Growth Model for Institutional Accountability 2017/18 Technical Report

April 2019

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

Contents

Figures

Tables

Introduction

This document describes the model used to measure student growth for institutional accountability in New York State for the 2017/18 school year and how three years of student growth results were combined to generate a three-year growth measure called the Growth Index. The Growth Index is new in 2017/18 and is used to make accountability determinations.[1]

The New York State Education Department (NYSED) reports both unadjusted and adjusted growth scores. Unadjusted growth scores include only prior achievement as a predictor variable while adjusted growth scores control for prior achievement and student characteristics as predictor variables.[2] Unadjusted scores are reported for informational purposes to educators and are used for institutional accountability in Grades 4—8 to calculate the Growth Index.

The Growth Index combines results from three years of growth models that yield growth scores for students, which are known as Student Growth Percentiles (SGP). In 2017/18, these growth models were implemented in Grades 4—8 ELA and mathematics and were based on assessing each student's change in performance between 2016/17 (and prior years) and 2017/18 on State assessments compared with students who have similar prior performance. For more information about how growth is used for institutional accountability purposes, see *Measuring Student Growth for Institutional Accountability in New York* (available here: http://www.p12.nysed.gov/accountability/essa.html).

Content and Organization of This Report

The results presented in this report are based on the 2017/18 student growth model and Growth Index results, with some comparison to prior-year results. This technical report contains four 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 more information about the 2015/16 and 2016/17 models, which also contribute to the 2017/18 Growth Index, see the Growth Model for Educator Evaluation 2015/16 Technical Report and the Growth Model for Educator Evaluation 2016/17 Technical Report.
[2] Details can be found in the *2017/18 Growth Model for Educator Evaluation Technical Report*, which is available on the NYSED Growth Measures Toolkits page.

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 in more detail the data used for model estimation.

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 schools of students in Grades 4—8. The Grades 4—8 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 Grades 3—8 State assessments 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. In 2017/18, the Department conducted a standards review process because the Grades 3—8 ELA and mathematics assessments were administered over the course of two days rather than over the course of three days as in previous years. Due to the State's new two-session test design and performance standards, the 2018 Grades 3—8 ELA and mathematics results are not directly comparable to prior-year results. While test scores cannot be compared to prior year scores, growth results may still be computed.

The New York Grades 4—8 institutional accountability growth model uses test scores in each subject area as a predictor for that area. Specifically, New York's Grades 4—8 institutional accountability growth model includes up to three prior year test scores (depending on the grade) 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. Two examples of how students would not have growth scores computed for them are:

1. Students without a prior-year test score (e.g., a 6th grade student with a valid 6th grade ELA test score in 2017/18 who did not have a valid ELA test score in 2016/17); or
2. Students with a prior-year test score for the same grade as the current year test score (e.g., a 6th grade student with a valid 6th grade ELA test score in 2017/18 who also had a 6th grade ELA test score in 2016/17).

Where applicable, missing data indicators were used for missing second and third year prior scores. These missing indicator variables allow the models 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 2017/18 analyses, data from 2017/18 were used as outcomes, with prior achievement predictors coming from the previous three years (going back to 2014/15). The specific tests used as predictors vary by grade and subject and are as follows (see also 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 up to three prior grade 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., 2017/18 Grade 6 students must have had a Grade 5 score in the same subject from 2016/17).

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 accountability 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 accountability 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 2017/18 test scores.

School Attribution

For the New York Grades 4—8 institutional accountability growth model, students were attributed to schools if they were continuously enrolled (i.e., enrolled in the same school on BEDS day and at the beginning of the State test administration in the spring). Table 2 shows attribution rates for schools for the 2017/18 model.

Table 2. Grades 4—8 School-Student Attribution Rates (2017/18 Single Year Model)

Grade	Valid Student Records	Valid Student Records Attributed to at Least One School	Attribution Rate
4	313,702	305,017	97%
5	307,109	299,566	98%
6	288,144	280,634	97%
7	273,963	267,436	98%
8	222,982	217,908	98%
Total	1,405,900	1,370,561	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.

More student records overall were attributed to schools in 2017/18 than in 2016/17 or 2015/16, but the attribution rate in 2017/18 (97%) was the same as in 2016/17 and 2015/16.

