COVID-Sim: Simulating Spring 2020 Assessment Scores

META-ANALYSIS APPENDICES

credo

The Center for Research on Education Outcomes
Stanford University
credo.stanford.edu

Appendix 1
Technical Summary of CREDO Simulation Methods

Appendix 2
Technical Summary of NWEA Methodology

Appendix 3
Student Subgroup Simulation Results for Best Three Approaches

Appendix 1

Technical Summary of CREDO Simulation Methods

The goal of this project is to estimate 2019-20 scaled scores based on historical state achievement tests by subject and grade or by End of Course exams. In this Technical Appendix, we describe the methods used to simulate 2017-18 achievement and form 2019-20 achievement proxies .

We simulated 2017-18 student achievement because we have actual 2017-18 student achievement data for comparison. The best approaches were then used with more recent data to build proxies.

- We have explored more than 20 different approaches in predicting student achievement.
- We have evaluated simulation quality using the mean absolute simulation error for the actual 2017-18 student achievement.
- Mean absolute simulation error is the average of the absolute value of the difference between simulated and actual student achievement across all students.
- A shortlist of five best-performing approaches was identified and used to compare 2017-18 achievement and produce detailed simulation diagnostics in every state, subject, grade, student subgroup, school attribute.

The best-performing methods at each tested grade/EOC have been used to create a full-year 2019-20 achievement proxy (to simulate a normal year of achievement).

- We create two *adjusted* proxies of 2019-20 achievement to a) gauge learning at the time of "school building closure" and b) accumulated "learning slide" associated with the coronavirus pandemic.
- Technical Appendix 2 describes the adjustment factors and their use with 2019-20 adjusted proxies.
- We use distribution parameters of state assessment scores in each state, subject, grade tested or EOC in 2018-19 to transform 2019-20 proxies of normalized scaled scores to scaled scores.

We also describe how we transform normalized scaled scores back to scaled scores.

Below are the five shortlisted simulation approaches used on 2017-18 achievement.

1. Do Nothing

Every student's predicted achievement in a given subject in 2017-18 is the average achievement across all students in his or her tested grade or EOC in the previous year (2016-17) in the same subject:

$$\widehat{S_{i,g,2018}} = \overline{S_{g,2017}} \text{ , } \forall \text{ student } i \text{ in expected tested grade/EOC } g \in [3,11] \text{ in 2017-18}$$

2. Copy Past Year's Score

Every student's predicted achievement in a *contiguous* tested grades between *4* and *11* in a given subject in 2017-18 is his or her actual achievement in the previous year (2016-17) in the same subject:

$$\widehat{S_{i,g,2018}} = S_{i,g-1,2017} \text{ , } \forall \text{ student } i \text{ in expected tested grade } g \in [4,11] \text{ in 2017-18}$$

Every student's predicted achievement in grade *3* is equal to the average actual achievement of all students in grade 3 in a given school in the previous year (2016-17).

Every student's predicted achievement in tested *non-contiguous* high school grades ("skip grades"), who did not test in 2016-17, is equal to the same student's achievement the last time he or she was tested in the same subject up to three years prior (i.e., up to 2014-15).

3. Bridging (requires 2018-19 assessment data)

Every student's predicted achievement in a *contiguous* tested grade between *4* and *10* in a given subject in 2017-18 is his or her average actual achievement in the previous year (2016-17) and the following year (2018-19):

$$\widehat{S_{i,g,2018}} = [\![\frac{S_{i,g-1,2017}+S_{i,g+1,2019}}{2}]\!] \text{, } \forall \text{ student } i \text{ in expected tested grade } g \in [4,10] \text{ in 2017-18}$$

Every student's predicted achievement in a tested *non-contiguous* high school grades ("skip grades"), who did not test in 2016-17, is equal to the average actual achievement of the same student in 2018-19 and 2015-16, if they took a test in the same subject in those years.

4. Linear Ordinary Least Squares (OLS) Regression with one prior

Every student's predicted achievement in a given subject in 2017-18 is the out-of-sample prediction from a regression model of actual achievement in the previous year (2016-17) on prior achievement (2015-16), student demographics, and enrolled grade indicators:

$$\widehat{S_{i,g,2018}} = \widehat{\beta_0} + \widehat{\beta_1} S_{i,g-1,2017} + \widehat{\beta_2} Poverty_{i,g-1,2017} + \widehat{\beta_3} ELL_{i,g-1,2017} + \widehat{\beta_4} SpEd_{i,g-1,2017} + \widehat{\beta_5} Female_{i,g-1,2017} + \widehat{\vec{\beta}_6} \overrightarrow{Race}_{i,g-1,2017} + \widehat{\vec{\beta}_7} \overrightarrow{Grade}_{i,g-1,2017},$$

$$\forall\ student\ i \text{ in expected tested grade } g \in [4,11] \text{ in 2017-18}$$

5. Linear Ordinary Least Squares (OLS) Regression with two priors

Every student's predicted achievement in a given subject in 2017-18 is the out-of-sample prediction from a regression model of actual achievement in the previous year (2016-17) on achievement in two prior years (2014-15 and 2015-16), student demographics, and enrolled grade indicators:

$$\widehat{S_{i,g,2018}} = \widehat{\beta_0} + \widehat{\beta_1} S_{i,g-1,2017} + \widehat{\beta_2} S_{i,g-2,2016} + \widehat{\beta_3} Poverty_{i,g-1,2017} + \widehat{\beta_4} ELL_{i,g-1,2017} + \widehat{\beta_5} SpEd_{i,g-1,2017} + \widehat{\beta_6} Female_{i,g-1,2017} + \widehat{\vec{\beta}_7} \overrightarrow{Race}_{i,g-1,2017} + \widehat{\vec{\beta}_8} \overrightarrow{Grade}_{i,g-1,2017},$$

$$\forall\ student\ i \text{ in expected tested grade } g \in [5,11] \text{ in 2017-18}$$

Simulation in **OLS with 1 prior**, **OLS with 2 priors** ("out-of-sample prediction") is done in two steps. In the first step, a model with 2016-17 achievement on the left-hand side and 2015-16 student information (achievement, demographics, grade level) on the right-hand side is estimated via OLS and estimated coefficients are retrieved. In the second step, predicted 2017-18 student achievement is calculated by inserting the retrieved estimated coefficients and 2016-17 student information on the right-hand side of the model.

