# Measurement of Academic Growth of Individual Students toward Variable and Meaningful Academic Standards[1]

S. Paul Wright
William L. Sanders
June C. Rivers

Value-Added Assessment and Research
SAS Institute Inc.
Cary, NC

## 1. Introduction

In recent years there has been much interest and effort extended in using longitudinal student achievement data to measure the influence of various educational entities on the rate of student academic progress for both formative and summative evaluations. These efforts have arisen from different historical academic quantitative initiatives. The most prominent of these approaches can be placed under two broad categories. Even though others have pursued various aspects of the work, the work of Raudenbush and Bryk (2002) and Goldstein (2003) with their approach based on hierarchical linear models and the work of Sanders, et al (1997) under the banner of value-added assessment represent these two major categories. In both categories there is an attempt to exploit the relationships that exist among test scores over time at the student level. However, there are major "real world" problems that exist even within the best-constructed longitudinally merged database that require the analyst to make important decisions that indeed determine the inference space to which the resulting estimates may be applied. Some of the major issues that must be dealt with are: 1. how are fractured student records to be used; 2. how are data from different test sources to be used; 3. how are non-vertically scaled test data to be used; 4. how are

[1] This paper was published in Lissitz, Robert (Ed.) (2006). *Longitudinal and Value Added Models of Student Performance.* Maple Grove, Minnesota: JAM Press.

data from students who move among buildings to be used; 5. at the classroom level how are team teaching, departmentalized and self-contained classroom instruction to be accommodated. It is recognized that not all of these issues are germane in every application of longitudinal modeling for assessment purposes, but they represent examples of important issues that must be addressed.

While measures of individual student progress already constitute an integral part of the process of estimating the effects of educational entities referred to above, there has been a growing interest in the measurement of individual student progress toward various academic standards for a variety of other purposes. Several different approaches have been proposed for doing this (some to be presented at this conference). It is this aspect that we wish to focus on using statistical models which, while longitudinal, might not be considered to be “growth models” in the traditional sense. We hope to show why these models are useful and perhaps superior, citing their advantages relative to some of the alternative growth models. We plan to emphasis the advantages in consideration of the “real world” problems mentioned above.

## 2. Some Longitudinal (Growth) Models

In the most common usage of the term presently, a “growth model” includes an explicit equation describing student academic growth over time — that is, “time” appears explicitly in the model (where, in educational applications, “time” typically represents Year and/or Grade). Because of the hierarchical nature of student data, recent growth models have mostly been hierarchical linear models (HLMs). For simplicity in describing these models, we will keep the hierarchy simple — students nested within schools — ignoring their nesting within classrooms within schools. This has been fairly common in practice because information on which teachers taught which students has often been unavailable to the analysts. Historically, because of software limitations, a “nested” model has been used, in which students were assumed to stay in the same school over time. Recent examples in which the nested model has been used include Kiplinger (2004), Doran and Izumi (2004), and Stevens (2005). More recently, software for “cross-classified” models has become available, allowing for the possibility that students may change schools from year to year. An example is Ponisciak and Bryk (2005). Both models are described in Raudenbush and Bryk (2002: chapter 8, p. 237-245; chapter 12, p. 389-396). We briefly describe each model.

## 2.1 Nested model

Shown below is a generic nested model using the notation of Raudenbush and Bryk (2002). In this model, "i" identifies the student, "j" identifies the school within which the student is nested, and "t" is "time" (Year/Grade), often coded as t = 0, 1, 2, … The explicit growth model of a student's scores over time is given in the level-1 equation; here a linear model is used. The level-2 equations account for variation in intercepts and slopes among students within schools. Variable X is a student-level characteristic; in practice there may be multiple X variables: minority status, poverty status, special education status, initial achievement level, etc. The level-3 equations account for variation among schools. Variable W is a school-level characteristic; in practice, there may be multiple W variables: percent minority, percent in poverty, mean initial achievement level, etc.

Level-1: $Y_{tij} = \pi_{0ij} + \pi_{1ij} \cdot t + \varepsilon_{tij}$.

Level-2: $\pi_{0ij} = \beta_{00j} + \beta_{01j} X_{ij} + r_{0ij}$,

$\pi_{1ij} = \beta_{10j} + \beta_{11j} X_{ij} + r_{1ij}$.

