Mathematics and Society
University of Isfahan
https://math-sci.ui.ac.ir

Mathematics and Society
ISSN (print): 2345-6493, ISSN (on-line): 2345-6507
Vol. 11 No. 1 (2026), pp. 109–131.

https://www.ui.ac.ir

EVALUATION OF THE MEASUREMENT ERROR OF IRANIAN LABOR FORCE INDICATORS DUE TO CHANGING DATA GATHERING MODE BASED ON THE MARKOV LATENT CLASS MODEL

LIDA KALHORI NADRABADI* AND PARIA TORABI KAHLAN

Abstract. One of the most important non-sampling errors is measurement error, which causes the measured value of the target variable to be different from its true value. The markov latent class model allows for the estimation of measurement error for categorical variables in panel surveys that have at least three repetition periods, with the advantage that it does not require re-interviews. In the Iranian labor force survey, which is conducted with rotational sampling panel with a 2-2-2 pattern, this model can be applied. In this article, the challenges of changing the survey mode in times of crisis and the advantages of telephone interviews compared to face-to-face interviews are reviewed. Also, in order to evaluate the results of the labor force survey in the context of the COVID-19 pandemic, when the survey mode was suddenly changed from face-to-face to telephone interviews, the markov latent class model was used to evaluate the measurement error using data from four survey periods from spring 2018 to summer 2021. The results indicate that the measurement error increased in the first period of changing the interview mode, but decreased in subsequent periods. Overall, it can be stated that telephone interviews can be a suitable method for collecting labor force survey data.

Keywords: Markov latent class model, Labor force survey, Measurement error

Article Type: Applied Paper.

Communicated by Majid Asadi.

*Corresponding author.

Received: 09-10-2024, Accepted: 20-05-2025, Published Online: 27-10-2025.

Cite this article: L. Kalhori Nadrabadi and P. Torabi Kahlan, Evaluation of the measurement error of Iranian labor force indicators due to changing data gathering mode based on the Markov latent class model, *Mathematics and Society*, **11** no. 1 (2026) 109–131. https://dx.doi.org/10.22108/msci.2025.142988.1700 .

1. Introduction

Iran's official statistics system, like other countries, is based on collecting data through filling out questionnaires by asking from sample households. Registered sources can be used is some cases as well. One of the most important challenges in this system is statistical error, especially measurement error, which has several sources. One of these source is respondent due to various reasons such as forgetting the real response, misunderstanding the question, not having enough information, or for intentional reasons.Therefore, accurate identification of error sources is required to control the survey error. The Labor Force Survey is one of the main surveys of the Statistical Center of Iran, which has been conducted quarterly and with a rotational sampling method since 2005.

The COVID-19 pandemic presented unprecedented challenges to national labor force surveys (LFS), as documented by the International Labour Organization (ILO, 2020). Given the variety of country contexts, appropriate solutions must ultimately be assessed at the country level and all challenges faced cannot be fully anticipated or prescribed. In addition, there may not be an ideal solution or perfect fit for the many challenges faced any given time. However, various nonetheless options can be considered to reduce the impact on generated statistics. Although there are many different sets of circumstances across countries, the main challenges facing most countries can be broadly categorized into two scenarios.

Scenario 1: Transitioning from fully face-to-face interviewing to telephone based interviewing. This scenario has been evaluated or implemented by many countries, some of whom have panel designs. Others have independent samples for each survey instance. These countries rely either on Computer Assisted Personal Interviewing (CAPI) or Paper and Pencil Interviewing (PAPI).

