Exploring for the change in change of employment status over time

Abstract

The study seeks to explore for presence of dissimilar developmental paths in employment status of American individuals. The empirical longitudinal exercise is conducted with hybrid latent class growth model and latent class growth mixtures. For testing the hypothesis of heterogeneous subgroups within employment status data three mainstream modelling variants of growth modelling are tested and elaborated in terms of model performance.Progressive employment status based typology is found robust to alternative model specifications voting for presence of clustered growth patterns in experiencing employment status over time .The resulted employment status based typology describes three clusters of active, inactive and mediocre active participants in terms of moving between various stages of employment status for the age span of 18years to 35 years i.e., from adult to post adult prime working life years. The presented approach to understand and explore employment status is novel perspective compared to classic survey based employment status data.

Keywords: *Hybrid Latent class growth model, growth mixtures, two level bivariate residuals, BLRT*

JEL Classification: *c49, c50, c51, c52.*

1.Introduction

When the aim is to explore trajectories of growth for heterogeneous sub-populations over the time or age and if the change is latent then we have three main stream approaches for finding developmental trajectories over time. These techniques stem from parametric or semi parametric distribution or hybrid distributions. In all kinds of growth models, change, development or progress in variable of interest over a given time span is calculated. Here change implies being dynamic, therefore contrary to static cross-sectional framework, repeated measured data or longitudinal data is prerequisite to utilize such models. The change can also be measured in continuous or discrete sense related to the theoretical query[26-27]. The continuous change in such approaches is measured in terms of mean and variance leading to quantitative difference measurement and discrete change is measured in terms of qualitative differences.

For continuous case, change from any one status to another status requires information for at least two time points and can be simply measured by score calculations. In conventional growth models average change is considered as a sufficient representative for the whole sample under study. The Conventional growth models hypothesize a single population for all individuals and a single growth path for estimating the change in parameter of interest. However, many real-life issues theoretically validate the categorical responses of individuals leading to distinctive sub-populations (e.g., socioeconomic classes and employment status categories). Daniel S Nagin and Tremblay (2005) consider using a single growth curve for whole sample as an over-simplified attempt for the possible complex growth patterns among members of different groups.To find discussion on limitations of standard or average growth approach see in (Karen L Nylund et al., 2007). Growth mixtures are a special version of regression mixtures where we find differential effects of time or age for bringing change in growth patterns of the sampled population in hand. Since these are specific to time or age predictor and serve additional purpose of finding change in parameters differences for different groups over time therefore this family of methods is distinctive from regression mixtures(Ishaq. S 2023)

In mainstream alternative parametric random effects or random coefficients approach , latent subgroups are assumed to exist within the data(Vermunt, 2017); (D. Nagin, 2009).To address and measure the difference in subsets of population based on continuous data ,difference in means,variances and co-variance statistics are compared for the subgroups. Naturally this implies the usual imposition of normally distributed data from which the subgroups may emerge. Unfortunately, the assumption of normal distributed data is always challenging and controversial to hold in case of real-life data sets since survey-based data sets mostly provide skewed responses. This limitation of parametric approach leads to hybrid latent class growth models (HLCM) proposed by (Vermunt, 2017).In HLCM approach, we may find class differences by including categorical latent variables along with continuous latent variable.This technique serves to find individual growth trajectories for each subclass with different means and slopes(Andruff et al., 2009).