Model

This section describes the statistical model used to measure student growth between two points in time on a single subject of a State assessment. New York's student growth model is run separately at the end of each school year and the resulting SGPs from the three most recent years are aggregated to create the Institutional Accountability Growth Index. This section describes the model used to create SGPs in the 2015/16, 2016/17, and 2017/18 school years and begins with a description of the statistical model used to form the comparison point against which students are measured, based on similar students. It then describes how SGPs are derived from the comparison point, followed by how the Growth Index is produced from these three years of SGPs.

At the core of the New York State institutional accountability growth model is the production of an SGP. This statistic characterizes each student's current year score relative to other students with similar prior test score histories. For example, an SGP equal to 75 denotes that a student's current year growth 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.

One common approach to estimating SGPs is to use a quantile regression model (Betebenner, 2009). This approach models the current year score as a function of prior test scores and finds the SGP by comparing the current year score to the predicted values at various quantiles of the conditional distribution.

The methods described here do not rely on the quantile regression method for two reasons. First, the typical implementation of the quantile regression makes no correction for measurement variance in the predictor variables or the outcome variable. Ignoring the measurement variance in the predictor variables yields bias in the model coefficients (e.g., Wei and Carroll, 2009). Further complicating the issue, the measurement variance in the outcome variable also adds to the bias in a quantile regression (Hausman, 2001), an issue that does not occur with linear regression.

A linear regression model is used to compute the SGPs each year for New York's growth model and is 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 growth 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:

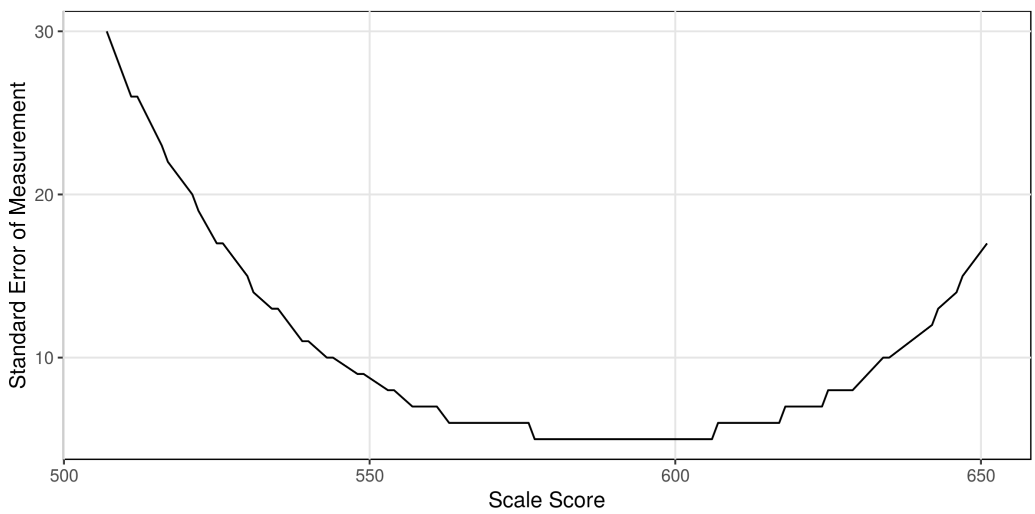
$$y_{ti} = \sum_{r=1}^{L} y_{t-r,i}\gamma_{t-r} + e_i$$

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 2017/18 Grade 8 ELA test in New York.

Figure 1. Conditional Standard Error of Measurement Plot (Grade 8 ELA, 2017/18)



Treating the observed scores as if they were the true scores introduces a bias in the regression that 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 the Grades 4—8 Growth Model

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 2015/16, 2016/17, and 2017/18 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 $f(x) \sim N(\hat{y}_i, \sigma^2_{yfi})$, although this is readily accomplished using the cumulative normal distribution function, $\Phi(\cdot)$.

