Growth Model for School Accountability 2016/17 Technical Report

November 2017

Prepared for the New York State Education Department by Education Analytics, Inc. under contract C-012798

Contents

Figures

Tables

Introduction

The New York State Education Department (NYSED) reports unadjusted growth scores that include only prior achievement as a predictor variable. NYSED also reports adjusted growth scores that control for prior achievement and student characteristics as predictor variables[1]. Unadjusted scores are reported for informational purposes to educators and are used for school accountability in Grades 4—8. For school accountability purposes, New York State uses a school's or subgroup's unweighted two-year average mean growth percentile (MGP) in ELA and mathematics.

This document describes the model used to measure student growth for institutional accountability in New York State for the 2016/17 school year. In 2016/17, growth models were implemented for institutional accountability in Grades 4—8 ELA and mathematics. All models are based on assessing each student's change in performance between 2015/16 (and prior years) and 2016/17 on State assessments compared with students who have similar prior performance. Revisions to the State-provided growth model will be considered in future years.

Content and Organization of This Report

The results presented in this report are based on 2016/17 and prior school years' data, with some comparison to prior-year results. This technical report contains four main sections:

1. **Data** – Description of the data used to implement the student growth model, including data processing rules and relevant issues that arose during processing.
2. **Model** – Description of the unadjusted statistical model.
3. **Reporting** – Description of reporting metrics.
4. **Results** – Overview of key model results aimed at providing information on model quality and characteristics.

[1] For information on the growth model used for educator evaluation, see the 2016/17 technical report.

Data

To measure student growth and attribute that growth to schools, at least two sources of data are required: student test scores that can be observed across time and information describing how students are linked to schools (i.e., identifying which school students attend for a tested subject).

The following sections describe the data used for model estimation in New York in more detail, including some of the issues and challenges that arose and how they were handled.

Test Scores

New York's student growth models drew on test score data from statewide testing programs in Grades 3—8 in ELA and mathematics for the growth model for schools of students in Grades 4—8. In Grades 4—8, institutional growth models are estimated separately by grade and subject using scores from each grade (e.g., Grade 5 mathematics) as the outcome.

State Tests in ELA and Mathematics (Grades 3—8)

The New York State tests at the elementary and middle school grade levels measure a range of knowledge and skills in mathematics and ELA. State tests in ELA and mathematics for Grades 3—8 are given in the spring. The 2016/17 school year was the fifth school year that the State tests were designed to measure the Common Core State Standards.

The New York Grades 4—8 institutional growth model uses test scores in each subject area as a predictor for that subject area (e.g., mathematics scores are used to predict mathematics scores). Specifically, New York's Grades 4—8 institutional growth model includes three prior test scores in the same subject area. If the immediate prior-year test score in the same subject was missing from the immediate prior grade, the student was not included in the growth measure for that subject. For example, students without a prior-year test score or with a prior-year test score for the same grade as the current year test score did not have growth scores computed for them.

For the other prior scores, missing data indicators were used. These missing indicator variables allow the model to include students who do not have the maximum possible test history and mean that the model results measure outcomes for students with and without the maximum possible assessment history. This approach was taken to include as many students as possible. For the 2016/17 analyses, data from 2016/17 were used as outcomes, with prior achievement predictors coming from the previous 3 years (going back to 2013/14). The specific tests used as predictors vary by grade and subject and are as follows and presented visually in Table 1:

- Grade 4 ELA and mathematics models used scores from Grade 3 in ELA and mathematics. Students were NOT included if they lacked Grade 3 scores from the immediate prior year in the same subject.
- Grade 5 ELA and mathematics models used scores from Grades 3 and 4 in ELA and mathematics. Students were NOT included if they lacked Grade 4 scores from the immediate prior year in the same subject.
- Grades 6—8 ELA and mathematics models used scores from Grades 3—7 in ELA and mathematics. Students were NOT included if they lacked the immediate prior-year score in the same subject (e.g., Grade 6 students must have had a Grade 5 score in the same subject from 2015/16).