Another extension to average growth scheme is growth mixture model. This approach is more comprehensive then HLCM since it imposes no structure on distribution of different latent classes of the data. The distribution for each class can therefore by normal or any according to variables of data. This approach like HLCM is based on testing of "change over time does not changes everyone equally" and the error in measurement of change is again measured through latent part with flexible distributions scope. The models under this scheme do not rely on the assumption that all individuals under study are drawn from a single population, which is the main limitation of conventional growth models. Compared to their cross-sectional counterpart 'latent class cluster model' these models in longitudinal version explore the clusters for which the rate of change or effect sizes are different. In this specific format these models are extension of latent class regression models where the predictor is mainly time or age .Thus, we may test longitudinal heterogeneity through the identification of unobserved sub-populations in the sample.The population heterogeneity in these models is captured by the inclusion of a categorical latent variable that identifies subgroups of individuals, typically referred to as latent classes. The growth mixture models (GMM) introduced by Muthén and Muthén (2000) are advocated for their effectiveness in categorizing homogeneous subgroups inside the grander heterogeneous group for finding important classes with respect to growth patterns and in addition for accounting measurement error in survey data (MacCallum & Austin, 2000) (Karen L Nylund et al., 2007). D. Nagin (2009) proposed to impose homogeneous individual growth trajectories within a class for GMM. This restricted version is known as Latent class growth mixture.According to Reinecke and Seddig (2011) the latent class Growth mixture is the most adjustable approach for incorporating inter individual differences in intra-individual change considering unobserved heterogeneity within a larger population, see into Andruff, Carraro, Thompson, Gaudreau, and Louvet (2009) and Agresti, Booth, Hobert, and Caffo (2000) and Ishaque (2023) for detailed review of growth models based on longitudinal data and mixture framework .

Considering the variants of growth models briefly now we justify their significance to explore employment status. For any economy, following and understanding employment status of individuals over the time is important for observing the dynamics of labor market. Since employment status choices over the age shape and effect individuals total. The average change in employment choices fails to depict the inside differences of individual experiences (Ishaque, forthcoming).Moreover, we know that the employment status data is collected in periodical surveys in USA like many economies, therefore the recorded responses contain measurement error and remain imperfect measure of employment status. Further the employment status experienced at the time of data collection is not the perfect measure for deciding about the actual employment situation immediately before or after the surveys. Therefore, like any economic variable, the latent variable based growth framework becomes particularly useful for incorporating measurement error in recorded variables as well for observing experience of change in change of employment status over the time. To the best of our knowledge, there is no empirical application utilizing growth approaches for exploring employment status in detail for any country, therefore the choice of empirical setup is novel from methodological and theoretical perspective. Reason for selection of USA was in continuation of employment analysis by (Ishaque, forthcoming) for adding more to historical insights of US labor market. In this study change in employment patterns between various classes of employed, unemployed, and out of labor force are explored and clustered for prime working age years of (16 to 35) using NLS-79 cohort data. For many socio-economic variables this data source provides longitudinal details of those who were around 16 years of age in 1979, and were around 35 in 1998. The research questions addressed in this study are: what kind of change exists between employment statuses across time? Do the qualitatively distinct subgroups within employment status data reveal any patterns in responses of employment choices over age?

2. Methodology

Absolute and relative fit statistics are utilized in growth models for finding the best model, for some background see into (Collins, Fidler, Wugalter, & Long, 1993). Longitudinal bivariate residuals (LBVR) is the new measure to evaluate model performance of longitudinal data (Jeroen K Vermunt & Jay Magidson, 2013). For concepts of other model evaluation statistics see into online technical appendix 1.

The data utilized in this study comes from NLS 1979 cohort beta version. This was the tested harmonized data for comparatives of two mainstream cohorts of NLS surveys. Some variables like age, employment status, education were compared for those who aged 14-16 years in 1979 and 1997. To access the data, go to https://www.nlsinfo.org/investigator and select "NLS Cross Cohort Beta.

In the following, we brief about growth modelling and its specific versions utilized in the analysis of employment status change and for differentials in growth trajectories of various classes. For sake of brevity the references for details are supplemented.

Underlying simple growth structure is the notion that all persons are drawn from a single population with shared parameters. This assumption is relaxed under the growth mixture framework and for mixture components varied growth parameters are feasible to calculate. This task of un-mixing the population in terms of different growth parameters is accomplished using latent categorical variables. These categorical variables allow to find trajectories or paths of change for subgroups for different groups of individual growth trajectories to vary around different group averages. The distinct growth models for each subgroup/cluster sometimes provide quite unique estimates of covariate effect.Since Latent class growth analysis (LCGA) is a restricted version of growth mixtures, the underlying difference of this model is the pre imposed homogenous structure of growth within each subgroup. Henceforth variance and covariance estimate for the growth factors within each class are assumed to be fixed to zero. For broad discussion of growth variants see into (D. S. Nagin & Land, 1993; Vermunt, 2017).

