CORE Growth Model Technical Considerations

CONTENTS

INTRODUCTION

This report describes technical aspects of the CORE growth model of 2015-16 estimated by Education Analytics. The report describes the statistical model and methods used to measure growth in the CORE districts. It also includes the parameters of the growth model and measures of the reliability of the growth estimates.

MODEL FRAMEWORK

The growth model used in CORE can be expressed statistically using the following equations:

Student achievement: $y_{1ij} = \zeta + \lambda y_{0ij} + \lambda^{alt} y_{0ij}^{alt} + X_i\beta + Z_j\gamma + \alpha_j + \varepsilon_{ij}$ (1)

Posttest measurement error: $Y_{1i} \ = y_{1i} \ + v_{1ij}$ (2)

Same-subject pretest measurement error: $Y_{0i} \ = y_{0ij} + v_{0ij}$ (3)

Other-subject pretest measurement error: $Y_{0ij}^{alt} = y_{0ij}^{alt} + v_{0ij}^{alt}$ (4)

where:

- y_{1ij} is true posttest achievement by student *i* in school *j*;
- y_{0ij} is true pretest achievement by student *i* in school *j* in the same subject as the posttest;
- y_{0i}^{alt} is true pretest achievement by student *i* in school *j* in a different subject (math in models of English language arts achievement, and vice versa) from the posttest;
- X_i is a vector of characteristics of student *i*;
- Z_j is a vector of characteristics of school *j*;
- α_j is the effect of school *j*;
- ε_{ij} is the unexplained component of true posttest achievement of student *i* in school *j*;
- Y_{1ij}, Y_{0i} , and Y_{0i}^{alt} are measured posttest, same-subject pretest, and other-subject pretest achievement for student *i* in school *j*; and
- v_{1ij}, v_{0i} , and v_{0i}^{alt} are measurement error in posttest, same-subject pretest, and other-subject pretest achievement for student *i* in school *j*.

Equation (1) models current student achievement as a linear function of prior student achievement, student characteristics, school characteristics, and school assignment. Equations (2) through (4) model the measurement error in the pretest. Substituting equations (2) through (4) into equation (1) yields the following model of measured student achievement:

$$\text{Measured achievement: } Y_{1ij} = \zeta + \lambda Y_{0ij} + \lambda^{alt} Y_{0i}^{alt} + X_i\beta + Z_j\gamma + \alpha_j + e_{ij} \quad (5)$$

where the error term e_{ij} is equal to:

$$\text{Residual of measured achievement: } e_{ij} = \varepsilon_{ij} + v_{1ij} - \lambda v_{0ij} - \lambda^{alt} v_{0ij}^{alt} \quad (6)$$

This error term only includes not only the unexplained component ε_{ij} of true posttest achievement y_{1ij}, but also the measurement error components v_{1ij}, v_{0ij}, and v_{0ij}^{alt} of the measured test scores Y_{1ij}, Y_{0ij}, and Y_{0i}^{alt}.

VARIABLES INCLUDED IN THE GROWTH MODEL

The posttest variable in the growth model is the Smarter Balanced (SBAC) assessment in math or English language arts. This is the outcome variable Y_{1ij} described in the modeling equations above. The model is estimated separately by subject and by the grade of the student at the time of the posttest, covering grades 4 through 8 and grade 11. In models in which the posttest is administered in grades 4 through 8, the pretests are SBAC assessments in math and English language arts administered in the previous grade. In models in which the posttest is administered in grade 11, the pretests are CAHSEE assessments in math and reading administered in grade 10. All models estimated include both math and reading/ELA pretests. These are the pretest variables Y_{0i} and Y_{0ij}^{alt} described in the modeling equations above.

All models also control for disability, English language learner, economic disadvantage, foster care, and homelessness at the student level. These are the variables that make up the vector X_i in the modeling equations above. English language learner status enters the model as four indicator variables: one for beginning and early intermediate level (ELD levels 1 and 2); another for intermediate level (ELD level 3); a third for early advanced and advanced levels (ELS levels 4 and 5), and a fourth for English language learners whose level is unavailable. Disability enters the model as two indicators: one for students with moderate disabilities (defined as learning disability and speech and language disability) and another for students with more severe disabilities (defined as all other types of disability). Economic disadvantage, foster care, and homeless all enter the model as single binary indicator variables.

The model also controls for the means by school of the pretests Y_{0ij} and Y_{0ij}^{alt} and the student characteristics X_i. These are the only school-level variables in the model, and are the school-level variables Z_j in the above modeling equations.

The Sample

The sample for any given model is made up of only students with measured scores for the posttest and both pretests. It is also limited to students who are continuously enrolled in a single school in the posttest year, with continuous enrollment defined as student enrollment from Fall Census Day (first Wednesday in October) to the first day of testing without a gap in enrollment of more than thirty consecutive days.

