Rivista Italiana di Economia Demografia e Statistica *Volume LXXIX n.1 Gennaio-Marzo 2025*

DOES MEASUREMENT ERROR BIAS EMPLOYMENT PATHWAYS? THE CASE OF ITALY[1]

Dimitris Pavlopoulos, Roberta Varriale, Silvia Loriga

Abstract. The exploration of employment trajectories over time may be significantly biased due to measurement errors in the data used for the analysis. This paper addresses this issue by employing a mixture hidden Markov model (MHMM) that detects and corrects for measurement errors. Specifically, we use an MHMM that includes two indicators for employment status, derived from linked data from the Italian Labour Force Survey and Administrative Data for the period 2017-2021.

1. Introduction

In recent years, the processual approach has been gaining momentum in labour market research (Abbott, 1983). Instead of 'merely' studying single transitions, a process-based method allows us to examine entire sequences and has the advantage of being able to provide a complete and detailed picture of entire trajectories, e.g. in the field of working careers (Aisenbrey and Fasang, 2017; Mattijssen and Pavlopoulos, 2019; Studer and Ritschard, 2016). A potential threat to this research is measurement error that exists in socio-economic data. In data coming from surveys, measurement error may be caused by cognitive processes, social desirability, design and implementation of the survey (Groves, 2004; Sudman *et al.*, 1996; Tourangeau *et al.*, 2000). In administrative data, measurement error can result from the misalignment between the definition of the administrative variable and the target one, administrative delays, wrong registration, or erroneous administrative procedures (Bakker and Daas, 2012; Oberski *et al.*, 2017). Research has shown that the measurement error can have severe effects on our estimates of mobility over time. Even a small magnitude of measurement error can bias considerably transition rates from unemployment to employment (Filipponi *et al.*, 2021) or from temporary employment to permanent employment (Pavlopoulos and Vermunt, 2015). The use of hidden Markov models (HMMs) is a promising approach for detecting and correcting measurement error in categorical longitudinal data (Biemer, 2011). The advantages of HMMs

[1] Any opinions and conclusion expressed are those of the authors and do not necessarily respect the views of the Italian national institute of statistics.

are that they are very flexible and do not assume the existence of a primary (error-free) source of information among those available (Filipponi et al., 2021). HMMs treat the 'true' employment status as a latent variable that is imperfectly measured by one or more observed indicators. Using linked data from different sources allows modelling realistic error specifications with an HMM. In more detail, when multiple data sources are available, HMMs can model both random and systematic measurement error. An extension of HMM is the mixture hidden Markov model (MHMM) that estimates the true employment status trajectories (formally: mixtures) (Vermunt, 2010). In this paper, we investigate the effect of measurement error on employment trajectories in Italy for the period 2017-2021. Furthermore, we aim to produce an error-corrected estimate of these trajectories. For this purpose, we apply a MHMM with two independent indicators for employment status. These indicators come respectively from the Italian Labour Force Survey (LFS) and Administrative Data (AD) managed by the Italian National Statistical Institute (Istat).

The remaining of this paper is organized as follows. In section 2, we describe the data that are used in the analysis. In section 3, we build the MHMM that is employed, and in section 4, we discuss the conclusions of our findings as well as the steps for further research.

2. The data

The Italian Labour Force Survey (LFS) follows the standards set by EU Regulation 2019/1700 of the European Parliament and the Council. LFS's primary goal is to provide employment statistics. LFS is a continuous household survey, conducted throughout the year and covers approximately 1.2% of the entire Italian population. The sample is selected according to a two-stage stratified sampling design: municipalities represent the primary statistical units and are stratified based on their population size; the final statistical units are the households, which are selected through a simple random sampling. The Italian LFS adopts a quarterly rotation scheme. Selected households are interviewed four times within a 15-months period. Each household is interviewed for two consecutive quarters, followed by a two-quarter break and then two consecutive survey quarters. Interviews are spread across all weeks of the quarter (reference weeks), maintaining the sample representativeness on monthly basis. The information on labour market participation refers to the reference week. For further details on the LFS contents, methodologies and organization see (Istat, 2006). Italian Administrative Data (AD) relevant to labour statistics are collected by the social security and tax authorities and are gathered and appropriately processed by Istat. These data go through different, source-specific editing and harmonization procedures (Baldi *et al.*, 2018).