The specifications of employed growth models for categorical variable employment status (ES) are given below:

2.1. Parametric latent class growth models.

A longitudinal model for categorical data that does model the individual differences is known as a generalized linear mixed model (Skrondal & Rabe-Hesketh, 2004).The same model is called random-effects, random-coefficients, mixed, or parametric latent growth-curve model(Vermunt & Van Dijk, 2001). The model is expressed as having two levels. Level 1 describes the unit change in latent responses at each time point, and at level 2 we describe the unit change over time. The Level 1 equations are:

$$r_{\mathrm{ti}}^{*} = \beta_{0i} + \beta_{1i} a_{\mathrm{ti}} + \varepsilon_{\mathrm{ti}} \tag{1}$$

At Level 2, individual differences in the random coefficients from Level are represented by variability (v_{0i}, v_{1i}) around the mean intercept β_{00} and mean slope β_{01}.The individual differences are modeled as a function of an individual-level, time-invariant covariate, y_i (multiple covariates are possible) quantified by regression coefficients β_{01} and β_{01} for intercept and slope, respectively. The conditional joint distribution of the intercept and slope is assumed to be multivariate normal. In the following base equations are described.For technical differences and detailed elaboration of the given models see into (Feldman, Masyn, & Conger, 2009).For semi parametric or hybrid version of the same model see into (Vermunt, Tran, & Magidson, 2008).(Vermunt, 2017).

$$\begin{aligned} \gamma_{0i} &= \beta_{00} + \beta_{01} y_i + v_{0i} \\ \gamma_{1i} &= \beta_{00} + \beta_{11} y_i + v_{1i} \end{aligned} \tag{2}$$

$$\mathrm{pr}(c_i = k \mid y_i) = \frac{\exp[\delta_k + \vartheta_k y_i]}{\sum_{h=1}^{K} \exp\left[\delta_h + \vartheta_h y_i\right]} \tag{3}$$

2.2. Latent Class Growth Mixtures.

From the standard growth model with the restriction of different growth curves of k subgroups or clusters we can add the subscript k in above sequence of equations where each class has its own variance covariance structure. By further imposing homogeneity of parameters change within each class we can acquire probablity based growth paths (Jung & Wickrama, 2008).

$$r_{\mathrm{kti}}^{*} = \gamma_{k0i} + \gamma_{k1i} a_{\mathrm{ti}} + \varepsilon_{\mathrm{kti}} \tag{4}$$

$$
\gamma_{k0i} = \beta_{k00} + \beta_{k01} y_i + v_{k0i}
$$
$$
\gamma_{k1i} = \beta_{k10} + \beta_{k11} y_i + v_{k1i} \tag{5}
$$

$$
\mathrm{pr}(c_i = k \mid y_i) = \frac{\exp[\delta_k + \vartheta_k y_i]}{\sum_{h=1}^{K} \exp\left[\delta_h + \vartheta_h y_i\right]} \tag{6}
$$

2.3. Longitudinal Bivariate Residuals (LBVR).

Since bivariate residuals (BVR) measure associations at multiple levels of group and individuals over the time, therefore we prefer to mention this measure of group connectedness here explicitly. The lowest score of BVR indicates better fit in terms of co-dependency addressed. Since the employment data had multilevel structure, where individuals were nested within time units over 16 years.So between-group differences and within-group similarities in responses to the change in employment status are measured by measuring presence of longitudinal associations group wise and pair wise see details in f (Nagelkerke, Oberski, & Vermunt, 2017).BVR-group is equivalent to the BVR obtained by using the group id variable also as a nominal covariate (with its effect set equal to 0). The BVR-pairs computes categorical indicators by setting up the two-way cross-tabulation for the responses of pairs of observations within groups. The estimated frequencies $E\left(n_{m,m'}\right)$are obtained as follows:

$$
E\left(n_{u,u'}\right) = \sum_{j=1}^{J} \sum_{i=1}^{I_j} \sum_{i'<i} we_i we_{i'} \sum_{l^g=1}^{M^g} \hat{P}(y_{jit} = u \mid l^g) \hat{P}\left(y_{ji't} = u' \mid l^g\right) \hat{P}(l^g \mid \mathbf{e}_j, \mathbf{y}_j)
$$

