Paper 273-2010

Building Latent Growth Models Using PROC CALIS: A Structural Equation Modeling Approach

Teck Kiang Tan, Trivina Kang, David Hogan
Centre for Research in Pedagogy and Practice, Nanyang Technological University, Singapore

ABSTRACT

This paper illustrates the structural equation modeling approach of building latent growth models (LGMs) using PROC CALIS. In the past decade, LGM has become one of the commonly used statistical models for analyzing longitudinal data analysis. Although recent years have seen the increase use of LGM to carry out research work in longitudinal analysis, there is limited work that has spelled out a systematic procedure of using PROC CALIS in modeling LGM. This paper serves to fill the gap for data analysts and researchers by showing the practical aspects and theoretical concerns of applying this modeling technique. This paper also includes a new conceptual idea of combining simplex approach and classical LGM to include autoregressive terms and moving average terms in order to improve the way we can conduct longitudinal analysis. The syntaxes of PROC CALIS are illustrated throughout the paper. Using data from a 4-wave longitudinal dataset of secondary school students, various LGMs are illustrated from the simplest of an unconditional LGM to conditional multivariate cross-Lag autoregressive LGM.

INTRODUCTION

Longitudinal analysis involves multiple responses taken in sequence on the same subject over time and we generally refer to these observations as repeated measures. Latent growth model (LGM) is one of the modeling techniques used to analyze repeated measures. Structural equation modeling (SEM) is a statistical approach to carry out overall hypotheses testing about relations among observed and latent variables. As LGM consists of observed variables to represent observations studied over time and their overall means and growth are specified in some particular a priori forms represented as latent variables, SEM naturally becomes one of the approaches in modeling LGM.

LGM has flourished over the last decade as one of the most commonly used statistical tool in analyzing longitudinal data. This is evidenced by the growing number of applications of this statistical technique in leading journals and books written specifically for this subject (e.g. Duncan, Duncan and Strychker, 2006; Bollen and Curran, 2006; Preacher, Wichman, MacCallum, and Briggs, 2008). Many researchers favor in LGM and argue for its superiority (Curran, 2000; Duncan, Duncan, Strycker, Li, & Alpert, 1999; Fan, 2003; McArdle & Bell, 2000). One of the greatest advantages of LGM probably lies in its ability to examine changes of interindividual 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”.

The PROC CALIS procedure is a general procedure for analyzing covariance structure using SEM approach. It provides the SEM estimated parameters and tests the appropriateness of the model. Just as in any SEM, the various goodness-of-fit indices of LGM are available in PROC CALIS. As this paper concentrates on the SEM approach of analyzing LGM, for the purpose of linking the model to the syntax and to provide a clearer explanation, path diagrams are graphed to show the various LGMs. The rationales behind of using these models are explained where appropriate.

DATA SOURCE AND SAMPLE

The data used in this paper to illustrate the various LGMs are obtained from the life pathway project conducted by the Centre for Research in Pedagogy and Research, Nanyang Technological University, Singapore. There are four waves of data in the dataset. The first wave started in the year 2005 when the students are in secondary one (Grade 7) and the survey continued annually till students reach their secondary four (Grade 9). For analytical purposes, two subject areas of analyses are examined to illustrate univariate and multivariate LGMs separately. The first set of data consists of 1,626 students whom we ask about their interests in politics. We use univariate LGMs to find out factors influencing these students’ political interests. The second analysis focuses on examining the relationships of student’s social skills and multi-literacy skills over time and factors affecting them. As there are two sets of constructs analyzed simultaneously over time, multivariate LGMs has been identified as appropriate for the second set of data and will be used to illustrate the various multivariate latent growth models. This second set of data consists of 2,036 students. Covariates use in both studies includes gender, race, stream, and housing type.

LINEAR LATENT GROWTH MODEL (LINEAR LGM)

Briefly, LGM investigates the longitudinal growth of a variable y of interest expresses by $\mathbf{y} = \boldsymbol{\Lambda\eta} + \boldsymbol{\varepsilon}$. The vector $\mathbf{y}$ contains the repeated measures of y over time. $\boldsymbol{\Lambda}$ is the matrix that contains the factor loadings specifying the hypothesized a prior growth pattern of y. $\boldsymbol{\eta}$ is a vector of factors capturing the facets of growth being modeled and $\boldsymbol{\varepsilon}$ contains the random normal residuals. For a linear LGM, $\boldsymbol{\eta}$ 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 2003; Willet and Sayer, 1994) when the factor scores of the intercept and slope are allowed to be correlated i.e. when we specify the covariance between the intercept and slope factors, linear LGM is in fact a special case of an oblique CFA. $\boldsymbol{\mu}_{\eta}$ is the vector of factor means and ζ is the residual term for the latent factors. As for the model-implied variance, $\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Psi\Lambda}' + \boldsymbol{\Theta}_{\varepsilon}$, $\boldsymbol{\Sigma}$ is the covariance matrix, $\boldsymbol{\Psi}$ represents the covariance matrix of the error equations ζ and $\boldsymbol{\Theta}_{\varepsilon}$ is the covariance matrix of the residual in $\boldsymbol{\varepsilon}$.

$$\mathbf{y} = \boldsymbol{\Lambda\eta} + \boldsymbol{\varepsilon}$$
$$\boldsymbol{\eta} = \boldsymbol{\mu}_{\eta} + \zeta$$
$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Psi\Lambda}' + \boldsymbol{\Theta}_{\varepsilon}$$

$\mathbf{y}$: Observations of Repeated Measures
$\boldsymbol{\Lambda}$: Factor Loadings
$\boldsymbol{\eta}$: Random Coefficients
$\boldsymbol{\mu}_{\eta}$: Latent Means
ζ : Random Effects
$\boldsymbol{\varepsilon}$: DisturbanceTerms
$\boldsymbol{\Sigma}$: Variance and Covariance of Observed Variables
$\boldsymbol{\Theta}_{\varepsilon}$: Disturbance Variance and Covariance
$\boldsymbol{\Psi}$: Disturbance Factor Variance and Covariance

Figure 1 shows the path diagram of a linear LGM for panel data with 4 occasions. In the linear LGM, consecutive measurements are modeled by a latent variable for the intercept, and a second latent variable for the slope of the curve. Y1 to Y4 are the observations of the response variable, political interests, at the 4 time points. In LGM, the expected score at time point zero is modeled by a latent intercept factor and the latent slope factor is the slope of the linear curve. As the intercept is constant over time, the factor loadings are constrained to one for all the time points. Similarly, as the latent slope represents the linear growth, it is constrained with loadings to be equal to an incremental of 1 from 0 in the first wave to 3 in the last wave (See Figure 1 for path diagram). We may choose any reference point for the initial and any metric for the growth but for the purpose of illustration, we fix wave 1 as the reference starting point.

The intercept mean indicates the average starting point of the variable of interest. The slope mean indicates the average rate of change over time. The intercept variance indicates the degree that people vary at the start of the study. The slope variance indicates the degree to which people vary in terms of their rate of change over time. The variances of the intercept and slope factors represent the individual deviations from the intercept and slope means respectively.

SAS® PROCEDURE CALIS

There are a number of procedures in SAS® that support LGM. These include PROC MIXED and PROC NLMIXED. However, this paper concentrates on the PROC CALIS (Covariance Analysis of Linear Structural Equations) because it is perhaps one of easiest ways in specification when we consider the structural equation model approach. The syntax of PROC CALIS becomes self-explanatory with the help of the path diagrams. The path diagrams are referred to for the various LGMs.