Estimating the Growth Model

Equation (5) includes student-level variables, school-level variables, and individual school effects among its right-hand-side variables. To accommodate this combination of variables, equation (5) is split into two separate equations:

Student-level: $Y_{1ij} = \zeta^{*} + \lambda Y_{0ij} + \lambda^{alt} Y_{0i}^{alt} + X_i\beta + \alpha_j^{*} + e_{ij}$ (5')

School-level: $\alpha_j^{*} = \zeta^{**} + Z_j\gamma + \alpha_j$ (5")

where the sum of the intercepts ζ^{*} and ζ^{**} is equal to the overall intercept ζ.

Equations (5') and (5") can be estimated in sequence to produce consistent estimates of the parameters in equation (5). First, equation (5') can be estimated as an errors-in-variables regression of measured posttest Y_{1ij} on measured pretests Y_{0i} and Y_{0ij}^{alt}, student characteristics X_i, and fixed school effects. This regression is estimated over a data set in which the student is the unit of observation. Second, equation (5") can be estimated as an ordinary least squares regression, with the fixed school effects estimates of $\hat{\alpha}_j^{*}$ from the previous regression as its left-hand-side variable and a vector of school characteristics Z_j as its right-hand-side variable. This regression can be estimated over a data set in which the school is the unit of observation, using the number of students in the school N_j as a weight. The residuals from this second regression are estimates of the school effects α_j. The standard error of these estimates (dubbed $\hat{\alpha}_j$) can be estimated by dividing an estimate of the variance of the error term e_{ij} across all students by the number of students in the school N_j. In practice, the model was estimated using a slightly different set of steps that yields identical results to the method described above.

Errors-in-Variables Regression

If the equation (5') is estimated using ordinary least squares, the coefficient estimates will be biased because some of its right-hand-side variables, specifically the pretests, are measured with error (Meyer, 1996; Meyer, 1999). As a result, an errors-in-variables model described in

Fuller (1987) is employed to produce consistent estimates of the coefficients and fixed effects in (5'). An errors-in-variables model uses estimates of the variance of measurement error to account for measurement error in right-hand-side variables. In models in which the posttests are administered in grades 4 through 8, the variance of measurement error for the SBAC pretests is computed as the average of the squared standard errors of measurement (SEMs) of the pretests across the regression sample. In models in which the posttests are administered in grade 11, both of the CAHSEE pretests are assumed to have a reliability of 0.9, which implies that the variance of pretest measurement error is equal to 0.1 times the variance of the pretest.

SHRINKAGE

The school fixed effect estimates $\hat{\alpha}_j$ from the estimated regression (5") are consistent estimates of the effects of individual schools on student achievement, controlling for prior achievement, student characteristics, and school characteristics. However, it will frequently be the case that smaller schools will be overrepresented among the highest and lowest estimated values of $\hat{\alpha}_j$. This is because the estimates for smaller schools are based on the growth of a smaller number of students and, as a result, will be more likely to be very high or very low as a result of randomness.

To account for this, an Empirical Bayes shrinkage approach was applied to the $\hat{\alpha}_j$ estimates. This shrinkage approach produces a prediction of the school effect $\tilde{\alpha}_j$ that minimizes expected mean squared error given both the fixed-effects estimate $\hat{\alpha}_j$ and the distribution across all schools of the school effects α_j. The practical effect of this shrinkage approach is to "shrink" the growth estimates of schools with smaller numbers of students toward the average school effect. The specific shrinkage approach used is a simple univariate formula for a variable with a mean of zero, $\tilde{\alpha}_j = [\hat{\omega}^2/(\hat{\omega}^2 + \hat{\sigma}_j^2)]\hat{\alpha}_j$, where $\hat{\omega}^2$ is an estimate of the variance of α_j across schools and $\hat{\sigma}_j^2$ is the square of the estimated standard error of $\hat{\alpha}_j$. The variance estimate $\hat{\omega}^2$ is computed by measuring the difference between the variance of the school fixed effects estimates $\hat{\alpha}_j$ and the mean of the squared standard errors $\hat{\sigma}_j^2$ across schools.

STANDARDIZATION

Both the unshrunk and shrunk growth measures are standardized by dividing them by the square root of the variance estimate, $\hat{\omega}^2$. This standardization puts school-level growth measure on a distribution where the mean school effect is zero and a unit is equal to a standard deviation of true growth α_j across schools. In practice, this puts the vast majority of growth measures on a scale between -2 and +2, with the growth measure of an average school equal to zero. We refer to this measure of growth as "tiered" growth.

A second standardization converted tiered, shrunk growth measures to percentile equivalents using the inverse standard normal cumulative distribution function. These percentile equivalents are equal to what the percentile rank of the growth measure would be assuming that school growth is normally distributed within grade and subject.

DIFFERENTIAL EFFECTS

Growth measures for student subgroups within schools are also produced in addition to the overall growth measures described above. Subgroup measures are produced by race (Asian, Black, Filipino, Hispanic, Pacific Islander, Multiracial, Native American, White, and race missing), English language learner (ELL, not ELL), disability (with disability, without disability) and economic disadvantage (economically disadvantaged, not economically disadvantaged). The analysis that produces subgroup growth measures is conducted separately for each class of subgroups. In other words, the subgroup analysis conducted for the set of race subgroups is separate from the analysis conducted for the set of disability subgroups.