Scenario 2: Multi-mode data collection is being used (e.g. CAPI plus Computer Assisted Telephone Interviewing (CATI)). In these cases face-to-face CAPI is generally used for the first interview of a panel design, plus a limited number of households for subsequent interviews. The challenge is to move these interviews to either telephone or Computer Assisted Web Interviewing (CAWI). Either scenario 1 or 2 (or others) could involve temporary disruptions (short or long) in data collection, which in reality creates a number of variations within each scenario. Under each scenario there will be proposals to guide countries on how to assess the validity/robustness of different approaches, as this needs to be an important consideration regardless of scenario or approach to change. The range of challenges and responses reported by countries attempting to maintain the operation of their LFS is very wide. The severity of these challenges also varies heavily. Countries fully utilize CATI/CAWI generally facing minimal disruption to data collection (although potentially still some). By contrast, some countries losed their capacity to collect data through existing methods entirely. The challenge facing these countries cannot be underestimated and it is feasible, if not likely, that data collection may need to be suspended, either briefly or for a longer period of time, and indeed this has already happened in many

https://dx.doi.org/10.22108/msci.2025.142988.1700

countries. Scenario 1 reported by a large number of countries, whereby restrictions on movement have caused all face-to-face interviews to be halted, and face-to-face interviewing was the only method in use (CAPI or PAPI). Under this scenario there can be different periodicities (quarterly continuous collection, quarterly periodic, annual continuous, annual periodic, etc.). Some of the countries in this situation utilized independent samples for each survey instance, while others used a panel design. The challenges were mostly the same but those with a panel design were able to potentially benefit from it in different ways. It recommended to consider ways to begin the move to telephone interviewing as it is less disrupted in case of restrictions of movement. The main challenges faced under this scenario are the following: a) retrieving contact details for the final sampling units (FSUs – generally households); b) enabling capacity to carry out interviews by phone; and c) assessing the suitability of PAPI/CAPI questionnaire for telephone interviews. Scenario 2 covers many countries in Europe and selected countries in other parts of the world. In almost all cases there is a panel design. The most typical scenario is a face-to-face interview for the first interview, with CATI (and possibly CAWI) for subsequent interviews. Some face-to-face interviewing will typically continue to take place for the subsequent interviews, but limited to those for whom CATI or CAWI are not possible, for example households with no phone, or respondents who state a preference for a face-to-face interview. As with those countries under scenario 1, the immediate response for the majority of these countries is to move any remaining face-to-face interviewing to one of the other modes, mostly CATI. Essentially the range of challenges and options is broadly be similar to those outlined for scenario 1, but some specific issues still warrant attention.

In Iran, as in many countries, the outbreak of COVID-19 caused the data collection method to change suddenly from face-to-face interviews to telephone interviews, which can affect data quality and typical error rates. In order to estimate the measurement error, gold-standard method, reliability analysis, latent class analysis, and markov latent class analysis are mainly used. Gold standard method measures the bias between a measurement with error and a measurement without error (the gold standard). The disadvantage of the gold standard method is that this value rarely exists, and in cases where it does, there may be latent (unobserved) error in the gold standard estimate, which can lead to invalid results. Reliability analysis decomposes the variance to determine the portion of the variance in the estimate that is due to measurement error. Latent class analysis uses the maximum likelihood method for estimates to determine the bias between the survey and maximum likelihood estimates.

Markov latent class models can be a good alternative for costly and time consuming methods such as re-interview, so as a result of which improvement in error estimation methodology and data quality are obtained In this article, markov latent class model ia applied as an alternative to the re-interview method, to estimate of the measurement error of the Iranian labor force survey which was conducted differently in aspect of data gathering mode under the conditions of the COVID-19 pandemic.

2. Main Results

With the sudden onset of the Covid-19 pandemic, the mode of collecting data for the Iranian Labor Force Survey suddenly changed from face-to-face interviews to telephone interviews due to health restrictions. This change could have affected the accuracy and quality of the data and increased the possibility of measurement errors. The markov latent class model is a statistical tool that allows for the estimation and analysis of measurement error for longitudinal categorical variables with at least three time periods. Employment status in laboure force survey is a categorical varible which is contain 3 level as Employed(e), Unemployed(U) and Not in the laboue force(N). The survey is conducted with a rotating panel design and the response variable is categorical, thus providing the conditions for using the hidden Markov class model. To apply the markov latent class model, data of four survey rounds from spring 2018 to summer 2020 are considered. Independence of classification error, time homogeneity of classification error for different time periods, and homogeneous error probabilities or different time periods are assumed as well as the Markov property $P(X_t|X_{t-1}) = P(X_{t+1}|X_t)$.

