CRESST REPORT 801

LATENT VARIABLE REGRESSION 4-LEVEL HIERARCHICAL MODEL USING MULTISITE MULTIPLE-COHORTS LONGITUDINAL DATA

JULY, 2011

Kilchan Choi

National Center for Research
on Evaluation, Standards, & Student Testing

UCLA | Graduate School of Education & Information Studies

Latent Variable Regression 4-Level Hierarchical Model Using Multisite Multiple-Cohorts Longitudinal Data

CRESST Report 801

Kilchan Choi
CRESST/University of California, Los Angeles

July, 2011

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
300 Charles E. Young Drive North
GSE&IS Building, Box 951522
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported by grant number R305U070004:01 from the U.S. Department of Education with funding to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

To cite from this report, please use the following as your APA reference: Choi, K. (2011). *Latent variable regression 4-level hierarchical model using multisite multiple-cohorts longitudinal data*. (CRESST Report 801). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

TABLE OF CONTENTS

LATENT VARIABLE REGRESSION 4-LEVEL HIERARCHICAL MODEL USING MULTISITE MULTIPLE-COHORTS LONGITUDINAL DATA

Kilchan Choi
CRESST/University of California, Los Angeles

Abstract

This report explores a new latent variable regression 4-level hierarchical model for monitoring school performance over time using multisite multiple-cohorts longitudinal data. This kind of data set has a 4-level hierarchical structure: time-series observation nested within students who are nested within different cohorts of students. These students are in turn nested within a school. Under the circumstance, this model attempts to estimate three performance indicators: initial status, growth rate, and educational gap parameters across different cohorts. Specifically, one can see a longitudinal pattern where each cohort of students within a school starts or ends, how much it gains/grows within a specific period of time, and how much the initial gap between initially-low performing students and initially-high performing students is magnified or diminished. Furthermore, these three performance indicators of different cohorts are also modeled as a function of cohort and school background characteristics, in order to examine the extent to which differences or fluctuations across different cohorts within schools. are related to differences in cohorts' and schools' characteristics. This longitudinal and multiple-cohorts perspective is important because particular school reform efforts or changes in school characteristics that may impact school-wide or particular grade level achievement might take place in some years and not others (e.g., increment of school budget, numbers of qualified teachers, student demographic composition, etc.). As such, this model is distinguished from the current value-added models in a way that provides us with a more comprehensive picture of school's performance based on student growth over time and the distribution of student growth across cohorts within a school.

Introduction

The No Child Left Behind Act (NCLB, 2002) establishes ambitious goals for increasing student learning and attaining equity in the distribution of student performance. Schools must assure that all students, including all significant subgroups, show adequate yearly progress (AYP) toward the goal of 100% proficiency by the year 2014. Though AYP tracks changes over time by setting annual growth targets toward an absolute benchmark of 100% proficiency, it is not a model that can accurately measure growth because changes are not based on the same individuals over time.

As an alternative approach, value-added models have been a prevailing measure of teacher/school effectiveness based on students' progress over time (Sanders, Saxton, & Horn, 1997; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004; Ponisciak & Bryk, 2005). These models analyze individual students' time-series repeated measures and can be defined as regression-based models typically using multilevel or mixed models. Though each of the value-added models have a slightly different focus and take a slightly different approach, the key similarity of these models is that teacher or school mean estimate of growth is of central interest, and the resulting school mean estimate of growth is viewed as a reflection of an overall school effect in the sampled population.

Moving beyond the single summary of school mean estimate of growth by attending to the relationship between students' initial status and their growth rate can help bring to light the distribution of achievement within schools (Choi & Seltzer, 2010; Choi, Seltzer, Herman, & Yamashiro, 2007; Seltzer, Choi, & Thum, 2003). For example, even in three schools with a similar school mean estimate of growth, we can observe three different patterns of distribution of student growth: the initial gap between high initial status students and low initial status students is either diminishing, widening, or remaining unchanged over time. Choi et al. (2007) identified some schools meeting the state criteria for AYP—widening the initial gap where they had above average students making substantial progress—but for below average students little to no progress was made. In contrast, other schools achieving AYP had below average students making adequate progress, but above average students making little gains; as a result, the initial gap diminished. These results raise questions about the meaning of ―adequate" progress—and to whom the notion of progress is referring. Thus, closely examining the distribution of student progress may provide an important supplementary or alternative measure of AYP, and single school mean estimate of growth obtained in value-added models.

This report attempts to extend Choi et al.'s (2007) approach by comparing the performance of different cohorts of students. Proposed here is a new value-added model and illustration of its use by analyzing a longitudinal multiple-cohorts data in an urban school district in a northwestern state. This data set includes student achievement scores from grades 3 to 5 in the years 1998 to 2004, in five different cohorts. For each of the five cohorts of students, the proposed model allows estimation of a school's value-added estimate of growth, and the educational gap parameter that captures how equitably student achievement is distributed within a school. Furthermore, the value-added estimates of growth and the educational gap parameters of five different cohorts are also modeled as a function of cohort and school background characteristics in order to examine differences or fluctuations across different cohorts within a school. This longitudinal and multiple-cohorts perspective is important because particular school reform efforts or changes in school characteristics that may impact school-wide or particular grade level achievement may take place in some years and not others (e.g., increment of school budget, numbers of qualified teachers, student demographic composition, etc.).

The analytic framework that will be applied for a multiple-cohort longitudinal study is to combine three different statistical modeling techniques—growth modeling, multilevel/hierarchical modeling, and latent variable modeling—into one integrated modeling framework (Choi & Seltzer, 2010; Choi et al., 2007). This model is based on a newly invented and cutting-edge statistical technique and is termed as a four-level latent variable regression hierarchical model. The major distinctive features of this model are as follows: (a) This model deals with multivariate outcomes and predictors which allows us to study structural relations among multiple latent and/or observed variables taking into account covariance among outcome measures; (b) This model incorporates the IRT-based measurement models, one of the benefits of which is that we will be able to take standard errors of measurement into account in our models; (c) The model incorporates structural equation modeling (SEM) features, in particular regression among latent variables; (d) The model takes into account the dependency of nested observations (e.g., time-series measures nested within students who are nested within classrooms, or experimental conditions) within a hierarchical modeling framework.

Four-Level Latent Variable Regression Hierarchical Model

For a heuristic purpose, a simple four-level latent variable regression hierarchical model (LVR-HM4) is presented, where latent variable regression is employed at levels two, three, and four, but no observed predictors are included at any level. Level-1 (within-student) model specifies a simple linear growth model. Y_{tijk} is the outcome score at measurement occasion t ($t = 1, 2, \ldots t_{ijk}$), for student i ($i = 1, 2, \ldots N_{jk}$), in cohort j ($j = 1, 2, .. J_k$), in school k ($k = 1, 2, \ldots K$). The time-metric variable, $Time_{tijk}$, is centered around the value of initial measurement occasion.

$$Y_{tijk} = \pi_{0ijk} + \pi_{1ijk}\, Time_{tijk} + \varepsilon_{tijk} \qquad (1)$$

Thus, the two growth parameters, π_{0ijk} and π_{1ijk} represent initial status and growth rate, respectively, for student i in cohort j, in school k.

