INDIVIDUAL'S MIXED GROWTH TRACK USING MULTILEVEL STRUCTURAL EQUATION MODEL

See-Heyon Jo, Chongju National University of Education
S. H. Jo, 135 Sugok Dong, Chongju, Korea 361-712

Key Words : Multilevel data, Mixed growth track

Recent times, statistical model for capturing an individual's growth track in the context of latent variables modeling has been an important research topic. Muthen and Khoo(1997) implied that empirical Bayes approach may be a new choice for the research problem. The initial application of empirical Bayes approach to individual's growth curve was presented by Strenio, Weisberg and Bryk(1983). When we observe real world, we may postulate that individuals have their own unique mixed growth patterns. For example, for a span of time some individuals may show linearly growth, for other span they may show nonlinear growth.

In this paper to capture individual's mixed growth track we extend the empirical Bayes approach to multilevel structural equation model(Jo, 1994). By employing multilevel structural equation model we can incorporate the measurement error of the latent variables. The method of estimation proposed in this paper is useful for unbalanced multilevel data because it does not require classifying groups into subgroups with the same number of lower-units. Although it often has been pointed out that latent variables modeling of growth is less effective in utilizing individual's background variables, casting the latent variables model into the framework of multilevel structural equation model enables us to flexibly employ the individual's time varying and time invarying covariates. For illustrations we analyzed a set of longitudinal data.

We define the vector of observed scores for an individual i across times p(p=1,..., P) as y_i. We also define the matrix of time-varying covariate(i.e, age) and its powers at the p-th occasion of measurement as A_i. That is,

$$A_i = \begin{bmatrix} a_{1i}^c & 0 & 0 & \dots & 0 \\ 0 & a_{2i}^c & 0 & \dots & 0 \\ 0 & 0 & a_{3i}^c & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & a_{pi}^c \end{bmatrix}$$

Where a_{pi}^c is a row vector of age at the p-th occasion and its powers. The individual's mixed growth track is captured by permitting the form of $a_p^c = \lfloor a_p^1 \quad a_p^2 \quad \dots \quad a_p^c \rfloor$ to be different across occasions based on the substantive theory or empirical observations. We define the matrix Λ, as shown by Strenio et al. (1983), and Muthen et al. (1997), to capture individual's linear growth track across P occasions of measurement as $\Lambda = [\lambda_1 \quad \lambda_2]$. Where λ_1 is a p-dimension column vector of units, and λ_2 is a P-dimension column vector, its p-th element is p-1. Then we have the following model.

$$y_i = A_i\xi + \Lambda\eta_i + \varepsilon_i \qquad (1)$$
$$\varepsilon_i \sim N(0, \Sigma)$$

To obtain more meaningful interpretation from the mixed growth track, we may specify the vector a_p^c as centered to the mean value of the sub-group in which individuals have in common the form of vector a_p^c are nested. On the second stage of multilevel modeling, the individual's linear growth vector of parameters becomes the outcome vector. That is,

$$\eta_i = Z_i \pi + u_i \qquad (2)$$

$$u_i \sim N(0, \mathrm{T}_\eta)$$

Where Z_i is a matrix of time-invarying covariates such as gender, ethnicity. Individual's initial status and growth rate without considering the effect of age are explained by his stable background variables. We note that the matrix T_η contains the information about the effect of individual's initial's status on the linear rate of growth without considering the influence of age.

If a test has a set of subtests, then we can incorporate the measurement error into the model by adding measurement model.

$$y_i = B\theta_i + \varepsilon_i \qquad (3)$$

$$\varepsilon_i \sim N(0, \Sigma)$$

$$\theta_i = A_i \xi + \Lambda \eta_i + v_i \qquad (4)$$

$$v_i \sim N(0, \mathrm{T}_\eta)$$

$$\eta_i = Z_i \pi + u_i \qquad (5)$$

$$u_i \sim N(0, \mathrm{T}_\eta)$$

The main distinction is the matrix B in the measurement model equation (3) called a factor loading matrix.