Table 1. Prior Year Same Subject Test Scores Included

ELA and Mathematics Model by Grade	Prior Year Same Subject Test Scores Included in the Model				
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
Grade 4	✓				
Grade 5	✓	✓			
Grade 6	✓	✓	✓		
Grade 7		✓	✓	✓	
Grade 8			✓	✓	✓

In addition to test scores, the New York Grades 4—8 institutional growth model also used the conditional standard errors of measurement of those test scores. All assessments contain some amount of measurement error, and the New York Grades 4—8 institutional growth model accounts for this error (as described in more detail in the Model section of this report). Conditional standard errors were obtained from published technical reports for the assessments' prior-year test scores, and the State's test vendor provided a similar table for the 2016/17 test scores.

School Attribution

For the New York Grades 4—8 institutional growth model, students were attributed to schools based on a continuous enrollment indicator. This variable describes whether a student was enrolled at the start and end of the year in a school or district (on BEDS day and at the beginning of the State test administration in the spring). Students who met this criterion were included in school-level MGPs. Unlike teacher attribution, student results were not weighted by attendance in determining a school MGP and growth score. The policy rationale for not using attendance weighting for schools (although it is used for teachers) is that school leaders may have more influence on student attendance, and on the integrity of attendance data, than do teachers. Table 2 shows attribution rates for schools.

Table 2. Grades 4—8 School-Student Attribution Rates

Grade	Valid Student Records	Valid Student Records Attributed to at Least One School	Attribution Rate
4	306,733	298,695	97%
5	293,617	286,709	98%
6	277,145	270,494	98%
7	264,155	258,719	98%
8	217,778	202,482	93%
Total	1,359,428	1,317,099	97%

Note. Student records are considered valid for the purposes of growth modeling when there are at least two consecutive years of valid assessment scores. Students can have as many as two valid records per year, one for ELA and one for mathematics.

The attribution rate at the school level in 2016/17 (97%) was the same as the value in 2015/16. More student records overall were attributed to schools in 2016/17 than in 2015/16.

Model

This section describes the statistical model used to measure student growth in New York between two points in time on a single subject of a State assessment. The section begins with a description of the statistical model used to form the comparison point against which students are measured, and follows with a description of how SGPs are derived from the comparison point. In addition, this section describes how MGPs and all variance estimates are produced.

At the core of the New York State institutional growth model is the production of an SGP. This statistic characterizes the student’s current year score relative to other students with similar prior test score histories. For example, an SGP equal to 75 denotes that the student’s current year score is the same as or better than 75% of the students in the State with prior test score histories and other measured characteristics that are similar. It does *not* mean that the student’s growth is better than that of 75% of all other students in the population.

The institutional model implemented for New York State is a linear regression model designed to account for measurement variance in the predictor variables, as well as the outcome variable, to yield unbiased estimates of the model coefficients. Subsequently, these model coefficients are used to form a predicted score, which is ultimately the basis for the SGP. Because the prediction is based on the observed score, it is necessary to account for measurement variance in the prediction as well. Hence, the model accounts for measurement variance in two steps: first in the model estimation and second in forming the prediction. The next section describes this model in detail.

Covariate Adjustment Model

The statistical model implemented as the MGP model is typically referred to as a *covariate adjustment model* (McCaffrey, Lockwood, Koretz, & Hamilton, 2004), as the current year observed score is conditioned on prior levels of student achievement as well as other possible covariates.

In its most general form, the model can be represented as follows:

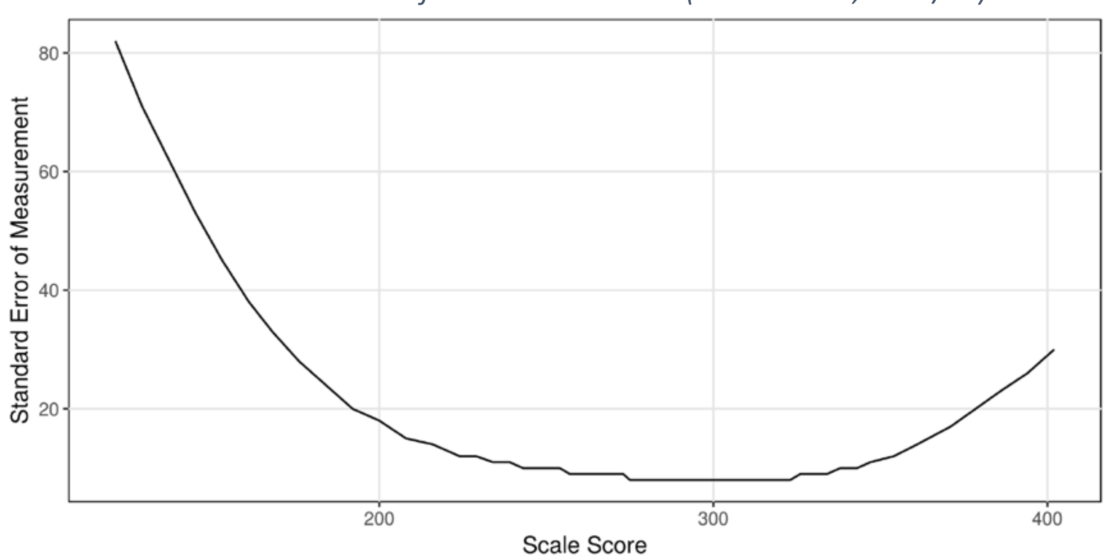
$$y_{ti} = \sum_{r=1}^{L} y_{t-r,i}\gamma_{t-r} + e_i$$

where y_{ti} is the observed score at time t for student i, y_{t-r} is the observed lag score at time $t-r$ $(r \in \{1,2,\ldots,L\})$ and γ is the coefficient vector capturing the effects of lagged scores.

Accounting for Measurement Variance in the Predictor Variables

All test scores are measured with variance, and the magnitude of the variance varies across the range of test scores. The standard errors (square roots of variances) of measurement are referred to as *conditional standard errors of measurement* (CSEMs) because the variance of a score is heteroscedastic and depends on the score itself. Figure 1 shows a sample from the Grade 8 ELA test in New York.

Figure 1. Conditional Standard Error of Measurement Plot (Grade 8 ELA, 2016/17)



Treating the observed scores as if they were the true scores introduces a bias in the regression, and this bias cannot be ignored within the context of a high-stakes accountability system

(Greene, 2003). In test theory, the observed score is described as the sum of a true score plus an independent variance component, $X = X^* + U$, where U is a matrix of unobserved disturbances with the same dimensions as X.

Our estimator accounting for the error in the predictor variables is derived in a manner similar to that of Goldstein (1995).

Specification for MGP Model for Grades 4—8

The preceding section provides details on the general modeling approach and specifically how measurement variance is accounted for in the model. The exact specification for the New York Grades 4—8 model in 2016/17 is described as follows:

$$y_{gi} = \mu + \sum_{l=1}^{K} \beta_l y_{g-r,i} + \sum_{s=1}^{M} \tau_s m_{si} + \varepsilon_i$$

where y_{gi} is the current year test scale score for student i in grade g, μ is the intercept, β_l is the set of coefficients associated with the three prior test scores, τ_s is the set of coefficients associated with the missing variable indicators, and ε_i is the student residual.

Student Growth Percentiles

The previously described regression models yield unbiased estimates of the coefficients by accounting for the measurement error in the observed scores. The resulting estimates are then used to form a student-level SGP statistic. For purposes of the growth model, a predicted value and its variance for each student are required to compute the SGPs as follows:

$$SGP_i = \Phi\left(\frac{y_i - \hat{y}_i}{\sqrt{\sigma^2_{yf,i}}}\right)$$

where SGP_i is the observed value of the outcome variable and $\hat{y}_i = w'\hat{\delta}$ where w' is the i^{th} row of the model matrix W, and the notation $\sigma^2_{yf,i}$ is used to mean the variance of the predicted value of y for the i^{th} student.