To produce the subgroup growth measures, we first produce individual growth measures equal to the sum of the unshrunk school growth estimate $\hat{\alpha}_j$ and the student-level residual estimate $\hat{\varepsilon}_{ij}$. These individual growth measures are averaged across students by school and subgroup to produce an unshrunk growth measure $\hat{\alpha}_{sj}$ for subgroup *s* within school *j*. Finally, a multivariate shrinkage approach that takes into account the correlation in growth across subgroups within schools is employed to produce shrunk growth measures $\tilde{\alpha}_{sj}$ for subgroup *s* within school *j*.

AGGREGATION

The steps described above yield results at the subject and grade level for each school. To compute overall, school-level measures, the unshrunk, tiered growth measures for each school and subject were averaged across grades to the elementary, middle-school, and high-school levels. These results, like the results for individual grades, were then shrunk using Empirical Bayes shrinkage, normalized to "tiered" growth measures, and converted to the percentile distribution using the inverse standard normal cumulative distribution function.

Subgroup results were normalized using analogous steps. Unshrunk, tiered growth measures for each school, subject, and subgroup were averaged across grades. These averaged results were shrunk using a multivariate shrinkage approach that took into account the correlation of growth across subgroups within schools. The shrunk results were then tiered and converted to the percentile distribution in the same way as the overall, non-subgroup results.

THREE METHODS

The growth model is measured using three different approaches, referred to as Models A, B, and C:

- Model A does not include controls for the student characteristics in X_i or the school characteristics in Z_j. The only controls in Model A are the previous year's achievement in math and ELA.
- Model B includes the controls for the student characteristics in X_i but not the classroom characteristics in Z_j.
- Model C includes both X_i and Z_j in the specification.

Note that $\alpha_j^* = \alpha_j$ in all approaches other than Model C, which is the only approach where the school-level equation (5") is necessary. The differences among the models can be illustrated using the following table:

	Includes pretests to control for prior achievement	**Includes student characteristics other than pretests**	**Includes school averages of pretests and student characteristics**
Model A	X		
Model B	X	X	
Model C	X	X	X

We calculated correlation results between all three models. Models A and B had the highest correlations. Both of these models' results also had high correlations with Model C, so the consistency of model results holds between model specifications.

Subject	**Grade**	**Correlation A-B**	**Correlation A-C**	**Correlation B-C**
ELA	4	0.997	0.969	0.980
ELA	5	0.997	0.987	0.994
ELA	6	0.998	0.981	0.984
ELA	7	0.998	0.926	0.935
ELA	8	0.999	0.972	0.975
ELA	11	1.000	0.909	0.914
ELA	**Overall**	**0.998**	**0.958**	**0.964**

Subject	Grade	Correlation A-B	Correlation A-C	Correlation B-C
Math	4	0.999	0.961	0.972
Math	5	1.000	0.961	0.966
Math	6	0.999	0.977	0.977
Math	7	0.999	0.950	0.958
Math	8	0.999	0.977	0.978
Math	11	0.999	0.820	0.842
Math	**Overall**	**0.999**	**0.944**	**0.951**
Overall	**Overall**	**0.999**	**0.951**	**0.958**

Model Selection Criteria

The next section describes some technical metrics for growth models along with some preliminary recommendations for what is high and low quality. In discussion with the CORE technical advisory committee, the decision was made to use a neutral model to ensure that growth was not correlated with any of the controlled variables. As the table below shows, model c has the lowest neutrality correlations of any of the models. Since any of the models satisfies the other criteria laid out in this report (reliability, predictive power, etc.) model c was chosen as the CORE growth model because of its neutrality properties.

Model Neutrality

We calculate correlations between growth estimates and school-level pretest/demographic covariates. This is a method for validating whether the variables we include on the right-hand side of our regression adequately control for school-level factors influencing growth percentile estimates. These correlations we deem “model neutralities”; the higher the correlation magnitude, the higher the level of “non-neutrality” for that particular covariate. In our modelling specification we do not include racial or ethnic demographics, but we do correlate school estimates with them to ensure model neutrality with respect to these variables.

We also draw a distinction between errors of covariate inclusion and exclusion: what we label errors of omission and commission. An error of omission is the same as omitted variable bias in classic regression analysis. Estimation results are inconsistent due to correlation between omitted right-hand side regressors and the model error term. This error can occur in Model A for not having included student-level demographics, or in Model A or Model B for not having included school-wide averages of prior achievement or demographics. For instance, if we do not control for school-wide averages of prior achievement, we could be potentially ignoring peer effects, i.e., the effect that having higher-achieving peers may have on student achievement.