Our sample in this study refers to the LFS cohorts that entered the survey from January 2017 to December 2021. From the LFS, information from all available survey waves in which these individuals participated is retained. The actual number of LFS observations in the data may be four or less than four in case of attrition or whether the LFS rotation scheme started before 2017 or ended after 2021. For the same set of individuals, quarterly information from the AD is retained, covering all quarters from January 2017 to December 2021. For each individual, there is a maximum of four observations from the LFS, whereas AD information are potentially available for all the twenty observations. The linkage between LFS and AD was conducted at the individual level through a pseudonymized code. We include, in our sample, individuals aged 25 to 55 who participated at least once in the LFS within this period. As our statistical model is computationally demanding, a 10% sample of units was randomly selected, stratifying the sample by the month of the first LFS interview. This procedure resulted in a sample of 39,847 individuals.

Our target variable on employment status is an aggregation of the International Classification of Status in Employment (ICSE-18), established by the International Labour Office and takes 4 values: (1) employees with a permanent contract (PE), (2) employees with a temporary contract (TE), (3) self-employed (SE), which encompasses employers, independent workers without employees, contributing family workers and dependent contractors and (4) not employed (NE). The creation of such target variable according to this simplified classification is relatively straightforward for LFS data, given that ICSE criteria inspired the definitions of LFS variables. However, in the AD, such a derivation presented several challenges as it uses different sources. In this respect, the administrative classification may not fully align with the statistical classification implemented in the LFS, as it is based on administrative concepts; on the other hand, LFS variables, based on self-reported information, may be affected by errors due to the subjective interpretation of the question by the respondent (especially in case of proxy answer). A potential source of discrepancies between LFS and AD is the different coverage: LFS aims at covering all employed population, including informal work, while AD contains information only for formal employment. Additional conceptual discrepancies between LFS and AD can be attributed to shortcomings in the data collection process. These include e.g. temporal misalignment of sources, particularly for occasional employment and discrepancies in the definitions of work signals across available sources. Tables 1 and 2 display the transition rates between the different employment categories in adjacent quarters in the LFS and AD. Despite the abovementioned sources of discrepancies, the disparities between the datasets are relatively minor.

Table 1 – *Observed transitions in the Labour Force Survey. Years 2017-2021.*

	Employment status *t*			
Employment status *t-1*	PE	TE	SE	NE
Permanent contract	0.962	0.012	0.006	0.020
Temporary contract	0.074	0.739	0.013	0.174
Self-employed	0.015	0.010	0.944	0.031
Not employed	0.019	0.063	0.013	0.906

Source: Istat, Labour Force Survey

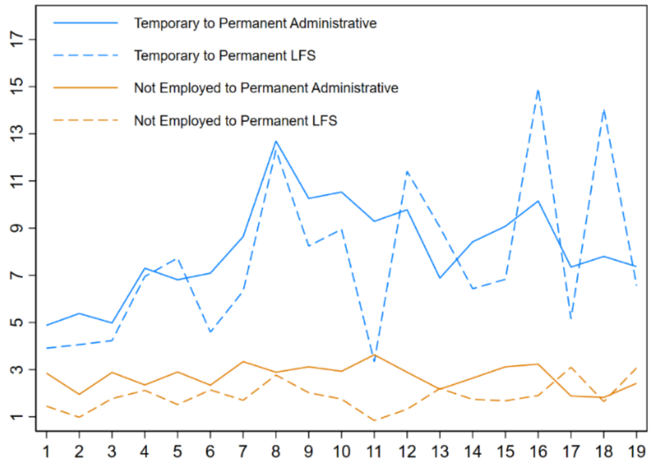
Table 2 – *Observed transitions in the Administrative Data. Years 2017-2021.*

	Employment status *t*			
Employment status *t-1*	PE	TE	SE	NE
Permanent contract	0.966	0.009	0.003	0.022
Temporary contract	0.078	0.717	0.017	0.187
Self-employed	0.011	0.012	0.956	0.022
Not employed	0.030	0.061	0.012	0.897

Source: Istat, Administrative Data

Figure 1 shows the observed transition rates between different employment categories across adjacent quarters from 2017 to 2021. The graph confirms the findings of Tables 1 and 2 that there are only minor differences in flow patterns between the LFS and AD. The largest transition rates occur from temporary to permanent employment, showing an increasing trend. This pattern is evident in both LFS and AD, suggesting that there is time dependence in transition probabilities.

