Journal of Applied Structural Equation Modeling

eISSN: 2590-4221 *Journal of Applied Structural Equation Modeling: 3(2),1-8, June 2019*

APPLICATION OF LATENT GROWTH MODELING ON MOTHER-REPORTED MONITORING

Ahmad Farooqi

CRCM, Department of Pediatrics, Wayne State University School of Medicine
Detroit, MI, 48201, United States of America

** afarooq@med.wayne.edu*

ABSTRACT

One of the basic observation in the social and behavioral sciences is that things are changed over the time. Longitudinal data analysis can yield valuable information about this change. Although many techniques have been developed to capitalize on these desirable features of longitudinal data, the structural equation modeling approach of building latent growth models (LGMs) has become one of the commonly used statistical models. A subset of data is taken from the National Longitudinal Survey 97, prepared by the Bureau of Labor Statistics, U.S. Four waves of mother monitoring reported by youth in the year 1997 to the year 2000 are used for the analysis. A total of 2675 adult respondents are used in our analysis. Mother monitoring scores reported by youth are used as a dependent variable. There are 52% male and 48% female in the data. Different linear, quadratic, autoregressive and moving average LGMs with gender as a covariate are used and compared to study the effects of mother monitoring over a 4 year period of time. It is found mom monitoring is increasing slowly over the period of time. An association was found between slop and intercept of fitted latent growth model and female has a significant effect on slop but not on the intercept of the fitted growth model. Five fit indices Chi-square, GFI, CFI, RMSEA, and AIC are used to select an appropriate model.

Keywords: *Structural equation modeling, Latent growth modeling, autoregressive, moving average, covariance analysis of linear structural equation*

INTRODUCTION

Latent growth modeling (LGM) is a statistical method used to study those research questions were all individuals in a given set of data are expected to change with respect to time. In other words LGM used in structural equation modeling framework to estimate growth trajectories. This statistical method needs measurement taken at three or more period of time to estimate growth. Tucker (1958) stated that "latent growth models represent repeated measures of dependent variables as a function of time and other measures. Such longitudinal data share the features that the same subjects are observed repeatedly over time, and on the same tests (or parallel versions), and at known times". Patrick, Snyder, Schrepferman, & Snyder (2005) stated that, "parental monitoring is a hypothetical psychological construct that has been used to explain a composite of parenting practice variables including awareness, communication, concern, supervision, and tracking of adolescent behavior. Poor monitoring is consistently associated with antisocial behavior in both cross-sectional and longitudinal studies". Suppose we want to study the effect of parental monitoring by mom (or dad) over a certain period of time. This can be studied by taking the measurements of parental monitoring taken at different periods of time reported by their children. Different statistical techniques can be used to study growth curve, which could be a positive or negative, linear or nonlinear or one can say autoregressive or moving average. Structural equation modeling (SEM) is one of the statistical techniques that can be used to study the latent growth modeling curve. In LGM first, we might want to see the shape of the growth cure if it is linear or nonlinear. Suppose the growth is linear then we have to estimate two parameters called intercept and slope. It is to be noted that the concept of this intercept is different as we used in the regression. It is the value at the start of a process and is the standard from which change is measured. In addition, to intercept, we have a slope. This parameter tells us how much the curve grows each year. One of the advantages of LGM probably lies in its ability to examine changes in inter-individual differences over time, as well as incorporate time-varying and time-invariant covariates into the model. Curran and Willoughby (2003) have well summarized it as "an intersection between variable-centered and person-centered analysis".

Research Questions & Purpose

Is there any change in the trajectory of mother monitoring reported by youth at given period of time? Is there any association between slop and intercept of fitted latent growth model? Does gender has a significant effect on slop and intercept of the fitted growth model?