An error of commission, conversely, ascribes too much predictive power to included covariates. It essentially papers over true differences in school quality by ascribing these differences to included covariates. This error can occur in Model C because of the presence of school effects in the second level of the regression. For instance, if we control for school-wide averages of economic disadvantage, we could potentially be ignoring the possibility that schools with higher proportions of economically disadvantaged students produce lower student growth in test scores than schools with lower proportions of economically disadvantaged youth. This may be due to more inexperienced teachers being assigned to schools in low-income areas, resources available at the school, or for other reasons. Crucially, these true differences in student growth will not be evident from the results in Model C because they are covered up by the school averages in a model that implicitly assumes that school-level averages and teacher assignments have nothing to do with each other.

Correlation of Model Variables

Subject	School Composition Variable	Model A	Model B	Model C
ELA	Economic Disadvantage	-0.092	-0.082	-0.001
ELA	ELL	-0.075	-0.054	-0.002
ELA	Foster	-0.048	-0.051	-0.014
ELA	Homeless	-0.019	-0.013	0.001
ELA	SPED Moderate	-0.034	-0.016	0.006
ELA	SPED Severe	-0.031	-0.023	-0.007
ELA	Pre test same subject	0.070	0.075	-0.001
ELA	Pretest Math	0.085	0.092	-0.002

Subject	School Composition Variable	Model A	Model B	Model C
Math	Economic Disadvantage	-0.180	-0.152	-0.001
Math	ELL	-0.122	-0.105	-0.002
Math	Foster	-0.041	-0.040	-0.006
Math	Homeless	-0.033	-0.032	0.003
Math	SPED Moderate	-0.055	-0.043	0.004
Math	SPED Severe	-0.021	-0.018	-0.009
Math	Pre test same subject	0.137	0.120	0.000
Math	Pretest ELA	0.180	0.163	0.001

Correlation with Demographics not In Model

Subject	School Composition Variable	Model A	Model B	Model C
ELA	% Asian	0.06	0.071	-0.009
ELA	% Black	-0.095	-0.109	-0.102
ELA	% White	0.045	0.035	-0.020
ELA	% Hispanic	-0.028	-0.021	0.063
ELA	% Native	-0.005	-0.01	-0.005
ELA	% Islander	-0.048	-0.049	-0.059
ELA	% Filipino	0.057	0.054	0.020
ELA	% Multi	-0.005	-0.011	-0.049
ELA	% Missing	0.026	0.026	0.026

Subject	School Composition Variable	Model A	Model B	Model C
Math	% Asian	0.189	0.182	0.091
Math	% Black	-0.106	-0.107	-0.085
Math	% White	0.107	0.087	-0.023
Math	% Hispanic	-0.141	-0.12	0.004
Math	% Native	-0.028	-0.029	-0.025
Math	% Islander	-0.006	-0.006	-0.005
Math	% Filipino	0.097	0.085	0.025
Math	% Multi	0.018	0.002	-0.064
Math	% Missing	-0.013	-0.012	-0.011

Model Fit

The purpose of using growth measures at the school level is to isolate the impact of schools on student achievement from other, non-school factors. The motivation is that controlling for prior student achievement and demographics will better pinpoint the effects of schools. We expect effective control variables to be good predictors of the student achievement outcome.

The table below presents an R-squared measure that specifically addresses the extent to which prior achievement and demographics predict current student achievement. Specifically, the R-squared measure presented below is equal to the proportion of the variance of achievement across students in the same schools that can be explained by variance in the student-level control variables in the model. A within-school fit measure is employed to specifically isolate the explanatory power of the control variables in a way that does not include the explanatory power of the school effects.

EA suggests thresholds of between 0.5 and 0.85 on measures of predictive power. If predictive power is too low (<0.5), it may be the case that the pretests and demographic controls are not sufficiently controlling for non-school factors to measure the impacts of schools. If predictive power is too high (>0.85), then it may be the case that the pretests and demographics are so predictive of the posttest that the posttest does not reflect the impacts of schools on student achievement.

The table below presents within-school R-squared measures for each grade and subject. In general, the fit of the model is very good. For example, about 76 percent of the variance of fourth-grade math achievement within schools is explained by variance in third-grade math and reading achievement and student demographics. Of particular interest are the model fit statistics for the 11th grade models, which use different assessments for the pretest (CAHSEE) and the posttest (SBAC). While the fit is lower in these models than in the elementary grades, it is still quite good. For example, about two-thirds of within-school variance in the 11th grade SBAC in English language arts can be explained with variance in the 10th grade CAHSEE and student demographics.

Note that within-school R-squared is the same between Model B and Model C. This is because the only difference between Models B and C is that Model C includes controls for across-school variation, while Model B does not. Since within-school R-squared measures model fit within schools, the values between Models B and C will be identical.