Figure 1 – *Transition flows from type of contracts by quarter, years 2017-2021.*



Sources: Istat, Labour Force Survey and Administrative Data

Table 3 presents the cross-classification of employment status from LFS and AD data. The diagonal cells concern cases where the two data sources agree on the classification. In contrast, off-diagonal values represent discrepancies in classification and indicate potential classification errors in at least one of the data sources. As Table 3 illustrates, the two data sources do not align for approximately 14.6% of the total number of cases. Beyond random classification errors, these discrepancies arise from distinct reasons, as suggested by Varriale and Alfó (2023) in their analysis of employment status. Errors in AD are typically attributable to miss-specifications of statistical concepts. For example, AD lack information on irregular work, or it may encounter difficulties in correctly identifying the reference period of the information. On the other hand, errors in the LFS survey may arise from miss-classification due to respondents providing incorrect answers or having an erroneous understanding of employment categories.

Table 3 – *Cross-classification of employment status, Administrative Data and Labour Force Survey, frequencies and percentages in brackets. Years 2017-2021.*

	Employment status, LFS				
Empl. status AD	PE	TE	SE	NE	Total
Perm. contract	41326 (41.1)	1993 (2.0)	1203 (1.2)	1290 (1.3)	45812 (45.6)
Temp. contr.	1217 (1.2)	5442 (5.4)	298 (0.3)	1268 (1.3)	8225 (8.2)
Self-employed	748 (0.7)	349 (0.3)	11033 (11.0)	1095 (1.1)	13225 (13.2)
Not employed	1942 (1.9)	1394 (1.4)	1876 (1.9)	28066 (27.9)	33278 (33.1)
Total	45233 (45.0)	9178 (9.1)	14410 (14.3)	31719 (31.5)	100540 (100)

Sources: Istat, Labour Force Survey and Administrative Data

Table 4 presents the same cross classification as Table 3 but reports the percentage distribution of employment status as measured by the LFS, conditional on the AD measurement. 90.2% of cases that are recorded as permanent contracts in AD are also classified as permanent contracts according to the LFS.

Table 4 – *Distribution of employment status from Labour Force Survey conditional on Administrative Data measurement. Years 2017-2021.*

	Employment category, LFS				
Empl. status AD	PE	TE	SE	NE	Total
Perm. contract	90.2	4.4	2.6	2.8	100
Temp. contract	14.8	66.2	3.6	15.4	100
Self-empl.	5.7	2.6	83.4	8.3	100
Not empl.	5.8	4.2	5.6	84.3	100
Total	45.0	9.1	14.3	31.5	100

Sources: Istat, Labour Force Survey and Administrative Data

The relevant percentages of classification agreement for the self-employed and not employed are also quite high (83.4% and 84.3%, respectively). In these categories, the only case with a high percentage of misclassification concerns the cases that are recorded as self-employed in AD as 8.3% of them are classified as not employed in the LFS. Major classification mismatches are only observed for temporary contracts. In fact, only 66.2% of those recorded as having a temporary contract in AD are observed as having a temporary contract in the LFS, while about 14.8% are classified as having a permanent contract and 15.4% as not employed. These findings are rather stable over time.

3. The mixture hidden Markov model

Let A_{it} and S_{it} measure the observed indicators of the employment status of individual i at time t according to AD and LFS, respectively. We use quarterly data from 2017-2021, meaning that t runs from 0 to T = 19, while N equals 66,127. The true (latent) employment status is indicated by X_{it}. All A_{it}, S_{it}, and X_{it} can take on four values representing the categories of employment status (permanent contract, temporary contract, self-employment and non-employment); a particular category of the observed indicators is denoted by a_t, s_t, and the latent variable x_t.