Educators and evaluators assessing the effectiveness of intervention program are more interested in the effects of the group's stable characteristics as well as the group's time-varying characteristics on individual's growth. Then the set of model equations are :

$$y_{ij} = B\theta_{ij} + \varepsilon_{ij} \qquad (6)$$

$$\varepsilon_{ij} \sim N(0, \Sigma)$$

$$\theta_{ij} = A_{ij}\xi_j + \Lambda \eta_{ij} + v_{ij} \qquad (7)$$

$$v_{ij} \sim N(0, T_\theta)$$

$$\eta_{ij} = Z_{ij}\zeta_j + u_{ij} \qquad (8)$$

$$u_{ij} \sim N(0, \mathrm{T}_\eta)$$

$$\xi_j = X_j \gamma + \varpi_j \qquad (9)$$

$$\varpi_j \sim N(0, T_\xi)$$

$$\zeta_j = W_j \pi + \delta_j \qquad (10)$$

$$\delta_j \sim N(0, T_\zeta)$$

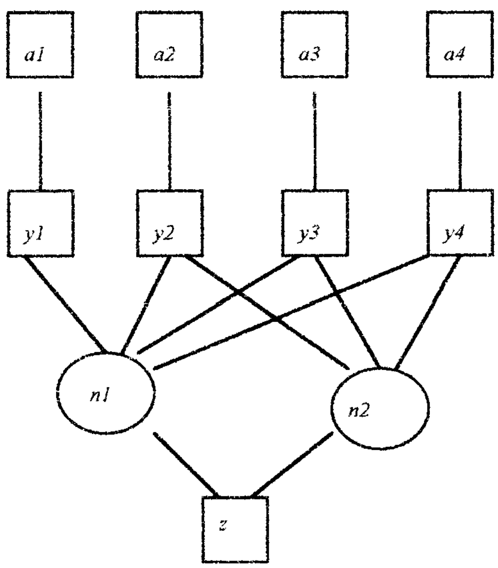
Where X_j is a matrix of group's time varying covariates. W_j is a matrix of group's time invarying covariates. At this point of model development we note that the influence of the time-varying latent trait on the next waves. Because the trait itself changes by the influence of the characteristics of him and his group, this changed trait influences the next growth trajectory of himself. Note that this influence process is occurred in a multilevel context. To investigate in a more detail we look into the variance-covariance matrix M, where

$$M = T_\theta + \Lambda Z T_\zeta Z^T \Lambda^T + \Lambda T_\eta \Lambda^T + A T_\xi A^T .$$

Then we obtain three matrices, a) T_θ, b)

$T_\theta + \Lambda Z T_\zeta Z^T \Lambda^T + \Lambda T_\eta \Lambda^T$, c) $T_\theta + A T_\xi A^T$.

These three matrices contain different types of information about the implicit influence of the changed trait on the next waves of growth. T_θ provides the amount of influence which does incorporate the individual's time-invarying information, group's time-invarying and time-varying information. Whereas from T_θ contained in the matrix (b), we obtain the amount of influence which incorporates only the individuals' time-invarying information. In addition, from T_θ contained in the matrix (c), we obtain the amount of influence which incorporates not only the individuals' time-varying information but also the group's time-varying information.



[Figure 1] A path-diagram for multilevel model

The more interesting model for educational evaluators is : $y_{ij} = \mathrm{B}\theta_{ij} + \mathrm{B}_b\theta_{bj} + \varepsilon_{ij}$ **(11)**

$$\theta_{ij} = \mathrm{A}_{ij}\xi_j + \Lambda\eta_{ij} + \nu_{ij} \qquad (12)$$

$$\theta_{bj} = A_{bj}\xi_b + \Lambda_b\eta_{bj} + \nu_{bj} \qquad (13)$$

The random vector $\xi_{ij}, \eta_{ij}, \eta_{bj}$ become the outcome vector as shown previously. All of the model shown in this paper can be capsuled into the general model equation (17).

Empirical Bayes Estimation via EM Algorithm

The aim of the empirical Bayes approach is to estimate the joint posterior modes of ξ, γ, η, π and ζ given $y, \Sigma, \mathrm{T}_\eta, \mathrm{T}_\theta, \mathrm{T}_\xi$ and T_ζ . This joint posterior distribution is given by

$$f(y, \xi, \kappa, \pi, \eta, \gamma \mid \Sigma, \mathrm{T}_\eta, \mathrm{T}_\theta, \mathrm{T}_\xi, \mathrm{T}_\pi) \propto$$

$$\begin{aligned} &f_1(y \mid \theta, \mathrm{B}, \Sigma) f_2(\theta \mid \xi, \eta, \mathrm{T}_\theta) f_3(\xi \mid \gamma, T_\xi) \\ &\times f_4(\eta \mid \pi, \mathrm{T}_\eta) f_5(\pi \mid \kappa, \mathrm{T}_\pi) f(\gamma) f(\kappa) \end{aligned} \qquad (14)$$

To determine the empirical Bayes estimator for $\xi, \gamma, \pi, \kappa, \eta$ the partial derivatives of (11) need to be taken with respect to $\xi, \gamma, \pi, \kappa, \eta$.