Here, the regression is of form

$$Y = W\delta + \epsilon$$

where

$$\epsilon \sim N(0, \sigma^2)$$

For this case, the classic variance of a predictor is

$$\sigma^2_{yf,i} = [1 + w_i'(w'w)^{-1}w_i]\hat{\sigma}^2_e$$

where $\hat{\sigma}^2_e$ is the variance of the predictor. However, in this case, we make two refinements to acknowledge the effect of measurement error on the residual variance. The first is to use the actual variance on y_i, called σ^2_{yi}, rather than the population variance on y_i, called $\bar{\sigma}^2_{yi}$, which is already included in $\hat{\sigma}^2_e$. This is done by subtracting the population variance and adding back the individual variance. Thus, the variance on the predictor becomes

$$\sigma^2_{yf,i} = [1 + w_i'(w'w)^{-1}w_i]\left[\sigma^2_e - \bar{\sigma}^2_{yi}\right] + \sigma^2_{yi}$$

The second refinement is to replace the population variance in w_i, called $\bar{\Sigma}$, with the individual variance in w_i, called Σ_i. This replacement is done in the same way as with the variance in y_i, so the variance estimate is now

$$\sigma^2_{yf,i} = [1 + w_i'(w'w)^{-1}w_i]\left[\sigma^2_e - \bar{\sigma}^2_{yi} - \delta'\bar{\Sigma}\delta\right] + \sigma^2_{yi} + \delta'\Sigma_i\delta$$

A predicted value for each student is used to compute the SGP. However, that prediction is based on the estimates of the fixed effects that were corrected for measurement variance but based on the observed score in vector w.

Figure 2 illustrates how the SGPs are found from the previously described approach. The illustration considers only a single predictor variable, although the concept can be generalized to multiple predictor variables, as presented earlier. For each student, we find a predicted value conditional on his or her observed prior scores and the model coefficients. To illustrate the concept, assume we find the prediction and its variance but do not account for the measurement variance in the observed scores used to form that prediction. We would form a conditional distribution around the predicted value and find the portion of the normal distribution that falls below the student's observed score. This is equivalent to

$$SGP_i = \int_{-\infty}^{y_i} f(x)dx$$

with $f(x) \sim N\left(\hat{y}_i, \sigma^2_{yfi}\right)$, although this is readily accomplished using the cumulative normal distribution function, $\Phi(\cdot)$.