In the MHMM, the joint probability of following a particular observed state path conditional on covariates V, t and t^2 is expressed as:

$$P(A_i = a_i, S_i = s_i | V_i, t) = \sum_{k=1}^{K} \sum_{x_0=1}^{4} \sum_{x_1=1}^{4} \dots \sum_{x_T=1}^{4} P(w_i = k) P(X_{i0} = x_0 | w_i = k)$$

$$\prod_{t=1}^{T} P\left(X_{it} = x_t | X_{i(t-1)} = x_{t-1}, w_i = k, t, t^2\right)$$

$$P(A_{i0} = c_0 | X_{i0} = x_0)$$

$$\prod_{t=1}^{T} P\left(A_{it} = a_t | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, A_{i(t-1)} = a_{t-1}\right)$$

$$\prod_{t=1}^{T} P(S_{it} = s_t | X_{it} = x_t, V_{it})^{\delta_{it}} \quad (1)$$

The latent employment status X_{it} follows a first-order Markov process. This means that the realization of the true employment status at time point t, X_{it}, is independent of all previous realizations of it conditional on its realization in the previous time point ($t - 1$).

The MHMM estimates four sets of probabilities: the initial state probabilities $P(X_{i0} = x_0)$, the latent transition probabilities $P(X_{it} = x_t | X_{i(t-1)} = x_{t-1})$ as well as the measurement error probabilities for the administrative data $P(A_{it} = a_t \vee X_{it} = x_t)$, and for the LFS $P(S_{it} = s_t | X_{it} = x_t)$. To model more realistic scenarios, the latent transition probabilities are assumed to be time-heterogeneous. In more detail, these probabilities depend on a quadratic specification of the quarter index (i.e. t and t^2), $P(X_{it} = x_t | X_{i(t-1)} = x_{t-1}, t, t^2)$). Finally, we estimate the probability of belonging to one of the K mixtures. Mixtures represent the employment trajectories as they are time-constant latent classes of individuals with similar initial latent state probabilities ($P(X_{i0} = x_0 \vee w_i = k)$) and latent transition probabilities ($P(X_{it} = x_t | X_{i(t-1)} = x_{t-1}, w_i = k, t, t^2)$).

The standard HMM maintains the Independent Classification Error assumption. This assumption means that the realizations of the observed states A_{it} (respectively, S_{it}) are independent of its previous and future realizations conditional on the value of the latent state X_{it}. However, this assumption is rather unrealistic as both survey and administrative employment data contain systematic measurement error (Pankowska *et al.*, 2021). To address this issue, we relax it by modelling error autocorrelation in both the administrative data and LFS indicator. In more detail, the measurement-error probabilities in LFS and administrative data are allowed to depend on the lagged observed and lagged latent employment status. As $X_{i(t-1)}$, $A_{i(t-1)}$ and $S_{i(t-1)}$ can take on 4 values, there are 16 ($4 * 4$) different sets of measurement-error probabilities in the administrative data and other 16 in the LFS. Each of these probabilities corresponds to each possible combination of lagged observed and latent employment status. To estimate a model with a meaningful number of probabilities as outcomes, we applied certain restrictions. For this purpose, we defined a constrained logit model that estimates one additional parameter when the same error can be made between adjacent time points. Similar restrictions have been previously applied in HMMs by others (Manzoni *et al.*, 2010; Pankowska *et al.*, 2021; Pavlopoulos and Vermunt, 2015).

The model is estimated with Maximum likelihood, while parameters are obtained using the forward-backward or Baum-Welch algorithm (Baum *et al.*, 1970), which is a variant of the Expectation-Maximization algorithm. Specifically, we use an extension of the Baum-Welch algorithm for MHMMs with covariates as described by Vermunt (2007). The estimation is done with the program Latent GOLD, version 6.0 (Vermunt and Magidson, 2016).