LINEAR LGM

We start with the syntax of linear LGM as shown below. Refer to Figure 1 for the path diagram of Linear LGM. The VAR statement states the variables used in the procedure. Since there are 4 time points of student's political interests, PIS1 to PIS4 represent the names of the manifest variables from wave 1 to wave 4. They are specified under the VAR statement. The LINEQS statements specify the causal relationships between the manifest political interests variables and the latent factors. The two latent factors: F_INT and F_SLP, represent the latent intercept and slope respectively. As the intercept factor has loading of 1 for all the 4 time points and the linear slope factor with pre-specified loadings from 0 to 3, there are four statements altogether stating in the LINEQS statements, each representing one path of the manifest variable in each year to the two latent factors. F1 to F4 state the residual terms of ε and D1 to D2 states the random effect terms of ζ. The STD and COV statements specify the name of the variances and the covariances to be estimated at the LINEQS statement. The covariance between the two latent factors is named as ZETA21.

```
* Linear LGM;
Proc CALIS Data=A;
Title "Linear Latent Growth Model";
LINEQS
  PIS1 = 1.0 F_INT + 0.0 F_SLP + F1,
  PIS2 = 1.0 F_INT + 1.0 F_SLP + F2,
  PIS3 = 1.0 F_INT + 2.0 F_SLP + F3,
  PIS4 = 1.0 F_INT + 3.0 F_SLP + F4,
  F_INT = Mean_INT INTERCEPT + D1,
  F_SLP = Mean_SLP INTERCEPT + D2;
STD
  D1-D2 = ZETA1-ZETA2,
  F1-F4 = 4*err:;
COV
  D1 D2=ZETA21;
VAR
  PIS1 PIS2 PIS3 PIS4;
Run;
```

QUADRATIC AND CUBIC 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 as shown in Figure 2. The mean of this quadratic slope represents the degree of quadratic curvature in the trajectory. In PROC CALIS, a third factor call F_Quad is added and the coefficient for the mean is called Mean_Quad. 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 specified in the COV statement (ZETA21, ZETA31, and ZETA32).

```
* Quadratic LGM;
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP + 0.0 F_Quad + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP + 1.0 F_Quad + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP + 4.0 F_Quad + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP + 9.0 F_Quad + F4,
  F_INT  = Mean_INT  INTERCEPT + D1,
  F_SLP  = Mean_SLP  INTERCEPT + D2,
  F_Quad = Mean_Quad INTERCEPT + D3;
STD
  D1-D3=ZETA1-ZETA3, F1-F4=4*err:;
COV
  D1 D2=ZETA21,
  D1 D3=ZETA31,
  D2 D3=ZETA32;
```

We could also fit a cubic LGM by adding a cubic latent factor. The LINEQS would be modified to include a cubic factor called F_Cubic and the means called Mean_Cubic. As the data for illustration contains only 4 time points, the cubic model is not identified and cannot be modeled. As such, the results would not be presented but the path diagram is shown in Figure 3 to illustrate a 5 time points model. The syntax of a cubic LGM is as follows:

```
* Cubic LGM;
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP +  0.0 F_Quad +  0.0 F_Cubic + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP +  1.0 F_Quad +  1.0 F_Cubic + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP +  4.0 F_Quad +  8.0 F_Cubic + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP +  9.0 F_Quad + 27.0 F_Cubic + F4,
  PIS5=1.0 F_INT + 4.0 F_SLP + 16.0 F_Quad + 64.0 F_Cubic + F5,
  F_INT   = Mean_INT   INTERCEPT + D1,
  F_SLP   = Mean_SLP   INTERCEPT + D2,
  F_Quad  = Mean_Quad  INTERCEPT + D3,
  F_Cubic = Mean_Cubic INTERCEPT + D4;
```

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 in the LINEQS statement that links the parameters of the manifest variable of current to previous time. For instance, to relate the influence of political interests at time 1 (PIS1) to political interests at time 2 (PIS2), we add in a parameter called ARlag1 and specify PIS1 after the parameter to relate it to the endogenous PIS2. Similar syntax to specify the parameters of later lags (ARlag2 and ARlag3) are shown below. The path diagram of AR LGM model is shown in Figure 4.

```
* Linear LGM - AR(1);
Proc CALIS Data=A;
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP + ARlag1 PIS1 + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP + ARlag2 PIS2 + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP + ARlag3 PIS3 + F4,
```

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, the syntax below use only ARlag1 to specify the equality of parameter. Figure 5 show the path diagram of AR LGM with equality constraint of the AR parameter.

```
* Linear LGM - AR(1);
* Equality of parameter;
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP + ARlag1 PIS1 + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP + ARlag1 PIS2 + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP + ARlag1 PIS3 + F4,
```

MOVING AVERAGE LGM

It is common to observe autocorrelated residuals in longitudinal data. Generally, LGMs assume temporal manifest variable errors are uncorrelated. This assumption may not be tenable as it has been long recognized that nuisance correlations among the manifest errors often emerge in longitudinal data analyses. The problem of not specifying the correlated errors in the model would result in bias the estimates (Marsh 1993, Marsh and Grayson, 1994a, 1994b; Sivo, Pan, & Brophy, 2004). Sivo (2001) indicates that MA LGM model helps to address the stability of a construct over time. The specification of MA LGM helps to relate current value of a time series as a function of the previous autocorrelated residuals. The following shows the SAS® syntax of the MA LGM by adding the three MA parameters (MAlag1, MAlag2 and MAlag3) to the left hand side of the respective manifest variable in each year. Figure 6 shows the path diagram of the MA linear LGM. Conceptually, it is more difficult to explain the MA specification as compare with AR specification in time series analysis. As such, it is more common in practical research to use AR instead of MA in explanation. Similar to time series analysis regarding the property and relationship between AR and MA, it was noted by Hamaker (2005) that algebraically both AR and MA LGM are equivalent if the autoregressive parameter is invariant over time and it lies between -1 and 1.

```
* Linear LGM - MA(1);
Proc CALIS Data=A UCOV Augment MaxIter=6000 MaxFunc=7000 all;
Title "Linear Latent Growth Model - AR(1)";
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP + MAlag1 F1 + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP + MAlag2 F2 + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP + MAlag3 F3 + F4,
```

Similar to the reason of setting equality in AR parameters in AR LGM, we could also limit the MA parameters by setting the equality constraint (See Figure 7). The syntax in the LINEQS statement is similar to that of the AR specification. In the example below, we use the same name MAlag1 to set equality of the estimate parameter. The mathematical limitation of setting of equality constraint of the AR and MA terms may not be so much a concern when we build multivariate LGM as there is more degree of freedom to play with. However, it is also noted that when the estimated coefficients of the AR and MA terms do not differ much, there is no harm to restrict them to equality as it generally makes the interpretation much easier and a reasonable theoretical explanation could well accompany with it.

```
* Linear LGM - MA(1);
* Equality of parameter;
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP + MAlag1 F1 + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP + MAlag1 F2 + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP + MAlag1 F3 + F4,
```

ARMA LGM

As stationary time series is evidenced by the presence of both AR and MA processes, it is not surprising that in the literature of LGM following the proposal of AR LGM and MA LGM, the ARMA LGM was subsequently suggested by Sivo et al (2005). ARMA LGM aims to account the AR for the correction of the autocorrelated observed scores, and at the same time, the MA for the autocorrelated residuals. One of the main reasons of including both the AR and MA parameters into the LGM is to filter out the effects of autocorrelation so that the estimates are more accurate (Sivo and Fan, 2008). Contrary to the principle of parsimony, Sivo and Fan (2008) argue and show that the fit of ARMA LGM is better than the simpler model of AR LGM and MA LGM. Sivo and Willson (2000) also note and suggest that an AR LGM may not sufficiently account for the effects of autocorrelation and it could either due to MA or AR or both. Following this line of argument, Sivo, Fan, and Witta (2009) have stated that it is not just warrant to specify the ARMA terms but should be mandatory to correct the bias estimate of not specifying them. Sivo, Fan, and Witta (2005) suggest that when there are at least four repeated measures, it is sensible to specify at least an AR component when modeling LGM. This is because it is not possible to model ARMA LGM with four time points due to identification. The specification in PROC CALIS of ARMA LGM is simple by including both the AR and MA parameters in LINEQS statement. The example below contains only the two terms, ARlag1 and MAlag1 to specify for the AR and MA parameters respectively. By restricting the equality of the parameter of both AR and MA terms, we could get an overidentified model, otherwise there is lack of degree of freedom for modeling a 4-wave data for ARMA LGM. Figure 8 shows the path diagram of ARMA linear LGM which is not possible to model due to underidentification whereas Figure 9 are constrained to give an over identified model.

```
* Linear LGM - ARMA(1,1);
* Equality of parameter in AR and MA;
LINEQS
  PIS1=1.0 F_INT + 0.0 F_SLP + F1,
  PIS2=1.0 F_INT + 1.0 F_SLP + ARlag1 PIS1 + MAlag1 F1 + F2,
  PIS3=1.0 F_INT + 2.0 F_SLP + ARlag1 PIS2 + MAlag1 F2 + F3,
  PIS4=1.0 F_INT + 3.0 F_SLP + ARlag1 PIS3 + MAlag1 F3 + F4,
```