For every shortlisted approach, we produced analyses of the mean raw, absolute, and quadratic simulation error for 2017-18 achievement at both the student and school level.

We have also reviewed diagnostics analyses for simulation of achievement in 2015-16 and 2016-17.

Simulation approaches producing smaller mean absolute error at the student level were preferred.

Below are the technical descriptions of the approaches used to form 2019-20 achievement proxies:

1. Linear Ordinary Least Squares (OLS) Regression with one prior

   Every student's predicted 2019-20 achievement in tested grade 4 in a given subject is the out-of-sample prediction from a regression model of actual achievement in the previous year (2018-19) on prior achievement (2017-18), student demographics, and enrolled grade indicators:

$$\widehat{S_{i,g,2020}} = \widehat{\beta_0} + \widehat{\beta_1} S_{i,g-1,2019} + \widehat{\beta_2} Poverty_{i,g-1,2019} + \widehat{\beta_3} ELL_{i,g-1,2019} + \widehat{\beta_4} SpEd_{i,g-1,2019} + \widehat{\beta_5} Female_{i,g-1,2019} + \widehat{\vec{\beta}_6}\overrightarrow{Race}_{i,g-1,2019} + \widehat{\vec{\beta}_7}\overrightarrow{Grade}_{i,g-1,2019},$$

$$\forall\ student\ i \text{ in expected tested grade } g = \{4\} \text{ in 2019-20}$$

2. Linear Ordinary Least Squares (OLS) Regression with two priors

   Every student's predicted 2019-20 achievement in *contiguous* tested grades between 5 and 11 in a given subject is the out-of-sample prediction from a regression model of actual achievement in the previous year (2018-19) on achievement in two prior years (2016-17 and 2017-18), student demographics, and enrolled grade indicators:

$$\widehat{S_{i,g,2020}} = \widehat{\beta_0} + \widehat{\beta_1} S_{i,g-1,2019} + \widehat{\beta_2} S_{i,g-2,2018} + \widehat{\beta_3} Poverty_{i,g-1,2019} + \widehat{\beta_4} ELL_{i,g-1,2019} + \widehat{\beta_5} SpEd_{i,g-1,2019} + \widehat{\beta_6} Female_{i,g-1,2019} + \widehat{\vec{\beta}_7}\overrightarrow{Race}_{i,g-1,2019} + \widehat{\vec{\beta}_8}\overrightarrow{Grade}_{i,g-1,2019},$$

$$\forall\ student\ i \text{ in expected contiguous tested grade } g \in [5, 11] \text{ in 2019-20}$$

3. School by Grade 3 Historical Average

Every student's predicted 2019-20 achievement in tested grade 3 is the average achievement of students taking the same assessment in the same school in the two prior school years (2017-18 and 2018-19):

$$\widehat{S_{i,s,g,2020}} = \overline{S_{s,g,\{2018,2019\}}} \text{ , } \forall \text{ student } i \text{ in school } s, \text{ expected to be in tested grade } g = \{3\} \text{ in 2019-20}$$

4. School by High School Grade Historical Average

Every student's predicted 2019-20 achievement in *non-contiguous* high school tested grades is the average achievement of students taking the same assessment in the same school in the two prior school years (2017-18 and 2018-19):

$$\widehat{S_{i,s,g,2020}} = \overline{S_{s,g,\{2018,2019\}}} \text{ ,}$$

$$\forall \text{ student } i \text{ in school } s, \text{ expected to be in a non} - \text{contiguous high school tested grade } g \text{ in 2019-20}$$

5. School by EOC Historical Average

Every student's predicted 2019-20 achievement in a given EOC exam is the average achievement of students taking the same exam in the same school in the two prior school years (2017-18 and 2018-19):

$$\widehat{S_{i,s,EOC,2020}} = \overline{S_{s,EOC,\{2018,2019\}}} \text{ , } \forall \text{ student } i \text{ in school } s, \text{ expected to take EOC exam } EOC \text{ in 2019-20}$$

We transformed proxies of 2019-20 normalized scaled scores to scaled scores by multiplying 2019-20 normalized scaled score proxies with the standard deviation of scores in the same state, subject and tested grade or EOC assessment in 2018-19 and then adding the average score in the same state, subject and tested grade or EOC assessment:

$$Scaled\ Score\ Proxy_{2019-20} = (Normalized\ Scaled\ Score\ Proxy_{2019-20} \times SD_{2018-19}) + \mu_{2018-19},$$

where $SD_{2018-19}$ and $\mu_{2018-19}$ are the standard deviation and mean, respectively, of scaled scores in the relevant state, subject, tested grade or EOC assessment.

For a small number of students for whom the 2019-20 scaled score proxy falls outside the range of the scaled scores appropriate for that state, subject, and tested grade or EOC, we truncated their scaled scores to the upper (lower) bound if their scaled score proxy is above (below) the upper (lower) bound of the relevant scale, specific to the state, subject, and tested grade or EOC.