4. Results

Model selection involved three stages. The first stage aimed at selecting the model with measurement-error specifications that best fit the data. In the second stage, we compared model specifications with a different number of mixtures (i.e. trajectories). To accommodate the computational intensity of the MHMMs, in this, we follow the two-step approach that was introduced by Bakk and Kuha (2018). This approach involves estimating models with a different number of mixtures where the measurement-error parameters have been fixed to the values estimated in the first step of the estimation process. The third and final stage involved testing whether including predictors to the MHMM can improve model fit.

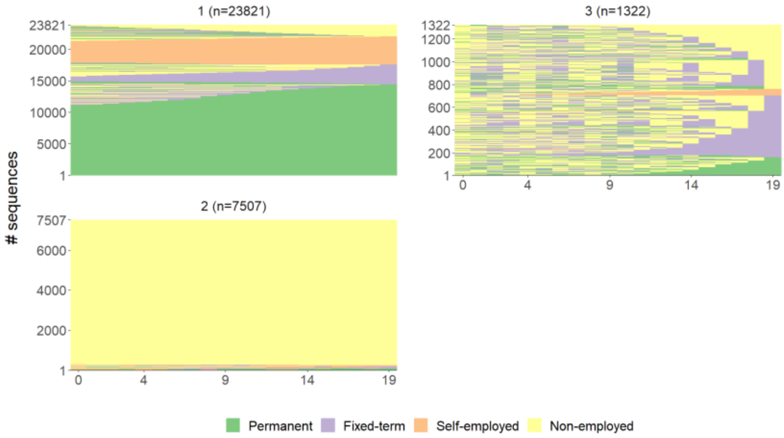
In total, in the first stage, we estimated 9 models[2]. Models 1-3 correct only for random measurement error. In more detail, Model 1 corrects for random measurement error in the indicator of the LFS, Model 2 in the indicator of the AD, and Model 3 in both indicators. Models 4-9 also add corrections for systematic measurement error. In more detail, Models 4-6 include an additional error coefficient for cases where an error was made in time point $t-1$ and can be repeated in t in the LFS indicator (Model 4), in the AD indicator (Model 5) or in both indicators (Model 6). Models 7-9 have a less restrictive specification as they always estimate an additional error coefficient for the cases where an error was made in time point $t-1$ in the LFS indicator (Model 7), the AD indicator (Model 8) or in both indicators (Model 9). Model fit measures (BIC and AIC) show that Model 5 performs best. In the next stage, we build on the selected model (Model 5) to find the optimal number of trajectories (mixtures). The two-step approach of Bakk and Kuha (2018) involves fixing the measurement model parameters to the values that we estimated by Model 5 and estimate models with an increasing number of mixtures, from 2 to 6. Model selection is, in this case, more complicated. As mixtures represent groups of individual trajectories, statistical criteria cannot provide alone the optimal solution. These statistical criteria (i.e. the AIC and the BIC) are combined with substantial considerations which refer to the size and the interpretation of trajectories. This aspect is somewhat described by the concepts of latent class separation and heterogeneity which are summarized by the measure of entropy. To sum up, besides looking at model fit measures, we select a solution with mixtures of substantial size, adequately separated with each other and having a meaningful interpretation. On the basis of these considerations, we selected an MHMM with 3 mixtures. This model has an entropy of 0.694 which, although not being particularly high, is acceptable. In the last stage, we also studied whether adding covariates (other than the indicator for the quarter) improved the model. This appeared not to be the case, so we

[2] We do not report the results of model fit indices for lack of space. These are available on request.

proceeded with the interpretation of the MHMM with 3 mixtures and only the quarter as a predictor of the latent employment state.

Figure 2 illustrates the index plot of the employment trajectories that were obtained by the Model including 3 mixtures. In this plot, every "line" represents an individual pathway across time. The status of each individual at each time point t represents actually the modal classification state according to the selected MHMM. In each of the trajectory plots, the x-axis runs from 0 to T = 19. The largest trajectory (trajectory 1 in Figure 3, which represents 75.4% of the sample) is a trajectory where employment is dominant. People in this category are most likely to be employed on a permanent contract, (less often) with a temporary contract or are self-employed. Transitions in this trajectory are rare and lead mostly to one of the three employment states. Trajectory type 2 (19.8% of the sample) is the non-employment states. Almost all pathways in this trajectory type consist exclusively of non-employment. Lastly, the smallest trajectory type (trajectory 3, 4.8% of the sample) is the trajectory with unstable and rather precarious pathways. This trajectory includes pathways with many transitions between temporary employment and non-employment.