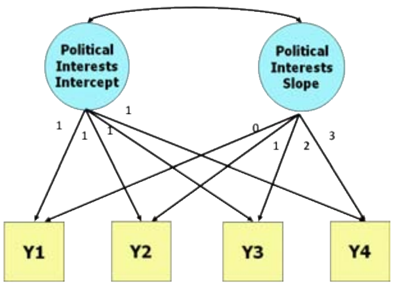

Figure 1 Unconditional Linear LGM

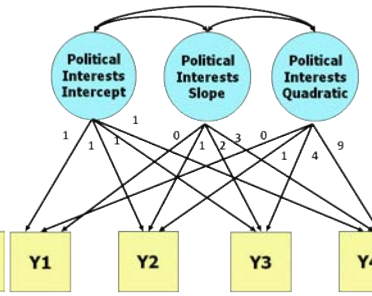

Figure 2 Unconditional Quadratic LGM

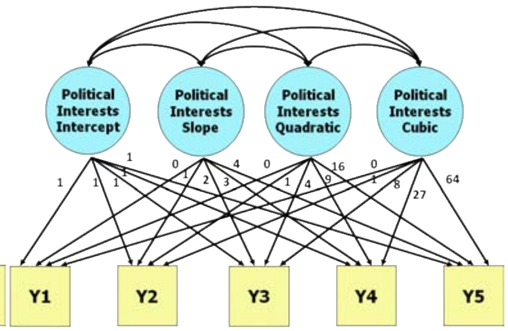

Figure 3 Unconditional Cubic LGM

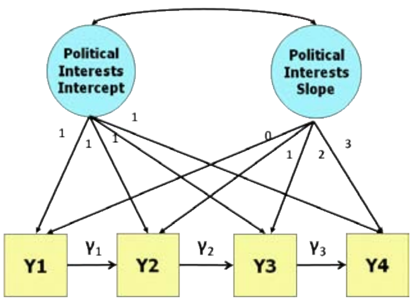

Figure 4 AR Linear LGM

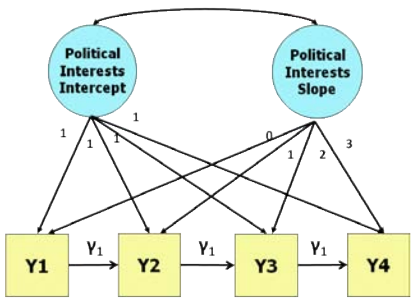

Figure 5 AR Linear LGM With Constraints

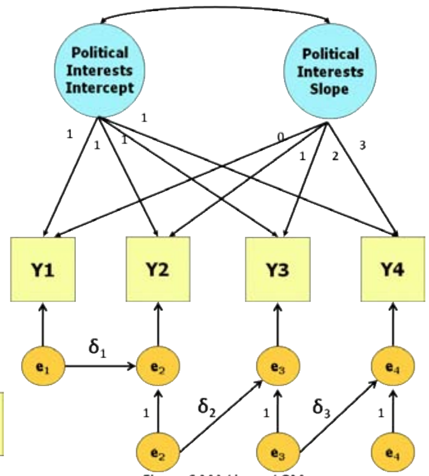

Figure 6 MA Linear LGM

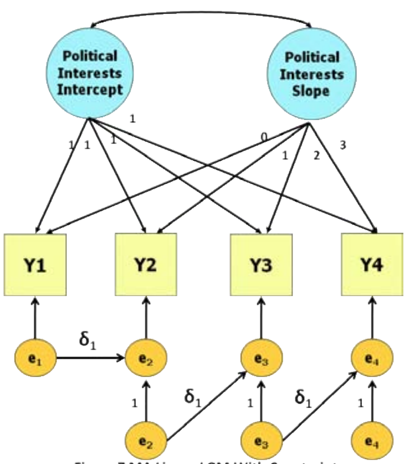

Figure 7 MA Linear LGM With Constraints

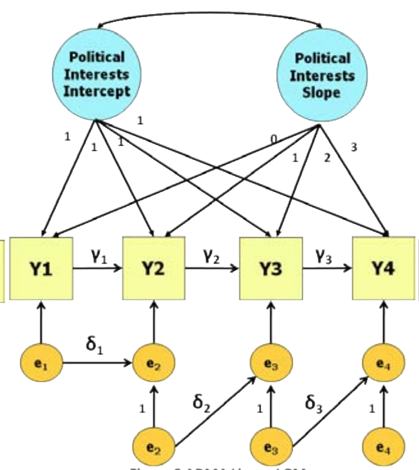

Figure 8 ARMA Linear LGM

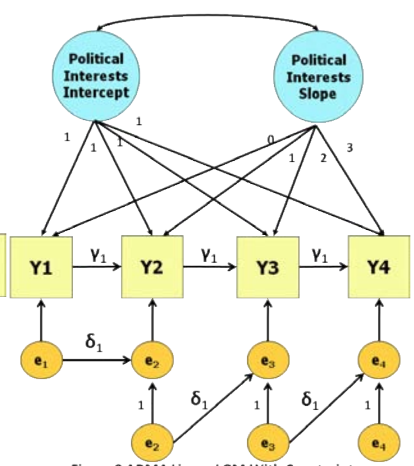

Figure 9 ARMA Linear LGM With Constraints

RESULTS - COMPARISON OF UNIVARIATE LGM MODELS

Seven models are illustrated using the survey data to examine the effects of student's political interests over time. The first model is the simplest LGM among all the models. The second model, Quadratic LGM, aims to account for the non-linear growth pattern. As there are only 4 waves of data, we cannot model cubic LGM as it is underidentified. The next five models add in the autoregressive and moving average parameters into the linear LGM to account for the temporal relationships of the manifest political interest variable over adjacent time. The first model with these lag terms is the AR LGM, follows by MA LGM. We have also put equality constraints on the AR LGM with the autoregressive parameters, and the MA LGM with equality constraint on the moving average parameters, as well as the ARMA LGM with both AR and MA equality constraints. As mentioned above, the main reason of setting equality constraint for the ARMA LGM is due to the identification problem. If they are all set free to be estimated, the model is underidentified.

Five fit indices are used for all the illustrated examples. They are Goodness of Fit Index (GFI), Comparative Fit Index (CFI), McDonald's Centrality Index (Mc), Standardized Root Mean Squared Residual Estimate (SRMR) and Root Mean Square Error of Approximation (RMESA). 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 and Mc are stand-alone indices that have a long history in the SEM literature. The CFI and incremental fit indices indicate the fit of a model improves on the nested null model. The SRMR summarizes the residual variation. The RMSEA is an estimate of misfit at the population rather than sample. We also report Chi-Square statistics although the literature has noted its limitations when there are violations of distributional assumption and large N. For comparison of nested model, the Chi-Square difference test is carried out.

Table 1 reports the results of the univariate LGMs. In term of goodness of fit, the quadratic model would be the model of choice rather than linear LGM. However, it has only 1 degree of freedom left. This makes further modeling difficult. More importantly, for a 4 time point longitudinal study, it would be less justifiable to choose a quadratic model as we may not be sure of the existing of a non-linear growth pattern with 4 observations and parsimony may give weight over model fit. So, we use linear LGM as the base for all the rest of the illustrated examples. The fit of the 5 models with AR or MA terms is better than the Linear LGM without the AR or MA parameters. Among the 5 AR or MA models, it turns out the ARMA LGM has the best fit. It has a non-significant Chi-square statistics which shows that the model fits well and this is supported by the other fit statistics as well.