Figure 2. Sample Growth Percentile from Model

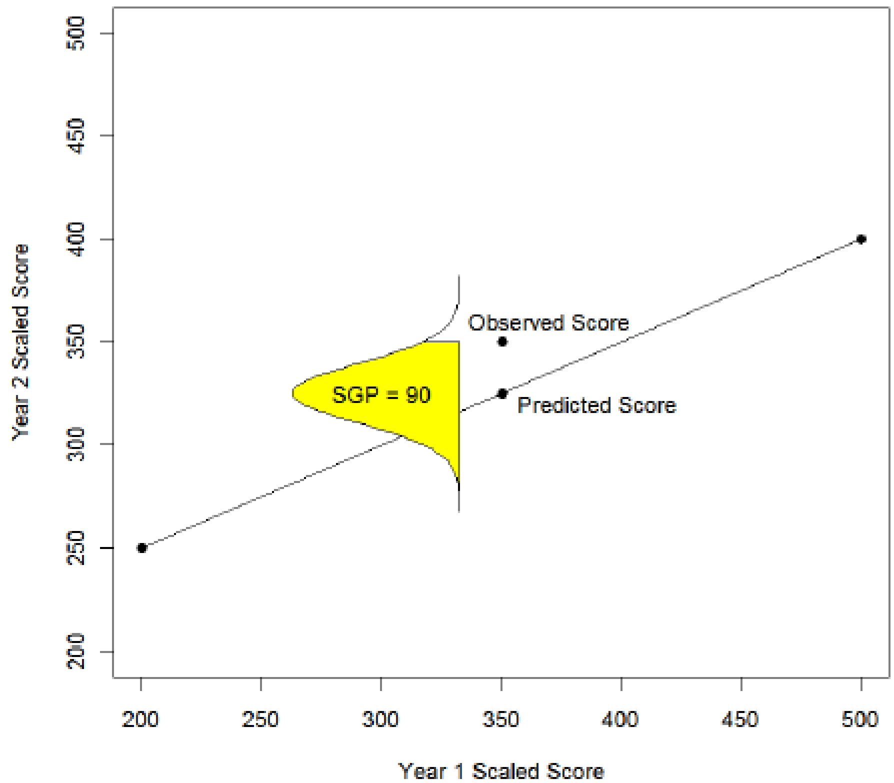
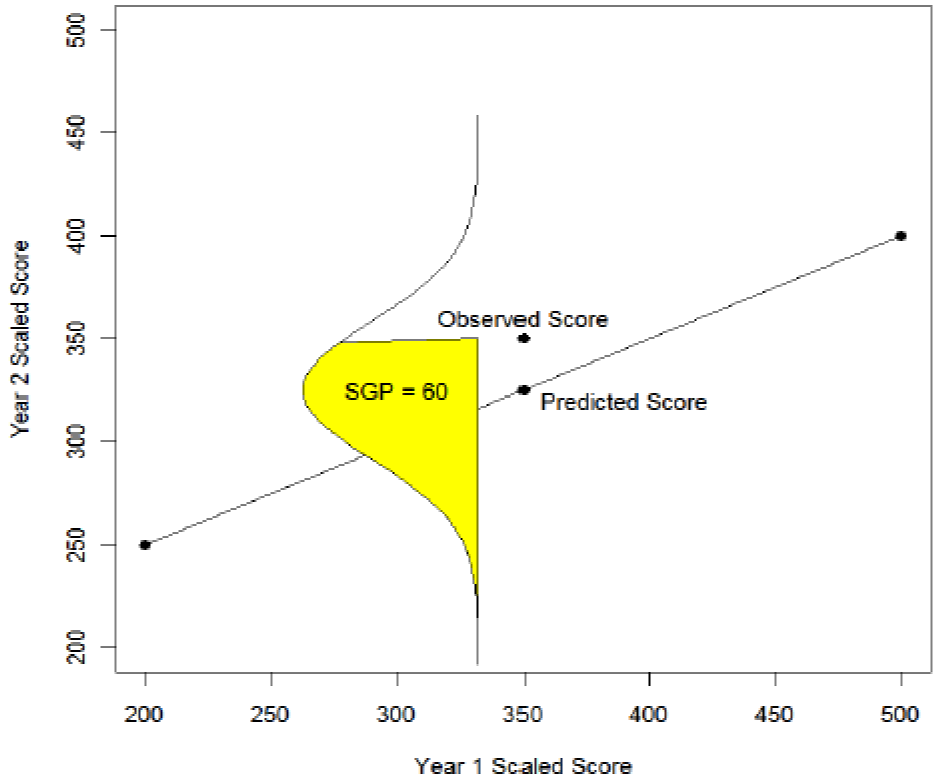


Figure 3 illustrates the same hypothetical student shown in Figure 2. Note that the observed score and predicted value are identical. However, the prediction variance is larger than in Figure 2. As a result, when we integrate over the normal from $-\infty$ to y_i, the SGP is 60, not 90 as in the previous example. This difference occurs because the conditional density curve has become more spread out, reflecting less precision in the prediction.

Figure 3. Sample Growth Percentile from Model



Mean Growth Percentiles

Once SGPs are estimated for each student, group-level (e.g., school-level) statistics can be formed that characterize the typical performance of students within a group. New York's growth model Technical Advisory Committee recommended using a mean SGP for educator scores. Hence, group-level statistics are expressed as the mean SGP within a group. This statistic is referred to as the *MGP*.

For each aggregate unit ($j \epsilon \{1,2, \ldots, J\}$), such as a school, the statistic of interest is a summary measure of growth for students within this group. Within group j, there are $\{SGP_{j(1)}, SGP_{j(2)}, \ldots, SGP_{j(N)}\}$. That is, there is an observed SGP for each student within group j.

Then the MGP for unit j is produced as the simple mean

$$\theta_j = mean(SGP_{j(1)})$$

As with all statistics, the MGP is an estimate, and it has a variance term. The following measures of variance are produced for the MGP.