Based on the 2-2-2 rotation pattern of the survey, sample units are included in the survey for two consecutive seasons, excluded from the survey for two seasons, and then included again in the survey for two consecutive seasons. Therefore, in order to satisfy the assumption of three time periods of data, which is necessary to implement the Markov model an additional assumption is made in the model. We suppose that no changes in the activity status occurs between the second and third periods. Therefore, those observations without changes in activity status from the second to third quarters are considered in the study. Accordingly, the average difference between the expected probability of stability and the observed probability of no change in the activity status between two consecutive quarters was considered as a criterion for assessing the amount of measurement error due to the change in the survey mode. Given that using weighted data does not have a significant effect on relative estimates and complicates calculations, unweighted data has been used.

In this study, data from four labor force survey periods (from spring 2018 to summer 2021) were analyzed with this model to examine the effect of changing the data collection mode on measurement error. Table 1 shows that in fact 96.2 percent of those employed in the spring of 2018 remained employed in the following quarter. Similarly, 70.2 percent of those unemployed in the spring of 2018 remained unemployed in the following quarter.

TABLE 1. Matrix of transition probabilities without classification error from spring to summer 2018

		Summer 2018		
		E	U	N
Spring 2018	E	0.962	0.007	0.032
	U	0.204	0.702	0.095
	N	0.023	0.005	0.973

https://dx.doi.org/10.22108/msci.2025.142988.1700

TABLE 2. Measurement error matrix of labor force status from spring to summer 2018

		Summer 2018		
		E	U	N
Spring 2018	E	0.984	0.006	0.066
	U	0.079	0.876	0.045
	N	0.009	0.003	0.988

According to Table 2, a person who is actually employed in the spring of 2018 has a 0.6 percent probability of being falsely reported as unemployed. There is also a 1.2 percent probability of being falsely reported as not labor force. Similarly, a person who is actually unemployed has a 7.9 percent probability of being falsely reported as employed and a 4.5 percent probability of being falsely reported as not labor force.

TABLE 3. Matrix of transition probabilities without classification error from spring to summer 2019

		Summer 2019		
		E	U	N
Spring 2019	E	0.950	0.009	0.041
	U	0.191	0.671	0.138
	N	0.016	0.004	0.979

Table 3 shows that in fact 95 percent of those employed in the spring of 2019 remained employed in the following quarter. Similarly, 67.1 percent of those unemployed in the spring of 2019 remained unemployed in the following quarter.

TABLE 4. Measurement error matrix of labor force status from spring to summer 2019

		Summer 2019		
		E	U	N
Spring 2019	E	0.983	0.008	0.008
	U	0.060	0.857	0.083
	N	0.007	0.002	0.991

According to Table 4, a person who is actually employed in the spring of 2019 has a 0.8 percent probability of being falsely reported as unemployed. There is also a 0.8 percent probability of being falsely reported as not labor force. Similarly, a person who is actually unemployed has a 0.6 percent probability of being falsely reported as employed and an 8.3 percent probability of being falsely reported as not labor force.

TABLE 5. Matrix of transition probabilities without classification error from spring to summer 2020

		Summer 2020		
		E	U	N
Spring 2020	E	0.958	0.006	0.036
	U	0.243	0.653	0.104
	N	0.021	0.006	0.973

https://dx.doi.org/10.22108/msci.2025.142988.1700

Table 5 shows that in fact 95.8 percent of those employed in the spring of 2020 remained employed in the following quarter. Similarly, 65.3 percent of those unemployed in the spring of 2020 remained unemployed in the following quarter.