BVR-pairs equals the resulting chi-squared value divided by M · (M − 1)/2 (the number of parameters of a symmetric association) and by the average group size see details in (Jeroen K Vermunt & Jay Magidson, 2013). BVR-time is equivalent to the BVR obtained by using the time variable as a nominal covariate (possibly with its effects set equal to 0). The estimated frequencies $E\left(n_{u,u'}\right)$are obtained as follows:

$$E\left(n_{u,u'}\right) = \sum_{i=1}^{I} we_i \sum_{t=2}^{T_i} \sum_{l_{t-1}^d=1}^{M^d} \sum_{l_t^d=1}^{M^d} \hat{P}\left(y_{it-1} = u \mid l_{t-1}^d\right) \hat{P}\left(y_{it'} = u' \mid l^d\right) \hat{P}\left(l_{t-1}^d, l_t^d \mid \mathbf{e}_j\right)$$

3.Results

The models were not directly nested in this study since each growth model has different assumptions therefore between the nested cases the selection criteria opted was relative fit between nested/non nested, parsimony, interpret ability (theoretical validity), low classification errors, ease of convergence, high entropy R2, and lowest two-level bivariate residuals. Initial diagnostic suggested us to vote between 3 or 4 class models as the best fit in each case(see Table 1). After handling the class enumeration problem, the selected model was further tested by bootstrap likelihood ratio test (BLRT). For technical details of BLRT see into (Ishaque, 2023) and technical appendix 1.

Table 1. Growth Models Specifications

Table 1 starts from standard homogeneous growth model and ends with the LCGM with 4 classes. Starting from the basic linear growth model we can observe a size able decrease in log likelihood based absolute and relative model fit statistics.It is to be noted that imposed structure is different, so models are not nested generally.Second entry shows hybrid latent class growth model where only intercept differs across 2, 3 and 4 class specifications. So, we can pick one best performing for this case of nested models. In second specification we vary only slopes or effect sizes across age for different possible class solutions. In third case inspired from first 2 specifications in favor of close options for 3 or 4 class as best fit we only tested for LCG variants for 3 and 4 classes. We can compare from the given summary table extent of bivariate residuals, entropy R2 and level of classification errors across models,and these are compared in detail from case to case in next pages.

In this study, one best fit model was not the objective since each of the models address development(change)in employment status under different assumptions so from three set of underlying assumed structures one best fit from each set was selected and compared in terms of interpretation of change over time. Henceforth, after selecting best from first set i.e., 3 class model, next models was found for included random slope to incorporate the change in growth of various classes around the mean value of change. We tested for whether the unique slope parameter brought further insight in understanding differential effect of age on given classes. Going back to the summary table 1 presented if we had to make only one choice of the most suitable representative model of given data then LCGM version models having fixed variance covariance structure within same class were best fit in terms of lowest value for relative fit statistics for four cluster case followed by 3 cluster case. Since clustering remained main objective so we gave more weightage to low classification errors otherwise in growth literature it is very much recommended to choose model based on relative fit criteria(Nylund, Asparouhov, & Muthén, 2007).

Table 2. Final selected model by BLRT

Further to make selection we applied bootstrapping to confirm final choice between LCGM 3 and LCGM 4. Bootstrapping based statistics revealed adding one more class does not add to understanding data better since the p-value is insignificant. see Table 2. In the following we discuss the model parameters of change for each specification.

3.1. Hybrid latent class growth(HLCG) model.

Following basic structure of these models repeated measurements on individuals are expressed as a function of time.We have measured individual differences in employment status when time equals zero and change in the various categories (employed, unemployed, out of work force) over time was modelled by permitting the intercept and slope coefficients to vary across individuals. The intercept and slope(s) are, therefore, referred to as random coefficients, random effects, or (latent) growth factors.To avoid over extraction of trajectories we started with random sets for all models since this option reduces the chance of over extraction of trajectories which is major issue encountered in growth mixtures(Vermunt, 2017).