Figure 1 shows a typical model of a latent growth curve. Each square box represents the variables we actually measure that are mother monitoring from the year 1997 to 2000. At the bottom of the figure, we have some e's (e1, e2, e3, and e4) represent each variable to contain some measurement error. Two ovals one of these is the intercept and the other is the slope. These are what we are most interested in. The intercept tells us the initial level at the start of the process and the slope tells us the rate of change. There might be more than one slope for complex nonlinear models. The four lines from the intercept to the four manifest variables are all fixed at a constant value of 1. The four lines from the slope to the four manifest variables that appear in the figure represent the linear growth with fixed values for the lines at 0, 1, 2, and 3 respectively. The two-headed arrow represents the correlation between the slope and intercept. Finally, the values (D1, D2) in the circles associated with slop and intercepts are the disturbance terms.

The main purpose of this study is to introduce readers to latent growth modeling and to provide a concrete application of how a statistical analysis can be performed which may serve a helpful guide to researcher and graduate students wishing to use this technique to explore a change in longitudinal data. The statement of the problem in this study "Is there any association between slop and intercept of fitted latent growth model? Does gender has a significant effect on slop and intercept of the fitted growth model?"

Significance of the Study

This study is conducted for the benefit of researcher and graduate students who would plane to make any related study precisely using latent growth model, under the umbrella of structural modeling.

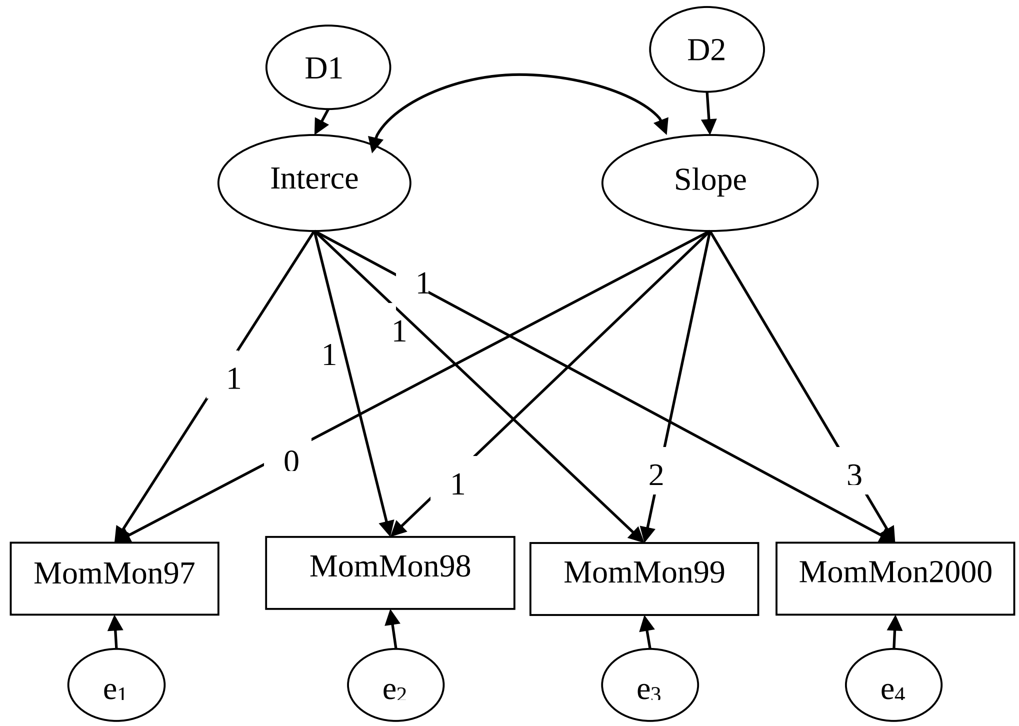

Figure 1: Typical Model of a Latent Growth curve

Linear Latent Growth Model (Linear LGM)

Latent growth modeling (LGM) is a statistical method used to study those research questions were all individuals in a given set of data are expected to change with respect to time. In other words LGM used in structural equation modeling framework to estimate growth trajectories. This statistical method needs measurement taken at three or more period of time to estimate growth. Suppose we want to study the effect of parental monitoring by mom (or dad) over a certain period of time. This can be studied by taking the measurements of parental monitoring taken at different periods of time reported by their children. Different statistical techniques can be used to study growth curve, which could be a positive or negative, linear or nonlinear or one can say autoregressive or moving average. Structural equation modeling (SEM) is one of the statistical techniques that can be used to study the latent growth modeling curve. In LGM first, we might want to see the shape of the growth cure if it is linear or nonlinear. Suppose the growth is linear then we have to estimate two parameters called intercept and slope.