In level-2 (between-student; within-cohort) model, initial status for each individual student i in cohort j, school k is modeled as a function of mean initial status for cohort j in school k (i.e., β_{00jk}) and its random effect. Furthermore, the growth rate is modeled as a function of student's initial status. Note that student's initial status is centered around his or her cohort's mean initial status in school k. Due to this centering, β_{10jk} represents mean growth rate for cohort j in school k.

$$\pi_{0ijk} = \beta_{00jk} + r_{\pi 0jk} \qquad r_{\pi 0jk} \sim N(0, \tau_{\pi 0jk}) \tag{2a}$$

$$\pi_{1ijk} = \beta_{10jk} + Bw_{jk}(\pi_{0ijk} - \beta_{00jk}) + r_{\pi 1jk} \qquad r_{\pi 1jk} \sim N(0, \tau_{\pi 1jk}) \tag{2b}$$

$$\mathrm{Cov}(r_{\pi 0jk}, r_{\pi 1jk}) = 0$$

The key parameter in the level-2 model is the latent variable regression coefficient, Bw_{jk} which captures the expected increase or decrease when student's initial status increases one unit. In this report, this random latent variable regression coefficient is termed as within-cohort and within-school initial status/rate of change slope (see Choi & Seltzer, 2010). As subscripts for cohort j and school k indicate, each cohort in each school has different within-cohort and within-school initial status/rate of change slope. Two random effects, $r_{\pi 0jk}$ and $r_{\pi 1jk}$ are assumed to be normally distributed with mean 0 and variances, $\tau_{\pi 0jk}$ and $\tau_{\pi 1jk}$. Note that these variances differ across cohorts and schools and the covariance between the random effects is equal to 0 (since growth parameter is conditional on the initial status parameter).

At level 3, differences across cohorts in terms of mean initial status, mean growth rate, and within-cohort-and-within-school initial status/rate of change slope are examined. This model can be viewed as a between-cohort within-school model.

$$\beta_{00jk} = \gamma_{000k} + U_{\beta 00jk} \qquad U_{\beta 00jk} \sim N(0, \tau_{\beta 00}) \tag{3a}$$

$$\beta_{10jk} = \gamma_{100k} + Bc_1(\beta_{00jk} - \gamma_{000k}) + U_{\beta 10jk} \qquad U_{\beta 10jk} \sim N(0, \tau_{\beta 10}) \tag{3b}$$

$$Bw_{jk} = Bw_{\_0k} + Bc_2(\beta_{00jk} - \gamma_{000k}) + U_{Bwjk} \qquad U_{Bwjk} \sim N(0, \tau_{Bw}) \tag{3c}$$

$$Cov(U_{\beta 00jk}, U_{\beta 10jk}) = 0,\ Cov(U_{\beta 00jk}, U_{Bwjk}) = 0$$

$$T_U = \begin{pmatrix} \tau_{\beta 00} & 0 & 0 \\ 0 & \tau_{\beta 10} & \tau_{Bw,\,\beta 10} \\ 0 & \tau_{\beta 10,\,Bw} & \tau_{Bw} \end{pmatrix} \tag{3d}$$

There are two latent variable regression coefficients in the model. First, Bc_1 presents the overall relationship between cohort mean initial status and cohort mean growth rate. In contrast to Bw_{jk}, i.e., within-cohort and within-school initial status/rate of change slope, this coefficient is termed as between-cohort initial status/rate of change slope. Thus, this coefficient captures the extent to which one unit increase of cohort's mean initial status results in increase or decrease in cohort mean growth. If the coefficient has a statistically significant positive value, it indicates that cohorts with higher mean initial status have higher growth. Second, Bc_2 captures the relationship between cohort mean initial status and within-cohort and within-school initial status/rate of change slope. Thus, this coefficient indicates the expected increase or decrease in Bw_{jk} when cohort mean initial status increases one unit. It is important to note that these two latent variable regression coefficients are not assumed to be random variables, while the mean within-cohort and within-school initial status/rate of change slope for school *k* vary across schools.

The fixed effect coefficients, γ_{000k} and γ_{000k} represent mean initial status and mean growth rate, respectively, across cohorts for school *k*. Likewise, $Bw_{\_0k}$ is mean within-cohort and within-school initial status/rate of change slope across cohorts for school *k*. As to random effects, $U_{\beta 00jk}$ captures each cohort's deviation from the school mean initial status in school *k*. $U_{\beta 10jk}$ and U_{Bwjk} are each cohort's deviations from the school mean growth rate and the school mean within-cohort initial status/rate of change slope, respectively, after controlling for cohort mean initial status in the corresponding equation. These random effects are assumed to be multivariate normally distributed with mean 0 and variance-covariance matrix, $\mathbf{T_U}$.

This level-3 model can be readily extended by including time-varying and time invariant cohort characteristics in the model. Suppose that we have five cohorts of students (J= 5) in each of the sample school *k*. Since these five successive cohort's mean, growth rate, and the within-cohort initial status/rate of change slope (i.e., cohort's growth parameters) can be considered as time-series quantities, for example, we can pose a various form of growth patterns, e.g., a simple growth, quadratic growth, a piecewise growth, or/and a general saturated form in order to model those parameters. By doing so, it is possible to examine the extent to which the cohorts' mean initial status, mean growth rate, and the within-cohort and within-school initial status/rate of change slope are different across cohorts, and to address what growth trajectories of cohorts' growth parameters look like. Furthermore, one can address a question such as the extent to which differences in the percentage of free/reduced lunch eligible students across cohorts are associated with the differences in the cohort's growth parameters. This kind of question relating cohort characteristics to cohort's growth parameters can be readily examined by including those characteristics in the left-hand side of Equations 3a, 3b, and 3c (see, Goldschmidt, Choi, Martinez, & Novak, 2010).

Lastly, the following equations present a level-4 (between-school) model where we specify relationships in a between-school level.

$$\gamma_{000k} = \theta_{0000} + V_{\gamma 000k} \qquad V_{\gamma 000k} \sim N(0, \tau_{\gamma 000}) \tag{4a}$$

$$\gamma_{100k} = \theta_{1000} + Bb(\gamma_{000k} - \theta_{0000}) + V_{\gamma 100k} \qquad V_{\gamma 100k} \sim N(0, \tau_{\gamma 100}) \tag{4b}$$

$$Bw_{\_0k} = Bw_{\_00} + Bw_{\_01}(\gamma_{000k} - \theta_{0000}) + V_{Bw\_0k} \qquad V_{Bw\_0k} \sim N(0, \tau_{Bw\_0}) \tag{4c}$$

Cov($V_{\gamma 000k}$, $V_{\gamma 100k}$) = 0, Cov($V_{\gamma 000k}$, V_{Bw_0k}) = 0

θ_{0000} is the grand mean initial status across cohorts and schools. There is another latent variable regression coefficient, Bb, capturing the relationship between school mean initial status and school mean growth rate. This coefficient is called as the beween-school initial status / rate of change slope in contrast to Bw_{jk} (Choi & Seltzer, 2010). In equation 4c, $Bw_{\_01}$ is a latent variable regression coefficient capturing the extent to which the difference in school mean initial status relates to difference in the within-cohort relationship between student's initial status and his or her growth. In other words, this coefficient indicates how much of an increase or decrease is expected in $Bw_{\_k}$ when school mean initial status increases by one unit.