Level-3: $\beta_{00j} = \gamma_{000} + \gamma_{001} W_j + u_{00j}$,

$\beta_{01j} = \gamma_{010} + \gamma_{011} W_j + u_{01j}$,

$\beta_{10j} = \gamma_{100} + \gamma_{101} W_j + u_{10j}$,

$\beta_{11j} = \gamma_{110} + \gamma_{111} W_j + u_{11j}$.

After exploratory modeling, some of the terms in the above model may be dropped (non-significant fixed-effects, random effects with negligible variability). For example, the $\gamma_{011} \cdot W_j$, $u_{01j}$, $\gamma_{111} \cdot W_j$, and $u_{11j}$ terms are often omitted, producing a final combined model such as the following:

$$Y_{tij} = (\gamma_{000} + \gamma_{001} W_j + \gamma_{010} X_{ij} + u_{00j} + r_{0ij}) + (\gamma_{100} + \gamma_{101} W_j + \gamma_{110} X_{ij} + u_{10j} + r_{1ij}) \cdot t + \varepsilon_{tij}.$$

See Raudenbush and Bryk (2002), Kiplinger (2004), Doran and Izumi (2004), and Stevens (2005) for additional information and examples.

## 2.2 Cross-classified model

Shown below is a simple, unconditional (no X or W variables), two-level cross-classified model. As in the nested model, the explicit growth model for a student is given by the level-1 equation. The level-2 equations contain the random variation among students in the intercepts and slopes of their growth curves (the "r" variables), and it also contains the "school effects" (the "u" variable). It is because schools and students are crossed rather than nested that they both appear in the same level of this hierarchical model. Conceptually, the "school effects" are expressed as "deflections" of the growth curve upward or downward; i.e., they affect the intercept (level) of the curve but not the slope (growth rate).

Level-1: $Y_{tij} = \pi_{0ij} + \pi_{1ij} \cdot t + \varepsilon_{tij}$.

Level-2: $\pi_{0ij} = \theta_0 + r_{0i} + u_{0j}$,

$\pi_{1ij} = \theta_1 + r_{1i}$.

Combined: $Y_{tij} = (\theta_0 + r_{0i} + u_{0j}) + (\theta_1 + r_{1i}) \cdot t + \varepsilon_{tij}$.

A problem with the above model is that the "school effect" disappears at the end of each school year. Shown below is a more realistic model which treats the school effects as cumulative. Such a model is conceptually no more complicated than the one above, but the notation is more challenging. For additional generality, the model below also includes one student-level characteristic ($X_i$) and one school-level characteristic ($W_j$). In this model, $D_{hij}$ is a dummy variable with $D_{hij}$=1 if student "i" was in school "j" at time "h", otherwise $D_{hij}$=0.

Level-1: $Y_{tij} = \pi_{0ij} + \pi_{1ij} \cdot t + \varepsilon_{tij}$.

Level-2: $\pi_{0ij} = \theta_0 + \beta_0 X_i + r_{0i} + \sum_j \sum_{h \le t} D_{hij} (\gamma_0 W_j + u_{0j})$,

$\pi_{1ij} = \theta_1 + r_{1i}$.

Combined: $Y_{tij} = [\theta_0 + \beta_0 X_i + r_{0i} + \sum_j \sum_{h \le t} D_{hij} (\gamma_0 W_j + u_{0j})] + [\theta_1 + r_{1i}] \cdot t + \varepsilon_{tij}$.

A common feature of these growth models, whether nested or cross-classified, is that the response variable (Y) represents a single characteristic that "grows" over time; that is, the test scores must be measured on a continuous vertically-linked scale. It is also necessary to specify an explicit growth function. Linear growth functions are quite popular because of their simplicity (which is why we have used them here), but non-linear functions can also be used. In contrast, the approach we describe below avoids these limitations.