Briefly, LGM investigates the longitudinal growth of a variable y of interest expresses by $y=\tau+\lambda\eta+e$. The vector y contains the repeated measures over the time and is a linear function of τ represents a model intercept and λ contains the factor loadings specifying the hypothesized a prior growth pattern of y. η (eta) is a vector of latent endogenous variable capturing the facets of growth being modeled and 'e' is an error term associated with y. For a linear LGM, η contains the a priori fixed values of factor loadings for the intercept and slope factors. As such, it could be considered as a special case of an oblique confirmatory 2-factor model (Molenaar 2008), when the factor scores of the intercept and slope are allowed to be correlated.

METHODOLOGY

Data Source, Participants, and Measures

The dataset used in this study was taken from the National Longitudinal Survey 97, prepared by the Bureau of Labor Statistics, U.S. Department of Labor. There are four waves of mother monitoring in the data set, started with mother monitoring reported by youth in the year 1997 and the survey continued annually till 2000. A total of 3578 adult respondents, in the year 2009, were participated in the survey. Due to incomplete data in the variables, a total of 2675 adult respondents in our analysis are used. Scores reported by youth ranges from 0 to 16 indicate the degree of parental monitoring by residential mother; higher scores indicate greater parental monitoring. In our study, the independent variable is gender and is distributed by male, female, and was coded 0= "male", 1= “female”. Mother monitoring scores reported by youth are used as a dependent variable; ranges from 0 to 16 indicate the degree of parental monitoring by a residential mother. It is to be noted that higher scores indicate greater parental monitoring.

In LGM, the two growth parameters are treated as latent variables (factors) and the repeated measurements are represented as indicators of the two factors. It is worth noting that the loadings of the time matrix are specified a priori in LGM to reflect the hypothesized “shape” of change. That distinguishes LGM from standard SEM models were usually loadings are estimated instead of being fixed. In the current study, the loadings of Slope are fixed as 0, 1, 2, 3, implying the shape of the growth trajectories is linear. Though the values of the factor loadings are arbitrary, the interpretation of the result might change according to the values the researchers choose (Stull 2008).

Linear LGM

We start with the linear LGM. Since there are 4-time points of mother monitoring, they represent the manifest variables from wave 1 to wave 4. We develop a causal relationship between the manifest mom monitoring variables and the latent factors. The two latent factors are generated represent the latent intercept and slope respectively. We fixed the intercept factor with a loading of 1 for all the 4-time points and the linear slope factor with pre-specified loadings from 0 to 3, each representing one path of the manifest variable in each year to the two latent factors. We denote e1 to e4 state the residual terms of $\boldsymbol{\varepsilon}$ and D1 to D2 states the random effect terms for latent variables, slope and intercept. The variances and covariance between the two latent factors are estimated. We then use Gender as a covariate in our model.

Quadratic LGM

LGM is not limited to linear function. When there are 3 or more points in repeated measures, we could incorporate nonlinear trajectories into LGM. One of the most common approaches to nonlinear trajectories is to use polynomials. The factor loadings can be fixed to represent a quadratic function of the observed time metric. The mean of this quadratic slope represents the degree of quadratic curvature in the trajectory. A third factor is added and the coefficient for the mean. The factor loadings for quadratic LGM are hence set to the power term starting from 0 in wave 1 and 9 in wave 4. As there are now three random terms for the three latent factors, three covariance terms are then specified.