The random effects in the above model (i.e., $V_{\gamma 000k}$, $V_{\gamma 100k}$, V_{Bw_0k}) are assumed multivariate normally distributed with mean **0** and variance-covariance matrix.

$$\mathbf{T}\mathbf{v} = \begin{pmatrix} \text{TM}_{\gamma 000} & 0 & 0 \\ 0 & \text{TM}_{\gamma 100} & \text{TM}_{\gamma 100,Bw\_0} \\ 0 & \text{TM}_{Bw\_0,\gamma 100} & \text{TM}_{Bw\_0} \end{pmatrix} \qquad (4d)$$

The variance term $\tau_{\gamma 000}$ captures the extent to which schools vary in their school mean initial status, and $\tau_{\gamma 100}$ and τ_{Bwk} are, respectively, residual variances for school mean gain and within-cohort initial status / gain slopes after taking into account school mean initial status. With respect to the off-diagonal elements of $\mathbf{T_v}$, we assume that *Cov*($V_{\gamma 000k}$, $V_{\gamma 100k}$) = 0 and *Cov*($V_{\gamma 000k}$, V_{Bw_0k}) = 0, since γ_{000k} is employed as a predictor in Equations 4b and c. $\tau_{\beta 10,Bw}$ captures the covariance between u_{10j} and u_{Bwj}.

Illustrative Example

Data

The four-level latent variable regression hierarchical model is illustrated using multiple-cohort longitudinal data from an urban district located in a northwestern state. As can be seen in Table 1 (see following page), the data includes 5 different cohorts and each cohort consists of longitudinal two-time points measures between the grade 3 and the grade 5 in 74 elementary schools. Among 74 schools, five schools have only four cohorts of data and one school has only three cohorts of data. The outcomes of interest are Iowa Test of Basic Skills (ITBS) in reading scale scores. These scale scores are vertically equated developmental scores. For each student in each cohort, ITBS reading scores at grade 3 and grade 5 provide the basis of estimating gains in reading achievement between two grades. Note that the standard error of measurement (SEM) connected with students' ITBS scores were included in the analyses. The conditional SEMs were provided for various ranges of test scores.

Table 1

Data Structure: Cohort, Grade and Year

	Year						
Cohort	1998	1999	2000	2001	2002	2003	2004
Cohort 1	Grade 3		Grade 5				
Cohort 2		Grade 3		Grade 5			
Cohort 3			Grade 3		Grade 5		
Cohort 4				Grade 3		Grade 5	
Cohort 5					Grade 3		Grade 5

Table 2 presents the descriptive statistics for the sample. The total number of students in the sample is 11,530, and the average number of students per cohort is 2,306. The mean score of ITBS reading at grade 3 for each cohort of students locates around 191 and its standard deviation is approximately 22. The observed gain between grades 3 and 5 in ITBS reading scores per cohort is around 29.5 points. As in mean scores at grade 3, average observed gain scores are very similar across five cohorts. In addition, the percentage of free/reduced price eligible students is approximately 39.5%, and each cohort's average percentage does not fluctuate much around the overall average percentage.

Table 2

Cohort-by-Cohort Descriptive Statistics of Observed Initial Status, Gain, and % of Free/Reduced Priced Lunch Students

		Status at grade 3			Observed gain (grades 3-5)		% free/reduced lunch	
Cohort	Years	*N*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	1998-2000	2501	188.4	23.5	30.5	16.1	41.6	4.9
2	1999-2001	2324	191.9	21.6	29.7	14.1	37.9	4.9
3	2000-2002	2054	191.7	23.6	28.7	14.4	40.2	4.9
4	2001-2003	2548	191.2	21.8	29.8	13.8	36.1	4.8
5	2002-2004	2103	191.8	23.3	29.0	14.0	40.1	4.9

Variability of initial status, observed gain, and percent of free/reduced lunch

One of the key questions of this cohort-to-cohort analysis is to what extent the variability of growth parameters (e.g., initial status and growth rate) across cohorts within a school is observed. In other words, is the between-cohort within school variability negligibly small compared to between-school variability?

Figure 1 shows variability of ITBS mean scores for students in grade 3 for each cohort (i.e., initial status) and each school. In the figure the circle represents the cohort's initial mean, while the triangle indicates school's overall mean across cohorts. School's initial mean ranges from approximately 170 to 225; however, the cohort's initial mean within a school varies from school to school. For example, cohort 1 in school # 40 has a very high initial mean compared to the rest of cohorts' means, whereas cohorts' initial mean in school # 34 are closely clustered around the overall school mean. When fitting this data into a simple two-level hierarchical model where outcomes are cohort's initial mean that are nested within school, the intra-class correlation is .13, which means that between-cohort variability is 13% and between-school variability is 87%.

