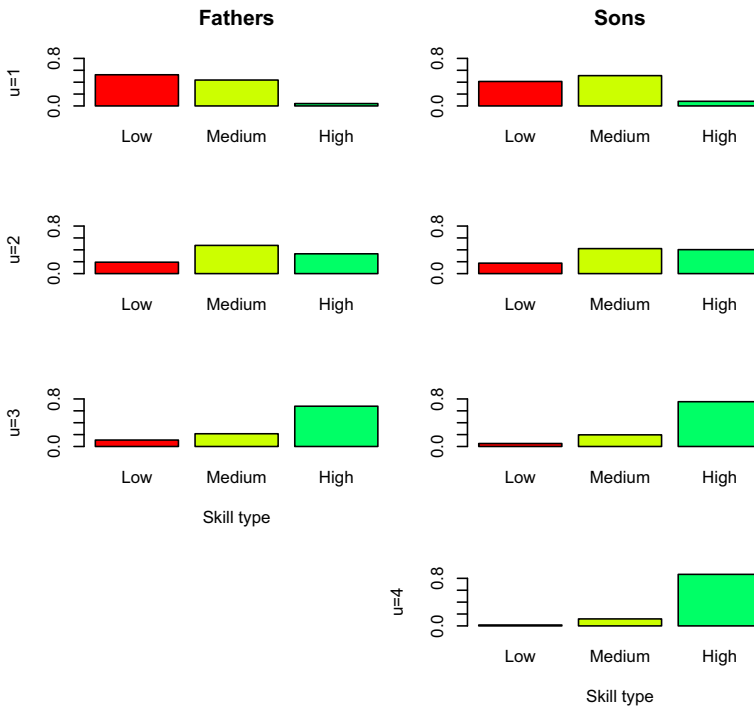


Fig. 3 State-dependent distributions: Occupation. *Notes:* The figure shows the probabilities of belonging to each occupational category (1: low-skilled, 2: medium-skilled, 3: high-skilled), conditional on each latent state, for both generations. Latent states are ordered by row

bottom is considerably higher than the persistence at the top, showing that a poverty trap may emerge.¹²

3.2.2 Race and ethnicity: mobility comparison

The ethnic and racial differences in the income distribution in the US have well-known foundations originating in the last centuries. These inequalities may seriously hamper the process of development if they are persistent over generations. Estimates of intergenerational mobility by ethnic group (or race) provide insights into whether racial differences in the United States are likely to be eliminated, and, if so, how long it might take. Other studies have analyzed the role that racial or ethnic origins play in terms of intergenerational mobility; see for example Mazumder (2014) and Bhattacharya and Mazumder (2011) who use NLSY data adopting different methodologies.¹³ The authors conclude that, in recent decades, black individuals have experienced substantially less (more) upward (downward) intergenerational mobility than white individuals.

¹²This 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.

¹³They adopt transition probabilities of relative income status and measures of *directional rank mobility*.

Table 5 Average posterior probabilities

Panel a: Fathers				
v	1	2	3	–
π_v^g				
White (w)	0.276 (0.056)	0.552 (0.043)	0.172 (0.055)	
Non-white (nw)	0.616 (0.062)	0.326 (0.052)	0.057 (0.025)	
$\pi_v^w - \pi_v^{nw}$	-0.341*** (0.022)	0.226 *** (0.041)	0.115*** (0.033)	
Panel b: Adult sons				
s	1	2	3	4
δ_s^g				
White (w)	0.221 (0.053)	0.470 (0.049)	0.247 (0.073)	0.062 (0.020)
Non-white (nw)	0.482 (0.073)	0.343 (0.039)	0.149 (0.044)	0.026 (0.014)
$\delta_s^w - \delta_s^{nw}$	-0.261*** (0.030)	0.127** (0.058)	0.098*** (0.035)	0.036*** (0.011)

Significance levels: *** $p < .01$ ** $p < .05$ * $p < .10$

Notes: The table shows the average posterior probability of belonging to each latent class, conditional on sons’ race/ethnicity. Row 3 of panels *a* and *b* show the differences in the probability of belonging to each class between white and non-white individuals

African-American and Hispanic individuals represent about 46% of our sample (Table 1). In the following, we partition the sample into $L = 2$ groups, based on sons’ racial or ethnic origins: white (w) and non-white (nw) individuals, where non-whites are Afro-american and Hispanic 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. Now, π_v^g may be interpreted as the probability of belonging to class v for a father whose son belongs to group g , with $g = w, nw$. 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 .