Figure 2. Sample Growth Percentile from Model

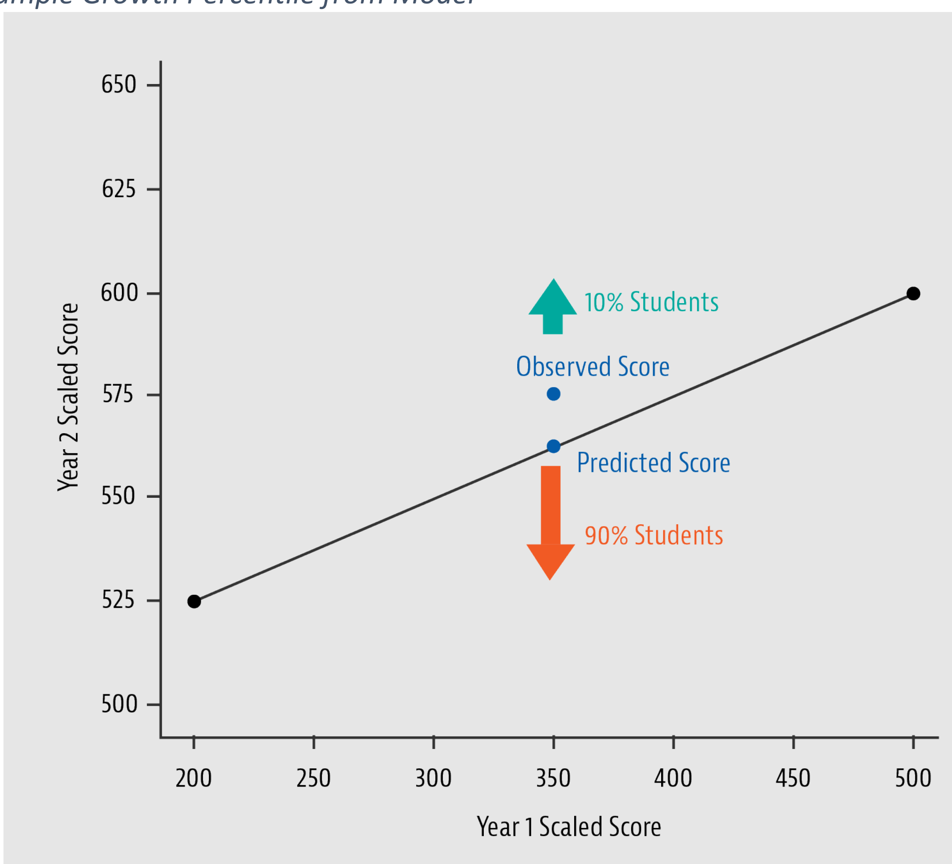
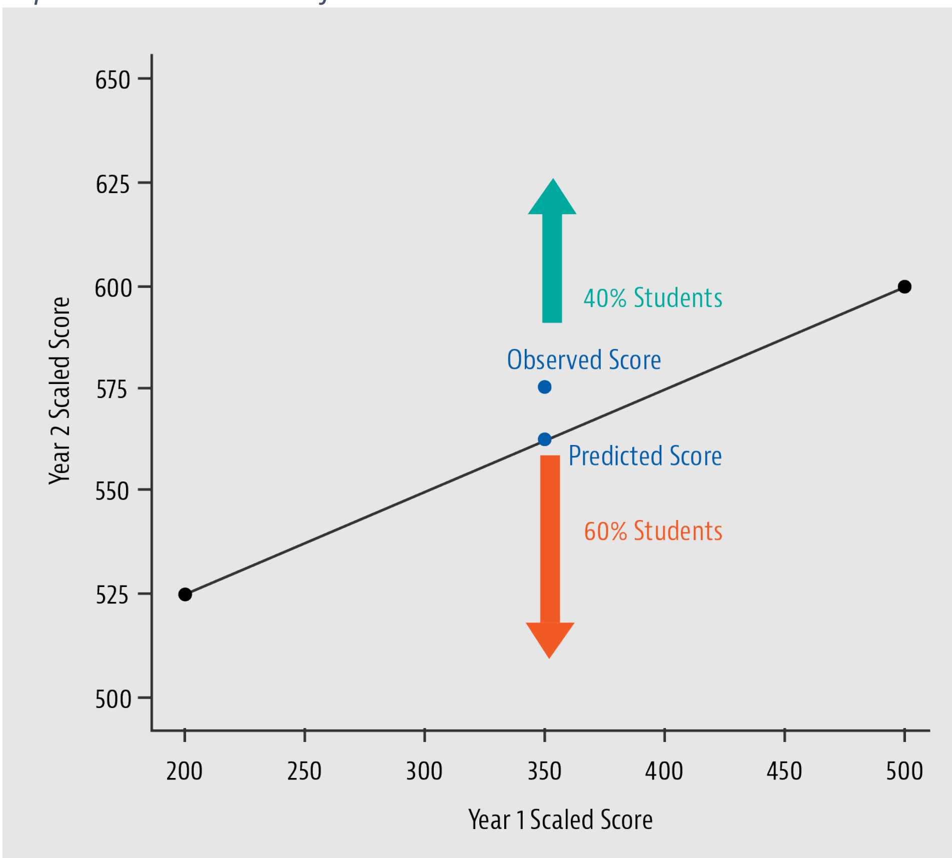