## 3. EVAAS® Projection Methodology—A different approach

**3.1 The methodology.** The purpose of the EVAAS projection methodology is to provide an estimate of an individual student's academic achievement level at some point in the future under the assumption that this student will have an average schooling experience in the future. (Note that in EVAAS applications we have occasion to obtain "predictions" both for future tests a student may take and for tests a student has already taken — residuals from the latter are useful for diagnostic purposes. It is often helpful to distinguish between the two by calling the former "projections" and the latter "predictions.") The basic methodology is simply to use a student's past scores to predict ("project") some future score. At first glance, the model used to obtain the projections appears to be no more complex than "ordinary multiple regression," the basic formula being:

$$\text{Projected\_Score} = M_Y + b_1(X_1 - M_1) + b_2(X_2 - M_2) + \ldots = M_Y + \mathbf{x}_i^T\, \mathbf{b}$$

where $M_Y$, $M_1$, etc. are estimated mean scores for the response variable (Y) and the predictor variables (Xs). However, several circumstances cause this to be other than a straightforward regression problem. (1) Not every student will have the same set of predictors; that is, there is a substantial amount of "missing data." (2) The data are hierarchical: students are nested within classrooms, schools, and districts, and the regression coefficients need to be calculated in such a way as to properly reflect this. (3) The mean scores that are substituted into the regression equation also must be chosen to reflect the interpretation that will be given to the projections. As noted above, in EVAAS applications a projection is the score that a student would be expected to make assuming that the student has *the average schooling experience* in the future. The means should therefore be those of an average school within the population of schools of interest. Also, given this interpretation, the nesting needs to be carried only to the school level (students within schools); it is not necessary to carry it to the classroom level.

The missing data problem can be solved by finding the covariance matrix of all the predictors plus the response, call it $\mathbf{C}$, with submatrices $\mathbf{C}_{XX}$, $\mathbf{C}_{XY}$ (and $\mathbf{C}_{YX} = \mathbf{C}_{XY}^T$), and $\mathbf{C}_{YY}$. The regression coefficients (slopes) can then be obtained as $\mathbf{b} = \mathbf{C}_{XX}^{-1}\, \mathbf{C}_{XY}$. For any given student, one can use the subset of **C** corresponding to that student's set of scores to obtain the regression coefficients for projecting that student's Y value. Because of the hierarchical nature of the data (the second problem), the covariance matrix **C** must be a pooled-within-school covariance matrix. We obtain this matrix by maximum likelihood estimation using an EM algorithm (to handle missing values) applied to school-mean-centered data. Means for an "average school" are obtained by calculating school-mean scores and averaging them over schools. For brevity, we refer to the elements of **C**, along with the vector of estimated means, as the "projection parameters." Generally, we obtain the projection parameters using the most recent year's data. That is, we use students who have a Y value in the most recent year and X values from earlier years to get the projection parameters. Projections are then obtained by applying these parameters to students who have X values in the current year (and earlier years) but no Y value.

Note that, unlike the growth models described above, the EVAAS methodology does not require vertically linked data nor does it need to assume a linear growth function (or any other specific growth function). Instead, what is required are good predictors of the response variable. The predictors need not be on the same scale with the response or with one another. Potentially, they could be test scores from different vendors and even in different subjects from the response. This gives the EVAAS methodology considerable flexibility.

**3.2 A connection to "growth models."** Consider the following simplified two-level nested linear growth model where "i" identifies a student and "t" is time:

Level-1: $Y_{ti} = \pi_{0i} + \pi_{1i} \cdot t + \varepsilon_{ti}$.

Level-2: $\pi_{0i} = \beta_{00} + r_{0i}$,

$\pi_{1i} = \beta_{10} + r_{1i}$.

Combined: $Y_{ti} = (\beta_{00} + r_{0i}) + (\beta_{10} + r_{1i}) \cdot t + \varepsilon_{ti} = (\beta_{00} + \beta_{10} \cdot t) + (r_{0i} + r_{1i} \cdot t + \varepsilon_{ti}) = \mu_t + \delta_{ti}$.