Within School and Between School Variation

Subject Grade	Within R^2 Model A	Within R^2 Models B and C
SBAC ELA Grade 04 2016 Spring	0.712	0.716
SBAC ELA Grade 05 2016 Spring	0.742	0.745
SBAC ELA Grade 06 2016 Spring	0.729	0.733
SBAC ELA Grade 07 2016 Spring	0.741	0.742
SBAC ELA Grade 08 2016 Spring	0.761	0.762
SBAC ELA Grade 11 2016 Spring	0.651	0.652
SBAC Math Grade 04 2016 Spring	0.754	0.755
SBAC Math Grade 05 2016 Spring	0.762	0.763
SBAC Math Grade 06 2016 Spring	0.764	0.768
SBAC Math Grade 07 2016 Spring	0.816	0.816
SBAC Math Grade 08 2016 Spring	0.786	0.787
SBAC Math Grade 11 2016 Spring	0.716	0.717

Reliability of Growth Measures

We refer to model reliability as the proportion of variance in measured growth that is attributable to the underlying effects of schools rather than statistical noise. This definition is akin to estimation of a signal-to-noise ratio, or the proportion of true variance to total variance.

We suggest a bottom threshold for reliability of 0.5 or greater. This threshold can be framed as at least half of the variance in growth results representing true differences in school quality. A reliability lower than 0.5 implies that there is more statistical noise, arising from factors like small schools or unreliable assessments, than there is measurement of true school effects in our data. A high result may imply a model failure in conjunction with low predictive power as described in the earlier discussion of model fit. In this case the model may be very reliably measuring attainment rather than school impact.

We define school variance as the amount of true differentiation between schools. We set thresholds of standard deviations of less than 0.05 and greater than 0.25 for schools. Results lower than 0.05 may result from the poor alignment of assessments and course curricula. It may also be evidence that schools do not follow the recommended sequence of course topics. Results that exceed the 0.25 threshold suggest outsized school impacts that do not fall in line with the literature on school value-added.

The table below presents measures of the reliability and variance of the growth measures produced by the CORE growth model. The first four columns of numbers are all measured in units of standard deviations of the posttest. The first column, the variance of estimates, is the variance of the estimated unshrunk growth measures across schools. This includes variance in true growth across schools as well as variance from estimation error in the growth measures. The second column is a measure of the variance from estimation error in the growth measures, and is equal to the average across schools in the squared standard errors of the growth measures. The third column is an estimate of the variance in true growth effects across schools, and is the difference between the first and second columns. The fourth column is the square root of the third column. The fifth column, the reliability of the growth measures, is equal to the proportion of total variance in growth measures across schools that is attributable to true differences across schools rather than to estimation error. It is equal to the third column divided by the first column.

Reliabilities

Model Description	Variance of estimates	Noise variance	Estimate of variance	Estimate of std. deviation	Reliability
ELA 04 2016 a	0.028	0.003	0.025	0.157	0.877
ELA 04 2016 b	0.028	0.003	0.024	0.156	0.877
ELA 04 2016 c	0.026	0.003	0.023	0.152	0.872
ELA 05 2016 a	0.025	0.003	0.021	0.146	0.866
ELA 05 2016 b	0.024	0.003	0.021	0.146	0.867
ELA 05 2016 c	0.024	0.003	0.021	0.144	0.865
ELA 06 2016 a	0.035	0.002	0.033	0.181	0.935
ELA 06 2016 b	0.035	0.002	0.033	0.181	0.937
ELA 06 2016 c	0.034	0.002	0.031	0.177	0.935
ELA 07 2016 a	0.022	0.001	0.021	0.145	0.947
ELA 07 2016 b	0.022	0.001	0.021	0.144	0.946
ELA 07 2016 c	0.019	0.001	0.018	0.133	0.938
ELA 08 2016 a	0.016	0.001	0.015	0.123	0.929
ELA 08 2016 b	0.016	0.001	0.015	0.123	0.930
ELA 08 2016 c	0.015	0.001	0.014	0.119	0.926
ELA 11 2016 a	0.037	0.002	0.035	0.188	0.958
ELA 11 2016 b	0.036	0.002	0.035	0.187	0.957
ELA 11 2016 c	0.030	0.002	0.029	0.169	0.948
Math 04 2016 a	0.032	0.003	0.029	0.171	0.905
Math 04 2016 b	0.032	0.003	0.029	0.169	0.904
Math 04 2016 c	0.030	0.003	0.027	0.164	0.898
Math 05 2016 a	0.032	0.003	0.029	0.170	0.907
Math 05 2016 b	0.032	0.003	0.029	0.170	0.906
Math 05 2016 c	0.030	0.003	0.027	0.163	0.899
Math 06 2016 a	0.034	0.002	0.032	0.179	0.939
Math 06 2016 b	0.034	0.002	0.032	0.179	0.941
Math 06 2016 c	0.033	0.002	0.031	0.175	0.938
Math 07 2016 a	0.017	0.001	0.016	0.128	0.939
Math 07 2016 b	0.017	0.001	0.016	0.127	0.938
Math 07 2016 c	0.016	0.001	0.015	0.121	0.933
Math 08 2016 a	0.019	0.001	0.017	0.132	0.934
Math 08 2016 b	0.019	0.001	0.017	0.132	0.934
Math 08 2016 c	0.018	0.001	0.017	0.129	0.931
Math 11 2016 a	0.027	0.001	0.025	0.159	0.947
Math 11 2016 b	0.025	0.001	0.024	0.155	0.944
Math 11 2016 c	0.018	0.001	0.016	0.128	0.921

Stability

Model results change from year to year. These changes are due to differences in the real growth effects of schools, perhaps because of staff turnover, new policy implementation or other structural issues within schools and across CORE. Growth estimates will also change between years because of measurement issues surrounding assessment and student-level data. These issues affect statistical noise and will contribute to differences in model results, year over year.