Table 1 Comparison of LGMs – Political Interests of Students (Grade 7 to Grade 10)

Parameter / Fit Statistics	Linear LGM	Quadratic LGM	AR Linear LGM	MA Linear LGM	AR Linear LGM With Constraint	MA Linear LGM With Constraint	ARMA Linear LGM With Constraint
Regression Weights							
Latent Intercept	2.5598	2.6090	2.6039	2.5634	2.5718	2.5619	2.6079
Latent Growth	0.3137	0.1778	0.1436	0.3138	0.3275	0.3133	0.3731
Latent Quadratic		0.0435					
Political Interest Grade 7			0.0310^		-0.0154^		-0.0647
Political Interest Grade 8			0.0945		-0.0154^		-0.0647
Political Interest Grade 9			0.1556		-0.0154^		-0.0647
Residual Political Interest G7				0.0779^		0.0845	0.1456
Residual Political Interest G8				0.0963		0.0845	0.1456
Residual Political Interest G9				-0.00693^		0.0845	0.1456
Variances / Covariances							
Intercept	1.09061	1.26856	1.04125	0.95745	1.13414	0.93171	1.01784
Slope	0.15118	0.63142	0.10782	0.14473	0.16060	0.11835	0.13902
Quadratic		0.03808					
Intercept and Growth	-0.09373	-0.31223	-0.11717	-0.05685	-0.10320	-0.03849	-0.04658
Intercept and Quadratic		0.05042					
Growth and Quadratic		-0.13290					
Residual Variances							
Political Interest Grade 7	1.22910	1.00819	1.27008	1.32185	1.19064	1.3665	1.28265
Political Interest Grade 8	1.21746	1.17306	1.27367	1.32903	1.20475	1.30806	1.31837
Political Interest Grade 9	1.04422	0.96322	1.21152	1.07462	1.03115	1.12821	1.12873
Political Interest Grade 10	0.76233	0.72623	0.89893	0.70085	0.74410	0.85972	0.83680
R^2							
Intercept	0.8574	0.8430	0.8669	0.8729	0.8537	0.8758	0.8699
Growth	0.3044	0.0477	0.1606	0.4051	0.4006	0.4535	0.5005
Quadratic		0.0473					
Fit Indices							
Chi-Square	21.7326	0.8704	5.9614	12.2669	20.8683	13.7364	3.8325
Df	5	1	2	2	4	4	3
p-value	0.0001	0.3508	0.0508	0.0022	0.0003	0.0082	0.2801
GFI	0.9947	1.0000	0.9986	0.9971	0.9949	0.9967	0.9991
CFI	0.9986	0.9999	0.9997	0.9992	0.9986	0.9992	0.9999
Mc	0.9950	1.0000	0.9988	0.9969	0.9950	0.9971	0.9998
SRMR	0.0155	0.0044	0.0023	0.0181	0.0115	0.0165	0.0040
RMSEA	0.0447	0.0000	0.0344	0.0553	0.0501	0.0381	0.0129

Note: ^ indicates p>.05

MULTIVARIATE LGM

Longitudinal studies and research generally involve more than one variable of interest. These repeated measures variables are generally correlated over time and may have different development trajectories. One variable may increase over time while others plateau or decrease. The multivariate LGM becomes handy for analyzing these variables simultaneously as it caters for different trajectories for the variables of interest and yet allows us to examine their interrelationships with respect to their latent factors such as intercepts and growths. The strength of SEM approach using PROC CALIS instead of PROC MIXED for LGM is that it caters for multivariate LGM. More importantly, the syntax of specifying multivariate LGM can be easily carried out by duplicating the syntax of univariate LGM with slight modification.

ASSOCIATIVE LGM

The associative LGM is perhaps a good starting model, one of the basic multivariate models, to look at before we proceed to more complicated LGMs. The pre-condition is that before we get into associative LGM, we have already found out the appropriate trajectories of the variables of interest for the separate univariate LGMs. For instance, when we examine two time series variables, we already establish that the two univariate LGMs fit well as linear LGMs. The associative LGM would examine the correlations among these four latent means and slopes for the two linear LGMs. If one of the variables fits well as linear LGM and the other as quadratic LGM, the associative LGM would examine the relationships of the five latent factors namely the two latent mean and slope for the first model and the three latent mean, slope and quadratic for the second model.

Specification of associative LGM in PROC CALIS is quite straight forward. What we have to do is to cut and paste syntax from the separate univariate LGMs and put them together in the appropriate sections of PROC CALIS. We use the second data set which contains two variables of interest namely social skills (SOCS) and multi-literacy skills (Multil) of students to illustrate this. The specification of these two variables within the LINEQS statement show similar syntax. The only addition is perhaps the number of variance and covariance terms for the multivariate LGM increase according in the STD and COV statement. Note that the standardized coefficients for the covariances are the correlations of the latent factors which are one of the main outcomes of interest when we model associative LGM. These correlation coefficients would enable us to examine the interrelationships of the latent factors for these two variables. Figure 10 depicts the associative LGM in graphical path diagram. The double arrows of the four constructs represent the covariance/correlations of the latent means and slopes of SOCS and Multil. These correlations help us to understand the relationships of the initial and growth means of SOCS and Multil. For instance, a positive coefficient of SOCS intercept factor and Multil intercept factor indicate if students have high social skills at the initial level would also have high multi-literacy skills at the initial level.

```
LINEQS
  SOCS_S1       = 1.0 F_SOCS_INT + 0.0 F_SOCS_SLP + F1,
  SOCS_S2       = 1.0 F_SOCS_INT + 1.0 F_SOCS_SLP + F2,
  SOCS_S3       = 1.0 F_SOCS_INT + 2.0 F_SOCS_SLP + F3,
  SOCS_S4       = 1.0 F_SOCS_INT + 3.0 F_SOCS_SLP + F4,
  F_SOCS_INT    = Mean_SOCS_INT INTERCEPT + D1,
  F_SOCS_SLP    = Mean_SOCS_SLP INTERCEPT + D2,
  Multil_S1     = 1.0 F_Multil_INT + 0.0 F_Multil_SLP + F5,
  Multil_S2     = 1.0 F_Multil_INT + 1.0 F_Multil_SLP + F6,
  Multil_S3     = 1.0 F_Multil_INT + 2.0 F_Multil_SLP + F7,
  Multil_S4     = 1.0 F_Multil_INT + 3.0 F_Multil_SLP + F8,
  F_Multil_INT = Mean_Multil_INT INTERCEPT + D3,
  F_Multil_SLP = Mean_Multil_SLP INTERCEPT + D4;

STD
  D1-D4=ZETA1-ZETA4, F1-F8=8*err:;
COV
  D1 D2 = ZETA21,
  D1 D3 = ZETA31,
  D1 D4 = ZETA41,
  D2 D3 = ZETA32,
  D2 D4 = ZETA42,
  D3 D4 = ZETA43;
```

FACTOR-OF-CURVES LGM

The factor-of-curves LGM fits factors with higher order to describe the lower order factors (McArdle, 1988). The illustrated example fits two higher order factors: One common intercept factor for the first-order intercept factor of SOCS and Multil and one common slope factor that incorporates both the first-order slope factor of SOCS and Multil. Figure 11 shows the path diagram of factor-of-curves LGM. For the syntax of PROC CALIS, we have to add in six additional equations to represent the relationship of the first-order factors and second-order factors and states the mean estimates of the second-order factors as shown below. The factor-of-curve LGM would be useful to summarize the lower factors as one high-order factor for easier interpretation. For instance, the two intercept factors of SOCS and Multil could be represented as one high order intercept factor. In the same vein, the two slope factors of SOCS and Multil could also grouped as a second order slope factor.

```
LINEQS
  F_SOCS_INT       =   Loading_SOCS_INT  F_INT + D1,
  F_SOCS_SLP       =   Loading_SOCS_SLP  F_SLP + D2,
  F_Multil_INT     =   Loading_Multi_INT F_INT + D3,
  F_Multil_SLP     =   Loading_Multi_SLP F_SLP + D4,
  F_INT            =   Mean_INT   INTERCEPT + D5,
  F_SLP            =   Mean_SLP   INTERCEPT + D6;
```