Recall that the "projection parameters" for the EVAAS methodology consist of a vector of estimated means plus an estimated covariance matrix. The final combined model above has this same structure. The collection of $\mu_t$ values constitutes a vector of means, and the "errors" ($\delta_{ti}$) are correlated. Specifically, the error covariance matrix for the i-th student is

$\mathbf{C}_i = \text{var}(\boldsymbol{\delta}_i) = \mathbf{Z}_i \, \mathbf{T} \, \mathbf{Z}_i^T + \mathbf{I}\sigma^2$ where

$\sigma^2 = \text{var}(\varepsilon_{ti})$, assumed to be the same for all "t" and "i";

$\mathbf{T} = \text{var}(\{r_{0i}, r_{1i}\})$, assumed to be the same for all "i";

$\mathbf{Z}_i$ has two columns: a column of "1"s (intercept column) and a column of "t"s.

However, there are important differences between the growth model and the EVAAS model in the nature of the means and covariances. (1) The estimated means in the EVAAS model need not fall along a straight line (or follow any other specific functional form); indeed, as already noted they need not be on the same scale or even in the same subject. (2) The covariance matrix in the EVAAS model is completely unstructured while the one from the linear growth model has a specific structure as shown above. Nevertheless, given the structural similarity of the two models in those cases where either model may be applied, it is of interest to compare their performance in making projections. This is addressed in the next section.

**3.3 Simulation results.** In order to better evaluate the differences between the EVAAS model and the linear growth model, a small simulation study was done. $Y_{ti}$ values were generated for 2500 students on 4 occasions (i = 1, …, 2500 and t = 0, 1, 2, 3) using the nested linear growth model of section 3.2 with $\beta_{00} = 400$ and $\beta_{10} = 100$, producing means ($\mu_t$) of 400, 500, 600, 700. In addition, $\sigma^2 = \text{var}(\varepsilon_{ti}) = 5^2 = 25$, $\tau_{00} = \text{var}(r_{0i}) = 15^2 = 225$, $\tau_{11} = \text{var}(r_{1i}) = 5^2 = 25$, $\tau_{01} = \text{cov}(r_{0i}, r_{1i}) = 0$. (Additional simulations were done in which $r_{0i}$ and $r_{1i}$ were either positively or negatively correlated; results were similar to those reported here.) Projections for t=3 were then obtained using the nested linear growth model and using the EVAAS model. For each model, the projected values were compared to the actual $Y_{3i}$ values using the mean prediction error (MPE), also called bias, calculated as $\sum [\text{projected}(Y_{3i}) - Y_{3i}] / 2500$, and using the mean squared prediction error (MSPE) calculated as $\sum [\text{projected}(Y_{3i}) - Y_{3i}]^2 / 2500$. Ideally, MPE should be zero, indicating unbiased projections; and a smaller MSPE indicates better performance.

Implementing the simulation highlights another difference between the EVAAS model and the growth model. In the EVAAS model, two different cohorts of students are required. One cohort, consisting of students who have already taken the t=3 test, is used to obtain the projection parameters. The parameters are then applied to a second cohort of students who (in actual applications) have not yet taken the t=3 test (although, in the simulation, we know their t=3 scores). Consequently, in the simulation, two different sets of 2500 scores were generated for the

EVAAS model. In contrast, the linear growth model requires slopes and intercepts for those students whose t=3 scores are going to be projected. Consequently the model parameter estimates must come from the same cohort of students whose scores are to be projected. This is the second cohort of students who would not yet have taken the t=3 test. In the simulation, this scenario was implemented by using the t=0, 1, 2 data from the second cohort to estimate the growth model parameters. These parameters were then used to project the score at t=3.

Results from this simulation are shown in the first row ("Linear-1") of Table 1 (the other rows are described below). What is noteworthy here is that the two models (labeled HLM and EVAAS in the table) performed equally well. Both produced essentially unbiased projections (MPE near zero), and the MSPEs were virtually identically. Indeed, the projections themselves were nearly identical. The "Max Abs Diff" column shows the magnitude of the largest difference between the projection from the growth model and the projection for the same student from the EVAAS model. In this case, the projections differed by less than half a point, at the most, in a data set in which the t=3 scores had a standard deviation of approximately 16.6.