Autoregressive Linear LGM

In Markov simplex modeling, the manifest variables are related to the previous period to show the extent of the temporal relationships between two-time points. This modeling strategy could be built into LGM by combining both the modeling approaches and this unites the features of LGM that take into account of the simplex models, the Autoregressive LGM (AR LGM). Bollen and Curran (2004, 2006) call it the autoregressive latent trajectory model. For AR LGM, additional AR terms are specified that links the parameters of the manifest variable of current to the previous time. For instance, to relate the mother monitoring at time 1 to mother monitoring at time 2, we add in a parameter called ARlag1 and specify mother monitoring at time 1 after the parameter to relate it to the endogenous mother monitoring at time 2. Sometimes, we would like to have a parsimonious model by specifying equality of the lag AR parameters. This can be done easily by giving the same name to the parameter. Instead of having different name ARlag1, ARlag2, ARlag3 for the above AR LGM, we use only ARlag1 to specify the equality of parameter. A gender is used as a covariate in the model. Figure 2 shows the path diagram of AR LGM with equality constraint of the AR parameter.

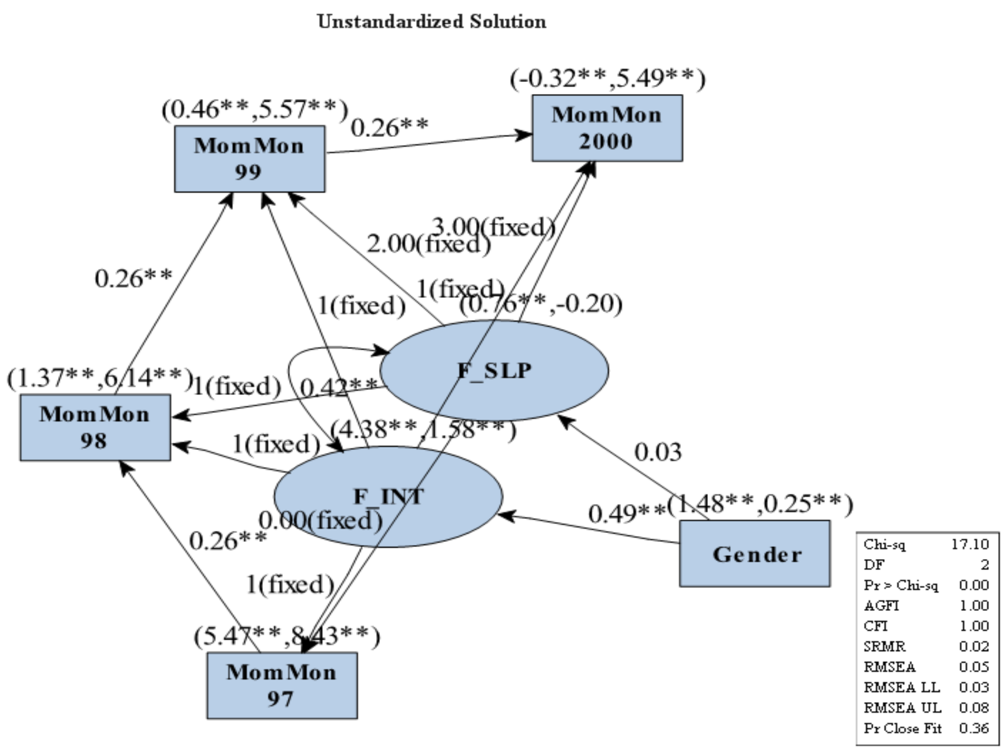

Figure: 2 Path Diagram of the Final AR LGM Model

Statistical Data Analysis

We analyzed and reported the categorical variables by numbers (%). The normality of the data is tested by Kolmogorov-Smirnov test. Median and Inter-Quartile Range are used to report the degree of mother monitoring. We use Six Endogenous (4 observed and 2 latent) and seven exogenous (4 errors, 2 disturbance, and 1 covariate) variables in the model. Maximum likelihood (ML) method of estimation is used to illustrate univariate Linear, Quadratic and Autoregressive Linear LGM models on 4 waves of longitudinal mother monitoring data. Gender is used as a covariate in this study. We used procedure CALIS in SAS (version 9.4, SAS Institute Inc. Cary, North Carolina) to perform statistical analyses. The significance level was set at 0.05.

