education analytics

TECHNICAL REPORT ON THE WISCONSIN VALUE-ADDED MODEL: TARGET GROUP UPDATE

ACADEMIC YEAR 2018-2019

Prepared by

Robert H. Meyer, CEO

Michael Christian, Research Scientist

- December 2020

110 E Main Street, Ste. 1000
Madison, WI 53703

608.466.4966

edanalytics.org

CONTENTS

INTRODUCTION

This report describes the value-added model used by Education Analytics to measure the effectiveness of Wisconsin public schools using assessment data from the Forward Exam, ACT Aspire, and ACT. The report was produced in the fall of 2020 but because of the cancellation of state assessments due to COVID-19 in the school year 2019-2020, DPI instructed Education Analytics to work on the creation of a target group metric using the 2018-2019 dataset. Therefore, some sections and numbers in this report are identical to the Technical Report 2018-2019.

The report is divided into three sections. The first section describes the data set used to produce the value-added estimates. The second section describes the model used to estimate value-added for schools in Wisconsin. Finally, the third section presents some properties of the value-added results.

Conceptually, value-added analysis is the use of statistical techniques to isolate the component of measured student knowledge that is attributable to schools from other factors. Such factors may include prior knowledge and student characteristics associated with growth in student achievement. In practice, value-added models focus on the improvement students make on annual assessments from one year to the next, considering differences in student characteristics. Value-added models often control for measurable student characteristics using available data, such as economic disadvantage and disability, to help isolate the impact of schooling.

The model used in Wisconsin includes the available set of student characteristics to identify the extent to which schools contribute to the improvement of student achievement outcomes. Once the school-level value-added results are calculated, these are averaged to obtain district scores. To calculate the final scores, up to three years of results are combined: 2016-2017, 2017-2018, and 2018-2019.

ANALYSIS DATA SET

Before estimation can take place, a substantial amount of work is required to assemble the analysis data sets used to produce the value-added estimates. A separate analysis data set is produced for each grade, subject, and test. In total, 16 analysis data sets are produced, covering grades 4 through 11 for English language arts (ELA) and math in 2018-19.

Each analysis data set includes students who have a test result in 2018-19 (the posttest) in the grade and subject being considered, test results in 2017-18 (the pretests) in both ELA and math, had full academic year (FAY) status in their school or district, and were tested in consecutive grades.

The model also includes students in voucher school programs (referred to as Private School Choice Programs in Wisconsin). In addition, privately run schools receiving voucher students were entitled to an optional value-added score that included all attending students, including those not receiving public funds.

Student-level variables

POSTTEST AND PRETEST VARIABLES

The test scores used are from the 2016-17, 2017-18, and 2018-19 administrations of the Forward, Aspire, and ACT assessments. The Forward assessment is administered to students in grades 4 through 8; the Aspire, to students in grades 9 and 10; and the ACT, in grade 11. The value-added system produces school-level measures for grades 4 through 11 in ELA and math based on performance on the 2018-19 assessment. The 2018-19 value-added in ELA uses the 2018-19 ELA score as the posttest, while the 2018-19 value-added in math uses the 2018-19 math score as the posttest. All value-added models include pretests in both ELA and math, from both one year before the posttest in 2017-18 and, when available, from two years before the posttest in 2016-17. The use of multiple lags of prior achievement is a new aspect of the model for 2018-19.

All test scores are transformed to the z-statistic scale with means equal to zero and standard deviations equal to 1 in each grade and subject. The Forward assessments are transformed to the z-statistic scale linearly, while the Aspire and ACT assessments were transformed to the z-statistic scale using rank-based z-statistics. The rank-based z-statistic transformation, which ranks scores and then assigns to them a z-statistic based on the value associated with that rank in the normal distribution, was made to transform the Aspire and ACT test scores to a normal distribution. Thus, in the value-added analyses, all test scores were measured relative to the state means, and in units of the statewide standard deviations of test scores in given grades and subjects.

RELIABILITY OF PRETEST VARIABLES

The reliability of an assessment is the proportion of variance in test scores that is a result of differences in student knowledge of the material covered by the assessment rather than of randomness. The reliability estimates of math and ELA pretest scores are available in the technical manual for the Forward exam prepared by the Wisconsin Department of Public Instruction. They range from 0.87 to 0.92 across years, grades, and subjects. Reliability estimates of the Aspire assessment are available in the ACT Aspire Technical Manual prepared by ACT Aspire. In the value-added analysis, a reliability of 0.93 was employed for the Aspire ELA and 0.90 was employed for the Aspire math assessments. All of these reliabilities suggest that the vast majority of the variance of these tests reflect tangible differences in student knowledge of the content area. These reliability estimates are used for a correction for measurement error in the pretests.

GENDER, RACE/ETHNICITY, ECONOMIC DISADVANTAGE, AND MIGRANCY

Gender, race/ethnicity, economic disadvantage, and migrancy are drawn from the Wisconsin Information System for Education data (WISEdata) elements. Specifically, the values for these variables are drawn from the Spring Demographic Snapshot of WISEdata captured on May 23, 2019.[1] In the analysis data set, students are assigned the gender, race/ethnicity, low-income status, and migrant status reported in the post-test year. Gender categories are male and female. Race categories are American Indian/Alaskan Native, Asian, Native Hawaiian/Pacific Islander, Black/African American, Hispanic/Latino, White, and multi-racial. The analysis employs an indicator for economically disadvantaged students and an indicator for migrant students.