TABLE 6. Measurement error matrix of labor force status from spring to summer 2020

		Summer 2020		
		E	U	N
Spring 2020	E	0.984	0.006	0.01
	U	0.061	0.891	0.047
	N	0.005	0.002	0.993

According to Table 6, a person who is actually employed in the spring of 2020 has a 0.6 percent probability of being falsely reported as unemployed. There is also a 1 percent probability of being falsely reported as not labor force. Similarly, a person who is actually unemployed has a 6.1 percent probability of being falsely reported as employed and a 4.7 percent probability of being falsely reported as not labor force. The total measurement error estimate for the period from spring 2018 to summer 2018, which is before the start of the COVID-19 pandemic, is 0.153, for the period from spring 2019 to summer 2019, which coincides with the outbreak of the COVID-19 pandemic and the sudden change in the survey mode, it is 0.168, and for the period from spring 2020 to summer 2019, it is 0.131. Since the spring of 2020, with the decrease in the prevalence of Covid-19 and nationwide vaccination due to the easing of restrictions, the survey was conducted in the previous method and face-to-face, therefore, the examination of the amount of measurement error due to the change in the survey mode from face-to-face interviews to telephone interviews is not relevant. Therefore, it seems that with the change in the survey mode, the measurement error initially increased, but with the continued implementation of the new mode, the measurement error decreased even compared to when the survey was only conducted face-to-face. Overall, the markov latent class model was introduced as an efficient and low-cost alternative to traditional error assessment methods (such as re-interview) and showed that telephone interviews can be a suitable option for collecting labor force data in crisis situations.

3. Conclusions

Sudden changes in the mode of data collection without the necessary opportunity to test and evaluate them can have significant effects on the quality of the survey results. One of the most important non-sampling errors is measurement error, which represents the difference between the response received from the respondent and the actual value of the under study variable. Obviously, due to the nature of this type of error and the lack of access to true values of responses, its calculation is difficult. In some special cases where administrative data are accessible or conditions of repeating part of the survey with very high accuracy is feasible in practice, the measurement error can be estimated. As an alternative the markov latent class model allows for the estimation of measurement error for

https://dx.doi.org/10.22108/msci.2025.142988.1700

categorical variables in panel surveys that have at least three repetition periods without the need for re-interview. Therefore in the Iranian labor force survey, which is conducted with rotational sampling with a 2-2-2 pattern, the markov latent class model is applicable to estimate the survey measurement error. In this article, in order to evaluate the results of the labor force survey during the Covid-19 pandemic, when the mode of conducting the survey was changed suddenly from face-to-face to telephone interviews or a combination of the two, the markov latent class model was used to evaluate the measurement error using data from four survey periods from spring 2018 to summer 2021. The results indicated that the measurement error increased in the first period of the change in the interview mode but decreased in the subsequent periods. In addition, it was observed that the share of employed and unemployed people joining the not in the labor force population was significant at the beginning of the COVID-19 pandemic, but it is interesting that as the epidemic has prolonged, more employed people have preferred to stay employed rather than become unemployed or not in the labor force and take the risk of looking for a better job. The share of unemployed people in the inactive population has also decreased, meaning that economic hardship and the state of the labor force are returning to pre-COVID-19. Based on the results of this study, it can be stated that telephone interviews, in addition to reducing costs, provide more accurate monitoring tools and can be a suitable alternative to face-to-face interviews for collecting labor force survey data. Using the telephone as a data collection tool allows interviewers to get access to employed individuals in order to fill out the questionnaire, thus reducing proxy response rates.

It should be noted that telephone surveys are methodologically different from surveys in which only interviews are conducted by telephone, and it is necessary to provide the necessary infrastructure for designing and implementing telephone surveys. For telephone surveys to be successful, attention must be paid to telephone coverage of the target population and appropriate sample design to prevent bias. The results show that, by observing the technical and methodological requirements, telephone surveys can maintain the quality of labor force data and be a reliable solution in critical situations. Also, the use of the markov latent class model is recommended for assessing and controlling measurement error.