Firstly, from table 1, for category-specific intercepts cases i.e., HLCG (2-4class), although 4 class solution could be selected on the base of lowest information criteria, high entropy R2 and low classification errors, but class 3 solution emerged as more parsimonious with highest entropy R2 (91%) and lowest classification error amongst competing alternatives. From table 3 we will see whether significance differential effects of age on shaping employment patterns of individuals do exist? This is located by finding mean change for various classes for various employment statuses. In case of perfect heterogeneity, development of various phases of employment overtime should be distinctive across subgroups of the selected individuals which is not the case here. Though the initial starting differences in employment status is seen by the different intercepts across classes.

Table 3. Regression Scores for 3class HLCG

From table 3, we find significant different random intercepts for each class implying initial position of employment status is different for each class. For issue of identifiability of parameters dummy coding was used. See details of coding in (Jeroen K Vermunt & Jay Magidson, 2013). First category labelled “employed” is taken as baseline from which the change in other categories is compared. Across the three classes we can see that over the given age duration of 16 years, class 1 individuals are most likely to be out of work force compared to baseline status of being employed, for this class mean change from employed to out of work force equals to 2.94. The change is also positive, big and significant for case of being unemployed compared to being employed. The chance of being unemployed after being employed is lowest for the class 3. In summary employment choices are picked differently across these three classes. The average effect of linear time/age is though significant to shape these trajectories but insignificant, therefore the extended cases of square or cubic trajectories are not calculated further.

Table 4. LBVR for 3c intercept HLCG

We find very high LBVR at second level of association for the individuals over time. This implies the present level of autocorrelation were not well accounted by these models. This also signals the poor performance of model for measuring long term associations. Therefore, we compared the other choices in terms on LBVR in next pages.

Table 5. Regression Scores for 3c rando,HLCG

For random slope version of HLCG, age lacks explanatory power for explaining class differences since the effect sizes of age are negligible. This implies the mean level of change is not much different across the three classes and random slopes is not suitable specification for this case. As far the conditional effects of time are concerned, we find some changed effect sizes naturally. For class 2 individuals we have highest likely change of being out of labor force over the time after being employed, this effect is large and significant for class 3 as well and lowest positive for class 2. Class 2 individuals are most unlike to be unemployed over the age and class 1 individuals have highest chance to be unemployed after being employed over the age , the effect size is around .5 and significant.

3.2. Latent class growth mixture results.

Table 6. Regression Score for 3class LCGM

In table 6, we see that the included continuous random effects for finding presence of heterogeneity within classes are highly significant and with very low standard errors. Variances of intercept shown through an aggregative u0 are high in magnitude compared to the u1 which measures average change around mean values for each cluster. Covariance is highly significant though negligible in size.The breakdown of continuous latent term in last decomposed matrix form shows the effects of considering different means and variance and covariance structure for the data in hand is significantly applicable. After the continuous part of mixture framework, we come to discuss the usual class element by reading the effect of distinct slopes and distinct intercepts for each class. We observe each of the class had different position to take development from one state to another. This effect is read through the conditional effect of time for each category for each class, also the effect of the only considered age (random slope) is negative and significant for each class though low in magnitude. This negative effect makes one thing clear that whatever the change faced for ES categories for the given three classes ,ultimately over time there were declines in affiliations to these patterns.

Table 7. LBVR Score for LCGM

The two level bivariate residuals value far below 2 indicates significant values of conditional independence at groups and individual level, this low reported value signals the model fit and suitability to study the measured change in structure (ES change). Since the higher values of these cross-dependence indicators implies model misfit in last versions.From comparing the results of above specification of latent class growth model to other 2 specifications in terms of model performance we concluded the last one stands as the best approach to observe changes in employment status over time for the given data.

4.Typology of change in ES

In this section, change in employment course over the age is discussed. The objective is to find and compare the patterns of change in above discussed modeling schemes and to infer whether the change in taking employment course over life depicts some pattern and that pattern is robust to modeling schemes or not? From above section we concluded that LCGM outperformed competing models in terms of most of model performance and for contributing to understand effect and change sizes. Second competing case was of HLCG 3class random intercept in terms of significant changes over time therefore, longitudinal paths or typologies of employment status/ES are discussed and compared for both models in the following.