Other computational methods reviewed for 2017-18 achievement simulation

In addition to the shortlisted approaches described in the presentation, we explored alternative *computational* and *statistical* approaches as foundational work.

The alternative approaches reviewed produced higher mean absolute simulation errors than the shortlisted approaches so we do not report their results in the main presentation.

For each of the alternative *computational* approaches reviewed, every student's simulated 2017-18 achievement was equal to:

a) a fixed achievement level: 0.1, 0.2, or 0.3 standard deviations.

b) the product of a fixed factor, 0.8, with the actual achievement in the previous year (2016-17).

c) the average achievement of students in the prior year with the same combination of state, grade, race, poverty status, ell status, special education status, and gender.

d) Every student's simulated 2017-18 achievement in *contiguous* tested grades between 6 and 10 in a given subject was equal to the actual achievement in the previous year (2016-17) plus the average change in achievement between the last two available growth periods for students in tested contiguous grades :

$$\widehat{S_{i,g,2018}} = S_{i,g-1,2017} + [\![\frac{(S_{i,g-1,2017} - S_{i,g-2,2016}) + (S_{i,g-2,2016} - S_{i,g-3,2015})}{2}]\!] ,$$

$$\forall \text{ student } i \text{ in expected contiguous tested grade } g \in [6,10] \text{ in 2017-18}$$

Every student's simulated 2017-18 achievement in *non-contiguous* tested grades in a given subject was equal to the actual achievement in 2015-16 plus the change in achievement between 2014-15 and 2015-16, provided the student tested in those years for students in tested non-contiguous grades (i.e., in high school), who did not take a test in a given subject in 2016-17:

$$\widehat{S_{i,g,2018}} = S_{i,g-1,2016} + (S_{i,g-1,2016} - S_{i,g-2,2015}) ,$$

$$\forall \text{ student } i \text{ in expected non} - \text{contiguous tested grade } g \text{ in 2017-18}$$

Every student's simulated 2017-18 achievement was equal to:

e) the average simulated 2017-18 achievement from shortlisted approaches 2, 3, and alternative computational method (d).

f) the average simulated 2017-18 achievement from shortlisted approaches 2 and alternative computational method (d).

g) the actual achievement in the previous year (2016-17) plus the average change in achievement between the last two growth periods:

$$\widehat{S_{i,g,2018}} = S_{i,g-1,2017} + (S_{i,g-1,2017} - S_{i,g-2,2016}),$$

$$\forall \text{ student } i \text{ in expected tested grade } g \in [6,10] \text{ in 2017-18}$$

For each of the alternative *statistical* approaches reviewed, every student's simulated 2017-18 achievement was equal to the out-of-sample prediction from a regression model of actual achievement in the previous year (2016-17) on student demographics, enrolled grade indicators and:

a) prior achievement level (2015-16) and quadratic prior achievement.

b) deciles of prior achievement (2015-16).

c) deciles of achievement in two prior years (2014-15 and 2015-16.

d) Out-of-sample prediction through student-level regression of achievement on student characteristics in shortlisted approaches 4, 5 and alternative statistical methods (a) through (c) has also been applied to samples of the student population stratified by grade.

Every student's simulated 2017-18 achievement was equal to the out-of-sample prediction from a regression model of actual student achievement in the previous year (2016-17) on school-level average achievement in 2015-16, categorical variables reflecting racial composition, the share of students in poverty, the share of ELL students, and the share of students in special education at the school level in 2015-16 and:

e) nothing else.

f) interactions of student demographics and school-level variables.

g) random intercepts at the school level.

Simulated average 2017-18 achievement at the school-by-tested grade/EOC level was equal to the out-of-sample prediction from a regression model of school-by-tested grade/EOC-level average actual achievement in the previous year (2016-17) on average prior achievement (2015-16) at the school-by-tested grade/EOC level, average demographics at the school-by-tested grade/EOC level, and grade indicator variables for 2015-16 and :

h) nothing else.

i) quadratic school-by-tested grade/EOC-level average prior achievement (2015-16).

Every school's simulated 2017-18 achievement was equal to the out-of-sample prediction from a regression model of school-level average actual achievement in the previous year (2016-17) on average prior achievement (2015-16) at the school level, school-level categorical variables capturing the demographics of the student population in 2015-16 and :

j) nothing else.

k) quadratic school-level average prior achievement (2015-16).

Appendix 2

Technical Summary of NWEA Methodology

- The COVID-related school building closures that swept the country in March 2020 likely caused learning loss in students through two pathways:

    1. School Building Closure: the reduction in learning due to school building closures around March 15, 2020 may lead to lower student achievement. This impact can be modeled directly with assessment data.

        - This also captures reduction in learning due to additional constraints some students are facing during this time, like lack of access to technology, or quiet spaces to concentrate on schoolwork, among other obstacles exacerbated or caused by the pandemic.

    2. Learning Slide: the use of remote instruction for the remainder of the school year may cause student achievement to erode from the point of school building closure, rather than at the start of the summer.

- To account for the learning slide impact of COVID-related school closures on student achievement, CREDO sought learning slide estimates from NWEA to apply to our 2019-20 adjusted March proxy, which accounts for the impact of school building closure on the 2019-20 normalized scale score proxy.

- NWEA's MAP Growth tests are administered to students in school districts across the country at three points during a school year: Fall, Winter, and Spring.
- NWEA can estimate students' learning loss over summer months by comparing students' Spring MAP test score with their Fall MAP test score in the subsequent school year (see Figure 1 of Kuhfeld, M. 2018 for a visual representation of NWEA's summer learning loss calculations).
- Estimates of students' summer learning loss can be used to understand the potential impact of COVID-related school closures on student achievement (see Kuhfeld, M., & Tarasawa, B. 2020 for NWEA's COVID-19 Slide Brief).