Figure 3 illustrates the same hypothetical student shown in Figure 2. Note that the observed score and predicted value are identical. However, the prediction variance is larger than in Figure 2. As a result, when we integrate over the normal from $-\infty$ to y_i, the SGP is 60, not 90 as in the previous 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



Combining SGPs Across Years to Generate a Growth Index

Once SGPs are estimated for each student, school- and district-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, or mean growth percentile (MGP). For accountability purposes, this three-year MGP is referred to as the Growth Index.

For NYSED's 2017/18 institutional accountability model, SGPs from 2015/16, 2016/17, and 2017/18 were combined to create a Growth Index for each accountability subgroup for public schools, charter schools, and districts.[3] To do so, three years of SGPs for continuously enrolled

[3] To be included in the Growth Index, a student must be continuously enrolled in a public school, charter school, or district that was open during the 2017/18 school year.

students were combined and a mean of SGPs was calculated for each accountability subgroup and for each school and district.

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 for each year within group j.

Then the Growth Index for unit j is produced as the simple mean

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

Many schools serve students from different grades and with results from different tested subjects. Because the SGPs are expressed as percentiles, they are free from scale-specific inferences and can be combined. Therefore, for the Growth Index, all SGPs of relevant students are pooled and the mean of the pooled SGPs is calculated.

A Growth Index is calculated for the All Students group and each of the accountability subgroups for which the count of SGPs is greater than or equal to 30 for the three year period: American Indian or Alaska Native, Black of African American, Hispanic or Latino, Asian/Pacific Islander, White, Multiracial, Students with Disabilities, English Language Learners, and Economically Disadvantaged.[4] The Growth Index is then rounded to the nearest tenth decimal place and assigned to one of four Levels based on the cut points described in Table 3.

Table 3. Growth Index to Growth Level

Growth Index	Growth Level
45 or less	1
45.1 to 50	2
50.1 to 54	3
Greater than 54	4

[4] When calculating the Growth Index for English language learners (ELLs), if there are 30 or more SGPs for the subgroup across the three years, then former ELLs are included in the Growth Index as the number of SGPs for former ELLs is less than half the number of ELLs in the current year. Former ELLs are students that were reported in at least one of the two previous reporting years but not in the current reporting year with a disability program service code. Similarly, if there are 30 or more SGPs for the Students with Disabilities subgroup, then former students with disabilities are included in the Growth Index. Former students with disabilities are students that were reported in at least one of the two previous reporting years but not in the current reporting year with a disability program service code.

Reporting

Institutional accountability growth results are provided for all students and disaggregated by subgroup as well as the student roster file are provided to Districts and charter schools.

Institutional Accountability Reports

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

- **Sum of SGPs –** The sum of the SGP results in ELA and in math for 2015/16, 2016/17, and 2017/18.
- **Number of Student Scores** – The number of SGP results in ELA and in math for continuously enrolled students for 2015/16, 2016/17, and 2017/18
- **Growth Index**– The mean of the SGPs for students attributed to the subgroup.
- **Growth Level** – The Growth Level associated with the reported Growth Index.

Results are presented at an aggregate level for the district and its schools and students separately.

Minimum Sample Sizes for Reporting

To report a Growth Index, a subgroup requires at least 30 student scores across three years. Table *4* presents the percentage of schools and districts that had at least one student attributed between 2015/16 and 2017/18 and the number/percent included in the institutional accountability model after applying the $n \geq 30$ rule.