A substantial difference exists in terms of the class compositions. In particular, most of non-white sons belong, on average, to the bottom classes (Table 5, panel *b*).

For instance, the probability of belonging to the bottom class is equal to 0.221 for white individuals, which is approximately 26 percentage points lower than the same probability for non-white individuals. This difference is statistically significant at the 1% level. On the other hand, white individuals have higher probabilities of belonging to class two, three and four (12.7, 9.8 and 3.6 percentage points, respectively). The same results hold for their fathers (panel *a*, third row).

Differences in transition patterns are less clear cut. Table 6 shows the transition probabilities for white individuals, while Table 7 displays the differences in these parameters among white and non-white individuals. For instance, there is a significant difference in the persistence in the lowest class. This probability is equal to 0.595 for white individuals, and

Table 6 Estimated transition matrix for white individuals

Fathers	White adult sons			
	First	Second	Third	Fourth
First	0.595 (0.110)	0.281 (0.093)	0.106 (0.033)	0.018 (0.017)
Second	0.094 (0.052)	0.609 (0.091)	0.273 (0.106)	0.024 (0.021)
Third	0.029 (0.027)	0.326 (0.125)	0.393 (0.139)	0.252 (0.098)

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

it is almost 10 percentage points lower than the corresponding probability for nonwhites (Table 7, top-left corner). White individuals also show a lower propensity to transition from the second to the lowest class (-0.062), whereas their probability of moving from the lowest to the second class is higher (0.068). On the other hand, persistence at the top is similar. It corresponds to the propensity of moving from the third to the fourth class, and it is equal to 0.252 for whites and to 0.239 for non-white individuals, where this difference is not statistically significant (Table 7, bottom-right corner). Upward mobility of white individuals is almost twice the upward mobility of non-white individuals, whereas downward mobility rate of non-white individuals is more than twice that of white individuals, although these values are low and not statistically different (top-right and bottom-left corners of Tables 6 and 7, respectively).

To provide formal evidence on the overall transition patterns, we introduce a set of summary measures of mobility, and we test the differences in these measures computed for the two groups. A mobility index is a function $M(\Pi)$ mapping Π into a scalar (Formby et al.

Table 7 Differences in transition probabilities among white and non-white individuals ($\pi_{s|v}^w - \pi_{s|v}^{nw}$)

Fathers	Adult sons			
	First	Second	Third	Fourth
First	-0.098*** (0.033)	0.068*** (0.025)	0.021 (0.015)	0.009 (0.009)
Second	-0.062** (0.025)	0.014 (0.031)	0.043** (0.019)	0.005 (0.009)
Third	-0.039 (0.036)	0.015 (0.052)	0.011 (0.057)	0.013 (0.061)

Significance levels: *** $p < .01$ ** $p < .05$ * $p < .10$

Notes: The table shows the differences in the estimated transition probabilities among white and non-white individuals. A negative (positive) figure denotes a higher probability for the non-white (white) individuals. Standard errors (in parentheses) are computed with a non-parametric bootstrap with 100 replications

2004). Several indices are available to researchers¹⁴ (Fields and Ok 1996; Checchi and Dardanoni 2002). 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 operators returning the maximum and minimum value, respectively, of the s th column of Π . The A index satisfies the normalization, immobility and perfect mobility axioms (Shorrocks 1978).¹⁵ 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. The 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 one for transitions to subsequent classes, equal to two for 2-step transitions, 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. Nonetheless, their interpretation is straightforward. They are bounded between 0 and 1, and the higher is their values, the greater is the mobility level of the matrix.

Table 8 reports the mobility indices computed on the estimated $\Pi_{3 \times 4}$ and matrices of the two groups. The point estimates suggest that the overall degree of mobility is higher for white individuals, and this holds true for all the considered indices. In particular, we have that the Anderson index equals 0.492 in the white group and 0.476 in the non-white group. However, these differences are not statistically different from zero (the last column of the table illustrates the test results). Overall, these results provide: (i) strong evidence of a racial/ethnic gap in class compositions, and (ii) mild evidence of a racial/ethnic gap in transition patterns.