However there is a more compact way of deriving empirical Bayes estimators, $\xi^*, \gamma^*, \pi^*, \kappa^*, \eta^*$ by using $\omega^*, \delta^*, v^*, u^*$. To derive $\omega^*, \delta^*, v^*, u^*$ we use a general model obtained by substituting the higher level model equations into the lower level model equations. The general model is :

$$\begin{aligned} y = &BAX\gamma + B\Lambda ZW\pi + BA\omega + \\ &B\Lambda Z\delta + B\Lambda u + Bv + \varepsilon \end{aligned} \qquad (15)$$

In a more compact form , we have

$$y = \begin{bmatrix} BAX & B\Lambda ZW \end{bmatrix} \begin{bmatrix} \gamma \\ \pi \end{bmatrix} + \begin{bmatrix} BA & B\Lambda Z \end{bmatrix} \begin{bmatrix} \omega \\ \delta \end{bmatrix} + \begin{bmatrix} B & B\Lambda \end{bmatrix} \begin{bmatrix} v \\ u \end{bmatrix} + \varepsilon \qquad (16)$$

The model equation can be simplified as follows.

$$y = \widetilde{Z}\vartheta_1 + Z^*\vartheta_2 + Z^+\vartheta_3 + \varepsilon \qquad (17)$$

and $\begin{bmatrix} BAX & B\Lambda ZW \end{bmatrix} = \widetilde{Z}$, $\begin{bmatrix} B & B\Lambda \end{bmatrix} = Z^+$

$\begin{bmatrix} BA & B\Lambda Z \end{bmatrix} = Z^*$, $\begin{bmatrix} \gamma \\ \pi \end{bmatrix} = \vartheta_1$, $\begin{bmatrix} \omega \\ \delta \end{bmatrix} = \vartheta_2$,

Based on the results of the previous research(Jo, 1994), the maximum likelihood estimators of the unknown parameters, $B, \Sigma, T_\theta, T_\xi, T_\eta, T_\zeta$ and the estimates of the conditional expectation of the random vector $\vartheta_1, \vartheta_2, \vartheta_3, \varepsilon$ are given as follows.

$$\begin{aligned}
\hat{B}_{ml} = & \left[y_{ij}\vartheta_1^{*T}\widetilde{Z}_{ij}^T + y_{ij}\vartheta_2^{*T}\widetilde{Z}_{ij}^{*T} + y_{ij}\vartheta_3^{*T}\widetilde{Z}_{ij}^{+T} \right] \\
& \Big[\widetilde{Z}_{ij}\vartheta_1^*\vartheta_1^{*T}\widetilde{Z}_{ij}^T + Z_{ij}^*\vartheta_{2j}^{*T}\widetilde{Z}\vartheta_1^{*T}\widetilde{Z}_{ij}^T + \widetilde{Z}_{ij}\vartheta_1^*\vartheta_{2j}^{*T}Z_{ij}^{*T} \\
& + Z_{ij}^{+T}\vartheta_{3ij}^*\vartheta_1^{*T}\widetilde{Z}_{ij}^T + \widetilde{Z}_{ij}\vartheta_1^{*T}\vartheta_{3ij}^{*T}Z_{ij}^{+T} + Z_{ij}^{+T}\vartheta_{3ij}^*\vartheta_{2j}^{*T}Z_{ij}^{*T} \\
& + Z_{ij}^*\vartheta_{2j}^*\vartheta_{3ij}^{*T}Z_{ij}^{+T} + Z_{ij}^+\vartheta_{3ij}^*\vartheta_{3ij}^{*T}Z_{ij}^{+T} + \widetilde{Z}_{ij}D_{\vartheta_1}^*\widetilde{Z}_{ij}^T \\
& + Z_{ij}^*D_{\vartheta_2}^*Z_{ij}^{*T} + Z_{ij}^+D_{\vartheta_3}^*Z_{ij}^{+T} + \widetilde{Z}_{ij}C_{\vartheta_1\vartheta_2}^*Z_{ij}^{*T} \\
& + Z_{ij}^*C_{\vartheta_1\vartheta_2}^{*T}\widetilde{Z}_{ij}^T + \widetilde{Z}_{ij}C_{\vartheta_1\vartheta_2}^*Z_{ij}^{+T} + Z_{ij}^+C_{\vartheta_1\vartheta_3}^*\widetilde{Z}_{ij}^T \\
& + Z_{ij}^+C_{\vartheta_1\vartheta_3}^{*T}\widetilde{Z}_{ij}^T + Z_{ij}^*C_{\vartheta_2\vartheta_3}^*Z_{ij}^{+T} + Z_{ij}^+C_{\vartheta_2\vartheta_3}^{*T}Z_{ij}^{*T} \Big]^{-1}
\end{aligned}$$