The analytic standard error of the unweighted MGP for schools is computed within unit j as

$$se(\theta_j) = \frac{sd(SGP_{ij})}{\sqrt{N_j}}$$

where $sd(SGP_{ij})$ is the sample standard deviation of the SGPs in group j, and N_j is the number of students in group j.

Combining Student Growth Percentiles Across Grades and Subjects

Many schools serve students from different grades and with results from different tested subjects. For evaluation purposes, there is a need to aggregate these SGPs and form a summary measure, in this case, mean growth percentiles (MGPs).

Because the SGPs are expressed as percentiles, they are free from scale-specific inferences and can be combined. For any aggregate-level statistics to be provided (MGPs), all SGPs of relevant students are pooled and the mean of the pooled SGPs is found.

Reporting

The main reporting metrics for schools of Grades 4—8 were as follows:

- **Number of Student Scores** – The number of SGPs included in an MGP.
- **Unadjusted MGP** – The mean of the SGPs for students attributed to the school based on similar prior achievement scores only, without taking into consideration ELL, disability, economic disadvantage, or other student characteristics.
- **Lower Limit and Upper Limit** – Highest and lowest possible MGP for a 95% confidence range.

MGPs disaggregated by grade and subject also are provided. Districts also are provided with student roster files. These files show which students were included in a school's MGP along with information about each student, such as whether the student has a disability or is identified as an ELL.

Minimum sample size requirements for reporting MGPs and growth ratings were determined to balance statistical reliability and availability of school growth scores. On one hand, setting no (or a low) minimum sample size will result in the greatest number of schools receiving information; on the other hand, the quality of the information they receive may be reduced. A minimum threshold of 16 student scores was implemented. Scores on any measure at any level based on fewer than 16 student scores were not reported.

After applying this rule, the fraction of schools with reported results is shown in Table 3 for Grades 4—8. The percentages of schools receiving results in 2016/17 were unchanged relative to the 2015/16 percentages.

Table 3. Grades 4—8 Reporting Rates

Number of Schools with at Least One Student Attributed	Number of Schools Meeting the Minimum Sample Size Requirement	Percentage of Schools Meeting the Minimum Sample Size Requirement
3,735	3,567	96%

For schools of Grades 4—8, the overall MGP (i.e., the MGP that combines information across all applicable grade levels and subjects outlined in the previous section) and upper and lower limit MGPs were used to determine growth ratings.

Results

This section provides an overview of the results of the 2016/17 growth model estimation. Some comparisons to earlier year growth model results are also included. A pseudo R-squared statistic and summary statistics characterizing the SGPs, MGPs, and their precision provide an overview of model fit.

Model Fit Statistics for Grades 4—8

The *R*-square value is a statistic commonly used to describe the goodness-of-fit for a regression model. Because the model implemented here is an error-in-variables (EiV) model[2], not a least squares regression, we refer to this as a *pseudo R*-square. Table 4 presents the pseudo *R*-square values for each grade and subject, computed as the squared correlation between the fitted values and the outcome variable.

Table 4. Grades 4—8 Unadjusted Model Pseudo R-Squared Values by Grade and Subject

Grade	ELA	Mathematics
4	0.63	0.68
5	0.68	0.73
6	0.71	0.74
7	0.72	0.74
8	0.68	0.61

Student Growth Percentiles for Grades 4—8

SGPs describe a student's current year score relative to those of other students in the data with similar prior academic histories. A student's SGP should not be expected to be higher or lower

[2] For additional information about the EiV approach see the Model section of the [2016/17 technical report].

based on his or her prior-year score. Table 5 shows the correlation between the prior-year scale score and SGP for each grade and subject. These correlations are usually negative as a result of using the EiV approach to account for measurement variance in the prior-year scale score; the correlation need not be zero. Squaring these values gives the percentage of variation in SGPs explained by prior-year scores for any grade and subject. Although prior-year test scores are generally good predictors of current year test scores, the prior-year test score is a poor predictor of current year SGPs. As shown in Table 5, prior-year test scores explain about 3% to 5% of the variation in SGPs[3]. Because SGPs are intended to allow students to show low or high growth no matter their prior performance, this result is as expected.