We estimate school stability by correlating school growth estimates between years. Correlations above 0.85 we deem too stable. High correlation magnitudes signal possible modeling issues related to our estimates not detecting true changes in school effects or measuring something other than school impact. Alternatively, correlations of school growth that fall below 0.2 may be considered unstable between years. Low magnitudes suggest that our data comprise more noise than information (or that there is severe instability in school impact in that sample).

Because of the two-year gap in testing, and the change from the CST to the testing format, EA and CORE did not calculate school stability correlations. The long gap in student testing combined with the switch in scales and formats we believe undermine any useful information about appropriate model selection that school stability correlations could impart.

Student Coverage

Because we expect a large percentage of "testable" students to have pretest and posttest scores, we check this expectation using our student-level test data. To construct student (and, subsequently, school) growth estimates, we require students in the model to have both a pretest and a posttest score.

In our experience, we are typically unable to include between 6% and 10% of our sample due to missing pretests or posttests. Thus, our threshold for good student coverage is 90% of students in our data having both pretests and posttests. If fewer than 80% of students have both posttests and pretests, the possibility is greater that the results of the growth model are affected by selection bias. This has the potential to distort the results if higher- or lower-growth students are disproportionately selecting out of the sample (for example, by opting out) by not taking the pretest or posttest. The table below shows the match rate for each of the models CORE-wide.

Student Coverage Table

Subject	Grade	Match Rate
ELA	4	94.5%
ELA	5	94.9%
ELA	6	93.9%
ELA	7	93.6%
ELA	8	94.1%
ELA	11	93.6%
ELA	**overall**	**94.1%**
Math	4	94.1%
Math	5	94.6%
Math	6	93.6%
Math	7	93.2%
Math	8	93.8%
Math	11	93.5%
Math	**overall**	**93.8%**
overall	**overall**	**94.0%**

Summary of Model Statistics and Criteria

Below is a summary of the criteria for each metric. We have further subdivided the criteria into a "green" within standard band and a "yellow" on the fringe of low quality band. Models that fail a particular metric in the "red" band are highly suspect. Models that fall into several yellow zones may be low quality.

	Metric	Red	Yellow	Green	Yellow	Red
Model Neutrality						
Controlled Variables	Correlation	NA	NA	NA	NA	NA
Not Controlled Variables	Correlation	NA	NA	NA	NA	NA
Predictive Power	R^2	.00 to .50	.50 to .55	.55 to .75	.75 to .85	.85+
Signal to Noise Ratio	Reliability	.00 to .50	.50 to .60	.60 to .90*	.90 to .95	.95+
School Variation	SD	.00 to .05	.05 to .08	.08 to .15	.15 to .25	.25+
School Stability	Correlation	.00 to .20	.20 to .40	.40 to .75	.75 to .85	.85+
Student Coverage	Proportion	.00 to .80	.80 to .90	.90+	NA	NA

* Reliability in theory has no upper bound on quality (more is better) when other metrics are also green. If other metrics are not green it could indicate a model failure if too high.

Model Coefficients

The tables below present the estimated coefficients on the student- and school-level variables in growth Models B and C. The coefficients on the student-level variables were the same across Models B and C since estimating Model C involves estimating Model B as a first stage. The coefficients on the school-level variables are only relevant to Model C, since they only enter into Model C. In the notation of the growth model equations at the beginning of this report, these are estimates of the coefficients λ, λ^{alt}, β, and γ. All coefficients are measured in units of standard deviations of the posttest. All coefficients on pretest variables measure the effect of a one standard deviation increase in the pretest in units of standard deviations of the posttest. R-squared measures are from the first-stage, student-level regression and include the school fixed effects in the explained component.