- NWEA and CREDO assume that the learning slide impact of the March 2020 school building closures is similar to students' typical 3-month learning loss during school closure over the summer.
- NWEA developed learning slide estimates for CREDO by using students' MAP Growth math and reading test events from NWEA's anonymized longitudinal data base.
- Typical growth rates were estimated across two school years (2017-18 and 2018-19) and the summer break in between using a series of multilevel growth models. (See Kuhfeld, M., & Tarasawa, B. 2020 for a technical description of NWEA's modeling strategy for calculating out-of-school learning loss.)

- NWEA's models also included aggregated school-level student characteristics supplied by CREDO:
  - level of students in poverty
  - level of students classified as English Language Learners, and
  - level of students receiving special education designations
- NWEA provided CREDO with measures of average learning slide for 504 unique combinations of grade, subject, and school demographic characteristics. If the base of schools administering MAP assessments in a state exceeded 10% of all schools, NWEA produced state-specific measures. They also created a set of measures based on a national sample of students for use in states where no state measures were possible.

- NWEA produces linking studies that allow for the prediction of students' proficiency levels on state summative assessments.
- According to NWEA, "Each study identifies the specific RIT scale (for Rasch Unit) scores from MAP that correspond to the various proficiency levels for each subject (reading, mathematics, etc.) and for each grade. These studies also estimate the probability that a student with a specific RIT score would achieve a status of "proficient" or better on her/his state test."

- The learning slide estimates provided by NWEA represent the amount of learning lost (in standard deviation units) during the three months of COVID-related school closures. They capture the 3-month slide of students' achievement, but do not account for the impact of school building closure on student learning.
- To account for the impact of school building closure, CREDO adjusted the 2019-20 normalized scale score proxies by deducting the average amount of learning lost (in standard deviation units) due to March 2020 building closures:

$$Adjusted\ March\ Proxy_{2019-20} = Normalized\ Scaled\ Score\ Proxy_{2019-20} - 0.1$$

- The resulting 2019-20 adjusted March proxy represents a student's estimated normalized achievement as of mid-March 2020.

- To account for the subsequent learning slide that students faced, CREDO applied the NWEA learning loss estimates to the 2019-20 adjusted March proxy:

$$Adjusted\ June\ Proxy_{2019-20} = Adjusted\ March\ Proxy_{2019-20} + Learning\ Loss\ Estimate_{2019-20}$$

- The resulting 2019-20 adjusted June proxy represents a student's estimated normalized achievement as of the typical end of the school year (assuming mid-June 2020).

- Therefore, the 2019-20 adjusted June proxy takes into account:

1. the knowledge a student did not acquire in 2019-20 due to the loss of classroom days prior to the end of the school year, and

2. the slide of a student's knowledge due to the reduced schooling for the three months from March to the end of the typical school year in June.

- If a learning slide estimate is missing for at least one subject within a grade due to insufficient NWEA data, we swap in the national estimates for both subjects within the grade.
    - For example, if a state-specific learning slide estimate for 4th grade math is missing, we replace the missing 4th grade math estimate **and** the state-specific 4th grade reading estimate with the corresponding national estimates.
- NWEA provided learning slide estimates for grades 3-8 and 10, as these grades correspond to their MAP Growth testing patterns.
    - For high school tested grades and EOCs in statewide assessments that did not have a specific learning slide estimate, we swapped in the grade 10 learning loss estimate for the relevant subject and school demographics.

Appendix 3

Student Subgroup Simulation Results for Best Three Approaches

Table A3.1: Simulation Diagnostics of 2017-18 Student Achievement in Reading by State and Student Subgroup for Scenario 3: *Bridging*

			Race/Ethnicity											
	Overall		Black		Hispanic		White		Students in Poverty		English Language Learners		Special Education Students	
State	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range
Arizona	0.377	1.084	0.375	1.076	0.370	1.057	0.386	1.112	0.371	1.061	0.365	1.050	0.372	1.070
Arkansas	0.337	0.974	0.340	0.985	0.336	0.961	0.336	0.975	0.342	0.984	0.342	0.963	0.351	0.996
District of Columbia	0.353	1.023	0.356	1.033	0.346	0.980	0.345	0.981	0.355	1.026	0.371	0.999	0.386	1.064
Illinois	0.396	1.144	0.413	1.201	0.401	1.155	0.388	1.120	0.406	1.174	0.426	1.203	0.427	1.225
Indiana	0.391	1.136	0.397	1.157	0.380	1.099	0.391	1.136	0.395	1.158	0.395	1.159	0.417	1.233
Kentucky	0.437	1.293	0.438	1.293	0.428	1.246	0.439	1.302	0.447	1.331	0.459	1.333	0.516	1.609
Louisiana	0.414	1.196	0.416	1.208	0.417	1.208	0.412	1.184	0.417	1.207	0.440	1.253	0.442	1.293
Michigan	0.375	1.106	0.393	1.162	0.374	1.097	0.371	1.095	0.388	1.143	0.375	1.083	0.399	1.200
Missouri	0.386	1.177	0.403	1.251	0.376	1.140	0.384	1.165	0.396	1.212	0.377	1.134	0.453	1.479
New Jersey	0.357	1.040	0.370	1.077	0.361	1.061	0.352	1.026	0.366	1.068	0.388	1.104	0.377	1.096
New Mexico	0.409	1.190	0.414	1.222	0.411	1.192	0.403	1.169	0.413	1.200	0.434	1.234	0.442	1.245
New York - Upstate	0.424	1.264	0.457	1.400	0.427	1.291	0.416	1.232	0.439	1.328	0.462	1.382	0.480	1.459
New York City	0.419	1.245	0.419	1.255	0.416	1.244	0.422	1.230	0.417	1.248	0.452	1.373	0.444	1.360
North Carolina	0.382	1.098	0.390	1.128	0.374	1.080	0.382	1.093	0.391	1.135	0.383	1.084	0.403	1.145
Rhode Island	0.387	1.126	0.384	1.108	0.387	1.119	0.387	1.130	0.386	1.110	0.389	1.093	0.382	1.080
South Carolina	0.356	1.041	0.343	1.002	0.343	0.988	0.368	1.076	0.350	1.020	0.337	0.969	0.323	0.945
Tennessee	0.415	1.252	0.436	1.361	0.416	1.289	0.408	1.208	0.434	1.350	0.490	1.645	0.497	1.714
Utah	0.384	1.139	0.416	1.216	0.399	1.179	0.380	1.129	0.399	1.181	0.426	1.241	0.430	1.274
Wisconsin	0.376	1.127	0.436	1.348	0.387	1.160	0.368	1.096	0.395	1.194	0.394	1.181	0.437	1.332
Minimum	0.337	0.974	0.340	0.985	0.336	0.961	0.336	0.975	0.342	0.984	0.337	0.963	0.323	0.945
Maximum	0.437	1.293	0.457	1.400	0.428	1.291	0.439	1.302	0.447	1.350	0.490	1.645	0.516	1.714