Table 8. Longitudinal profile of employment status for HLCG

In profile table 8 ,we have the advantage to compare the likely pattern of change in course of all 3 employment choices for each of three classes. Class 1 reveals the change (0.22 % to 0.51 %) in having first status (employed) over the 16 years. (We are explaining in terms of first and last point's change otherwise it is possible to read the change in each class for each 3 categories of employment status for each year). For class 2 this change margin is initially quite higher (0.81 to 0.93) compared to class 3 which have individuals who likely had the probability to remain in this category like class 1 but with higher range of change in their status of being employed over time (0.28 to 0.71). Class 3 emerges as different for category 1 status change over the time since its individual rise gradually for likely to be employed over the time. Similarly other categories reveal major differences over the prime age life course for this class. For category of 'out of work force' we had class 1 reporting highest proportion of likely cases that is 68% followed by steady decline in this status up to 41 % in last reported years. In summary the response patterns for growth of various employment status categories suggests class 1 (24 % size) have more likely cases who had grown over the time for being employed, and more likely cases who initially and finally ended up with being out of labor force whereas class 2 has individuals more likely to steadily remain employed around whole life span considered followed by class 3.

On the basis of distinctive patterns for three classes on the base of common response patterns as Mediocre active, Mostly Inactive, Active

Table 9. Longitudinal profile of employment status for LCGM

We can see that the posterior probability based results are somehow similar to the previous discussed pattern of growth over time for various categories of Employment status. To summarize the likely cases of growth for ES category 1 for class 1, it is changing positively over time suggesting the rate of being employed is positive. For unemployed category there are more likely cases for whom growth in being unemployed is low. For 'out of work force' we observe somehow similar pattern whereas being inactive the reported rates are low at initial youth years to middle years and finally declined growth rate.This suggests overall more individuals of this class are economically active. For class 2, unemployed categories are steady over the time with low starts and ends whereas out of work force individuals remained part of this group for more than 50 % more or likely all the time.

Class 3 had opposite developmental course for individuals being fully employed compared to other two classes, it had highest reported likely cases of being active labor at youth years and had cases of such individuals from initial probability of 76 % with the positive change upto 90% being employed over time. This naturally suggested decline in growth of other 2 categories for this group. As we can see for category out of workforce there was persistent decline in terms of size of 15% to 1 % over age. We label the classes in row on the base of common response patterns as Mediocre active, Mostly Inactive, persistent Active.

Conclusion and Policy Implications

For any economy, following and understanding employment status of individuals over the time is important for observing the dynamics of labor market. We know that the employment status data is collected in periodical surveys in USA, therefore the recorded responses contain measurement error being imperfect measure of employment status. Further the employment status experienced at the time of data collection is not the perfect measure for deciding about the actual employment situation immediately before or after the surveys. Therefore, like any economic variable, the latent variable based growth framework becomes particularly useful for incorporating measurement error in recorded variables as well for observing experience of change in change of employment status over the time. By employing longitudinal data over the prime working age of American labour class we attempted to explore growth differences in observed employment course. To serve the objective and for testing the hypothesis of heterogeneous subpopulation within the larger population we employed methods which could measure inter individual differences in intra individual change over time. Three mainstream modelling variants of growth modelling were tested and elaborated in terms of model performance. Lastly status typology was built based on the consensus of model variants.

Though class sizes appeared different under above discussed growth variants but typically three patterns of change were observed under each specification. The highest proportion had those likely cases that remained and grow to be employed and least likely cases of being inactive or unemployed over age. That cluster was called **persistent active**. The relative smaller cluster had initially lower reported cases of being employed which persistently growing in likely to be employed and highest cases of individual to be unemployed following remarkable slows in such status over age and distinctive highest and consistent cases of being out of labour force over the age. On part of such contributors, the second cluster was labelled as mostly **inactive**. Third segment had lowest number of likely cases with different response patterns of giving lowest employment starters cases rising over age to be employed but low in proportion to other clusters so this cluster was named as **mediocre active.**

The growth modeling utilized in this study sets an exemplary case study ,a non-conventional and rather more theoretically sound approach for looking into subject of employment status statistics by policy makers of any economy. Since latent growth approach can split measurement error from recorded survey items, plus it is unique statistically endorse the presence or absence of sub classes within total employment status and respective changes in employment statuses over time. The three clusters found in this study though had no stark differences but they led us to think about the possibilities of various employment status patterns leading to demand subjective policy measures accordingly.