	Math 4		Math 5		Math 6	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Student-level covariates:						
Math pretest	0.832	(0.004)	0.812	(0.004)	0.744	(0.005)
Reading/ELA pretest	0.064	(0.004)	0.105	(0.005)	0.161	(0.005)
ELL (ELD levels 1, 2)	-0.038	(0.006)	0.026	(0.007)	-0.115	(0.010)
ELL (ELD level 3)	-0.021	(0.005)	-0.012	(0.005)	-0.053	(0.006)
ELL (ELD levels 4, 5)	0.034	(0.005)	0.004	(0.005)	0.003	(0.005)
ELL (level unavailable)	0.040	(0.018)	0.066	(0.018)	0.015	(0.020)
Disability (moderate)	-0.029	(0.006)	-0.028	(0.006)	-0.162	(0.006)
Disability (severe)	-0.051	(0.009)	-0.065	(0.009)	-0.170	(0.009)
Economic disadvantage	-0.046	(0.005)	-0.026	(0.004)	-0.039	(0.005)
Foster care	0.018	(0.020)	-0.001	(0.021)	-0.081	(0.023)
Homeless	0.002	(0.010)	0.006	(0.009)	-0.020	(0.010)
School-level averages:						
Math pretest	-0.147	(0.036)	-0.208	(0.031)	-0.224	(0.055)
Reading/ELA pretest	0.105	(0.040)	0.184	(0.036)	0.241	(0.061)
ELL (ELD levels 1, 2)	-0.102	(0.084)	-0.113	(0.118)	0.494	(0.283)
ELL (ELD level 3)	0.045	(0.080)	0.146	(0.076)	0.227	(0.167)
ELL (ELD levels 4, 5)	-0.030	(0.069)	0.129	(0.074)	0.072	(0.111)
ELL (level unavailable)	-0.823	(0.311)	0.065	(0.313)	1.068	(0.517)
Disability (moderate)	-0.060	(0.112)	0.006	(0.111)	0.189	(0.188)
Disability (severe)	0.122	(0.170)	-0.199	(0.162)	-0.341	(0.248)
Economic disadvantage	-0.145	(0.037)	-0.172	(0.035)	-0.099	(0.051)
Foster care	-0.543	(0.461)	-0.146	(0.431)	-0.731	(0.731)
Homeless	0.054	(0.116)	-0.125	(0.101)	-0.398	(0.160)
R-squared	0.86		0.86		0.87	
No. of students	103,522		101,582		92,757	
No. of schools	1,319		1,325		802	

	Math 7		Math 8		Math 11	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Student-level covariates:						
Math pretest	0.940	(0.005)	0.932	(0.006)	0.868	(0.005)
Reading/ELA pretest	0.004	(0.005)	0.033	(0.006)	0.026	(0.006)
ELL (ELD levels 1, 2)	0.036	(0.009)	0.131	(0.010)	0.123	(0.013)
ELL (ELD level 3)	-0.039	(0.006)	0.068	(0.008)	0.033	(0.010)
ELL (ELD levels 4, 5)	-0.005	(0.006)	0.033	(0.006)	0.019	(0.009)
ELL (level unavailable.)	-0.009	(0.013)	0.084	(0.022)	0.131	(0.023)
Disability (moderate)	-0.010	(0.006)	-0.011	(0.007)	0.036	(0.009)
Disability (severe)	0.023	(0.009)	-0.044	(0.010)	0.059	(0.013)
Economic disadvantage	-0.022	(0.004)	-0.022	(0.005)	-0.032	(0.005)
Foster care	-0.026	(0.024)	0.014	(0.027)	-0.042	(0.034)
Homeless	0.004	(0.010)	0.013	(0.010)	-0.005	(0.013)
School-level averages:						
Math pretest	-0.081	(0.051)	-0.108	(0.047)	0.076	(0.050)
Reading/ELA pretest	0.060	(0.053)	0.170	(0.050)	0.067	(0.067)
ELL (ELD levels 1, 2)	-0.526	(0.279)	-0.156	(0.323)	0.626	(0.226)
ELL (ELD level 3)	0.205	(0.217)	0.340	(0.258)	-0.172	(0.302)
ELL (ELD levels 4, 5)	-0.045	(0.146)	-0.146	(0.125)	-0.296	(0.203)
ELL (level unavailable)	-0.451	(0.126)	0.076	(0.632)	0.244	(0.572)
Disability (moderate)	-0.181	(0.214)	-0.024	(0.236)	0.228	(0.250)
Disability (severe)	0.264	(0.287)	-0.129	(0.287)	0.042	(0.316)
Economic disadvantage	-0.090	(0.047)	0.111	(0.049)	-0.144	(0.044)
Foster care	-1.987	(1.011)	-0.273	(0.854)	-0.547	(1.105)
Homeless	0.098	(0.157)	0.213	(0.159)	-0.711	(0.192)
R-squared	0.93		0.92		0.85	
No. of students	95,505		95,711		83,523	
No. of schools	450		454		416	