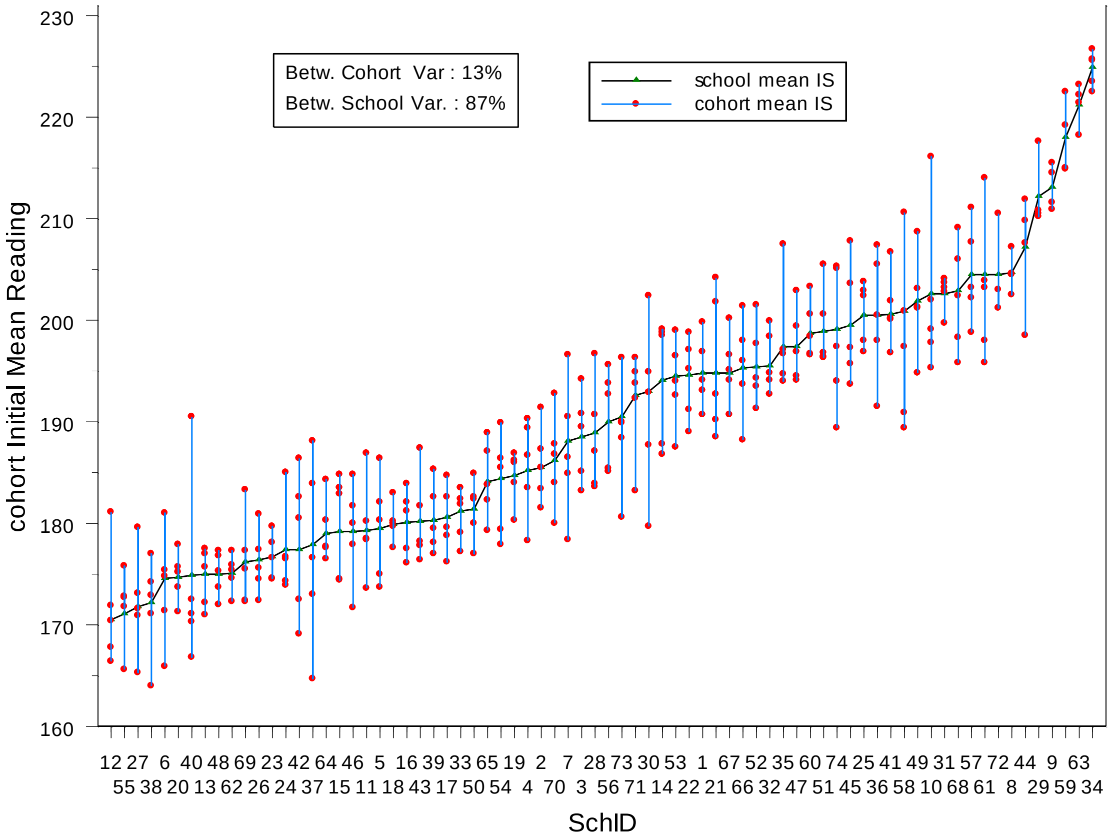

Figure 1. Observed initial ITBS reading mean scores by cohort and school.

In contrast to the cohort initial mean, the observed mean gains between grades 3 to 5 show far greater variability in between-cohort than in between-school. School mean gains of the sampled 74 schools range from approximately 17 to 35. But the variability between cohorts in some schools is very large, as much as 40 points in school #40. The intra-class correlation based on a two-level HM using cohort's mean gain data is .69, which indicates that the between-cohort variability is more than twice the between-school variability.

As for variability in the percentage of free/reduced price lunch eligible students, the between-cohort variability is approximately 17%, and between-school variability is 83%. As such, we can see that there is a great deal of between-cohort variability in terms of mean gain, whereas cohort's initial mean and mean percentage of free/reduced lunch students are relatively homogeneous between cohorts within a school.

Results

Model 1: Unconditional Four-Level Hierarchical Model

In order to examine the magnitude of variability for each level—student, cohort, and school, an unconditional four-level hierarchical model is posed as follows. In order to incorporate the SEMs, left hand and right hand sides of Equation 1 are re-scaled by inverse of SEMs at time *t*, for student *i*, cohort *j* in school *k*.

$$Y^{*}_{tijk} = \pi_{0ijk} + \pi_{1ijk}\, Time^{*}_{tijk} + \varepsilon^{*}_{tijk} \qquad \varepsilon^{*}_{tijk} \sim N(0, 1) \qquad (5)$$

By re-scaling the outcome and time metric based on an estimate of the precision associated with each ITBS reading score, (1/ SEM(Y_{tijk})), ε^{*}_{tijk} is now assumed to be normally distributed with mean 0, but its variance is now 1. $Time^{*}_{tijk}$ takes value of 0 for test score at the 3rd grade and 1 for test score at the 5th grade so that π_{0ijk} and π_{1ijk} represent status at the 3rd grade (i.e., initial status) and gain between the 3rd and the 5th grade, respectively, for student *i* in cohort *j*, school *k*.

$$\pi_{0ijk} = \beta_{00jk} + r_{\pi 0jk} \qquad r_{\pi 0jk} \sim N(0, \tau_{\pi 0}) \qquad (6a)$$

$$\pi_{1ijk} = \beta_{10jk} + r_{\pi 1jk} \qquad r_{\pi 1jk} \sim N(0, \tau_{\pi 1}) \qquad (6b)$$

$$\beta_{00jk} = \gamma_{000k} + U_{\beta 00jk} \qquad U_{\beta 00jk} \sim N(0, \tau_{\beta 00}) \qquad (7a)$$

$$\beta_{10jk} = \gamma_{100k} + U_{\beta 10jk} \qquad U_{\beta 10jk} \sim N(0, \tau_{\beta 10}) \qquad (7b)$$

$$\gamma_{000k} = \theta_{0000} + V_{\gamma 000k} \qquad V_{\gamma 000k} \sim N(0, \tau_{\gamma 000}) \qquad (8a)$$

$$\gamma_{100k} = \theta_{1000} + V_{\gamma 100k} \qquad V_{\gamma 100k} \sim N(0, \tau_{\gamma 100}) \qquad (8b)$$

Equations 6a and 6b specify level-2 (between-student) model. β_{00jk} and β_{10jk} represent cohort *j* in school *k*'s initial status and gain, respectively. In Equation 7a and 7b γ_{000k} and γ_{100k} are mean initial status and gain for school *k*. Lastly, θ_{0000} and θ_{1000} are grand mean initial status and grand mean gain, respectively.

Table 3

Model 1: Unconditional Four-Level Model

Fixed effects		Estimate	*SE*	95% Interval	Median
Grand mean initial status (θ_{0000})		189.10	1.46	(186.20, 192.00)	189.10
Grand mean gain (θ_{1000})		30.32	.50	(29.34, 31.29)	30.32
Variance components:					
1. Level-2 variance: Between-student					
Initial status ($\tau_{\pi 00}$)	*61.5%	264.10	4.40	(255.60, 272.90)	264.10
Gain ($\tau_{\pi 10}$)	**75.4%	81.28	2.46	(76.51, 86.16)	81.26
Cov ($r_{\pi 0}$, $r_{\pi 1}$)		25.23	2.40	(20.46, 29.86)	25.26
2. Level-3 variance: Between-cohort					
Initial status ($\tau_{\beta 00}$)	*2.8%	11.75	1.77	(8.55, 15.49)	11.65
Gain ($\tau_{\beta 10}$)	**11.6%	12.56	1.60	(9.68, 15.95)	12.47
Cov ($u_{\beta 00}$, $u_{\beta 10}$)		-3.82	1.24	(-6.37, -1.48)	-3.79
3. Level-4 variance: Between-school					
Initial status ($\tau_{\gamma 000}$)	*35.7%	153.20	26.40	(109.60, 212.80)	150.40
Gain ($\tau_{\gamma 100}$)	**13.0%	14.01	3.00	(9.10, 20.80)	13.68
Cov ($V_{\gamma 000}$, $V_{\gamma 100}$)		30.57	7.21	(18.43, 46.78)	29.89

Note. * Initial status variance % for each of the three levels (between-student, between-cohort, between-school)
** Gain variance % for each of the three levels (between-student, between-cohort, between-school).

As can be seen in Table 3, the grand mean initial status in this sample is approximately 189.1 and the grand mean gain is 30.3. The key interest of this model is the extent to which the variability of each level differs across levels. As for variability of initial status, the percentage of between-student, between-cohort within school, and between-school variances over the total variance are, respectively, 61.5%, 2.8%, and 35.7%. This indicates that there is a very sizable variability between schools, but an ignorable variability in between cohorts. In contrast, between-cohort and between-school variances in gain take 11.6% and 13.0% of the total variance. These results suggest that initial status across cohorts is very homogeneous, but most of variability comes from between schools, whereas between-cohort variability in gain is as large as between-school variability.