Notes:

(1) Columns titled *Mean* and *95% Range* show the average and 95%-error value range (97.5th percentile-2.5th percentile), respectively, of the *absolute simulation error* overall and by subgroup in standard deviation units for the presented simulation scenario.

(2) Row titled *Minimum (Maximum)* shows the minimum (maximum) of the absolute simulation error statistic represented in each column.

(3) Shortlisted scenario presented:

Bridging: each student's simulated score is equal to the average of his or her actual achievement in the previous year and the following year. For details and special cases, please see Appendix 1.

Table A3.2: Simulation Diagnostics of 2017-18 Student Achievement in Reading by State and Student Subgroup for Scenario 4: *Regression with 1 Prior*

			Race/Ethnicity											
	Overall		Black		Hispanic		White		Students in Poverty		English Language Learners		Special Education Students	
State	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range
Arizona	0.428	1.239	0.426	1.232	0.419	1.210	0.440	1.276	0.421	1.212	0.412	1.201	0.425	1.220
Arkansas	0.384	1.108	0.387	1.117	0.381	1.093	0.383	1.108	0.390	1.122	0.389	1.094	0.406	1.152
District of Columbia	0.396	1.143	0.397	1.143	0.390	1.119	0.388	1.154	0.397	1.143	0.422	1.174	0.419	1.180
Illinois	0.430	1.253	0.444	1.295	0.431	1.246	0.425	1.247	0.437	1.270	0.446	1.256	0.460	1.321
Indiana	0.425	1.236	0.433	1.246	0.415	1.194	0.425	1.241	0.430	1.251	0.433	1.262	0.462	1.348
Kentucky	0.475	1.406	0.478	1.390	0.462	1.349	0.477	1.414	0.487	1.445	0.493	1.468	0.574	1.760
Louisiana	0.466	1.367	0.470	1.384	0.478	1.388	0.462	1.355	0.470	1.382	0.509	1.455	0.504	1.469
Michigan	0.426	1.247	0.440	1.291	0.423	1.223	0.424	1.238	0.440	1.289	0.425	1.245	0.448	1.335
Missouri	0.430	1.329	0.453	1.416	0.422	1.289	0.426	1.315	0.444	1.372	0.430	1.309	0.509	1.657
New Jersey	0.414	1.229	0.420	1.230	0.418	1.228	0.413	1.241	0.421	1.235	0.448	1.246	0.437	1.271
New Mexico	0.450	1.313	0.454	1.310	0.452	1.315	0.443	1.300	0.454	1.322	0.468	1.309	0.485	1.366
New York - Upstate	0.460	1.422	0.501	1.572	0.465	1.437	0.451	1.385	0.482	1.503	0.522	1.566	0.532	1.669
New York City	0.452	1.382	0.453	1.397	0.450	1.394	0.452	1.349	0.453	1.395	0.512	1.569	0.492	1.527
North Carolina	0.444	1.292	0.458	1.335	0.440	1.278	0.439	1.275	0.461	1.347	0.474	1.326	0.512	1.399
Rhode Island	0.423	1.231	0.422	1.191	0.427	1.243	0.420	1.228	0.428	1.233	0.429	1.213	0.426	1.237
South Carolina	0.394	1.138	0.379	1.096	0.382	1.096	0.406	1.169	0.388	1.119	0.376	1.086	0.364	1.059
Tennessee	0.497	1.525	0.515	1.645	0.498	1.562	0.490	1.474	0.515	1.627	0.603	2.100	0.582	2.019
Utah	0.437	1.298	0.466	1.379	0.447	1.322	0.434	1.292	0.450	1.326	0.471	1.360	0.485	1.408
Wisconsin	0.428	1.274	0.483	1.479	0.435	1.285	0.421	1.247	0.446	1.330	0.441	1.313	0.492	1.471
Minimum	0.384	1.108	0.379	1.096	0.381	1.093	0.383	1.108	0.388	1.119	0.376	1.086	0.364	1.059
Maximum	0.497	1.525	0.515	1.645	0.498	1.562	0.490	1.474	0.515	1.627	0.603	2.100	0.582	2.019

Notes:

(1) Columns titled *Mean* and *95% Range* show the average and 95%-error value range (97.5th percentile-2.5th percentile), respectively, of the *absolute simulation error* overall and by subgroup in standard deviation units for the presented simulation scenario.