4 Conclusions

This paper aims at proposing a new approach based on a multivariate framework to the study of intergenerational mobility. The adopted model may be cast in the class of latent

¹⁴In our framework, the unequal dimensions of the matrix imply that perfect immobility cannot be characterized by an identity matrix. Thus, not all the standard indices may be applied.

¹⁵This index is also suited for studies on mobility comparison across countries and over time, as it may be computed irrespective of the matrix dimensions.

Table 8 Mobility comparison among racial or ethnic groups

	Whole sample	White	Non-white	Group difference
<i>A</i>	0.474 (0.055)	0.492 (0.054)	0.476 (0.062)	0.016 (0.033)
<i>B</i> ₁	0.155 (0.028)	0.163 (0.028)	0.144 (0.031)	0.020 (0.014)
<i>B</i> ₂	0.387 (0.013)	0.390 (0.015)	0.382 (0.011)	0.008 (0.009)

Significance levels: *** $p < .01$ ** $p < .05$ * $p < .10$

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 (ethnic) groups. The last column (Group difference) reports the differences of the indices computed on the group-specific matrices. Standard errors (in parentheses) are computed with non-parametric bootstrap with 100 replications

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. 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 show that the level of upward and downward mobility in the US is low. Moreover, non-white individuals show a higher persistence at the bottom of the distribution. We further look to the more comprehensive mobility indices encompassing all the movements among latent classes. We find that overall mobility is not statistically different among the two groups. However, a significant discrepancy is found in terms of class belonging.

Given its characteristics, the model may be also applied to compare mobility across societies. The resulting latent transition matrices may be adopted in comparative studies even though the number of latent classes may differ between countries. While it would be impossible to test the difference in each single transition probability, due to non-comparable class structures, comparable mobility indices are available in the literature. Future possible extensions of the model involve a more structured parametrization of initial and transition probabilities, with the inclusion of continuous covariates affecting the mobility process (e.g., years of schooling). Finally, future work might aim to adapt the model to study intragenerational mobility and the transition between social classes over the individual life-cycle.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10888-022-09554-6>.

Data Availability The data analyzed in this study is publicly available at the following link: <https://www.bls.gov/nls/nlsy79.htm>