Model 2: Four-Level Latent Variable Regression Hierarchical Model (LVR- HM4)

Model 2 examines the extent to which initial status is consequential to amount of gain in three different levels: student, cohort and school. Specifically, the research questions addressed are the following.

- Do students with high initial status gain more than those with low initial status?
- Do cohorts with high initial status have a higher gain than those with low initial status?
- Do schools with high initial status gain more compared to those with low initial status?
- Are the relationships between students' initial status and his or her gain different depending upon cohort's initial status?
- Are the relationships between students' initial status and his or her gain different depending upon school's initial status?

Model 2 consists of four-level models. Level-1 model is specified as in Equation 5, and levels 2, 3, and 4 are specified as in the previous section, Equations 2a, 2b, 3a, 3b, 3c, 4a, 4b, and 4c. This model employs latent variable regression at levels 2, 3, and 4. Furthermore, the within-cohort-and-within-school initial status / gain parameter is treating as a random variable at levels 3 and 4.

Along with the research questions addressed above, of central interest is estimating growth parameters for each cohort in each school. First, initial status or final status depending upon coding scheme in time-metric variable for each cohort in each school provides a series of longitudinal information how the cohort-to-cohort performance changes over time.

Second, gain estimate for each cohort in each school also provides important pieces of longitudinal information about how much growth for each cohort of students is achieved. From a school's perspective, these cohort-to-cohort gain estimates can be considered as a productivity indicator.

Third, this model uniquely provides a gap indicator for each cohort in each school. Questions regarding whether a student who initially starts high gains more than one who initially starts low are related to the distribution of student growth within each cohort in each school. The gap indicator in this LVR-HM4 tells us the extent to which the initial gap between students, say 30 points initially, is magnified or diminished over time for each cohort of students in each school. The gap indicator is referred to the key latent variable regression coefficient, within-cohort and within-school initial status/gain slope.

Table 4

Model 2: Four-Level Latent Variable Regression Model –Estimating Within-School Initial Status / Gain Slopes (Bw_{jk}), Cohort Initial Status / Gain Slope (Bc_1), Cohort Initial Status / Bw_{jk} Slope, and a Between-School Initial Status / Gain Slope (Bb)

Item	Estimate	*SE*	95% Interval		Median
Fixed effects:					
1. Model for between-school:					
Model for school mean initial status (γ_{000k}):					
Grand mean (θ_{0000})	189.300	1.480	(186.40,	192.30)	189.300
Model for school mean gain (γ_{100k}):					
Grand mean (θ_{1000})	30.200	.470	(29.30,	31.20)	30.200
School mean initial status (Bb)	.194	.033	(.13,	.26)	.194
Model for within-cohort-and-within-School initial status/gain slope (Bw_k):					
Mean initial status/gain slope (Bw_{00})	.022	.017	(-.011,	.055)	.022
School mean initial status (Bw_{01})	-.005	.001	(-.008,	-.003)	-.005
2. Model for between-cohort within-school:					
Cohort initial status/ gain slope (Bc_1)	-.321	.131	(-.59,	-.07)	- .319
Cohort initial status / Bw_{jk} slope (Bc_2)	-.001	.006	(-.012,	.010)	- .001
Variance components:					
1. Level-3 variance: Between-cohort					
Initial status ($\tau_{\beta 00}$)	8.190	1.700	(5.17,	11.82)	8.090
Gain ($\tau_{\beta 10}$)	6.350	1.240	(4.21,	9.01)	6.260
Initial/gain Slope (τ_{BW})	.003	.001	(.002,	.006)	.003
Cov ($u_{\beta 00}$, $u_{\beta 10}$)	-.008	.031	(-.070,	.053)	-.008
2. Level-4 variance: Between-school					
Initial status($\tau_{\gamma 000}$)	158.400	28.190	(112.30,	222.10)	155.400
Gain ($\tau_{\gamma 100}$)	7.010	1.789	(4.076,	11.050)	6.820
Initial/gain slope (τ_{BWk})	.007	.002	(.000,	.012)	.007
Cov ($V_{\gamma 000}$, $V_{\gamma 100}$)	.058	.046	(-.026,	.155)	.056

Table 4 presents all the parameters' estimates, SD, 95% interval, and median values in Model 2. The grand mean initial status is 189.3 and the grand mean gain is 30.2. As to the latent variable regression coefficients, the between-school initial status/gain slope is .194 and its 95% interval is between .13 and .26. This indicates that on average schools with higher initial status gain more than schools with lower initial status. At the cohort level, the relationship is the opposite. The between-cohort initial status/gain slope is -.321, and its 95% interval does not capture 0. The overall student's initial status/gain slope ($Bw_{\_00}$) is close to 0 and it is not statistically different from 0. However, the statistically significant negative point estimate of $Bw_{\_01}$ capturing the relationship between school mean initial status and the within-school initial status/gain slope indicates that low initial status students tend to gain more than high initial status students when they are in a high mean initial status school.

The graphical summary of the results in Model 2 are presented in Figure 2. This figure presents the expected growth parameters for each cohort in three different schools—schools with mean initial status values that are, respectively, approximately 26 points below (i.e., approximately 2 SD), close to the mean, and 28 points above the mean. Note that all the growth parameters for these three schools are estimated within a Gibbs sampler so that the point estimate and its 95% interval for each growth parameter are plotted.

The first thing to note in Figure 2 is that more gain is obviously observed for a school with higher initial status than a school with lower initial status. However, when initial status change is compared to gain change (compared the first low to the second low plot), cohort initial status and cohort gain go against each other. In other words, high initial status cohort goes with low gain, whereas low initial status cohort goes with high gain. If we flip the cohort gain plot over, then it becomes extremely similar with cohort initial status plot. Aforementioned significant negative between-cohort initial status/gain slope (i.e., Bc_1) captures this strong relationship.

The third low of plots present the gap indicator, which raises the question: What is the final gap between two students initially 30 points apart (i.e., 15 point above and 15 point below the school mean initial status)? The gains were estimated for these two students in cohort *j* in school *k*, and compared the final status to each other. The results are plotted in the circle line. As can be seen, the 30 point initial gap becomes larger by as many as approximately 5 points in the first school (i.e., low mean initial status school). In the second school, the initial gap is maintained at the end. Finally, in the third school (i.e., high mean initial status school), the initial 30 point gap is reduced by up to 25 points. Another interesting point is that gap indicators fluctuate minimally across cohorts, since the latent variable regression coefficient capturing the relationship between cohort mean initial status and within-cohort and within-school initial status/gain slope, Bc_2, is very close to 0.