$$\begin{aligned}
\hat{\Sigma}_{ml} = \frac{1}{N} \Big[\sum\sum \{ & \varepsilon_{ij}^*\varepsilon_{ij}^{*T} + B\widetilde{Z}_{ij}D_{\vartheta_1}^*(B\widetilde{Z}_{ij})^T \\
& + BZ_{ij}^*D_{\vartheta_2}^*(BZ_{ij}^*)^T + BZ_{ij}^+D_{\vartheta_3}^*(BZ_{ij}^+)^T \\
& + B\widetilde{Z}_{ij}C_{\vartheta_1\vartheta_2}^*Z_{ij}^{*T}B^T + BZ_{ij}^*C_{\vartheta_1\vartheta_2}^{*T}\widetilde{Z}_{ij}^TB^T \\
& + B\widetilde{Z}_{ij}C_{\vartheta_2\vartheta_3}^*Z_{ij}^{+T}B^T + BZ_{ij}^{+T}C_{\vartheta_2\vartheta_3}^*\widetilde{Z}_{ij}^TB^T \\
& + B\widetilde{Z}_{ij}C_{\vartheta_1\vartheta_2}^*Z_{ij}^{*T}B^T + BZ_{ij}^+C_{\vartheta_1\vartheta_2}^{*T}\widetilde{Z}_{ij}^TB^T) \Big]
\end{aligned}$$

Where $\varepsilon_{ij}^* = y_{ij} - B\widetilde{Z}_{ij}\vartheta_1^* - BZ_{ij}^*\vartheta_{2j}^* - BZ_{ij}^+\vartheta_{3ij}^*$,

$\hat{T}_{\xi(ml)} = USDIAG(\hat{T}_{2ml})$, USDIAG(.) means the upper subdiagonal part. $\hat{T}_{\zeta(ml)} = LSDIAG(\hat{T}_{2ml})$, LSDIAG(.) means lower subdiagonal part. Where

$$\hat{T}_{2ml} = \frac{1}{J}\left[\sum (\vartheta_{2j}^*\vartheta_{2j}^T + D_{\vartheta_{2j}}^*) \right]$$

And $\hat{T}_{\theta(ml)} = USDIAG(\hat{T}_{3ml})$, USDIAG(.) means the upper subdiagonal part. $\hat{T}_{\eta(ml)} = LSDIAG(\hat{T}_{3ml})$, LSDIAG(.) means lower subdiagonal part. Where

$$\hat{T}_{3ml} = \frac{1}{N}\left[\sum\sum (\vartheta_{3ij}^*\vartheta_{3ij}^T + D_{\vartheta 3_{ij}}^*) \right]$$

The formula for computing $\vartheta_1^*, \vartheta_2^*, \vartheta_3^*, D_{\vartheta_1}^*$ and $D_{\vartheta_2}^*, D_{\vartheta_3}^*, C_{\vartheta_1\vartheta_2}^*, C_{\vartheta_1\vartheta_3}^*, C_{\vartheta_2\vartheta_3}^*$ are given in Jo(1994).

The empirical Bayes estimators for an individual's mixed growth track is given as $\hat{y}_i = A_i\xi^* + \Lambda\eta_i^*$ for the model with only one group. And $\hat{\theta}_i = A_i\xi^* + \Lambda\eta_i^* + v_i^*$ for the model incorporating measurement error. And $\eta_i^* = Z_i\pi^* + u_i^*$.

Finally $\hat{\theta}_{ij} = A_{ij}\xi_j^* + \Lambda\eta_{ij}^* + v_{ij}^*$ for the model with many groups. $\eta_{ij}^* = Z_{ij}(W_j\pi^* + \delta_j^*) + u_{ij}^*$ and $X_j\gamma + \varpi_j^*$ for ξ_j^*.