Disclaimer

This paper is an extended version of a Thesis document of the same author.

The Thesis document is available in this link: https://file-thesis.pide.org.pk/pdf/phd-econometrics-2016-saima-ishaque--comparison-and-evaluation-of-methods-for-handling-data-clusters-in-regression-models.pdf

Declarations

The author declares no conflict of interest.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc have been used during writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

1.

2.

3.

References

1. Agresti, A., Booth, J. G., Hobert, J. P., & Caffo, B. (2000). Random‐effects modeling of categorical response data. *Sociological Methodology*

2. *30*(1), 27-80.
3. Andruff, H., Carraro, N., Thompson, A., Gaudreau, P., & Louvet, B. (2009). Latent class growth modelling: a tutorial. *Tutorials in quantitative methods for psychology, 5*(1), 11-24.
4. Collins, L. M., Fidler, P. L., Wugalter, S. E., & Long, J. D. (1993). Goodness-of-fit testing for latent class models. *Multivariate Behavioral Research, 28*(3), 375-389.
5. Feldman, B. J., Masyn, K. E., & Conger, R. D. (2009). New approaches to studying problem behaviors: a comparison of methods for modeling longitudinal, categorical adolescent drinking data. *Developmental psychology, 45*(3), 652.
6. Ishaq. S , S. (2023). Testing for homogenous or heterogenous doers in Longitudinal latent class regression framework. *Empirical Economic Review, 6(1)*

7. 20-46.
8. Ishaque, s. (2023). *Comparison And Evaluation Of Methods For Handling Data Clusters In Regression Models.* (PhD doctorate). PIDE, Islamabad.
9. Ishaque, S. (forthcoming). Exploring for competing latent class clusters for US employment

10. data. *Evaluation Review*.
11. Jung, T., & Wickrama, K. A. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and personality psychology compass, 2*(1), 302-317.
12. Muthén, B., & Muthén, L. (2000). Integrating person‐centered and variable‐centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research*

13. *24*(6), 882-891.
14. Nagelkerke, E., Oberski, D. L., & Vermunt, J. K. (2017). Power and type I error of local fit statistics in multilevel latent class analysis. *Structural Equation Modeling: A Multidisciplinary Journal*

15. *24*(2), 216-229.
16. Nagin, D. (2009). *Group-based modeling of development*: Harvard University Press.
17. Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology, 31*(3), 327-362.
18. Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(4), 535-569.
19. Reinecke, J., & Seddig, D. (2011). Growth mixture models in longitudinal research. *Advances in Statistical Analysis, 95*(4), 415-434.
20. Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*: Chapman and Hall/CRC.
21. Vermunt, J. K. (2017). *Growth models for categorical response variables: standard, latent-class, and hybrid approaches*: Routledge.
22. Vermunt, J. K., & Magidson, J. (2013). Technical guide for Latent GOLD 5.0: Basic, advanced, and syntax. *Belmont, MA: Statistical Innovations Inc*.

23. Vermunt, J. K., & Magidson, J. (2013). Technical guide for Latent GOLD 5.0: Basic, advanced, and syntax. *Statistical Innovations Inc*.
24. Vermunt, J. K., Tran, B., & Magidson, J. (2008). Latent class models in longitudinal research. *Handbook of longitudinal research: Design, measurement, and analysis*, 373-385.
25. Vermunt, J. K., & Van Dijk, L. (2001). A nonparametric random-coefficients approach: The latent class regression model. *Multilevel Modelling Newsletter, 13*(2), 6-13.

26. Klug K, Bernhard-Oettel C, Mäkikangas A, Kinnunen U, Sverke M. Development of perceived job insecurity among young workers: A latent class growth analysis. International archives of occupational and environmental health. 2019 Aug 1;92:901-18.

27. Pedersen C, Halvari H, Solstad BE, Bentzen M. Longitudinal trajectories of physical activity among employees participating in a worksite health promotion intervention: A latent class growth approach. Psychology of Sport and Exercise. 2019 Jul 1;43:311-20.