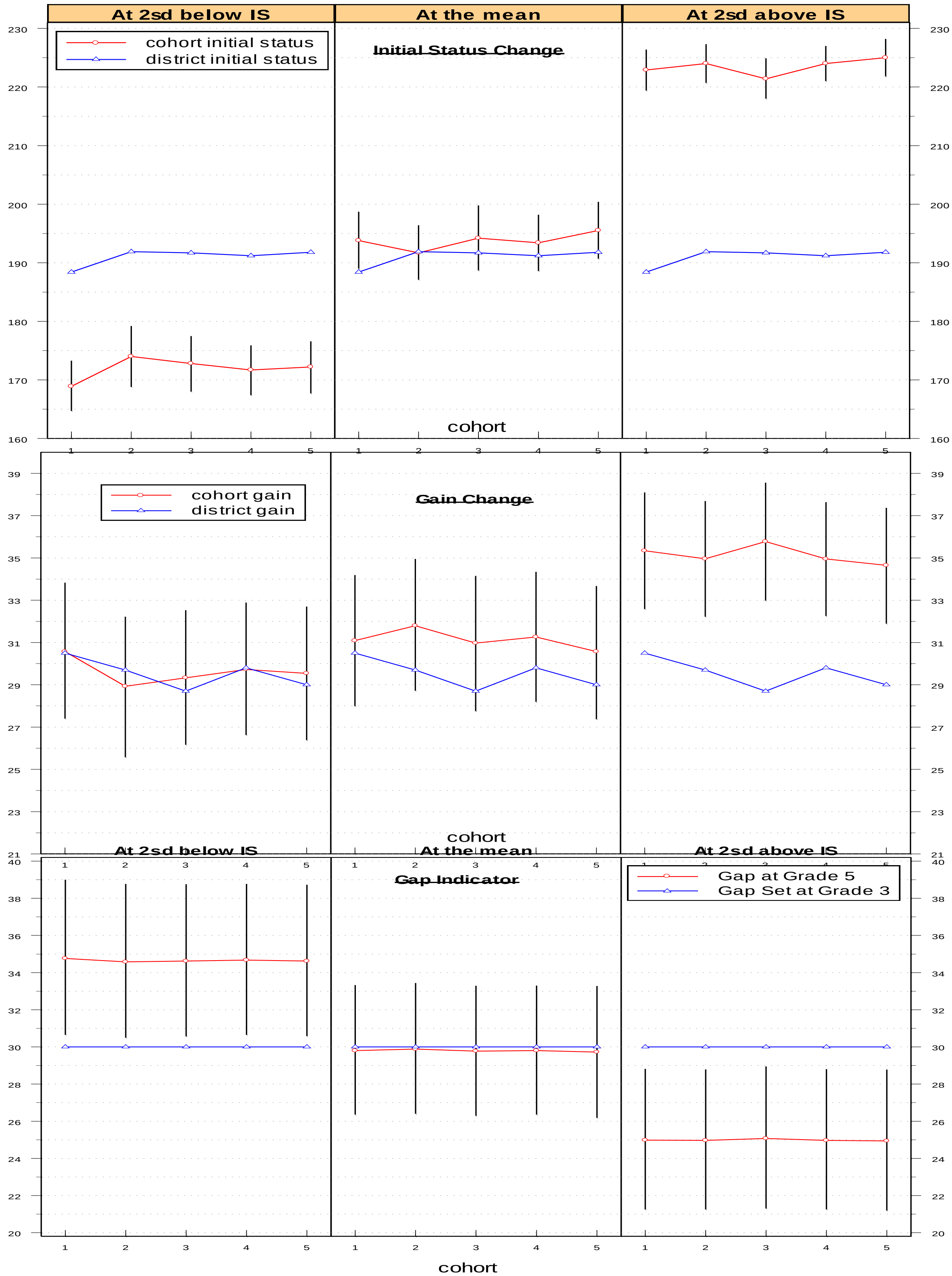

Figure 2. School's three growth performance indicators: Cohort initial status, cohort gain, and cohort gap indicator (based on 4-level latent variable regression hierarchical model).

Model 3: Conditional Four-Level Latent Variable Regression Hierarchical Model

Model 3 is expanded by including student-level, cohort-level, and school-level variables in the model. Level-1 model is the same as in Model 2. A student characteristic variable flagging whether a student is eligible for free/reduced price lunch (*FRL*) is included in the following level-2 model. This variable is centered around the school average of this variable.

$$\pi_{0ijk} = \beta_{00jk} + \beta_{01}(FRL_{ijk} - \overline{FRL..}) + r_{\pi 0jk} \qquad r_{\pi 0jk} \sim N(0, \tau_{\pi 0jk}) \qquad (9a)$$

$$\pi_{1ijk} = \beta_{10jk} + \mathrm{Bw}_{jk}(\pi_{0ijk} - \beta_{00jk}) + \beta_{11}(FRL_{ijk} - \overline{FRL..}) + r_{\pi 1jk} \qquad r_{\pi 1jk} \sim N(0, \tau_{\pi 1jk}) \qquad (9b)$$

One time-varying variable, $NCLB_{jk}$,is included in the following level-3 model. This variable takes a value of 0 for the first four cohorts and 1 for the last cohort. Thus, we can contrast the growth parameters between the first four cohorts of students, which is before the NCLB era and the last four cohort of students which is after the NCLB era.

$$\beta_{00jk} = \gamma_{000k} + \gamma_{001}NCLB_{jk} + U_{\beta 00jk} \qquad U_{\beta 00jk} \sim N(0, \tau_{\beta 00jk}) \qquad (10a)$$

$$\beta_{10jk} = \gamma_{100k} + \gamma_{101}NCLB_{jk} + \mathrm{Bc}_1(\beta_{00jk} - \gamma_{000k}) + U_{\beta 10jk} \qquad U_{\beta 10jk} \sim N(0, \tau_{\beta 10jk}) \qquad (10b)$$

$$\mathrm{Bw}_{jk} = \mathrm{Bw}_{\_0k} + \mathrm{Bw}_{\_1}NCLB_{jk} + \mathrm{Bc}_2(\beta_{00jk} - \gamma_{000k}) + U_{\mathrm{Bw}jk} \qquad U_{\mathrm{Bw}jk} \sim N(0, \tau_{\mathrm{Bw}jk}) \qquad (10c)$$

Note, however, that $NCLB_{jk}$ variable is centered around the group mean, so that the meaning of intercept remains the same as in Model 2. For example, γ_{000} represents mean initial status for school k and γ_{001} captures the difference between before-NCLB and after-NCLB cohort in terms of mean initial status.

In a school-level model below (Equations 11a, 11b, and 11c), an indicator variable is included that flags whether or not a school met ―adequate yearly progress" (AYP) criteria. Only six schools within in the sample did not meet AYP. As such, AYP_k takes a value of 0 for AYP schools and 1 for non-AYP schools.

$$\gamma_{000k} = \theta_{0000} + AYP_k\theta_{0001} + V_{\gamma 000k} \qquad V_{\gamma 000k} \sim N(0, \tau_{\gamma 000}) \qquad (11a)$$

$$\gamma_{100k} = \theta_{1000} + \mathrm{Bb}(\gamma_{000k} - \theta_{0000}) + AYP_k\theta_{1001} + V_{\gamma 100k} \qquad V_{\gamma 100k} \sim N(0, \tau_{\gamma 100}) \qquad (11b)$$

$$\mathrm{Bw}_{\_k} = \mathrm{Bw}_{\_00} + \mathrm{Bw}_{\_01}(\gamma_{000k} - \theta_{0000}) + \mathrm{Bw}_{\_02}AYP_k + V_{\mathrm{Bw}\_0k} \qquad V_{\mathrm{Bw}\_0k} \sim N(0, \tau_{\mathrm{Bw}\_0}) \qquad (11c)$$

Table 5

Model 3: Conditional 4-Level Latent Variable Regression Model – Comparing Performances Between the Before-and-After NCLB Era Cohorts and AYP School Versus Non-AYP School.