Illustrative Data and Results

We apply the proposed model to a set of longitudinal data. Note that in this example we do not concerned about the reliability of measurements, and all of the individuals are nested within only one group. This is more realistic for a setting of a classroom teaching and for an educational program implemented at a school site. The raw data composed of four psychological test scores for individuals. The number of individuals is 30. The data and descriptive statistics of the sample are shown in table 1. For our modeling the matrix A_i is a polynomial of ages.

$$A_i = \begin{bmatrix} a_{1i}^c & 0 & 0 & 0 \\ 0 & a_{2i}^c & 0 & 0 \\ 0 & 0 & a_{3i}^c & 0 \\ 0 & 0 & 0 & a_{4i}^c \end{bmatrix}$$

Where $a_p^c = \left| a_p^1 \quad a_p^2 \quad a_p^3 \right|$.

The estimates shown in Table 1 through 6.

TABLE 1. Illustrative Data

Subject	v1	v2	v3	v4	age	gender
1	32.39	145.15	120.76	215.70	3.58	1
2	39.54	213.11	158.33	265.00	4.44	1
3	34.42	159.11	127.63	222.26	3.80	1
4	32.78	148.98	119.10	211.94	3.65	1
5	38.78	269.78	196.32	323.74	5.04	0
6	25.35	87.90	83.71	149.96	2.70	0
7	38.76	236.03	174.37	289.91	4.71	0
8	25.65	100.17	89.03	165.18	2.90	0
9	37.77	245.73	178.70	298.29	4.81	0
10	22.75	105.25	89.17	172.13	3.01	0
11	30.71	122.47	105.66	189.88	3.29	1
12	38.16	219.10	165.67	277.69	4.48	1
13	20.31	67.92	67.57	126.76	2.15	1
14	44.56	252.97	183.78	302.84	4.81	1
15	30.29	186.37	143.95	248.68	4.18	0
16	24.65	101.14	90.87	170.70	2.96	0
17	44.73	358.27	240.16	386.73	5.79	0
18	32.92	181.73	139.71	239.55	4.12	0
19	32.25	149.63	122.74	212.17	3.68	0
20	30.79	147.54	118.36	209.50	3.67	0
21	43.41	297.80	207.03	336.58	5.27	1
22	16.49	58.15	59.73	117.05	1.97	1
23	32.95	176.82	141.92	244.97	4.03	1
24	39.88	258.34	187.43	311.18	4.90	1
25	24.43	99.92	87.37	164.95	2.97	0
26	32.77	183.57	138.57	243.79	4.16	0
27	37.39	242.08	174.06	294.64	4.78	0
28	43.13	321.74	223.72	362.26	5.50	0
29	43.94	322.49	220.36	358.12	5.49	0
30	40.08	284.03	202.56	330.37	5.14	0
Mean	33.73	191.44	145.27	248.08	4.066	0.4
SD	7.60	81.97	49.61	73.25	1.015	0.49

From the Table 2 we find that for the first two spans, the growth track for the group is positively quadratic. However for the third span, the growth track is negatively quadratic. For the fourth span, the growth track is purely linear. From the Table 3 we find that when the age is null, the pure initial status for female is −31.6368, the gap between female and male is 3.221. While the linear growth rate across 4 time points for female is 19.8768, the gap in linear growth rate between female and male is 0.1413. the estimate of parameters $\eta_i^* = [\eta_\alpha^*, \eta_\beta^*]$, $A_i\xi$, and $\hat{y}$ for each individual are given in the Table 6.

TABLE 2. Estimates of Parameters ξ^*

ξ_{11}^*	ξ_{12}^*	ξ_{13}^*	ξ_{21}^*	ξ_{22}^*
35.7049	-7.3229	0.6041	6.9415	-0.1484
ξ_{23}^*	ξ_{31}	ξ_{32}^*	ξ_{33}^*	ξ_{41}
0.7091	-4.3631	0.9588	0.2023	41.7282
ξ_{42}^*	ξ_{43}			
0.6501	0.5373			

TABLE 3 Estimates of Parameter π

π_1^*	π_2^*	π_3^*	π_4^*
31.6368	3.2210	19.8768	0.1413

TABLE 4 Estimates of Parameters T_η

	η_α^*	η_β^*
η_α^*	0.9417	-0.3213
η_β^*	-0.3213	0.9650

TABLE 5 Estimates of Parameter Σ

	y_1	y_2	y_3	y_4
y_1	1.6620			
y_2	0.6576	1.2502		
y_3	0.9118	0.3334	2.4886	
y_4	-1.0552	-0.5784	0.5677	3.1839