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence,

and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adermon, A., Lindahl, M., Palme, M.: Dynastic human capital, inequality and intergenerational mobility. *Amer. Econ. Rev.* **111**, 1523–48 (2021)
- Anderson, G.: Measuring aspects of mobility, polarization and convergence in the absence of cardinality: Indices based upon transitional typology. *Soc. Indic. Res.* **139**(3), 887–907 (2018)
- Anderson, G., Pittau, M.G., Zelli, R.: Assessing the convergence and mobility of nations without artificially specified class boundaries. *J. Econ. Growth* **21**, 283–304 (2016)
- Anderson, G., Farcomeni, A., Pittau, M.G., Zelli, R.: A new approach to measuring and studying the characteristics of class membership: Examining poverty, inequality and polarization in urban China. *J. Econ.* **191**, 348–356 (2016)
- Anderson, G., Farcomeni, A., Pittau, M.G., Zelli, R.: Rectangular latent Markov models for time-specific clustering, with an analysis of the well being of nations. *J. R. Stat. Soc. (Series C)* **68**, 603–621 (2019)
- Arcidiacono, P., Jones, J.B.: Finite mixture distributions, sequential likelihood and the EM algorithm. *Econometrica* **71**, 933–946 (2003)
- Atkinson, S.E.: The performance of standard and hybrid em algorithms for ml estimates of the normal mixture model with censoring. *J. Stat. Comput. Simul.* **44**, 105–115 (1992)
- Bartholomew, D.J. *Stochastic Models for Social Processes*, 3rd edn. Wiley, London (1982)
- Bartolucci, F.: Likelihood inference for a class of latent Markov models under linear hypotheses on the transition probabilities. *J. R. Stat. Soc.: Series B (Stat. Methodol.)* **68**, 155–178 (2006)
- Bartolucci, F., Farcomeni, A.: A multivariate extension of the dynamic logit model for longitudinal data based on a latent Markov heterogeneity structure. *J. Am. Stat. Assoc.* **96**, 816–831 (2009)
- Bartolucci, F., Farcomeni, A., Pennoni, F.: *Latent Markov Models for Longitudinal Data*. Chapman and Hall/CRC Press, Boca Raton (2013)
- Baum, L., Petrie, T., Soules, G., Weiss, N.: A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164–171 (1970)
- Bhattacharya, D., Mazumder, B.: A nonparametric analysis of black–white differences in intergenerational income mobility in the United States. *Quant. Econ.* **3**, 335–379 (2011)
- Björklund, A., Jäntti, M.: Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. *Res. Soc. Stratif. Mobil.* **70**, 100455 (2020)
- Black, S., Devereux, P.: *Recent Developments in Intergenerational Mobility*, 1st edn., vol. 4B, chapter 16, pp. 1487–1541. Elsevier (2011)
- Bollen, K.A.: *Structural Equations with Latent Variables*. Wiley, New York (1989)
- Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Jäntti, M., Eriksson, T., Österbacka, E.: Nonlinearities in intergenerational earnings mobility: Consequences for cross-country comparisons. *Econ. J.* **117**, C72–C92 (2007)
- Cameron, S., Heckman, J.J.: The dynamics of educational attainment fo Black, Hispanic, and White Males. *J. Polit. Econ.* **109**, 455–499 (2001)
- Cecchi, D., Dardanoni, V.: Mobility comparisons: Does using different measures matter? *Research on Economic Inequality*, 9 (2002)
- Chetty, R., Hendren, N., Kline, P., Saez, E.: Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quart. J. Econ.* **129**, 1553–1623 (2014)
- Collins, L.M., Lanza, S.T.: *Latent class and latent transition analysis: with applications in the social, behavioral, and health sciences*. Wiley, Hoboken (2010)
- Corak, M., Heisz, A.: The intergenerational earnings and income mobility of Canadian men: Evidence from longitudinal income tax data. *J. Hum. Resour.* **34**, 504–556 (1999)
- Corak, M., Lindquist, M.J., Mazumder, B.: A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Econ.* **30**, 185–200 (2014)
- Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.: Series B (Stat. Methodol.)* **39**, 1–38 (1977)