CONDITIONAL FACTOR-OF-CURVES LGM

The above LGMs discussed so far are referred to as unconditional models as they are without predictors to explain the differences in the mean values of latent initial and growth. Although unconditional LGMs permit us to chart the course of variable of interest over time by means of their initial value and growth, many research questions are concerned about the factors that contribute to the development and growth of the variable in question. When models incorporate covariates, we call them as conditional LGMs. Figure 12 shows the path diagram of conditional factor-of-curves LGM with one time-independent explanatory variable - gender. This is represented by the path of gender (male as the dummy variable) to the common intercept and slope factors. The syntax of including the covariates into the model is straight forward by specifying the name of the estimated coefficient and the variable name after the common intercept and slope factors (F_INT and F_SLP respectively) to measure the effect of the covariate to the common initial and slope respectively.

```
  F_INT           = Mean_INT     INTERCEPT     +  Male_INT     Male   +  D5,
  F_SLP           = Mean_SLP     INTERCEPT     +  Male_SLP     Male   +  D6;
```

Additional covariates can be added after the gender covariate as shown below. The Malay is the race dummy variable for the ethnic group Malay.

```
  F_INT           = Mean_INT     INTERCEPT     +  Male_INT     Male
                                               +  Malay_INT    Malay
                                               +  …                   +  D5,
```

AR, MA, ARMA FACTOR-OF-CURVES LGM

Similar to the argument and conceptions of the earlier modeling approach of incorporating AR, MA, and ARMA terms into the LGM, the multivariate LGM with multiple variables can also build in these parameters. Figure 13, 14, and 15 display the path diagram of AR, MA, and ARMA factor-of-curves LGM respectively. The conditional ARMA factor-of-curves LGM is shown in Figure 16 with one covariate, MALE, added to the model. The PROC CALIS syntax of these models follows the syntax described earlier so they are not duplicated.

CROSS-LAG AR FACTOR-OF-CURVES LGM

In longitudinal analysis, we might be interested to examine the temporal relationships across different measures. For instance, in our example, we would like to know the effect of the first wave social skills on second wave of multi-literacy skills. This could be done by introducing the cross-lag terms into the model as shown by Figure 17. We could retain the AR parameters for measuring the effect on the same variable over time and at the same time specifying the cross-lag terms to examine the effect over different measures. Figure 18 extends the cross-lag model by incorporating the covariates. The syntax of LINEQS statement of including the cross-lag terms are shown below. For instance, ARMultiCrosslag1 is the name given to the coefficient to measure the effect of multi-literacy skills in wave one on social skills in wave 2.

```
LINEQS
  SOCS_S1        = 1.0 F_SOCS_INT   + 0.0 F_SOCS_SLP                                                      + F1,
  SOCS_S2        = 1.0 F_SOCS_INT   + 1.0 F_SOCS_SLP   + ARSOCSlag1  SOCS_S1 + ARMultiCrosslag1 Multil_S1 + F2,
  SOCS_S3        = 1.0 F_SOCS_INT   + 2.0 F_SOCS_SLP   + ARSOCSlag2  SOCS_S2 + ARMultiCrosslag2 Multil_S2 + F3,
  SOCS_S4        = 1.0 F_SOCS_INT   + 3.0 F_SOCS_SLP   + ARSOCSlag3  SOCS_S3 + ARMultiCrosslag3 Multil_S3 + F4,
  Multil_S1      = 1.0 F_Multil_INT + 0.0 F_Multil_SLP                                                    + F5,
  Multil_S2      = 1.0 F_Multil_INT + 1.0 F_Multil_SLP + ARMultilag1 Multil_S1 + ARSOCSCrosslag1  SOCS_S1 + F6,
  Multil_S3      = 1.0 F_Multil_INT + 2.0 F_Multil_SLP + ARMultilag2 Multil_S2 + ARSOCSCrosslag2  SOCS_S2 + F7,
  Multil_S4      = 1.0 F_Multil_INT + 3.0 F_Multil_SLP + ARMultilag3 Multil_S3 + ARSOCSCrosslag3  SOCS_S3 + F8,
```

For the 3 conditional multivariate models, 5 demographic variables are used in the study to examine their effects on student's social skills and multi-literacy skills. The demographic variables include gender, race, parent's marital status, stream, and types of housing. There are three main ethnic/racial groups in Singapore, namely Chinese, Malay, and Indian and they are included in the analysis as standalone categories. Students from other ethnic/racial groups are grouped into one named as "other ethnic group" since they are very small in terms of percentage and number. Students' stream/track in secondary school consists of gifted/special, express, normal academic, and normal technical (list from the highest to the lowest streaming). Gender, race, and stream are dummy coded. Since most of the students come from families with parents who are married and stayed together, it is also dummy coded as a dichotomous variable.