TABLE 6. Estimates of Parameters $\eta^*, A_i\xi, \hat{y}$

Subject	η_α intercept	η_β slope	$A_i\xi$ time1	time2	time3	time4
1	-28.11	21.19	61.25	152.64	107.16	177.97
2	-28.14	18.54	66.62	221.85	149.47	239.92
3	-28.49	19.40	62.67	168.61	117.17	192.85
4	-28.42	19.40	61.71	157.60	110.29	182.64
5	-31.80	21.18	70.76	281.71	184.24	289.76
6	-31.34	19.57	54.28	99.54	72.30	124.24
7	-30.97	19.75	68.38	247.54	164.56	261.62
8	-31.02	20.24	56.12	110.18	79.52	135.70
9	-31.48	19.85	69.07	257.57	170.38	269.95
10	-32.04	20.35	57.06	116.38	83.66	142.19
11	-28.66	19.45	59.24	133.3	94.77	159.29
12	-28.48	21.07	66.88	225.53	151.65	243.06
13	-28.16	20.46	48.07	74.17	54.49	94.64
14	-26.44	20.10	69.07	257.57	170.38	269.95
15	-33.10	20.53	65.01	198.96	135.77	220.08
16	-31.57	20.37	56.64	113.53	81.76	139.23
17	-32.26	19.39	77.74	371.53	234.21	360.92
18	-32.05	19.22	64.64	193.92	132.72	215.65
19	-31.25	20.09	61.90	159.76	111.64	184.66
20	-31.48	19.37	61.84	159.03	111.19	183.98
21	-28.97	18.90	72.63	307.42	198.77	310.48
22	-28.55	19.90	45.60	67.03	49.28	85.51
23	-29.66	21.21	64.09	186.54	128.23	209.09
24	-28.84	20.54	69.71	266.83	175.72	277.60
25	-31.94	18.90	56.72	114.10	82.14	139.82
26	-31.51	18.71	64.89	197.27	134.75	218.60
27	-31.16	19.22	68.86	254.53	168.62	267.44
28	-31.75	20.15	74.72	334.73	214.00	332.16
29	-31.04	19.48	74.62	333.51	213.32	331.20
30	-31.61	21.31	71.55	292.70	190.47	298.66

Subject	$\hat{y}$ time1	time2	time3	time4
1	33.13	145.72	121.43	213.44
2	38.47	212.24	158.41	267.40
3	34.17	159.51	127.48	222.55
4	33.29	148.59	120.68	212.44
5	38.95	271.09	194.80	321.50
6	22.93	87.77	80.11	151.61
7	37.41	236.32	173.11	289.92
8	25.09	99.40	88.98	165.41
9	37.59	245.94	178.61	298.04
10	25.01	104.69	92.33	171.22
11	30.57	124.10	105.02	189.00
12	38.39	218.11	165.31	277.79
13	19.90	66.47	67.26	127.88
14	42.63	251.22	184.14	303.81
15	31.90	186.38	143.73	248.57
16	25.06	102.34	90.94	168.79
17	45.47	358.66	240.73	386.84
18	32.59	181.10	139.13	241.28
19	30.65	148.60	120.58	213.69
20	30.35	146.92	118.45	210.62
21	43.66	297.35	207.62	338.24
22	17.04	58.38	60.54	116.67
23	34.43	178.09	140.99	243.07
24	40.87	258.52	187.96	310.37
25	24.78	101.06	88.01	164.60
26	33.37	184.47	140.67	243.23
27	37.70	242.58	175.90	293.94
28	42.96	323.13	222.56	360.88
29	43.57	321.94	221.24	358.60
30	39.93	282.39	201.49	330.99

In sum the proposed multilevel structural equation model and estimation method has a potential for a variety of problems in analyzing longitudinal data. The multilevel structural equation model is useful whenever multiple indicators are necessary to incorporate measurement error in "growth" study using longitudinal data.

References

Jo, S. H.(1994). Empirical Bayes estimation for unbalanced multilevel structural equation models via the EM algorithm. *Unpublished doctoral dissertation*, Michigan State University.

Muthen, B.O., Khoo, S. T.(1997). Longitudinal studies of achievement growth using latent variable modeling. *Graduate school of Education & Information Studies,*