- Dotto, F., Farcomeni, A., Pittau, M.G., Zelli, R.: A dynamic inhomogeneous latent state model for measuring material deprivation. *J. R. Stat. Soc.: Series A (Stat. Soc.)* **182**, 495–516 (2018)
- Eide, E., Showalter, M.: Factors affecting the transmission of earnings across generations: A quantile regression approach. *J. Human Resour.* **34**(2), 253–267 (1999)
- Erikson, R., Goldthorpe, J.: *The Constant Flux: A Study of Class Mobility in Industrial Countries*. Oxford University Press, New York (1992)
- Fields, G., Ok, E.: The meaning and measurement of income mobility. *J. Econ. Theory* **71**(2), 349–377 (1996)
- Flachaire, E., Núñez, O.: Estimation of the income distribution and detection of subpopulations: An explanatory model. *Comput. Stat. Data Anal.* **51**, 3368–3380 (2007)
- Fletcher, J., Han, J.: Intergenerational mobility in education: Variation in geography and time. *J. Hum. Cap.* **13**, 585–634 (2019)
- Formby, J.P., Smith, W.J., Zheng, B.: Mobility measurement, transition matrices and statistical inference. *J. Econ.* **120**, 181–205 (2004)
- Goodman, L.A.: Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61**, 215–231 (1974)
- Grawe, N.: Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints. *J. Hum. Resour.* **39**, 813–827 (2004)
- Haider, S., Solon, G.: Life-cycle variation in the association between current and lifetime earnings. *Amer. Econ. Rev.* **96**, 1308–1320 (2006)
- Halliday, T.J., Mazumder, B., Wong, A.: The intergenerational transmission of health in the United States: A latent variables analysis. *Health Econom* **29**, 367–381 (2020)
- Jäntti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Österbacka, E., Bjorklund, A., Erikson, T.: 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 (2006)
- Jäntti, M., Jenkins, S.: Income Mobility. IZA Discussion Papers 7730 (2013)
- Jöreskog, K.G., Goldberger, A.S.: Estimation of a model with multiple indicators and multiple causes of a single latent variable. *J. Amer. Stat. Assoc.* **70**(351a), 631–639 (1975)
- Krishnakumar, J.: Multidimensional measures of poverty and well-being based on latent variable models. In: Kakwani, N., Silver, J. (eds.) *Quantitative Approaches to Multidimensional Poverty Measurement*, Chapter 7, pp. 118–134. Palgrave Macmillan, London (2008)
- Krishnakumar, J., Nagar, A.L.: On exact statistical properties of multidimensional indices based on principal components, factor analysis, MIMIC and structural equation models. *Soc. Indic. Res.* **86**, 481–496 (2008)
- Lazarsfeld, P.F., Henry, N.W.: *Latent Structure Analysis*. Houghton Mifflin, Boston (1969)
- Lee, G., Scott, C.: EM algorithms for multivariate Gaussian mixture models with truncated and censored data. *Comput. Stat. Data Anal.* **56**, 2816–2829 (2012)
- Li Donni, P., Rodriguez, J., Rosa Dias, P.: Empirical definition of social types in the analysis of inequality of opportunity: A latent classes approach. *Soc. Choice Welf.* **44**, 673–701 (2015)
- Lindsay, B., Clogg, C.C., Grego, J.: Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *J. Am. Stat. Assoc.* **86**, 96–107 (1991)
- Long, J., Ferrie, J.: Intergenerational occupational mobility in Great Britain and the United States since 1850. *Amer. Econ. Rev.* **103**, 1109–1137 (2013)
- Lubotsky, D., Wittenberg, M.: Interpretation of regressions with multiple proxies. *Rev. Econ. Stat.* **88**, 549–562 (2006)
- Markandya, A.: Intergenerational exchange mobility and economic welfare. *Eur. Econ. Rev.* **17**, 307–324 (1982)
- Mazumder, B.: Black–white differences in intergenerational economic mobility in the United States. *Econ. Perspect.*, 38 (2014)
- Mazumder, B., Acosta, M.: Using occupation to measure intergenerational mobility. *The ANNALS of the American Academy of Political and Social Science* **657**, 174–193 (2015)
- McLachlan, G.J., Jones, P.N.: Fitting mixture models to grouped and truncated data via the EM algorithm. *Biometrics* **44**, 571–578 (1988)
- McLachlan, G.J., Peel, D.: *Finite mixture models*. Wiley, New York (2000)
- Modalsli, J.: Multigenerational persistence: evidence from 146 years of administrative data. *Journal of Human Resources*, 1018–9825R2 (2021)
- Mroz, T.A.: Discrete factor approximation in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *J. Econ.* **92**, 233–274 (1999)
- O’Neill, D., Sweetman, O., Van de Gaer, D.: The effects of measurement error and omitted variables when using transition matrices to measure intergenerational mobility. *J. Econ. Inequal.* **5**, 159–178 (2007)

- Schluter, C., Van de Gaer, D.: Upward structural mobility, exchange mobility, and subgroup consistent mobility measurement: U.S.-German mobility rankings revisited. *Rev. Income Wealth* **57**, 1–22 (2011)
- Schwarz, G.: Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978)
- Shorrocks, A.F.: The measurement of mobility. *Econometrica* **46**, 1013–1024 (1978)
- Solon, G.: Intergenerational income mobility in the United States. *Amer. Econ. Rev.* **82**, 393–408 (1992)
- Tullio, F., Bartolucci, F.: Causal inference for time-varying treatments in latent Markov models: An application to the effect of remittances on poverty dynamics. *Ann. Appl. Stat.* **16**, 1962–1985 (2022)
- Vosters, K.: Is the simple law of mobility really a law? Testing Clark's hypothesis. *Econ. J.* **128**, F404–F421 (2018)
- Vosters, K., Nybom, M.: Intergenerational persistence in latent socioeconomic status: Evidence from Sweden and the United States. *J. Labor Econ.* **35**, 869–901 (2017)
- Welch, L.R.: Hidden Markov models and the Baum-Welch algorithm. *IEEE Inform. Theory Soc. Newslett.* **53**, 1–13 (2003)
- Zimmerman, D.: Regression toward mediocrity in economic stature. *Amer. Econ. Rev.* **82**, 409–29 (1992)
- Zucchini, W., MacDonald, I.L.: *Hidden Markov Models for Time Series: An Introduction Using R*. Chapman and Hall/CRC press, Boca Raton (2009)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.