Table 5. Grades 4—8 Unadjusted Model Correlation Between SGP and Prior-Year Scale Score

Grade	ELA	Mathematics
4	-0.171	-0.137
5	-0.153	-0.191
6	-0.136	-0.148
7	-0.152	-0.219
8	-0.150	-0.222

Reliability of Unadjusted MGPs

It is useful to examine the reliability statistic to assess the precision of the school-level MGPs, specified here as ρ:

$$\rho = 1 - \left(\frac{\bar{\sigma}}{sd(\hat{\theta}_j)}\right)^2$$

where $\bar{\sigma}$ is the mean standard error of the MGP, and $sd(\hat{\theta}_j)$ is the standard deviation between school MGPs. In theory, the highest possible value is one, which would represent complete precision in the measure. When the ratio is zero, the variation in MGPs is explained entirely by sampling variation. Larger values of ρ are associated with more precisely measured MGPs.

Table 6 provides the weighted mean standard errors, the weighted standard deviations, and the values of weighted ρ for the unadjusted model for schools, using the number of SGPs as weights. Higher values of ρ are associated with more precisely measured MGPs.

---

[3] You can measure the strength of the relation between SGP and prior-year scale score by squaring the correlation and multiplying by 100. This gives you the amount of variation in SGPs explained by prior-year scale score.

Table 6. Grades 4—8 Weighted Unadjusted Model Mean Standard Errors, Standard Deviation, and Value of ρ by Grade for Schools, Weighted by Number of SGPs

Grade	Weighted Mean Standard Error	Weighted Standard Deviation	Weighted Reliability Statistic (ρ)
4	2.278	8.129	0.913
5	2.321	7.843	0.903
6	1.929	8.133	0.934
7	1.926	7.094	0.913
8	2.112	6.999	0.893

Table 7 provides the share of schools whose combined unadjusted MGPs are significantly above or below the State mean, using the 95% confidence intervals. In all cases, the percentage exceeding the mean is larger than what would be expected by chance alone, indicating the model distinguishes between schools (2.5% of schools would be expected to be above or below the mean by chance alone).

Table 7. Grades 4—8 Unadjusted Model School Combined MGPs Above or Below the Mean at a 95% Confidence Level

	Below Mean		Above Mean	
Grade	N	%	N	%
4	737	31%	595	25%
5	570	25%	595	26%
6	451	28%	484	30%
7	352	24%	428	29%
8	318	22%	385	27%

Neutrality of Unadjusted MGPs

Given that a primary claim for the use of MGPs in institutional accountability is that all schools can demonstrate growth, regardless of the academic starting point of students, it is necessary to determine if there is a strong relationship between MGPs and average prior achievement for students in a school. To that end, Table 8 shows the correlations between MGPs and average prior achievement, which are low to moderate across all grades and subjects. These correlations illustrate that the MGPs are substantially neutral to prior achievement.

Table 8. Correlation Between Unadjusted Overall MGP and Average Prior Achievement Across Grades and Subjects

Measure of Prior Achievement		
Subject	**Grade**	**Correlation Between Unadjusted Overall MGP and Prior Achievement**
ELA	Grade 4	0.088
	Grade 5	0.041
	Grade 6	0.053
	Grade 7	0.067
	Grade 8	-0.100
Mathematics	Grade 4	0.159
	Grade 5	-0.045
	Grade 6	0.046
	Grade 7	-0.022
	Grade 8	0.003

References

Goldstein, H. (1995). *Multilevel statistical models.* Bristol, UK: University of Bristol. Retrieved from http://www.bristol.ac.uk/cmm/team/hg/multbook1995.pdf

Greene, W. H. (2003). *Econometric analysis* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Hausman, J. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives, 15*(4), 57–67.

Henderson, C. R. (1953). Estimation of variance and covariance components. *Biometrics, 9,* 226–252.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2004). *Evaluating value-added models for teacher accountability.* Santa Monica, CA: RAND.

McCaffrey, D. F., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education, Finance and Policy, 4*(4), 572–606.

Wei, Y., & Carroll, R. J. (2009). Quantile regression with measurement error. *Journal of the American Statistical Association, 104,* 1129–1143.