The initial ("Linear-1") simulation represents a "best case scenario" in that all the assumptions of the models were met. Additional simulations were done to examine the consequences of violating the model assumptions. An examination of the pattern of scores across grades (along the $50^{th}$ or any other percentile) for any number of large scale, vertically linked standardized tests reveals that nonlinearity is the rule rather than the exception. Thus it seemed important to examine the consequences of nonlinearity. The data for these additional simulations were generated with using the model

$$Y_{ti} = \mu_t + r_{0i} + \varepsilon_{ti}.$$

As before, $\sigma^2 = \mathrm{var}(\varepsilon_{ti}) = 25$ and $\tau_{00} = \mathrm{var}(r_{0i}) = 225$. For $\mu_t$, three different specifications were used in order to control the degree of nonlinearity. First, as a reference, a linear growth curve was used with $\mu_t = \{400, 500, 600, 700\}$, the same as in the previous simulation. This is referred to as "Linear-2" in Table 1. It differs from "Linear-1" in that the data were generated with random intercepts but without random slopes, resulting in less overall variation. This resulted in a smaller MSPE for both models, but otherwise the results were comparable to the "Linear-1" results.

Second, a "negligibly nonlinear" growth pattern, labeled "Nonlinear-1," was used with $\mu_t$ = {400, 505, 605, 700}. If these values are plotted against time, the nonlinearity is not evident to the naked eye. The results for the EVAAS model were the same as for "Linear-2." This was expected since the EVAAS model does not assume linearity. For the HLM linear growth model, however, the results were quite different. The projections were biased, with the projected scores averaging about 8 points higher than the actual scores; and the MSPE was about three times as large as for the EVAAS model.

Third, a "modestly nonlinear" growth pattern, labeled "Nonlinear-2," was used with $\mu_t$ = {400, 510, 610, 700}. Again, as expected, the EVAAS model results were the same as for "Linear-2." For the HLM linear growth model, the amount of bias in the projections doubled to over 16 points, and the MSPE grew to nearly ten times the size of the MSPE for the EVAAS model.

Table 1. Growth Curve Simulation Results

| | MPE HLM | MPE EVAAS | MSPE HLM | MSPE EVAAS | Max Abs Diff |
|---|---|---|---|---|---|
| Linear-1 | −0.19 | −0.11 | 65.1 | 65.1 | 0.47 |
| Linear-2 | −0.19 | +0.14 | 32.0 | 32.0 | 0.55 |
| Nonlinear-1 | +8.15 | +0.14 | 98.3 | 32.0 | 8.37 |
| Nonlinear-2 | +16.48 | +0.14 | 303.5 | 32.0 | 17.05 |
| Nonlinear-1 / Linear-2 | −0.19 | −3.02 | 32.0 | 41.1 | 3.09 |
| Nonlinear-2 / Linear-2 | −0.19 | −6.19 | 32.0 | 70.2 | 6.25 |

Because the HLM growth model is a special case of the EVAAS model, and because nonlinear growth is so commonplace, it seemed most reasonable to focus on the consequences of nonlinearity. Nevertheless, during the question-and-answer period following the public presentation of this paper, the question was raised as to whether we had "tilted" our simulations to favor EVAAS. In response, we have added two additional simulations which favor the linear growth model. While, in general, the linear growth model makes more restrictive assumptions than the EVAAS model, there is one assumption that the EVAAS model makes that is not made by the linear growth model. Recall that the EVAAS methodology requires two cohorts of students; one cohort is used to obtain the parameter estimates to be used in making projections in the second cohort. The linear growth model, in contrast, uses the same cohort (the second cohort) for both parameter estimation and projection. The assumption, in using the EVAAS methodology, is that the same parameters (means and covariances) apply in both cohorts.

In the two additional simulations, this assumption was violated by using a model to generate cohort one that differed from that used to generate cohort two. Cohort two was generated using the "Linear-2" specifications. As a result, the assumptions of the linear growth model were met; and as the last two rows of Table 1 show, the linear growth model results matched those of the "Linear-2" simulation (row two of Table 1). Cohort one, however, was generated from a nonlinear model, either "negligibly nonlinear ("Nonlinear-1") or "modestly nonlinear" ("Nonlinear-2"). Thus, the EVAAS parameter estimates were "anticipating" nonlinear growth, but the students to whom they were applied to get projections displayed linear growth. As a consequence, the EVAAS model results were biased, but the amount of bias was smaller than that of the linear growth model when applied to nonlinear data. Also, as expected, the MSPEs for the EVAAS model were higher than for the linear growth model; but, again, the increase in MSPE was much smaller than that which occurred with the linear growth model under nonlinearity. Specifically, the EVAAS MSPE was larger than the linear growth model MSPE by about 28% (for "Nonlinear-1") to 220% ("Nonlinear-2"). Recall that in the "reverse" situation when cohort two had nonlinear growth (rows 3 and 4 of Table 1), the linear growth model MSPE was larger than the EVAAS MSPE by about 307% ("Nonlinear-1") to 948% ("Nonlinear-2").