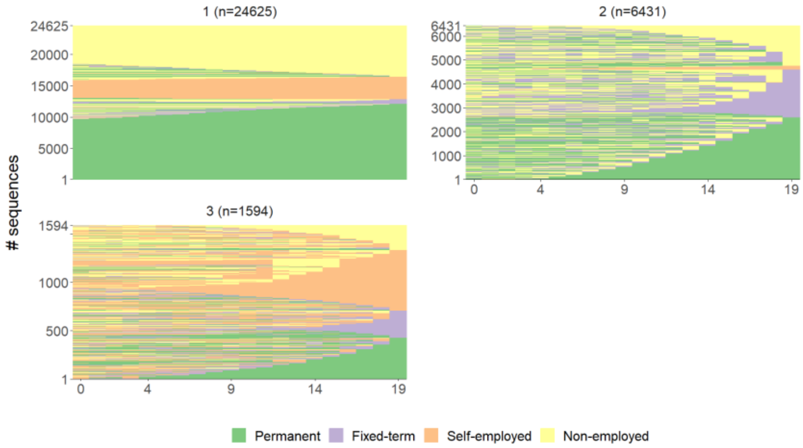
Figure 2 – *Employment trajectories with measurement-error correction.*



Sources: Istat, Labour Force Survey and Administrative Data

To evaluate whether measurement error biases the trajectories that we find in this analysis, we estimate the same model but this time we assume that the observed indicators are error free. In other words, we estimate a mixture Markov model with 3 mixtures. This solution has an entropy of 0.732. We present the 3 trajectories that were produced by this model in Figure 3. The size of the 3 trajectories is similar to those with measurement error correction, 73.2%, 21.2% and 5.6%. However, what these trajectories represent differs considerably from trajectories that were extracted by the model that includes measurement error correction. Failing to correct for measurement error leads to the estimation of more heterogeneous trajectories that include many more transitions than what really exists. In more detail, the largest trajectory includes mostly stable pathways of individuals in permanent employment, self-employment and non-employment. The second largest trajectory includes unstable pathways where individuals move often between temporary employment, permanent employment and non-employment. The smallest trajectory groups also unstable pathways but contrary to the second trajectory, these pathways often include self-employment as well.

Figure 3 – *Employment trajectories without measurement-error correction.*



Sources: Istat, Labour Force Survey and Administrative Data

5. Conclusion

In the article, we estimated a mixture hidden Markov model (MHMM) to study the effect of measurement error on the typology of employment trajectories in Italy. For this purpose, we used linked data from administrative sources and the Labour Force Survey for the years 2017-2021. Our MHMM corrected for both random and systematic measurement error in both data sources.

Our results indicate that measurement error has a severe biasing effect on employment trajectories. Although the size of the estimated mixtures is not severely biased, the typologies coming out of the two analyses (i.e. with and without measurement error correction by employing an MHMM instead of a mixture Markov model) are substantially different. In this respect, research applying a processual approach to study social phenomena should not be ignorant of measurement error.

Applying error-correction is necessary to reveal existing typologies of employment or other social processes.

Acknowledgements

This paper is part of the project DYNANSE which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 864471).

References

ABBOTT A. 1983. Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, Vol. 16, No. 4, pp. 129-147.

AISENBREY S., FASANG A.E. 2017. The Interplay of Work and Family Trajectories over the Life Course: Germany and the United States in Comparison, *American Journal of Sociology*, Vol. 122, No. 5, pp. 1448-1484.

BAKK Z., KUHA J. 2018. Two-Step Estimation of Models Between Latent Classes and External Variables, *Psychometrika*, Vol. 83, No. 4, pp. 871-892.

BAKKER B.F.M., DAAS P.J.H. 2012. Methodological challenges of register-based research, *Statistica Neerlandica*, Vol. 66, No. 1, pp. 2-7.