Appendix A. Model Coefficients

The tables that follow display regression model coefficients (labeled as "Effects") for the New York growth model in each grade and subject. For the Grades 4—8 model, these model coefficients represent the predicted change in current year test scores for one unit of change in each variable shown in the table, holding other variables constant. For example, in Table A 1, the predicted change in a student's current year ELA test score given a one-point increase in a student's prior grade ELA test score is 0.903. The interpretation of a one-unit change varies by variable type. For yes/no variables, model coefficients represent the predicted change in current year test scores given a change from no to yes. Missing flags are yes/no variables set to yes if the noted variable is missing and no otherwise.

Because of the differences in model and variable types, it is important to keep in mind that effect sizes cannot be compared directly across different types of variables.

Table A 1. Grade 4 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	27.053	0.577	0.000
Prior-Grade ELA Scale Score	0.903	0.002	0.000

Table A 2. Grade 5 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	-24.820	0.618	0.000
Prior-Grade ELA Scale Score	0.839	0.004	0.000
Two-Grades-Prior ELA Scale Score	0.233	0.004	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	68.815	1.182	0.000

Table A 3. Grade 6 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	32.410	0.514	0.000
Prior-Grade ELA Scale Score	0.602	0.004	0.000
Two-Grades-Prior ELA Scale Score	0.187	0.004	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	54.236	1.307	0.000
Three-Grades-Prior ELA Scale Score	0.108	0.004	0.000
Missing Flag: Three-Grades-Prior ELA Scale Score	32.448	1.196	0.000

Table A 4. Grade 7 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	45.430	0.491	0.000
Prior-Grade ELA Scale Score	0.630	0.004	0.000
Two-Grades-Prior ELA Scale Score	0.143	0.004	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	39.955	1.154	0.000
Three-Grades-Prior ELA Scale Score	0.106	0.004	0.000
Missing Flag: Three-Grades-Prior ELA Scale Score	32.112	1.081	0.000

Table A 5. Grade 8 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	31.061	0.571	0.000
Prior-Grade ELA Scale Score	0.734	0.005	0.000
Two-Grades-Prior ELA Scale Score	0.144	0.005	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	39.825	1.384	0.000
Three-Grades-Prior ELA Scale Score	0.038	0.004	0.000
Missing Flag: Three-Grades-Prior ELA Scale Score	12.203	1.199	0.000

Table A 6. Grade 4 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	12.932	0.544	0.000
Prior-Grade Mathematics Scale Score	0.951	0.002	0.000

Table A 7. Grade 5 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	20.129	0.525	0.000
Prior-Grade Mathematics Scale Score	0.729	0.003	0.000
Two-Grades-Prior Mathematics Scale Score	0.213	0.004	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	65.057	1.159	0.000

Table A 8. Grade 6 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	-27.058	0.607	0.000
Prior-Grade Mathematics Scale Score	0.723	0.005	0.000
Two-Grades-Prior Mathematics Scale Score	0.217	0.005	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	65.154	1.491	0.000
Three-Grades-Prior Mathematics Scale Score	0.139	0.005	0.000
Missing Flag: Three-Grades-Prior Mathematics Scale Score	41.836	1.364	0.000

Table A 9. Grade 7 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	12.964	0.534	0.000
Prior-Grade Mathematics Scale Score	0.728	0.004	0.000
Two-Grades-Prior Mathematics Scale Score	0.144	0.004	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	44.216	1.302	0.000
Three-Grades-Prior Mathematics Scale Score	0.086	0.004	0.000
Missing Flag: Three-Grades-Prior Mathematics Scale Score	26.761	1.091	0.000

Table A 10. Grade 8 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p value
Constant Term	-17.107	0.873	0.000
Prior-Grade Mathematics Scale Score	0.823	0.007	0.000
Two-Grades-Prior Mathematics Scale Score	0.162	0.008	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	47.712	2.208	0.000
Three-Grades-Prior Mathematics Scale Score	0.047	0.006	0.000
Missing Flag: Three-Grades-Prior Mathematics Scale Score	15.902	1.699	0.000