To summarize: When the assumptions of the models are met, they performed equally well (in these simulations). When the assumptions of one model were violated, but the assumptions of the other model were met, the "correct" model performed better. However, in these simulations, the EVAAS model seemed to be more robust in the presence of a violation of assumptions than did the linear growth model. Finally, as to the realism of the violations of assumptions, nonlinear growth seems to be commonplace in most of the vertically scaled tests with which we are familiar. Thus, the consequences of nonlinearity are of particular concern. On the other hand, the possibility that the means and covariances might change non-negligibly from one cohort to the next seems to us less plausible, at least in the case of a large district with a relatively stable student population. For a smaller district undergoing considerable demographic change, this would be more of a concern and would merit careful monitoring.

**3.4 EVAAS Projection Advantages.** There are a number of features of the EVAAS projection methodology that are attractive.

- Unlike growth curve models, *there is no requirement that the tests scores (Ys and Xs) be vertically linked*; indeed they need not even be from the same test company or even in the same subject! The important feature is that the X-values be good predictors of the Y-value. This provides an enormous amount of flexibility in the choice of what could be projected and which predictors (Xs) to use in making the projections (see Section 4).
- Even in the case when the Xs and Ys are vertically linked, *there is no assumption required about the overall shape of the growth curve*. Use of the covariance matrix (**C**) does carry with it the implicit assumption of linear relationships between pairs of scores but no assumption of linearity, or any other shape, over time.
- Missing values are easily handled so that different students can have different sets of predictors.
- Massive data sets are readily accomodated. For example, for one of our applications (for the state of Tennessee), we are able to provide projections for every student in the state *to a variety of endpoints, ranging from next year's test scores to high school end-of-course test scores to college entrance exam scores*.

## 4. Using Projections to Enhance NCLB and Other Important Educational Objectives

The motivation for using projections in conjunction with NCLB is this: having students who are currently below proficiency but who are "on track" to be proficient at some future point should not be held against a school (or district), especially if that school has done a good job of accelerating students toward proficiency. In effect, this carries the "adequate yearly progress" idea to the student level. Projections provide a convenient way to identify whether or not a student is on track to be proficient. However, several decisions must be made in order to use projections for this purpose.

First, specific future assessments must be chosen as the basis for projecting eventual proficiency. These assessments must include Mathematics and Reading (and eventually Science) since the focus of NCLB is proficiency in those subjects. As an example of such assessments, in Tennessee there are high school "gateway" tests in Mathematics (Algebra I), Reading (English

II) and Science (Biology I) which students must pass to graduate. These provide reasonable future points at which to assess projected proficiency. Because of Tennessee's long history of annual testing, it is possible to begin projecting student results on these gateway tests as early as fourth grade!

Second, "eventual proficiency" must be defined. A simple definition is: if the projected score is above the proficiency cutpoint, the student is considered to be on track to being proficient. Third, one must decide how projected proficiency could get counted in the NCLB percent proficient calculations. Here is one possibility that demonstrates how this could be applied. If a student's projected score is above the proficiency standard for an approved academic endpoint, then this student will be deemed to be "proficient." After all students' projected scores are evaluated in this manner, then all of the approved AYP rules can be applied, including the requirements for meeting the standards for all subgroups.