(2) Row titled *Minimum (Maximum)* shows the minimum (maximum) of the absolute simulation error statistic represented in each column.

(3) Shortlisted scenario presented:

Regression with 1 Prior: each student's simulated score comes from a statistical model that includes past year's score. For details and special cases, please see Appendix 1.

Table A3.3: Simulation Diagnostics of 2017-18 Student Achievement in Reading by State and Student Subgroup for Scenario 5: *Regression with 2 Priors*

			Race/Ethnicity											
	Overall		Black		Hispanic		White		Students in Poverty		English Language Learners		Special Education Students	
State	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range
Arizona	0.397	1.154	0.397	1.154	0.389	1.126	0.408	1.187	0.391	1.132	0.383	1.102	0.394	1.135
Arkansas	0.352	1.025	0.355	1.025	0.351	1.008	0.351	1.029	0.358	1.038	0.358	1.012	0.372	1.064
District of Columbia	0.372	1.089	0.375	1.114	0.362	1.050	0.359	1.032	0.373	1.095	0.406	1.110	0.404	1.164
Illinois	0.404	1.186	0.419	1.225	0.405	1.178	0.401	1.179	0.412	1.203	0.425	1.204	0.432	1.243
Indiana	0.394	1.155	0.400	1.165	0.385	1.126	0.394	1.155	0.399	1.177	0.408	1.205	0.429	1.285
Kentucky	0.434	1.286	0.440	1.284	0.422	1.243	0.435	1.294	0.447	1.326	0.451	1.349	0.535	1.678
Louisiana	0.431	1.258	0.430	1.263	0.430	1.242	0.431	1.255	0.433	1.267	0.458	1.321	0.464	1.348
Michigan	0.396	1.165	0.408	1.197	0.396	1.147	0.394	1.159	0.410	1.202	0.399	1.160	0.419	1.236
Missouri	0.401	1.241	0.419	1.311	0.394	1.195	0.398	1.231	0.414	1.281	0.404	1.212	0.475	1.510
New Jersey	0.391	1.172	0.398	1.173	0.395	1.169	0.390	1.185	0.399	1.179	0.434	1.235	0.410	1.196
New Mexico	0.423	1.235	0.423	1.228	0.425	1.241	0.415	1.216	0.428	1.248	0.446	1.251	0.455	1.285
New York - Upstate	0.425	1.304	0.462	1.453	0.425	1.291	0.417	1.278	0.443	1.366	0.471	1.374	0.493	1.540
New York City	0.416	1.251	0.416	1.263	0.413	1.259	0.419	1.228	0.415	1.256	0.457	1.417	0.452	1.396
North Carolina	0.402	1.178	0.416	1.226	0.398	1.167	0.397	1.156	0.418	1.241	0.430	1.211	0.450	1.264
Rhode Island	0.396	1.156	0.395	1.106	0.398	1.175	0.394	1.156	0.402	1.166	0.382	1.087	0.396	1.170
South Carolina	0.372	1.086	0.355	1.037	0.361	1.051	0.386	1.124	0.365	1.066	0.356	1.041	0.336	0.996
Tennessee	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Utah	0.408	1.222	0.429	1.307	0.418	1.241	0.405	1.218	0.421	1.243	0.440	1.258	0.451	1.300
Wisconsin	0.396	1.181	0.447	1.353	0.405	1.201	0.390	1.158	0.413	1.235	0.410	1.229	0.455	1.364
Minimum	0.352	1.025	0.355	1.025	0.351	1.008	0.351	1.029	0.358	1.038	0.356	1.012	0.336	0.996
Maximum	0.434	1.304	0.462	1.453	0.430	1.291	0.435	1.294	0.447	1.366	0.471	1.417	0.535	1.678

Notes:

(1) Columns titled *Mean* and *95% Range* show the average and 95%-error value range (97.5th percentile-2.5th percentile), respectively, of the *absolute simulation error* overall and by subgroup in standard deviation units for the presented simulation scenario.

(2) Row titled *Minimum (Maximum)* shows the minimum (maximum) of the absolute simulation error statistic represented in each column.

(3) Shortlisted scenario presented:

Regression with 2 Priors: each student's simulated score comes from a statistical model that includes scores from two prior years. Regression with 2 priors is not applicable in Tennessee because of the lack of 2015-16 test scores and the associated simulation diagnostics are not available (n.a.). For details and special cases, please see Appendix 1.

Table A3.4: Simulation Diagnostics of 2017-18 Student Achievement in Math by State and Student Subgroup for Scenario 3: *Bridging*