References

[1] P. Biemer, *Latent class analysis of survey error*, Wiley Series in Survey Methodology, John Wiley & Sons, Inc., Hoboken, NJ, 2011.

[2] P. Biemer and M. Bushery, On the validity of markov latent class analysis for estimating classification error in labor force data, *Surv. Methodol.*, **26** no. 2 (2000) 139–152.

[3] International Labour Organization (ILO), COVID-19: Guidance for labour statistics data collection, Guidance to data producers to maintain labour force survey data collection, *ILO Technical Note*, Geneva: ILO, 2020.

[4] L . Kalhori Nadrabadi and Z. Fallah Mohsekhani, Using Markov Latent Class Models in Estimating the Classification Error of Iranian Labor Flow Statistics, *JSS*, **16** no. 1 2022 189–207. [In Persian]

[5] K. Khalagi, M. A. Mansournia, A. Rahimi-Movaghar, K. Nourijelyani, M. Amin-Esmaeili, A. Hajebi, V. Sharif, R. Radgoodarzi, M. Hefazi and A. Motevalian, Assessing measurement error in surveys using latent class analysis: application to self-reported illicit drug use in data from the Iranian Mental Health Survey, *Epidemiology and Health*, **38** (2016) 1–11.

[6] F. Mehran, M. Nourmohammadi, Z. Rezaee, Z. Fallah Mohsenkhani, L. Kalhori Nadrabadi, A. Izadi, M. Reyhani, T. Baghfalaki, M. Naghikhani, N. Yousefi and R. Aliakbari Saba, *Identifying and controlling non-response and measurement errors in survey designs*, Statistical Research and Training Center, Tehran, Iran, 2017. [In Persian]

[7] B. D. Meyer and N. Mittag, An empirical total survey error decomposition using data combination, *J. Econometrics*, **224** no. 2 (2021) 286–305.

[8] H. Navvabpour, T. Chegini and A. Safarnejadborojeni, An application of the estimation of total survey error in a small-scale survey, *Journal of Official Statistics of Iran*, **25** (2014) 47–68. [In Persian]

[9] H. Navvabpour, T. Chegini and A. Safarnejadborojeni, An application of measurement error evaluation using latent class analysis, *Andisheye Amari*, **22** (2017) 85–96. [In Persian]

[10] H. Navvabpour, M. Safakish and A. Izadi, *Offitial Statistics*, Statistical Research and Training Center, Tehran, 2018. [In Persian]

[11] C. A. Poulsen, *Latent structure analysis with choice modeling applications*, Arhus School of Business Administration and Economics, Arhus, Denmark, 1982.

[12] F. Van de Pol and R. Langeheine, Mixed markov latent class models, in C.C. Clogg ed., *Sociological Methodology*, **20** (1990) 213–247.

[13] F. Van de Pol and J. De Leeuw, A latent markov model to correct for measurement error, *Sociological Methods and Research*, **15** no. 1-2 (1986) 118–141.

[14] J. K. Vermunt, *LEM: a general program for the analysis of categorical data*, Department of Methodology and Statistics, Tilburg University, 1997.

[15] L. M. Wiggins, *Panel analysis*, latent probability models for attitude and behavior processing, Elsevier SPC, Amsterdam, 1973.

Lida Kalhori Nadrabadi

Department of Technical Designs and Statistical Methods, Statistical Research and Training Center, Tehran, Iran

Email: Kalhori@srtc.ac.ir

Paria Torabi Kahlan

Department of Department of Technical Designs and Statistical Methods, Statistical Research and Training Center, Tehran, Iran

Email: torabiparya@yahoo.com

https://dx.doi.org/10.22108/msci.2025.142988.1700