Item	Estimate	*SE*	95% Interval	Median
Fixed effects:				
1. Model for between-school:				
Model for school mean initial status (γ_{000k}):				
AYP school mean (θ_{0000})	189.80	1.120	(187.60, 192.20)	189.800
AYP vs. Non-AYP school (θ_{0001})	-6.63	3.970	(-14.41, 1.20)	-6.630
Model for school mean gain (γ_{100k}):				
AYP school mean (θ_{1000})	30.600	0.450	(29.70, 31.50)	30.600
AYP vs. Non-AYP school (θ_{1001})	-1.490	1.510	(-4.40, 1.49)	-1.500
School mean initial status (Bb)	.192	.047	(.099, .285)	.191
Model for within-school-and-within-school Initial status/gain				
Slope (Bw_k):				
Mean initial status/gain slope (Bw_{00})	.000	.017	(-.035, .032)	.000
School mean initial status (Bw_{01})	-.007	.002	(-.010, -.004)	-.007
AYP vs. non-AYP school (Bw_{02})	-.046	.059	(-.162, .072)	-.046
2. Model for between-cohort within-school:				
Cohort initial status/ gain slope (Bc_1)	-.440	.165	(-.777, -.134)	-.434
Cohort initial status / Bw_{jk} slope (Bc_2)	-.002	.007	(-.015, .011)	-.002
Before vs. after NCLB era:				
Cohort initial status difference (θ_{001})	1.826	.574	(0.70, 2.94)	1.8300
Cohort gain difference (θ_{101})	.331	.642	(-.98, 1.62)	.3190
Cohort Bw_{jk} difference (Bw_{1})	.022	.030	(-.036, .081)	.0220
Variance components:				
1. Level-3 variance: Between-cohort				
Initial status ($\tau_{\beta 00}$)	6.060	1.400	(3.40, 9.09)	5.9800
Gain ($\tau_{\beta 10}$)	6.040	1.210	(3.91, 8.66)	6.0000
Initial/gain slope (τ_{BW})	.003	.001	(.002, .006)	.0030
Cov ($u_{\beta 00}$, $u_{\beta 10}$)	-.021	.030	(-.080, .036)	-.0200
2. Level-4 variance: Between-school				
Initial status($\tau_{\gamma 000}$)	82.300	15.160	(57.6, 116.40)	80.500
Gain ($\tau_{\gamma 100}$)	7.070	1.830	(4.07, 11.20)	6.860
Initial/gain slope (τ_{BWk})	.006	.002	(.004, .010)	.006
Cov ($V_{\gamma 000}$, $V_{\gamma 100}$)	.033	.043	(-.049, .123)	.031

The results for Model 3 are presented in Table 5. First, AYP schools have higher school mean initial status than non-AYP schools by 6.63, but there is no statistically significant difference, since the 95% interval for θ_{0001} includes value of 0. Likewise, the school mean gain between AYP and non-AYP schools is not statistically different. Second, as to comparison of the performance between before-NCLB cohorts and after-NCLB cohorts, before-NCLB cohorts have statistically higher mean initial status than after-NCLB cohort by approximately 1. However, there is no statistical difference in terms of gain between two groups. Third, the estimates and 95% intervals of all the latent variable regression coefficients are very similar to ones in Model 2.

In summary, a step-wise modeling process was adopted to analyze the multiple cohort ITBS data. At first, an unconditional four-level HLM (Model 1) was fitted to estimate the variance components from different levels. By decomposing the total variability, it was found that the major variability in gain came from between-school and between-cohort. Next, the initial status was added (Model 2) to each level of regression models. By doing these, the connection between initial status and gains on both student levels and school levels was explored. In this step, it was found that relationships between students' initial status and gains differed and depended on overall school performance. Finally, student and school background variables (Model 3) were added to test whether certain covariates (e.g., free lunch, indicator of adequate yearly progress) affected students' learning outcome variables. The results showed that there were differences in the initial status for different student/school groups. However, the student and school covariates did not lead to statistically significant differences on students' test gains.

Summary and Discussion

Through analyzing multisite multiple-cohort longitudinal data, two distractive features of MMCGM are illustrated: First, it envisions the change or stability of value-added estimates of different cohorts of students within a school. Unlike the other value-added models that essentially provide only one overall value-added estimate, cohort-to-cohort estimate of student growth allows us to have supplementary information of how a school has performed across years and cohorts. Second, moving beyond the single summary of school mean estimate of growth, the gap parameter that captures the relationship between students' initial status and their growth rate can help bring to light the distribution of achievement within schools (Choi & Seltzer, 2010; Choi, et al., 2007; Seltzer, Choi, & Thum, 2003;). For example, even in three schools with a similar school mean estimate of growth, we can observe three different patterns of distribution of student growth: the initial gap between high initial status students and low initial status students is either diminishing, widening, or remaining unchanged over time. Choi et al. (2007) identified some schools meeting the Washington state criteria for AYP—widening the initial gap where they had above average students making substantial progress—but for below average students little-to-no progress was made. In contrast, other schools achieving AYP had below average students making adequate progress, but above average students making little gains; as a result, the initial gap diminished. These results raise questions about the meaning of “adequate” progress—and to whom the notion of progress is referring. Thus, we believe that closely examining the distribution of student progress may provide an important supplementary or alternative measure of AYP, and single school mean estimate of growth obtained in value-added models.

In intervention-based experimental studies, it is not uncommon to encounter higher-level hierarchically nested data. Suppose that we are interested in comparing end-of-year performance of Title I students who receive an innovative remedial reading program to students who receive other more traditional remedial reading programs, using a pretest-posttest design. The students are sampled from many schools across districts to increase external validity of the study across schools and districts. The resulting data for this study has a 4-level nested structure: students are nested within classrooms and schools, which are nested within different school districts. The LVR-HM4 framework presented in this report enables researchers to broaden their research questions and help the actual implementation of a higher level of hierarchical models.