			Race/Ethnicity									
	Overall		Black		Hispanic		White		Students in Poverty		English Language Learners		Special Education Students	
State	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range
Arizona	0.376	1.090	0.371	1.087	0.375	1.089	0.377	1.094	0.376	1.094	0.387	1.141	0.370	1.077
Arkansas	0.424	1.247	0.421	1.208	0.436	1.298	0.422	1.238	0.427	1.251	0.453	1.372	0.454	1.350
District of Columbia	0.388	1.176	0.380	1.158	0.380	1.118	0.445	1.369	0.380	1.143	0.384	1.106	0.414	1.240
Illinois	0.362	1.085	0.395	1.194	0.370	1.102	0.349	1.037	0.380	1.137	0.405	1.206	0.436	1.305
Indiana	0.345	1.056	0.370	1.161	0.348	1.062	0.339	1.028	0.355	1.096	0.374	1.157	0.397	1.251
Kentucky	0.403	1.204	0.382	1.114	0.385	1.119	0.407	1.229	0.401	1.203	0.393	1.138	0.468	1.563
Louisiana	0.404	1.183	0.414	1.211	0.403	1.188	0.396	1.153	0.412	1.204	0.425	1.261	0.446	1.340
Michigan	0.335	1.027	0.380	1.173	0.347	1.043	0.323	0.983	0.361	1.116	0.368	1.120	0.409	1.276
Missouri	0.396	1.457	0.495	1.985	0.416	1.575	0.373	1.300	0.441	1.719	0.454	1.770	0.568	2.175
New Jersey	0.379	1.125	0.398	1.192	0.395	1.173	0.367	1.078	0.397	1.184	0.428	1.265	0.414	1.238
New Mexico	0.443	1.313	0.457	1.397	0.447	1.327	0.428	1.249	0.448	1.329	0.467	1.390	0.496	1.467
New York - Upstate	0.386	1.300	0.485	1.609	0.436	1.473	0.356	1.167	0.440	1.483	0.523	1.693	0.514	1.631
New York City	0.418	1.403	0.452	1.492	0.439	1.471	0.367	1.211	0.429	1.434	0.496	1.618	0.500	1.605
North Carolina	0.399	1.152	0.409	1.166	0.404	1.162	0.393	1.144	0.411	1.174	0.402	1.145	0.411	1.190
Rhode Island	0.352	1.027	0.373	1.080	0.373	1.068	0.339	0.993	0.368	1.064	0.406	1.140	0.411	1.156
South Carolina	0.379	1.151	0.348	1.027	0.361	1.074	0.405	1.238	0.359	1.076	0.359	1.061	0.321	0.949
Tennessee	0.411	1.359	0.469	1.603	0.430	1.437	0.388	1.240	0.449	1.502	0.519	1.682	0.512	1.637
Utah	0.341	1.109	0.391	1.256	0.376	1.220	0.332	1.072	0.370	1.206	0.403	1.255	0.403	1.274
Wisconsin	0.391	1.490	0.564	1.973	0.450	1.751	0.361	1.261	0.455	1.768	0.522	1.934	0.575	1.979
Minimum	0.335	1.027	0.348	1.027	0.347	1.043	0.323	0.983	0.355	1.064	0.359	1.061	0.321	0.949
Maximum	0.443	1.490	0.564	1.985	0.450	1.751	0.445	1.369	0.455	1.768	0.523	1.934	0.575	2.175

Notes:

(1) Columns titled *Mean* and *95% Range* show the average and 95%-error value range (97.5th percentile-2.5th percentile), respectively, of the *absolute simulation error* overall and by subgroup in standard deviation units for the presented simulation scenario.

(2) Row titled *Minimum (Maximum)* shows the minimum (maximum) of the absolute simulation error statistic represented in each column.

(3) Shortlisted scenario presented:

Bridging: each student's simulated score is equal to the average of his or her actual achievement in the previous year and the following year. For details and special cases, please see Appendix 1.

Table A3.5: Simulation Diagnostics of 2017-18 Student Achievement in Math by State and Student Subgroup for Scenario 4: *Regression with 1 Prior*

			Race/Ethnicity											
	Overall		Black		Hispanic		White		Students in Poverty		English Language Learners		Special Education Students	
State	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range
Arizona	0.430	1.246	0.423	1.220	0.429	1.239	0.434	1.262	0.430	1.244	0.437	1.272	0.424	1.225
Arkansas	0.469	1.352	0.454	1.300	0.474	1.380	0.471	1.357	0.467	1.351	0.482	1.415	0.477	1.417
District of Columbia	0.438	1.338	0.433	1.298	0.432	1.291	0.470	1.501	0.435	1.304	0.432	1.240	0.452	1.363
Illinois	0.410	1.222	0.436	1.299	0.414	1.223	0.400	1.190	0.422	1.252	0.435	1.274	0.474	1.404
Indiana	0.395	1.194	0.439	1.308	0.404	1.194	0.386	1.164	0.411	1.235	0.444	1.342	0.454	1.388
Kentucky	0.449	1.356	0.421	1.227	0.428	1.277	0.454	1.377	0.445	1.353	0.434	1.335	0.522	1.715
Louisiana	0.453	1.334	0.460	1.353	0.455	1.337	0.445	1.310	0.459	1.352	0.481	1.439	0.492	1.457
Michigan	0.392	1.188	0.435	1.312	0.404	1.205	0.381	1.147	0.419	1.272	0.433	1.301	0.470	1.420
Missouri	0.448	1.571	0.537	2.033	0.470	1.657	0.426	1.441	0.489	1.783	0.514	1.894	0.611	2.255
New Jersey	0.414	1.222	0.431	1.269	0.425	1.251	0.403	1.190	0.429	1.266	0.462	1.359	0.456	1.366
New Mexico	0.479	1.410	0.491	1.418	0.481	1.407	0.474	1.412	0.481	1.415	0.490	1.435	0.529	1.546
New York - Upstate	0.439	1.509	0.542	1.778	0.493	1.679	0.408	1.383	0.495	1.672	0.587	1.846	0.572	1.778
New York City	0.465	1.580	0.502	1.649	0.485	1.637	0.414	1.418	0.476	1.610	0.554	1.763	0.551	1.733
North Carolina	0.450	1.298	0.469	1.318	0.459	1.323	0.437	1.278	0.469	1.332	0.479	1.346	0.473	1.305
Rhode Island	0.419	1.259	0.436	1.296	0.441	1.297	0.406	1.238	0.433	1.282	0.482	1.329	0.477	1.339
South Carolina	0.429	1.309	0.387	1.153	0.412	1.241	0.461	1.406	0.404	1.210	0.410	1.234	0.356	1.062
Tennessee	0.506	1.773	0.584	2.078	0.541	1.888	0.474	1.593	0.558	1.950	0.702	2.170	0.644	2.091
Utah	0.392	1.304	0.452	1.388	0.434	1.405	0.380	1.266	0.425	1.397	0.461	1.411	0.469	1.448
Wisconsin	0.433	1.579	0.582	2.045	0.482	1.864	0.408	1.374	0.488	1.889	0.548	2.043	0.593	2.080
Minimum	0.392	1.188	0.387	1.153	0.404	1.194	0.380	1.147	0.404	1.210	0.410	1.234	0.356	1.062
Maximum	0.506	1.773	0.584	2.078	0.541	1.888	0.474	1.593	0.558	1.950	0.702	2.170	0.644	2.255