BALDI C., CECCARELLI C., GIGANTE S., PACINI S., ROSSETTI F. 2018. The Labour Register in Italy: The New Heart of the System of Labour Statistics, *Rivista Italiana Di Economia Demografia e Statistica*, Vol. LXXII, No. 2, pp. 95-105.

BAUM L.E., PETRIE T., SOULES G., WEISS N. 1970. A Maximization Technique Occuring in the Statistical Analysis of Probabilistic Functions of Markov Chains, *The Annals of Mathematical Statistics*, Vol. 41, No. 1, pp. 164-171.

BIEMER, P. 2011. *Latent Class Analysis of Survey Error*. New Jersey: John Wiley Sons.

FILIPPONI D., GUARNERA U., VARRIALE R. 2021. Latent Mixed Markov Models for the Production of Population Census Data on Employment. In PERNA C., SALVATI N., SCHIRIPPA S. (Eds.) *Book of short papers SIS 2021*, Pearson, pp. 112-117.

GROVES R.M. 2004. Survey errors and survey costs. New Jersey: Wiley.

ISTAT 2006. La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione, *Istat, Metodi e Norme*, Vol. 32, pp. 173-196.

MANZONI A., VERMUNT J. K., LUIJKX R., MUFFELS R. 2010. Memory Bias in Retrospectively Collected Employment Careers: A Model-Based Approach to

Correct for Measurement Error, *Sociological Methodology*, Vol. 40, No. 1, pp. 39-73.

MATTIJSSEN L., PAVLOPOULOS D. 2019. A Multichannel Typology of Non-standard Employment Careers, *Social Science Research*, Vol. 77, pp. 101-114.

OBERSKI D.L., KIRCHNER A., ECKMAN S., KREUTER F. 2017. Evaluating the Quality of Survey and Administrative Data with Generalized Multitrait-Multimethod Models, *Journal of the American Statistical Association*, Vol. 520, pp. 1477-1489.

PANKOWSKA P., BAKKER B., OBERSKI D., PAVLOPOULOS, D. 2021. Dependent interviewing: A remedy or a curse for measurement error in surveys?, *Survey Research Methods*, Vol. 15, No. 2, pp. 135-146.

PAVLOPOULOS D., VERMUNT J.K. 2015. Measuring Temporary Employment. Do Survey or Register Data Tell the Truth?, *Survey Methodology*, Vol. 41. No. 1, pp. 197-214.

STUDER M., RITSCHARD G. 2016. What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures, *Journal of the Royal Statistical Society. Series A: Statistics in Society*, Vol. 179, No. 2, pp. 481-511.

SUDMAN S., BRADBURN N.M., SCHWARZ N. 1995. Thinking about answers: The application of cognitive processes to survey methodology, *Quality of Life Research*, Vol. 12, pp. 719-720.

TOURANGEAU R., RIPS L.J., RASINSKI K.A. 2000. The psychology of survey response. Cambridge: University Press.

VARRIALE R., ALFO' M. 2023. Multi-source statistics on employment status in Italy, a machine learning approach, *METRON*, Vol. 81, No. 1, pp. 37-63.

VERMUNT J.K. 2007. Growth models for categorical response variables: standard, latent-class, and hybrid approaches. In VAN MONTFORT K., OUD J., SATORRA A. (Eds.) *Longitudinal models in the behavioral and related science*, Erlbaum, pp. 139-158.

VERMUNT J.K. 2010. Longitudinal Research Using Mixture Models. In MONTFORT K., OUD J.H., SATORRA A. (Eds.), *Longitudinal Research with Latent Variables*, Springer, pp. 119-152.

VERMUNT J.K., MAGIDSON J. 2016. Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax. Belmont, MA: Statistical Innovations Inc.

Dimitris PAVLOUPOLOS, Vrije Universiteit Amsterdam, d.pavlopoulos@vu.nl
Roberta VARRIALE, Sapienza University of Rome, roberta.varriale@uniroma1.it
Silvia LORIGA, Italian National Institute of Statistics, siloriga@istat.it