RESULTS AND DISCUSSION

The current study addressed three primary research questions. First, is there any change in the trajectory of mother monitoring reported by youth for a given period of time? Second, is there any association between slop and intercept of fitted latent growth model? and finally, does gender has a significant effect on slop and intercept of the fitted growth model? First, we have check whether our dependent variable is distributed normally. It can be seen from Table1, we found a skewness ranges from -0.59 to -0.47 and kurtosis from -0.08 to 0.02 for all four waves of mother monitoring. Since both measurements fall within the limits of ±1 criteria of normality, therefore non-normality is not an issue in our case of fitting latent growth modeling. The mean mom monitoring scores from the year 1997 to 2000 are 10, 10, 9, and 9 respectively with a stander deviation of almost 3 each year. There are 52% male and 48% female in our study. It can be seen from Table 2, that the estimated correlation between the manifest variables is positive and significant.

Table 1: Descriptive Statistics

Variable	N	Mean	Standard Deviation	Skewness	Kurtosis
MomMon97	2675	10.57	3.14	-0.59	0.02
MomMon98	2675	10.07	3.18	-0.58	0.01
MomMon99	2675	9.84	3.29	-0.52	-0.08
MomMon2000	2675	9.81	3.20	-0.47	-0.05

Table 2: Pearson's Correlation Coefficient Matrix

	MomMon97	MomMon98	MomMon99	MomMon2000
MomMon97	1.00	0.44 (<.0001)	0.38 (<.0001)	0.36 (<.0001)
MomMon98	0.44 (<.0001)	1.00	0.59 (<.0001)	0.49 (<.0001)
MomMon99	0.38 (<.0001)	0.59 (<.0001)	1.00	0.59 (<.0001)
MomMon2000	0.36 (<.0001)	0.49 (<.0001)	0.59 (<.0001)	1.00

Four models are illustrated using National Longitudinal Survey 97 data to examine the effects of mother monitoring over time. The first model is the simplest LGM among all the models. In second LGM, we introduce gender as a covariate to get some more degrees of freedom. The third model, Quadratic LGM with gender as a covariate, and the aim are to account for the non-linear growth pattern. As there are only 4 waves of data, we cannot model cubic LGM due to under identification. The next models we add in the autoregressive parameters into the linear LGM to account for the temporal relationships of the manifest mom monitoring variable over adjacent time. The model lag term is the AR LGM, with equality constraints on the AR LGM with the autoregressive parameters. The MA LGM and ARMA LGM cannot be modeled due to the under-identification problem. We only provide the interpretation of the estimates for the best-fitted model. It can be seen from the Table 3 that the estimated mean intercept is 4.38 and is significant. The estimated mean slope is 0.80 which is also significant, indicating the average increase of mom monitoring score over one year. The influence of mother monitoring at any current year to mother monitoring of its previous year is estimated at 0.26 and is significant. The estimated variance of intercept is 1.58, suggesting the youth reported are not homogeneous on this trait. The variance of slop was negative and set to zero. The latent intercept and slope of female were found 0.49 and 0.03 respectively. The significant correlation between intercept and slope is 0.42, indicating that the youth's initial score is positively correlated with the growth rate. All four residual variances are also significant.

Table 3: Comparisons of Univariate LGM Models

Parameters	**Linear LGM**	**Linear LGM with Gender**	**Quadratic LGM with Gender**	**AR LGM With Gender**
Latent Intercept	6.31*	6.55*: (0.51)*	6.00*: (0.48)*	4.38* : (0.49)*
Latent Slope	1.43*	0.86*: (0.09)*	1.42*: (0.20)	0.80* : (0.03)
Latent Quadratic			-0.07*: (-0.04)	0.26*
Mom Monitoring 1998				0.26*
Mom Monitoring 1999				0.26*
Mom Monitoring 2000				
Variance/Covariance				
Intercept	5.18*	5.11*	5.78*	1.58*
Slope	0.42*	0.42*	4.78*	0.00
Quadratic	-0.35*	-0.36*	0.34*	0.42*
Intercept and Growth			-1.55*	
Intercept and Quadratic			0.26*	
Growth and Quadratic			-1.21*	
Residual Variance				
Mom Monitoring 1997	4.68*	4.68*	4.00*	8.42*
Mom Monitoring 1998	4.68*	4.68*	4.00*	6.12*
Mom Monitoring 1999	4.68*	4.68*	4.00*	5.58*
Mom Monitoring 2000	4.68*	4.68*	4.00*	5.49*