Table 4. *Grades 4-8 Reporting Rates (2017/18 Growth Index)*

Measure	Level	Number with at Least One Student Attributed	Number Meeting the Minimum Sample Size Requirement	Percentage Meeting the Minimum Sample Size Requirement
2017/18 Growth Index (3-year MGP)	School	3,598	3,491	97%
	District	718	712	99%

Results

This section provides an overview of the results for the single-year 2017/18 growth model estimation followed by the three-year Growth Index results. A pseudo R-squared statistic and summary statistics characterizing the SGPs, MGPs, and their precision provide an overview of model fit.

Single-Year 2017/18 Growth Model

The Growth Index incorporates three years of growth scores from single-year models that are estimated separately for each grade and subject. In addition to being combined with 2015-16 and 2016-17 results to contribute to the 2017-18 Growth Index, results of the 2017-18 single-year model were shared with schools and districts to show a one-year MGP for informational purposes only. The following results are related to the single-year 2017-18 growth model.[5]

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)[6] model, not a least squares regression, we refer to this as a *pseudo R*-square. Table 5 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 5. Grades 4—8 Unadjusted Model Pseudo R-Squared Values by Grade and Subject (2017/18 Single Year Model)

Grade	ELA	Mathematics
4	0.61	0.67
5	0.66	0.73
6	0.68	0.73
7	0.70	0.74
8	0.68	0.65

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 and other measured characteristics. A student's SGP should not be expected to be higher or lower based on his or her prior-year score. Table 6 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

[5] For more information about the 2015/16 and 2016/17 models, see the Growth Model for Educator Evaluation 2015/16 Technical Report and the Growth Model for Educator Evaluation 2016/17 Technical Report.
[6] EiV regression is a method to estimate consistent coefficients when variables are measured with error, such as assessment scores. EiV regression allows us to acknowledge and account for that error.

test scores, the prior-year test score is a poor predictor of current year SGPs. As shown in Table 6, prior-year test scores explain about 2% to 5% of the variation in SGPs. Because SGPs are intended to allow students to show low or high growth no matter their prior performance, this result is as expected.

Table 6. Grades 4—8 Unadjusted Model Correlation Between SGP and Prior-Year Scale Score (2017/18 Single Year Model)

Grade	ELA	Mathematics
4	-0.154	-0.172
5	-0.154	-0.161
6	-0.150	-0.184
7	-0.140	-0.206
8	-0.136	-0.263

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:

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

where $\bar{\sigma}$ is the mean standard error of the MGP, and $sd\left(\hat{\theta}_j\right)$ 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 7 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. These results are based upon the one-year MGPs for the 2017/18 model. The values shown below are very similar to what was reported for the 2015/16 and 2016/17 models.

Table 7. Grades 4—8 Weighted Unadjusted Model Mean Standard Errors, Standard Deviation, and Value of ρ by Grade for Schools, Weighted by Number of SGPs (2017/18 Single Year Model)

Grade	Weighted Mean Standard Error	Weighted Standard Deviation	Weighted Reliability Statistic (ρ)
4	2.175	7.837	0.910
5	2.273	8.496	0.922
6	2.237	7.423	0.899

Grade	Weighted Mean Standard Error	Weighted Standard Deviation	Weighted Reliability Statistic (ρ)
7	1.909	7.828	0.930
8	1.822	8.039	0.940

Table 8 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 (i.e., 2.5% of schools would be expected to be above or below the mean by chance alone).

Table 8. Grades 4—8 Unadjusted Model School Combined MGPs Above or Below the Mean at a 95% Confidence Level (2017/18 Single-Year Model)

	Below Mean		Above Mean	
Grade	N	%	N	%
4	777	32%	627	26%
5	597	26%	570	24%
6	400	24%	505	31%
7	408	28%	457	31%
8	323	22%	420	29%

Growth Index

The following provides results of the Growth Index, which includes growth models for 2015-16, 2016-17, and 2017-18.

Neutrality of Growth Index

It is helpful to consider the relationship between the Growth Index and school characteristics, to identify any relationships that suggest non-neutrality. The scatter plots in Figures 4 through 8 provide a visual representation of the correlation between the school Growth Index and five school characteristics: the percent of students who are ELL, the percent of SWD, the percent of students in poverty or with economic disadvantage (ED), and the mean prior ELA or mathematics score of the students.