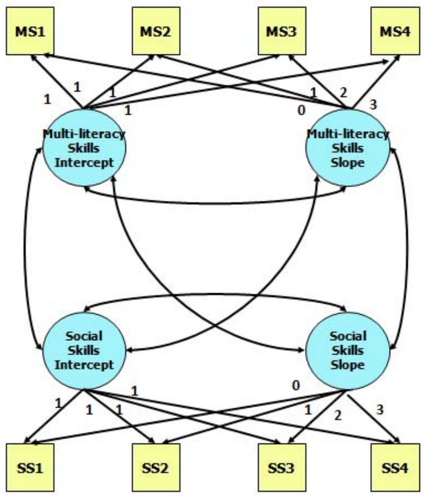

Figure 10 Associative LGM

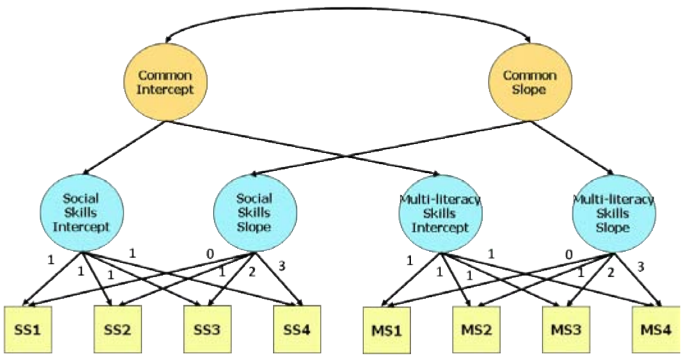

Figure 11 Factor-of-Curves LGM

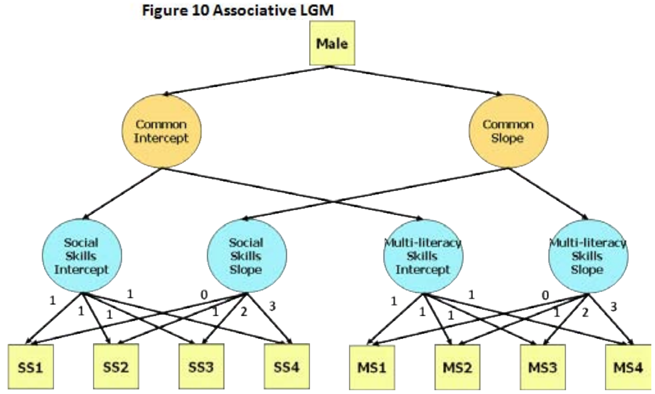

Figure 12 Conditional Factor-of-Curves LGM

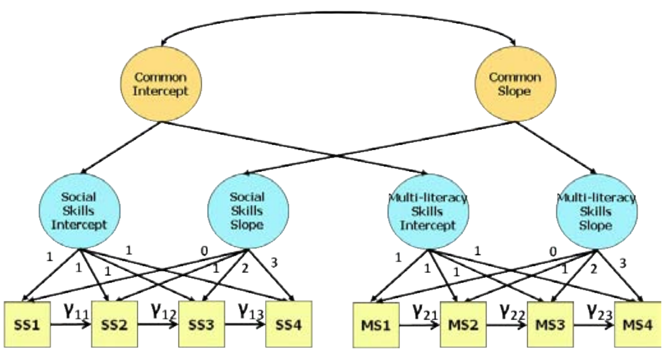

Figure 13 AR Factor-of-Curves LGM

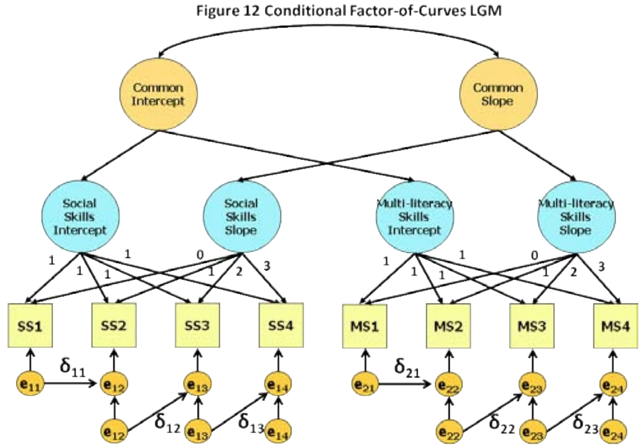

Figure 14 MA Factor-of-Curves LGM

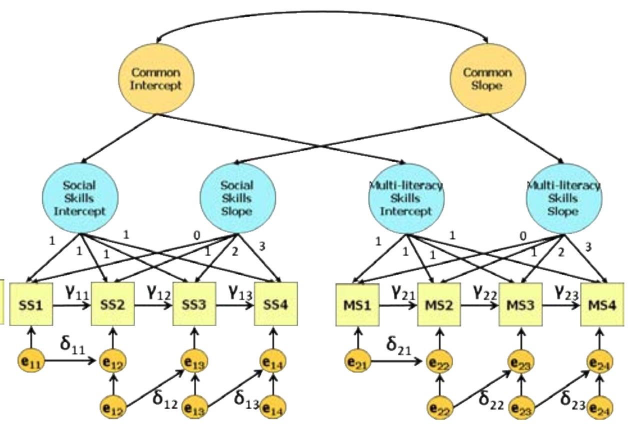

Figure 15 ARMA Factor-of-Curves LGM

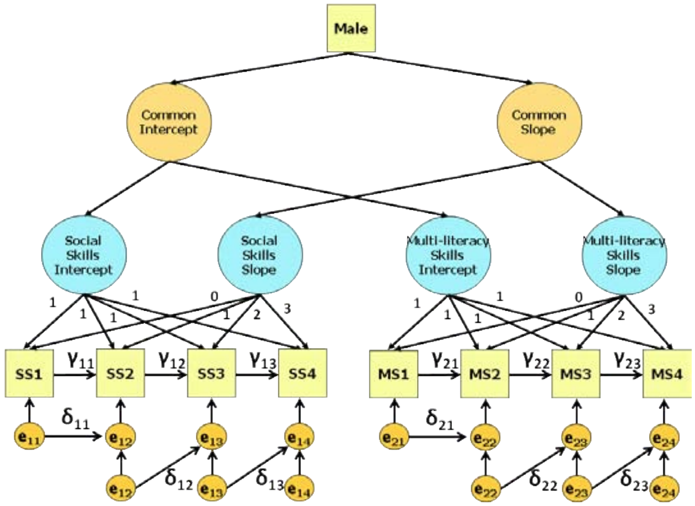

Figure 16 Conditional ARMA Factor-of-Curves LGM

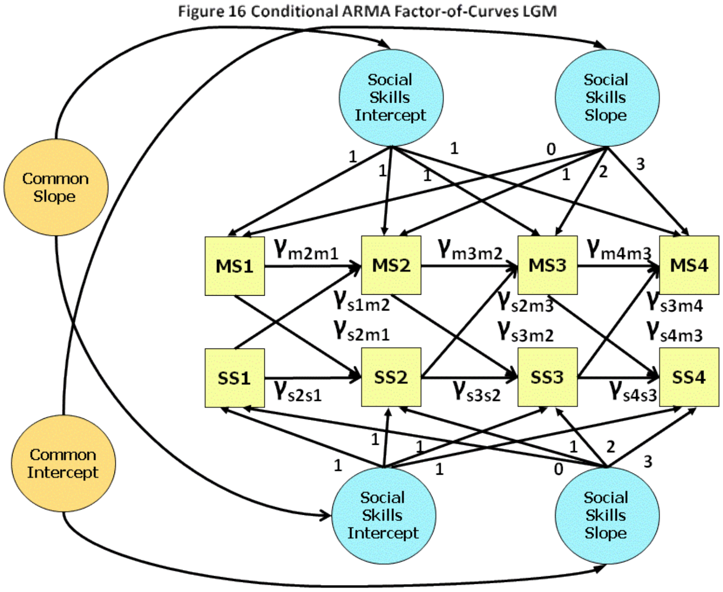

Figure 17 Cross-Lag AR Factor-of-Curves LGM

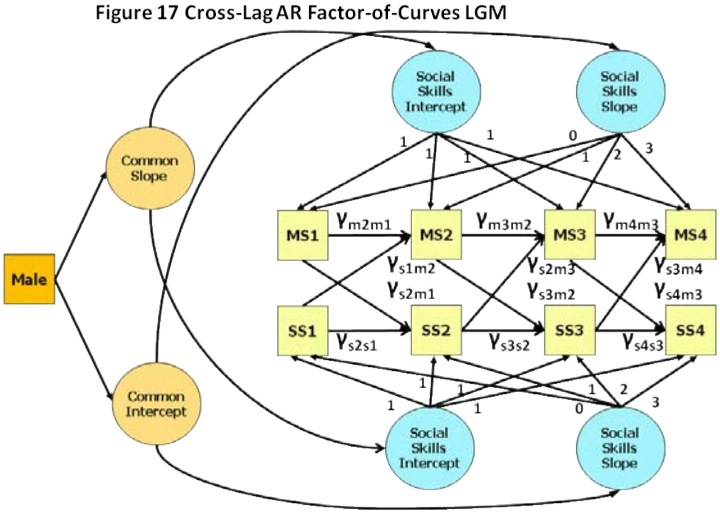

Figure 18 Conditional Cross-Lag AR Factor-of-Curves LGM

Table 2 Comparison of Multivariate LGMs – Social/Leadership Skills and Multi-literacy Skills of Students (Grade 7 to Grade 10)