Furthermore, LVR-HM4 provides very important methodological strategies and substantive implications regarding the differential program effects in a longitudinal study of program effectiveness. Consider a multi-site intervention study regarding the effectiveness of two remedial reading programs over time and the resulting data has a 4-level hierarchical structure (i.e., time series observations nested within students, which in turn are nested within different schools and districts.). In these settings, interest often centers on the difference in growth rates between students in treatment sites and students in comparison sites. However, it is also important to consider whether those students who are most in need of help are those who are benefiting most. We might see that in some sites, program A might be more effective for students with extreme reading difficulties, while program B might be more successful in the case of students with milder initial reading difficulties. However, in other sites, among students with extreme reading difficulties, rates of progress may be more rapid for students in Program B, whereas among students with milder difficulties, rates of progress may be more rapid for students in Program A. Exploring this issue entails regressing rates of change on the initial status in each site (level 2). Then in a between-site (level 3) model, in addition to examining whether site mean growth rates tend to be more rapid in treatment sites, we can also explore whether initial status/rate of change coefficients tend to be negative, for example, in treatment sites versus positive in comparison sites. The LVR-HM4 can be readily applied to address those questions. In a level-2 (between-individual or within-site) model, we estimate a latent variable regression coefficient, and we regress those coefficients on site-level variable at level 3.

References

Choi, K., & Seltzer, M. (2010). Modeling heterogeneity in relationships between initial status and rates of change: Treating latent variable regression coefficients as random coefficients in a three-level hierarchical model. *Journal of Educational and Behavioral Statistics,35*(1), 54-91.

Choi, K., Seltzer, M., Herman, J., & Yamashiro, K. (2007). Children Left Behind in AYP and Non-AYP Schools: Using Student Progress and the Distribution of Student Gains to Validate AYP. *Educational Measurement: Issues and Practice*, *26*(3), 21-32.

Goldschmidt, P., Choi, K., Martinez, F., & Novak, J. (2010). Using growth models to monitor school performance: comparing the effect of the metric and the assessment. *School Effectiveness and School Improvement*, 1-21McCaffrey, D., Lockwood, J. R., Koretz, D., Louis, T., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29* (1), 67-101.

Ponisciak, S. M., & Bryk, A. (2005). Value-added analysis of the Chicago public schools: An application of hierarchical models. In R. Lissitz (Ed.), *Value-added modeling: Issues with theory and applications* (pp.40–81). Maple Grove, MN: JAM Press.

Sanders, W. L., Saxton, A. M., & Horn, S. P. (1997). The Tennessee value-added assessment system: A quantitative outcomes-based approach to educational assessment. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 137-162). Thousand Oaks, CA, Corwin Press, Inc.

Seltzer, M., Choi, K., & Thum, Y.M. (2003). Examining relationships between where students start and how rapidly they progress: Implications for conducting analyses that help illuminate the distribution of achievement within schools. *Educational Evaluation and Policy Analysis* , *25*(3), 263-286.

Appendix

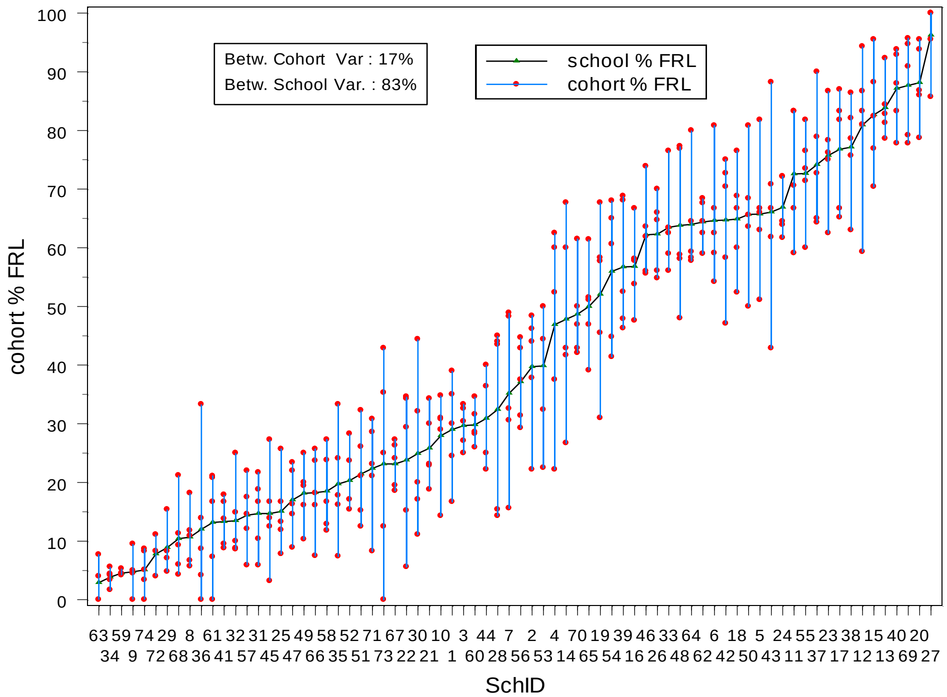

Figure A1. The percentage of students (by cohort and school) eligible to receive free/reduced price lunch. .

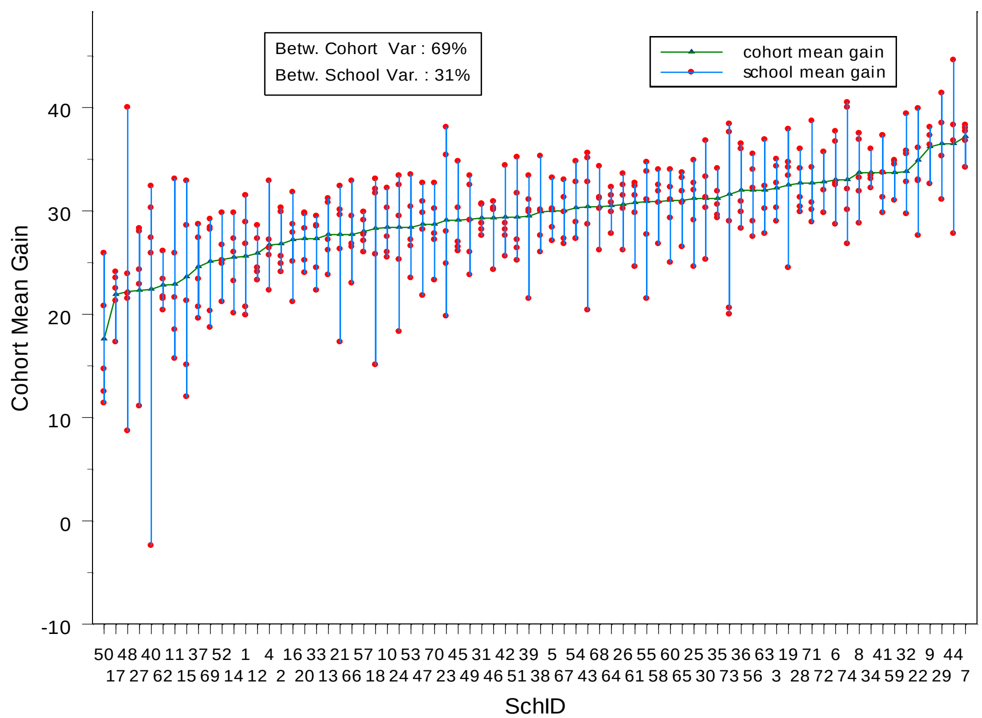

Figure A2. Observed mean gain scores by cohort and school.