Figure 4. Growth Index Scores by Percentage of ELL Students in School

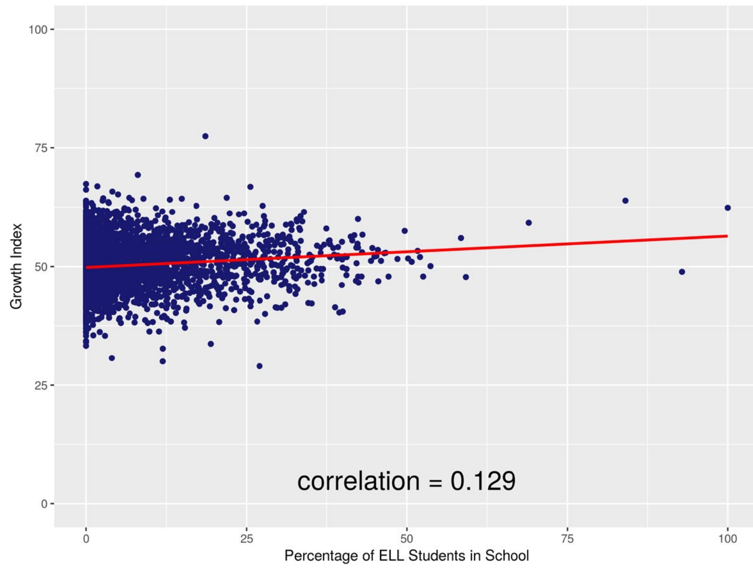


Figure 5. Growth Index Scores by Percentage of SWD Students in School

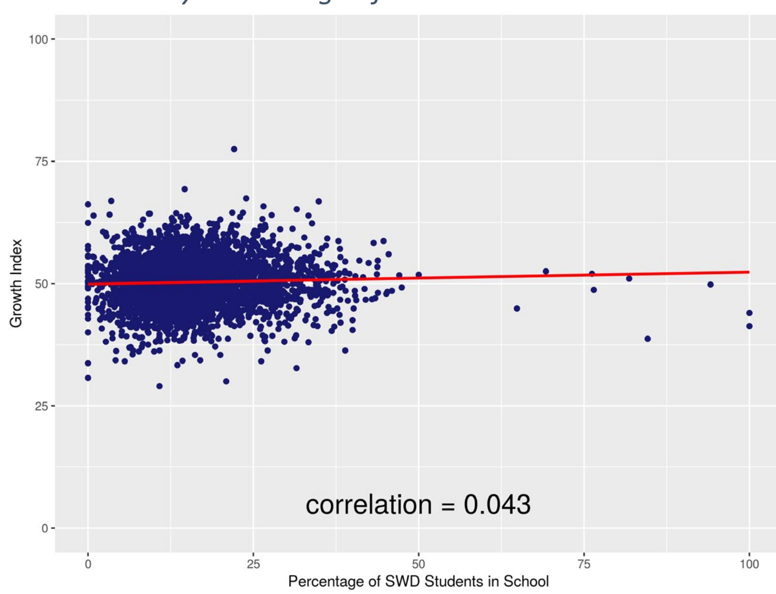
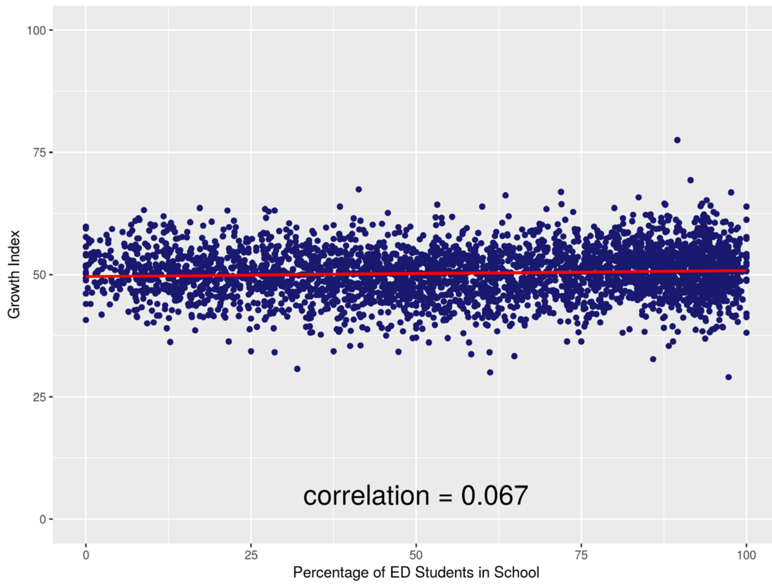