Parameter / Fit Statistics	Associative LGM	Factor-of-Curve LGM	Conditional Factor-of-Curve LGM	AR Factor-of-Curve LGM	MA Factor-of-Curve LGM	ARMA Factor-of-Curve LGM	Conditional ARMA Factor-of-Curve LGM	Cross-Lag AR Factor-of-Curve LGM	Conditional Cross-Lag AR Factor-of-Curve LGM
Regression Weights									
Intercept –Social Skills (SS)	4.0140	0.9969*	0.9970*	0.9974*	0.9967*	0.9938*	0.9945*	0.9982*	0.9979*
Growth – SS	0.1029	1.0000*	1.0000*	1.0000*	-1.0000*	1.0000*	1.0000*	1.0000*	1.0000*
Intercept - Multi-literacy Skills (MS)	4.0712	0.9978*	0.9979*	0.9973*	0.9982*	0.9961*	0.9969*	0.9978*	0.9983*
Growth - MS	0.0287	0.7314*	0.7148*	0.8486*	-0.7172*	0.9840*	0.9742*	0.7756*	0.9370*
Common Intercept	-	4.0387	3.8796	4.0429	4.6485	4.0430	3.8908	4.0159	3.9496
Common Growth	-	0.2817	0.6235	0.0183^	-1.0646	1.4691	1.6576	0.3235	0.5306
SS Grade 7	-	-	-	0.0329	-	-0.3178	-0.2209	0.0776	0.0673
SS Grade 8	-	-	-	0.0313^	-	-0.6482	-0.4610	0.0123^	-0.0112^
SS Grade 9	-	-	-	0.0731	-	-0.9452	-0.6646	0.0599^	0.0011^
Residual SS G7	-	-	-	-	0.0058^	0.2985	0.1766	-	-
Residual SS G8	-	-	-	-	0.0098^	0.6736	0.4881	-	-
Residual SS G9	-	-	-	-	-0.1604	1.0000	0.6726	-	-
MS Grade 7	-	-	-	0.0163^	-	-0.2098	-0.1372	-0.0735	0.0483
MS Grade 8	-	-	-	0.0060^	-	-0.4396	-0.2970	-0.0953	0.0160^
MS Grade 9	-	-	-	0.0216^	-	-0.6505	-0.4346	-0.1165	0.0129^
Residual MS G7	-	-	-	-	-0.0382^	0.1461	0.0848	-	-
Residual MS G8	-	-	-	-	0.1050	0.5304	0.3900	-	-
Residual MS G9	-	-	-	-	0.0632^	0.7083	0.4899	-	-
Cross-Lag SS Grade 7	-	-	-	-	-	-	-	0.0440	-0.0786
Cross-Lag SS Grade 8	-	-	-	-	-	-	-	0.0133^	-0.0998
Cross-Lag SS Grade 9	-	-	-	-	-	-	-	0.0058^	-0.1263
Cross-Lag MS Grade 7	-	-	-	-	-	-	-	-0.0943	-0.0951
Cross-Lag MS Grade 9	-	-	-	-	-	-	-	-0.0779	-0.0761
Cross-Lag MS Grade 9	-	-	-	-	-	-	-	-0.1348	-0.1084
Initial on Male	-	-	0.0332^	-	-	-	0.0219^	-	0.0361^
Initial on Malay	-	-	0.2923	-	-	-	0.3015	-	0.2949
Initial on Indian	-	-	0.6356	-	-	-	0.6349	-	0.6476
Initial on Others	-	-	0.3684	-	-	-	0.3713	-	0.3683
Initial on Married	-	-	-0.0352^	-	-	-	-0.0232^	-	-0.0325^
Initial on Gifted/Special	-	-	0.2591	-	-	-	0.2560	-	0.2618
Initial on Normal Academic	-	-	-0.1263	-	-	-	-0.1099^	-	-0.1285
Initial on Normal Technical	-	-	-0.0709^	-	-	-	-0.0554^	-	-0.0723^
Initial on 4-Room	-	-	0.0137^	-	-	-	0.0303^	-	0.0149^
Initial on 5-Room	-	-	0.0907^	-	-	-	0.1159	-	0.0890^
Initial on Private Property	-	-	0.2120	-	-	-	0.2217	-	0.2171
Growth on Male	-	-	-0.0911^	-	-	-	-0.0197^	-	-0.0334^
Growth on Malay	-	-	-0.1751	-	-	-	0.0088^	-	-0.0461^
Growth on Indian	-	-	-0.2058	-	-	-	0.1239	-	-0.0300^
Growth on Others	-	-	-0.0609^	-	-	-	0.0967^	-	0.0127^
Growth on Married	-	-	-0.0319^	-	-	-	-0.0414^	-	-0.0224^
Growth on Gifted/Special	-	-	0.0928^	-	-	-	0.1396	-	0.0592^
Growth on Normal Academic	-	-	-0.3606	-	-	-	-0.2330	-	-0.1508
Growth on Normal Technical	-	-	-0.1556	-	-	-	-0.1176	-	-0.0693
Growth on 4-Room	-	-	-0.0452^	-	-	-	-0.0343^	-	-0.0221^
Growth on 5-Room	-	-	-0.0858^	-	-	-	-0.0353^	-	-0.0285^
Growth on Private Property	-	-	-0.1928	-	-	-	-0.0287^	-	-0.0623^
Correlation									
SS Intercept, SS Slope	-0.36651	-	-	-	-	-	-	-	-
SS Intercept, MS Intercept	0.80975	-	-	-	-	-	-	-	-
SS Slope, MS Intercept	-0.42406	-	-	-	-	-	-	-	-
SS Slope, MS Slope	-0.36314	-	-	-	-	-	-	-	-
SS Slope, MS Slope	0.87691	-	-	-	-	-	-	-	-
MS Intercept, MS Slope	-0.33077	-	-	-	-	-	-	-	-
Common Slope and Intercept	-	-0.43959	-0.45782	-0.49464	0.40857	0.14982	-0.0322^	-0.30368	-0.32111

Parameter / Fit Statistics	Associative LGM	Factor-of-Curve LGM	Conditional Factor-of-Curve LGM	AR Factor-of-Curve LGM	MA Factor-of-Curve LGM	ARMA Factor-of-Curve LGM	Conditional ARMA Factor-of-Curve LGM	Cross-Lag AR Factor-of-Curve LGM	Conditional Cross-Lag AR Factor-of-Curve LGM
Variances									
SS Intercept	0.44769	0.10321	0.10169	0.08651	0.11050	0.20538	0.18170	0.05874	0.07084
SS Slope	0.02856	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
MS Intercept	0.44350	0.07350	0.07041	0.09101	0.05998	0.13367	0.10489	0.07406	0.05891
MS Slope	0.02814	0.01329	0.01357	0.00785^	0.01269	0.02890	0.02236	0.00469^	0.00802
Common Intercept	-	0.35486	0.30971	0.35373	0.46209	0.36823	0.32769	0.36618	0.34251
Common Slope	-	0.35063	0.56730	0.12635	5.31344	0.08896	0.13839	0.06938	0.09492
Residual Variances									
SS Grade 7	0.38725	0.37906	0.38371	0.39142	0.37427	0.26528	0.29739	0.40443	0.40103
SS Grade 8	0.40634	0.40544	0.40512	0.41010	0.41811	0.40948	0.40274	0.39956	0.39441
SS Grade 9	0.36169	0.46211	0.35995	0.37877	0.33115	0.37849	0.36392	0.36419	0.34656
SS Grade 10	0.29427	0.27239	0.27235	0.30915	0.20160	0.32875	0.30813	0.30748	0.28403
MS Grade 7	0.29081	0.29503	0.29738	0.29051	0.29874	0.20657	0.23631	0.30037	0.30439
MS Grade 8	0.29490	0.29722	0.29732	0.29323	0.31427	0.29462	0.29478	0.28787	0.28878
MS Grade 9	0.27518	0.27571	0.27515	0.27991	0.29672	0.29523	0.29360	0.26817	0.27085
MS Grade 10	0.21781	0.21769	0.21919	0.21858	0.22976	0.23598	0.23470	0.22529	0.22626
R^2									
SS Intercept	0.9730	0.9938	0.9939	0.9948	0.9935	0.9877	0.9887	0.9965	0.9957
SS Growth	0.2707	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
MS Intercept	0.9740	0.9956	0.9958	0.9946	0.9958	0.9922	0.9937	0.9956	0.9965
MS Growth	0.0284	0.5350	0.5110	0.7201	0.4792	0.9684	0.9534	0.9282	0.8779
Common Intercept	-	0.9787	0.9815	0.9788	0.9790	0.9780	0.9808	0.9778	0.9803
Common Growth	-	0.1846	0.2259	0.0026	0.1785	0.9604	0.9518	0.6015	0.6713
Fit Indices									
Chi-Square	250.72	290.33	390.64	205.71	267.63	180.50	284.34	131.58	237.65
Df	22	25	90	18	18	12	78	12	78
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
GFI	0.9726	0.9684	0.9805	0.9772	0.9713	0.9807	0.9859	0.9848	0.9877
CFI	0.9961	0.9954	0.9958	0.9968	0.9957	0.9971	0.9971	0.9979	0.9978
Mc	0.9452	0.9368	0.9287	0.9548	0.9404	0.9594	0.9505	0.9710	0.9615
SRMR	0.0107	0.0132	0.0072	0.0013	0.0124	0.0009	0.0024	0.0014	0.0025
RMSEA	0.0716	0.0723	0.0406	0.0717	0.0827	0.0832	0.0361	0.0701	0.0318

Note: ^ indicates p>.05; SS. denotes Social Skills; MS. denotes Multi-literacy Skills; * Standardized Coefficient; Growth slope factor loadings to common slope factor set to 1 due to negative variance.