Notes:

(1) Columns titled *Mean* and *95% Range* show the average and 95%-error value range (97.5th percentile-2.5th percentile), respectively, of the *absolute simulation error* overall and by subgroup in standard deviation units for the presented simulation scenario.

(2) Row titled *Minimum (Maximum)* shows the minimum (maximum) of the absolute simulation error statistic represented in each column.

(3) Shortlisted scenario presented:

Regression with 1 Prior: each student's simulated score comes from a statistical model that includes past year's score. For details and special cases, please see Appendix 1.

Table A3.6: Simulation Diagnostics of 2017-18 Student Achievement in Math by State and Student Subgroup for Scenario 5: *Regression with 2 Priors*

			Race/Ethnicity											
	Overall		Black		Hispanic		White		Students in Poverty		English Language Learners		Special Education Students	
State	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range	Mean	95% Range
Arizona	0.404	1.172	0.394	1.128	0.401	1.160	0.410	1.194	0.402	1.162	0.398	1.145	0.391	1.124
Arkansas	0.442	1.287	0.431	1.245	0.449	1.328	0.443	1.286	0.442	1.288	0.458	1.371	0.442	1.321
District of Columbia	0.428	1.293	0.424	1.267	0.422	1.271	0.450	1.422	0.426	1.274	0.427	1.242	0.443	1.327
Illinois	0.386	1.156	0.411	1.223	0.390	1.161	0.375	1.126	0.397	1.183	0.424	1.243	0.450	1.328
Indiana	0.370	1.126	0.410	1.239	0.385	1.147	0.360	1.095	0.388	1.178	0.449	1.343	0.427	1.317
Kentucky	0.412	1.243	0.390	1.138	0.394	1.157	0.416	1.265	0.410	1.245	0.392	1.186	0.482	1.630
Louisiana	0.422	1.245	0.430	1.269	0.418	1.216	0.415	1.225	0.429	1.269	0.447	1.320	0.466	1.374
Michigan	0.365	1.112	0.404	1.213	0.376	1.125	0.355	1.083	0.390	1.186	0.405	1.226	0.440	1.326
Missouri	0.421	1.428	0.495	1.775	0.440	1.462	0.404	1.325	0.458	1.608	0.480	1.662	0.565	2.038
New Jersey	0.387	1.152	0.405	1.196	0.399	1.183	0.377	1.122	0.404	1.196	0.435	1.264	0.427	1.277
New Mexico	0.453	1.331	0.459	1.374	0.455	1.332	0.446	1.321	0.457	1.339	0.470	1.367	0.498	1.463
New York - Upstate	0.408	1.403	0.515	1.695	0.463	1.587	0.377	1.263	0.465	1.575	0.565	1.740	0.538	1.656
New York City	0.441	1.516	0.481	1.591	0.465	1.582	0.383	1.330	0.453	1.545	0.541	1.699	0.533	1.663
North Carolina	0.418	1.214	0.433	1.234	0.429	1.239	0.407	1.193	0.434	1.245	0.444	1.266	0.424	1.196
Rhode Island	0.387	1.150	0.401	1.203	0.411	1.206	0.374	1.108	0.405	1.201	0.460	1.240	0.446	1.248
South Carolina	0.410	1.252	0.368	1.096	0.396	1.202	0.441	1.341	0.385	1.164	0.393	1.191	0.336	1.004
Tennessee	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Utah	0.370	1.246	0.420	1.321	0.415	1.346	0.358	1.209	0.403	1.338	0.452	1.364	0.446	1.364
Wisconsin	0.407	1.435	0.555	1.900	0.457	1.686	0.383	1.266	0.463	1.736	0.533	1.911	0.566	1.930
Minimum	0.365	1.112	0.368	1.096	0.376	1.125	0.355	1.083	0.385	1.162	0.392	1.145	0.336	1.004
Maximum	0.453	1.516	0.555	1.900	0.465	1.686	0.450	1.422	0.465	1.736	0.565	1.911	0.566	2.038

Notes:

(1) Columns titled *Mean* and *95% Range* show the average and 95%-error value range (97.5th percentile-2.5th percentile), respectively, of the *absolute simulation error* overall and by subgroup in standard deviation units for the presented simulation scenario.

(2) Row titled *Minimum (Maximum)* shows the minimum (maximum) of the absolute simulation error statistic represented in each column.

(3) Shortlisted scenario presented:

Regression with 2 Priors: each student's simulated score comes from a statistical model that includes scores from two prior years. Regression with 2 priors is not applicable in Tennessee because of the lack of 2015-16 test scores and the associated simulation diagnostics are not available (n.a.). For details and special cases, please see Appendix 1.