Figure 6. Growth Index Scores by Percentage of ED Students in School



Figures 7 and 8 show the relationship between the Growth Index and the Z-score of student prior achievement. A Z-score represents the number of standard deviations above or below the mean. Since the assessment scales are not designed to be averaged directly across grades or testing regimes, the Z-score provides a way to represent multiple grades and years of test scores together.

Figure 7. Growth Index Scores by Mean Prior ELA Z-Score Students in School

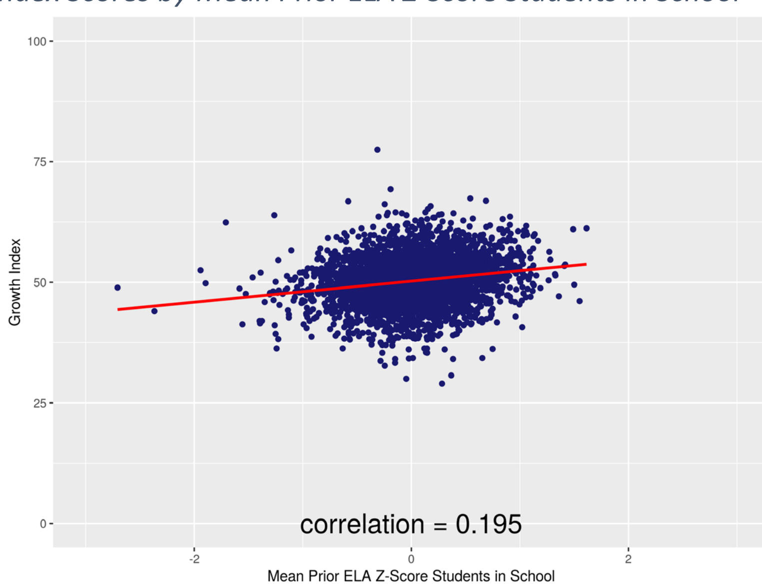
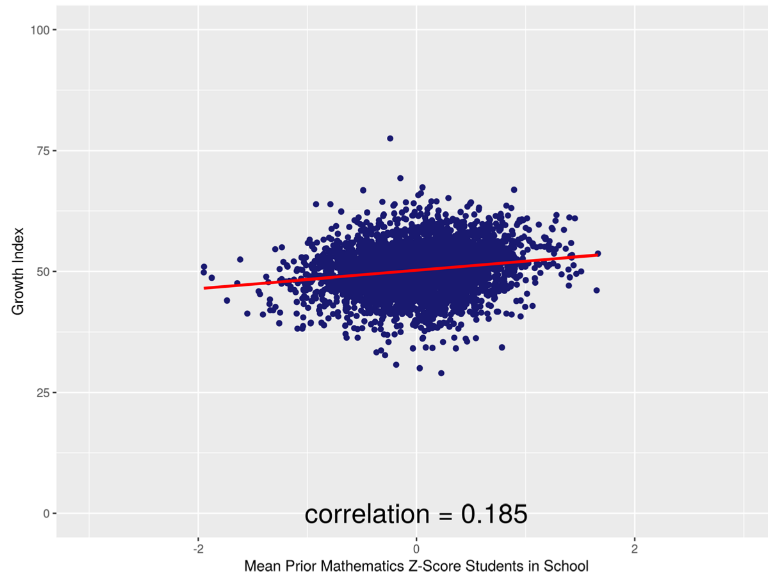


Figure 8. Growth Index Scores by Mean Prior Math Z-Score Students in School



The scatter plots above show that the Growth Index has a low to moderate correlation with respect to school demographic and pretest characteristics. The low correlation means that the Growth Index can be considered to be neutral with respect to these school characteristics and this neutrality means that schools can demonstrate growth, regardless of the academic starting point or characteristics of their students.