Although the individual student projections are increasingly applicable to the NCLB safe harbor discussion, one of the original intentions for providing this student level information was to encourage educators to consider the academic needs of individual students rather than groups of students. The availability of projections to *varying* endpoints, serves as a reminder to educators that some students are underserved educationally if the only expectation for them is minimal proficiency. Proficiency in the next grade may be the direst academic need for a student or a school with disproportionately more students at lower achievement levels, but it should never be interpreted as a blanket expectation for all students within those schools. Students with demonstrated higher levels of academic attainment deserve a schooling experience that enables appropriate academic progress, even beyond the minimum proficiency determination. Students whose academic growth is sustained each year at an appropriate level are better prepared for advanced high school course work and have a greater likelihood of college or entry level work success.

One specific use of the individual student projections is an indicator for determining course assignment. In urban areas, access within the district to the projections can accelerate the guidance necessary to enroll students who transfer frequently within the district. For example, students who demonstrate sufficiently high probabilities of success in algebra in grade six should be encouraged to complete algebra before grade nine and to continue with rigorous coursework across the high school years. Particularly in schools serving poor or minority students, lack of availability of advanced curricula for adequately prepared students in middle grades contributes to widening achievement gaps. Sixth grade student projections to algebra proficiency provide a heads-up regarding the number of algebra classes necessary to meet the academic needs of the school's population.

Another contributor to widening achievement gaps has to do with students' inability to view themselves as potentially successful college students some time in the future. Special mentoring programs to enhance counseling support are another use of the student projections. In these instances, the student projections provide a communication vehicle for principals, teachers and guidance counselors to demonstrate to students and parents the importance of rigorous courses for students who are academically prepared to be successful in these courses.

## 5. Discussion/Conclusions

The terms "growth model" and "value added model" are often loosely used interchangeably. However, a clearer distinct between the two terms can be drawn if the intent of the use of the analytical results from the subsequent data analyses are implied. If the intent is to use student longitudinal data to account for prior academic achievement levels to enable a fairer, more objective measure of the influence of various educational entities on the rate of student progress, then the term "value-added model" is most often used. If however, the intent is to use the longitudinal analyses to provide estimates of future performance for individual students then the terms "growth model" or "projection model" would be favored.

Regardless of the use of the longitudinal data, either for "growth or projection models" or for "value-added models," several non-trivial analytical problems have to be addressed:

- How to accommodate fractured student records without introducing major biases in the resulting analyses by either eliminating the data for students with missing data, or by using overly simplistic imputation procedures?
- How to exploit all of the longitudinal data for each student when all of the historical data are not on the same scale?

- How to provide educational policy makers more flexibility in the use of historical data when testing regimes have changed over time? *Note: As we have worked with many districts within many states, very few have maintained the same testing regime consistently over years.*

Considering all of these factors, we have deliberately chosen to pursue our projection modeling efforts because many of the other proposed growth models lack the flexibility and robustness to accommodate the reality of the data structures that presently exist and are likely to be present in the future. Additionally, it is with this same recognition that all of our value-added models have been engineered to have this same flexibility and robustness.

**References**


Doran, H. C. and Izumi, L. T. (2004). Putting Education to the Test: A Value-Added Model for California. Pacific Research Institute. Available at www.pacificresearch.org/press/edu/papers.html.

Goldstein, H. (2003). *Multilevel Statistical Methods, 3rd Edition.* London: Hodder Arnold.

Kiplinger, V. L. (2004). Longitudinal Study of Student Growth in Reading, Writing and Mathematics Achievement in Thirty-six Colorado Public School Districts, Phase II, Final Report. Available at http://matrix10.d20.co.edu/lssg/phase2.html.

Lissitz, R. W., editor (2005). *Value Added Models in Education: Theory and Applications.* Maple Grove, Minnesota: JAM Press.

Ponisciak, S, and Bryk, A. S. (2005). Value-Added Analysis of the Chicago Public Schools: An Application of Hierarchical Models. In Lissitz (2005).

Raudenbush, S. A. and Bryk, A. S. (2002). *Hierarchical Linear Models, Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage Publications.

Sanders, W. L., Saxton, A. M., and Horn, S. P. (1997). The Tennessee Value-Added Assessment System: A Quantitative, Outcome-Based Approach to Educational Assessment. In J. Millman (ed.), *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?*. p. 137-162. Thousands Oaks, CA: Corwin Press.

Stevens, J. (2005). The Study of School Effectiveness as a Problem in Research Design. In Lissitz (2005).