Five fit indices are used for the study. They are Chi-Square (χ^2), Goodness of Fit Index (GFI), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMESA) and Akaike Information Criterion (AIC). As these indices have their relative merits, there is ease for comparison if we put them side by side to assess the overall fit of the models. The GFI index has a long history in the SEM literature. The CFI index indicates the fit of a model improves on the nested null model. The RMSEA is an estimate of a misfit at the population rather than a sample. Table 4 reports the results of the univariate LGMs.

Table 4: Comparison of Univariate LGM Models

Fit Indices	**Linear LGM**	**Linear LGM With Gender**	**Quadratic LGM With Gender**	**AR LGM With Gender**
Chi-Square	165	168	18	17
Df	4	6	1	2
P-value	<.0001	<.0001	<.0001	<.0001
GFI	0.98	0.99	0.99	0.99
CFI	0.95	0.95	0.99	0.99
RMSEA	0.12	0.10	0.08	0.05 (0.03, 0.08)
AIC	185	197	55	53

The Significance of chi-square probability indicated the overall rejection of the model. However, GFI (0.99) and CFI (0.99) indicated an acceptable model fit, as values greater than 0.95 indicate a good fit (Hu and Bentler, 1999). Also, RMSEA (0.05) indicated a very good fit. A RMSEA less than 0.06 indicate acceptable model fit (Hu and Bentler, 1999). The AIC (53) has a minimum value as compared to other models, which shows that the AR LGM model is best among the fitted models.

CONCLUSION

This study has covered the basics of structural equation modeling (SEM) which is one of the techniques that can be used to study the latent growth curve. SEM is a highly versatile tool heavily used in the psychology research literature to investigate a variety of problems. The SEM analysis provides flexibility in determining the relationships between variables. Direct as well as indirect relationships between variables can be specified and estimated. The SEM method, using SAS procedure CALIS, estimated rates of change, variances of measured variables and latent variables, mean initial value and mean rate of change. We illustrate and carrying out various latent growth models from a simpler to a more complex model. Path diagram for the fitted model is presented. This will help readers to have a better idea of what the model is trying to achieve. Although the current paper presents the different LGMs and makes comparisons among them, it does not intend to answer the question of whether the mixing of simplex methods and LGM is a better choice.

Limitations of the Study

Like other studies, this study also has limitations. One of the limitations of the study was the missing observations. In a total of 3578 adult respondents, observations were missing in mother's monitoring and gender, after dropping these missing observation we left a total of 2675 adult respondents in our analysis. Since the sample size was quite large, dropping missing observation does not have much effect on our analysis.

ACKNOWLEDGMENT

I am thankful to National Longitudinal Survey 97 prepared by the Bureau of Labor Statistics, U.S. Department of Labor, using their data for my analysis.

REFERENCES

Bollen, K. A., and Curran, P. J. (2004) Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods and Research*, 32, 336-383.

Bollen, K. A., & Curran, P.J. (2006) *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.

Curran, P. J., Bauer, D. J., & Willoughby, M. T. (2003) Testing and probing main effects and interactions in latent trajectory models. Manuscript submitted for publication.

Hu, L. & Bentler, P. M. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.

Molenaar, P. C. M., and Campbell, C. G. (2008) Discussion of the special issue on growth models for longitudinal data in educational research. *Educational Research and Evaluation*, *14(4)*, 377-390.

Patrick, M. R., Snyder, J., Schrepferman, L. M., & Snyder, J. (2005) The joint contribution of early parental warmth, communication and tracking, and early child conduct problems on monitoring in late childhood. *Child Development*, 75(5), 999-1014.

Stull, D. E. (2008) Analyzing growth and change: latent variable growth curve modeling with an application to clinical trials. *Quality of Life Research*, 17, 47-59.

Tucker, L.R. (1958) Determination of parameters of a functional relation by factor analysis. *Psychometrika* 23, 19-23.