RESULTS - COMPARISON OF MULTIVARIATE LGM MODELS

The first model in the list is the associative LGM. The correlation of the latent initial of SS (Social Skills) and MS (Multi-literacy Skills) are high with correlation coefficient of 0.81 (Table 2). Similarly, the correlation of latent slope of SS and MS is also high (r=0.88). The correlations of initial and slope are all negatively and moderately correlated. Since both the intercept factors and slope factors are highly correlated, it is logically to build a factor-of-curves LGM to collapse these four first-order factors into two common second-order factors for easier modeling. This also makes more sense for interpretation when we include covariates later on into the models as relating the covariates to two second-order factors rather than four first-order factors make the analysis much easier. The fit of the factor-of-curves LGM is reasonable. Because the latent slope of SOCS to the common slope has negative variance, the error is set to zero.

The conditional factor-of-curves LGM shows that there is race, stream, type of housing effect on the initial level of both the social and multi-literacy skills represented by the common initial factor. These covariates also show significant effects on the common growth factor.

The next three models include the AR, MA, and ARMA terms as indicated as the AR, MA, and ARMA factor-of-curves LGM respectively. The Chi-square difference tests of comparing the ARMA LGM over the AR and MA LGMs show significant improvement on the ARMA LGM. Similarly, the conditional ARMA factor-of-curve LGM compared to conditional factor-of-curve LGM without ARMA parameters also show significant improvement for the Chi-square difference test for the conditional ARMA LGM. In short, the ARMA models outperform the other models.

The cross-lag AR factor-of-curve LGM has additional 6 parameters for the cross-lag compare to the AR factor-of-curve LGM. The fit of the cross-lag model shows significant improvement for the Chi-square difference test.

USEFUL HINTS FOR MODEL DIAGNOSTIC

There are a few options in the PROC CALIS that are useful while modeling LGM. In any modeling process, we sometimes would like to check the modeling details when the model is not producing the expected outcome. One way is to specify the ALL option after the PROC CALIS statement to print all optional output. If we have a rough idea of the estimated value of the parameter, we can put parentheses on the right side of the parameter to initialize the starting value. For instance COV D1 D2 = ZEWTA21 (-0.45). While parameters are within certain range we might want to use the BOUNDS statement so that we restrict them to the range we want the model to estimate. For instance, we would like to specify in the ARMA LGM to restrict the autoregressive parameter within the -1 and 1 boundary.

CONCLUSION AND SUGGESTIONS

This paper illustrates the use of PROC CALIS in carrying out the various latent growth models. Their syntaxes are given for all the models and differences in the syntax from a simpler to a more complex model are highlighted. Path diagrams for all the models are presented and referred to in the section of the models being discussed and their syntaxes elaborated on so that reader has a clear idea of the connection of linking the syntax to the pictorial presentation of the model. This will help readers to have better idea of what the various models are 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. The fit indices of the mix approach do show better fit for the combined approach when we compare them using Chi-square difference test for the nested model. The question of the incremental advantages in practical research to address research questions of using ARMA in LGM probably would need further research in this area.

ACKNOWLEDGEMENTS

This paper acknowledges the Centre for Research in Pedagogy and Practice (CRPP), National Institute of Education, Singapore for making the data available from the research project “Life Pathway Project” (CRP 46/08 DH). We would also like to thank Melvin Chen, Vicente Chua, Ridzuan and Alex Yeung who took time to go through the paper and give their comment.

RECOMMENDED READING

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. (2000). A latent curve framework for the study of developmental trajectories in adolescent substance use. In J. S. Rose, L. Chassin, C. C. Presson,&S. J. Sherman (Eds.), Multivariate applications in substance use research: New methods for new questions (pp. 1–42). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Curran, P. J., and Willoughby, M. T. (2003). Implications of latent trajectory models for the study of developmental psychopathology. *Development and Psychopathology, 15*, 581-612.

Duncan, T.E., Duncan, S.C., Strycker, L.A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Duncan, T. E., Duncan, S. C. and Strychker, L. A. (.2006). *An introduction to latent variable growth curve modeling*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fan, X. (2003). Power of latent growth modeling for detecting group differences in linear growth trajectory parameters. Structural Equation Modeling, 10, 380–400.

Hamaker, E. L. (2005). Conditions for the equivalence of the autoregressive latent trajectory model and a latent growth curve model with autoregressive disturbances. *Sociological Methods and Research, 33(3)*, 404-416.

McArdle, J. J. (1988). Dynamic structuring equation modeling of repeated measures data. In R. B. Cattell and J. Nesselroade (Eds.), *Handbook of Multivariate Experimental Psychology* (2nd Edition, pp. 561-614). New York: Plenum.

Marsh, H. W. (1993). Stability of individual differences in multiwave panel studies: Comparison of simplex models and one factor models. Journal of Educational Measurement, 30, 157–183.

Marsh, H. W., & Grayson D. (1994a). Longitudinal confirmatory factor analysis: Common, time-specific, item-specific, and residual-error components of variance. Structural Equation Modeling, 1, 116–145.

Marsh, H.W.,&Grayson D. (1994b). Longitudinal stability of latent means and individual differences: A unified approach. Structural Equation Modeling, 1, 317–359.

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.

McArdle, J. J., & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), Modeling longitudinal and multilevel data (pp. 69–108). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., and Briggs, N. E. (2008). *Latent Growth Curve Modeling*. SAGE Publications, Inc.

Sivo, S. and Fan, X. (2008). The latent curve ARMA(p,q) panel model: Longitudinal data analysis in education research and evaluation. *Educational Research and Evaluation, 14(4)*, 363-376.

Sivo, S. A. (2001). Multiple indicator stationary time series models. Structural Equation Modeling, 8(4), 599–612.

Sivo, S. A., Pan, C. C. and Brophy, J. (2004). Temporal cross-lagged effects between subjective norms and students' attitudes regarding the use of technology. *Journal of Educational Media and Library Sciences, 42(1)*, 63-73.

Sivo, S.A., Fan, X., & Witta, E.L. (2005). The biasing effects of unmodeled ARMA time series processes on latent growth curve model estimates. Structural Equation Modeling, 12(2), 215–232.

Sivo, S.A., & Willson, V.L. (1998). Is parsimony always desirable? Identifying the correct model for a longitudinal panel data set. Journal of Experimental Education, 66(3), 249–255.

Sivo, S.A., & Willson, V.L. (2000). Modelling causal error structures in longitudinal panel data: A Monte Carlo study. Structural Equation Modeling, 7(2), 174–205.

Sivo, S. A., Fan X., & Witta, L. (2009). The Biasing Effects of Unmodeled ARMA Time Series Processes on Latent Growth Curve Model Estimates, *Structural Equation Modeling, 12(2)*, 215–231

CONTACT INFORMATION

Teck Kiang Tan
Centre for Research in Pedagogy and Practice
Nanyang Technological University
1 Nanyang Walk, Block 5, Basement 3
Singapore 637616
Work Phone: 065-62196277
Email: teckkiang.tan@nie.edu.sg

Trivina Kang
Centre for Research in Pedagogy and Practice
Nanyang Technological University
1 Nanyang Walk, Block 5, Basement 3
Singapore 637616
Work Phone: 065-62196277
Email: trivina.kang@nie.edu.sg

David Hogan
Office of Education Research
Nanyang Technological University
1 Nanyang Walk, Block 5, Level 3
Singapore 637616
Work Phone: 065-62196277
Email: david.hogan@nie.edu.sg