Growth Levels

As noted above, for accountability purposes, the Growth Index is translated to a Growth Level. Table 9**Error! Reference source not found.Error! Reference source not found.**describes the observed distribution of Growth Levels for schools and districts for the All Students subgroup based on their 2017/18 Growth Index score.

Table 9. School and District Level Distributions

Output Level	Level 1	Level 2	Level 3	Level 4
School	14%	34%	31%	22%
District	10%	51%	32%	7%

Note: Because of rounding, percentages may not add to 100 percent.

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 0.505. 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	444.361	0.329	0.000
Prior-Grade ELA Scale Score	0.505	0.001	0.000

Table A 2. Grade 5 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	437.257	0.317	0.000
Prior-Grade ELA Scale Score	0.381	0.002	0.000
Two-Grades-Prior ELA Scale Score	0.150	0.002	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	45.430	0.670	0.000

Table A 3. Grade 6 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	446.873	0.331	0.000
Prior-Grade ELA Scale Score	0.335	0.002	0.000
Two-Grades-Prior ELA Scale Score	0.119	0.003	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	35.100	0.838	0.000
Three-Grades-Prior ELA Scale Score	0.054	0.002	0.000
Missing Flag: Three-Grades-Prior ELA Scale Score	15.864	0.609	0.000

Table A 4. Grade 7 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	448.533	0.294	0.000
Prior-Grade ELA Scale Score	0.383	0.002	0.000
Two-Grades-Prior ELA Scale Score	0.080	0.002	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	22.520	0.654	0.000
Three-Grades-Prior ELA Scale Score	0.045	0.002	0.000
Missing Flag: Three-Grades-Prior ELA Scale Score	13.132	0.605	0.000

Table A 5. Grade 8 ELA Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	433.378	0.340	0.000
Prior-Grade ELA Scale Score	0.414	0.003	0.000
Two-Grades-Prior ELA Scale Score	0.097	0.003	0.000
Missing Flag: Two-Grades-Prior ELA Scale Score	27.374	0.793	0.000
Three-Grades-Prior ELA Scale Score	0.034	0.002	0.000
Missing Flag: Three-Grades-Prior ELA Scale Score	10.116	0.622	0.000

Table A 6. Grade 4 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	458.903	0.265	0.000
Prior-Grade Mathematics Scale Score	0.457	0.001	0.000

Table A 7. Grade 5 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	455.280	0.247	0.000
Prior-Grade Mathematics Scale Score	0.356	0.002	0.000
Two-Grades-Prior Mathematics Scale Score	0.117	0.002	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	35.605	0.510	0.000

Table A 8. Grade 6 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	441.461	0.291	0.000
Prior-Grade Mathematics Scale Score	0.368	0.002	0.000
Two-Grades-Prior Mathematics Scale Score	0.082	0.002	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	24.685	0.645	0.000
Three-Grades-Prior Mathematics Scale Score	0.065	0.002	0.000
Missing Flag: Three-Grades-Prior Mathematics Scale Score	19.917	0.582	0.000

Table A 9. Grade 7 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	457.200	0.270	0.000
Prior-Grade Mathematics Scale Score	0.380	0.002	0.000
Two-Grades-Prior Mathematics Scale Score	0.053	0.002	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	16.024	0.678	0.000
Three-Grades-Prior Mathematics Scale Score	0.034	0.002	0.000
Missing Flag: Three-Grades-Prior Mathematics Scale Score	10.098	0.572	0.000

Table A 10. Grade 8 Mathematics Unadjusted Model Coefficients

Effect Name	Effect	Standard Error	p-value
Constant Term	431.316	0.439	0.000
Prior-Grade Mathematics Scale Score	0.502	0.004	0.000
Two-Grades-Prior Mathematics Scale Score	0.045	0.004	0.000
Missing Flag: Two-Grades-Prior Mathematics Scale Score	12.927	1.034	0.000
Three-Grades-Prior Mathematics Scale Score	0.014	0.003	0.000
Missing Flag: Three-Grades-Prior Mathematics Scale Score	4.265	0.756	0.000