	ELA 4		ELA 5		ELA 6	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Student-level covariates:						
Math pretest	0.165	(0.004)	0.143	(0.005)	0.127	(0.005)
Reading/ELA pretest	0.711	(0.005)	0.759	(0.005)	0.750	(0.005)
ELL (ELD levels 1, 2)	-0.113	(0.006)	-0.102	(0.008)	-0.129	(0.010)
ELL (ELD level 3)	-0.044	(0.005)	-0.044	(0.005)	-0.055	(0.007)
ELL (ELD levels 4, 5)	0.044	(0.006)	0.013	(0.005)	-0.011	(0.005)
ELL (level unavailable)	-0.023	(0.019)	-0.017	(0.019)	-0.045	(0.021)
Disability (moderate)	-0.125	(0.006)	-0.121	(0.006)	-0.153	(0.007)
Disability (severe)	-0.147	(0.010)	-0.191	(0.009)	-0.158	(0.010)
Economic disadvantage	-0.045	(0.005)	-0.023	(0.005)	-0.041	(0.005)
Foster care	-0.045	(0.021)	-0.047	(0.022)	-0.046	(0.024)
Homeless	-0.017	(0.010)	-0.005	(0.010)	-0.023	(0.010)
School-level averages:						
Math pretest	0.019	(0.034)	-0.010	(0.028)	-0.085	(0.055)
Reading/ELA pretest	-0.080	(0.037)	-0.040	(0.033)	0.065	(0.062)
ELL (ELD levels 1, 2)	-0.213	(0.079)	-0.105	(0.107)	0.265	(0.288)
ELL (ELD level 3)	0.083	(0.075)	0.041	(0.069)	0.182	(0.169)
ELL (ELD levels 4, 5)	-0.127	(0.065)	0.162	(0.067)	0.041	(0.113)
ELL (level unavailable)	-0.809	(0.285)	-0.106	(0.282)	1.435	(0.525)
Disability (moderate)	-0.020	(0.106)	-0.056	(0.100)	0.054	(0.190)
Disability (severe)	0.198	(0.160)	-0.119	(0.145)	-0.584	(0.252)
Economic disadvantage	-0.125	(0.034)	-0.100	(0.032)	-0.118	(0.052)
Foster care	-0.640	(0.416)	-0.249	(0.390)	-0.963	(0.741)
Homeless	0.141	(0.109)	-0.119	(0.091)	-0.382	(0.163)
R-squared	0.83		0.85		0.83	
No. of students	103,615		101,657		92,836	
No. of schools	1,324		1,327		805	

	ELA 7		ELA 8		ELA 11	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Student-level covariates:						
Math pretest	0.213	(0.005)	0.193	(0.005)	0.179	(0.005)
Reading/ELA pretest	0.706	(0.006)	0.742	(0.005)	0.698	(0.006)
ELL (ELD levels 1, 2)	0.000	(0.010)	-0.074	(0.010)	0.117	(0.014)
ELL (ELD level 3)	-0.059	(0.007)	-0.040	(0.008)	-0.012	(0.011)
ELL (ELD levels 4, 5)	-0.038	(0.007)	-0.020	(0.006)	-0.037	(0.009)
ELL (level unavailable)	-0.052	(0.014)	-0.006	(0.021)	0.097	(0.024)
Disability (moderate)	-0.038	(0.007)	-0.088	(0.007)	-0.008	(0.009)
Disability (severe)	-0.031	(0.010)	-0.100	(0.010)	0.060	(0.014)
Economic disadvantage	-0.041	(0.005)	-0.014	(0.004)	-0.007	(0.005)
Foster care	-0.058	(0.026)	0.005	(0.026)	-0.188	(0.035)
Homeless	-0.005	(0.010)	-0.008	(0.010)	-0.022	(0.014)
School-level averages:						
Math pretest	0.202	(0.057)	0.017	(0.044)	0.257	(0.065)
Reading/ELA pretest	-0.269	(0.058)	-0.048	(0.047)	-0.142	(0.086)
ELL (ELD levels 1, 2)	-0.924	(0.305)	-0.773	(0.301)	0.422	(0.293)
ELL (ELD level 3)	-0.057	(0.235)	0.307	(0.239)	-0.586	(0.382)
ELL (ELD levels 4, 5)	0.405	(0.159)	0.205	(0.116)	-0.242	(0.260)
ELL (level unavailable)	-0.459	(0.136)	-0.251	(0.591)	-0.614	(0.739)
Disability (moderate)	-0.102	(0.232)	-0.394	(0.220)	0.833	(0.328)
Disability (severe)	0.357	(0.315)	0.122	(0.273)	-0.420	(0.405)
Economic disadvantage	-0.088	(0.051)	0.035	(0.045)	-0.004	(0.057)
Foster care	-2.089	(0.984)	-0.565	(0.828)	-0.313	(1.193)
Homeless	0.169	(0.173)	0.133	(0.149)	-0.942	(0.245)
R-squared	0.84		0.86		0.77	
No. of students	95,691		95,890		84,241	
No. of schools	453		455		417	

Conclusion

This report describes the statistical model underpinning the growth measures produced for CORE by Education Analytics and presents summary results about the estimated model and the estimated growth measures. It also provides model diagnostics that illustrate why Model C was chosen for the school model.

References

Fuller, Wayne (1987). Measurement error models. New York: John Wiley and Sons.

Meyer, R. H. (1996). Value-added indicators of school performance. In Hanushek, E. and Jorgenson, W. (Eds.), Improving America's schools: The role of incentives, pp. 197–223. Washington, DC: National Academy Press.

Meyer, R. H. (1999). The production of mathematics skills in high school: What works In Mayer, S. and Peterson, P. (Eds.), Earning and learning: How schools matter, pp. 169–204. Washington, DC: The Brookings Institution.