ENGLISH LANGUAGE PROFICIENCY CLASSIFICATION

There are seven indicators for English-language proficiency (ELP) included in the analysis dataset. Students with ELP classifications of 1 through 5 are considered to be English-language learners in ascending levels of proficiency. Students with an ELP classification of 6 are those that were formerly classified as having limited English proficiency. ELP classification is drawn from the WISEdata Snapshot.

DISABILITY

The analysis includes five indicators for students with disabilities according to their primary disability code. There are separate indicators for emotional/behavioral disability (EBD), learning or intellectual disability (LD/ID), autism (A), and speech/language disability (SL). All other disability codes are grouped into a single indicator for other disabilities. Disability status is based on a student having an active IEP or ISP between December 1 and June 30.

[1] WISEdata is a dynamic data delivery system. Snapshots capture a static version of the data as it was delivered to Wisconsin DPI on a given date. The Spring Demographic Snapshot taken near the end of the school year was for the purpose of supplying demographic characteristics to associate with student assessment results.

School enrollment

Students that have full academic year (FAY) status at a single school are assigned to that school using the school enrollment data. For the purpose of Wisconsin accountability systems and therefore value-added modeling, FAY is defined as being enrolled from the beginning of the year through completion of required statewide testing. Some students have FAY status in a single district but not at a single school because of mobility within the district. These students are included in the district growth measures but not in the school growth measures.

Voucher students

The analysis set includes test scores for voucher students attending private schools. All such schools receive a value-added score based on voucher students only. In addition, these private schools with voucher students are given the option to receive a second report card in the Wisconsin accountability system (including a value-added score) which includes non-voucher students as well as voucher students. Such schools are denoted as “opt-in” schools because they opted to receive the second non-compulsory score. Growth measures for "opt-in" schools that include non-voucher students are computed using a parallel analysis that applies the parameters of the estimated value-added model to a data set that includes both voucher and non-voucher students.

Descriptive statistics of analysis samples

Tables 1 and 2 describe the sample used for the 2018-19 school year. Note that the sample includes students from public schools and private schools participating in one of the Private School Choice programs in Wisconsin. The private school students include non-voucher students attending schools that opted in to receive a score for all their students.

Table 1. Math Sample

Variable	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
Number of Students	59,425	60,705	61,101	59,760	59,031	57,929	57,573	56,006
Number of Public School Students	56,408	57,697	58,060	56,863	56,434	55,651	55,150	53,863
Number of Voucher Students	2,453	2,414	2,463	2,310	1,980	1,775	1,795	1,481
Number of Non-Voucher Private School Students	236	265	245	251	271	151	361	358
Total Number of Private School Students	2,689	2,679	2,708	2,561	2,251	1,926	2,156	1,839
Number of Public Schools	1,089	1,038	688	651	650	527	542	555
Number of Private Schools	132	133	131	131	121	59	62	56

Variable	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
Number of Public School District Codes	425	426	426	425	425	386	385	384
Posttest Mean	578.637	602.772	612.339	627.186	646.407	426.446	428.461	19.687
Posttest Standard Deviation	51.114	52.442	57.663	59.772	57.111	9.221	9.595	5.239
Math Pretest Mean	557.297	578.143	600.415	613.951	624.976	647.428	427.09	429.007
Math Pretest Standard Deviation	50.144	52.324	55.744	56.620	64.556	59.058	8.939	9.445
ELA Pretest Mean	557.705	582.026	601.888	611.087	629.212	634.002	426.726	427.960
ELA Pretest Standard Deviation	46.293	51.276	47.856	49.600	55.645	58.425	7.209	7.433
Proportion in ELP Level 1	0.005	0.002	0.002	0.003	0.003	0.003	0.001	0.002
Proportion in ELP Level 2	0.015	0.006	0.005	0.008	0.008	0.006	0.004	0.004
Proportion in ELP Level 3	0.042	0.029	0.020	0.026	0.020	0.017	0.015	0.015
Proportion in ELP Level 4	0.026	0.038	0.031	0.017	0.013	0.013	0.013	0.011
Proportion in ELP Level 5	0.001	0.002	0.002	0.001	0.001	0.001	0.001	0.001
Proportion in ELP Level 6 (former English learners)	0.014	0.027	0.042	0.046	0.048	0.048	0.049	0.048
Proportion Female	0.488	0.493	0.488	0.487	0.488	0.486	0.493	0.496
Proportion Asian	0.042	0.039	0.039	0.040	0.038	0.039	0.038	0.038
Proportion African American	0.103	0.103	0.100	0.097	0.094	0.083	0.064	0.064
Proportion Hispanic	0.136	0.137	0.135	0.135	0.127	0.121	0.115	0.109
Proportion Native American	0.010	0.010	0.011	0.011	0.011	0.010	0.009	0.010
Proportion Native Hawaiian or Other Pacific Islander	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Proportion Two or More Races	0.044	0.042	0.041	0.038	0.035	0.033	0.029	0.028
Proportion Special Education: Emotional Behavioral	0.015	0.016	0.017	0.016	0.017	0.015	0.012	0.011
Proportion Special Education: Learning/Intellectual	0.038	0.042	0.046	0.048	0.048	0.049	0.044	0.044
Proportion Special Education Autism	0.013	0.013	0.012	0.012	0.012	0.012	0.010	0.010
Proportion Special Education: Speech/Language	0.032	0.021	0.014	0.010	0.006	0.003	0.003	0.002
Proportion Special Education: Other	0.036	0.034	0.036	0.036	0.035	0.036	0.030	0.028
Proportion with Economic Disadvantage	0.457	0.456	0.444	0.429	0.404	0.377	0.336	0.323
Proportion Migrant	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 2. English Language Arts (ELA) Sample

Variable	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
Number of Students	59,443	60,720	61,108	59,765	59,064	56,572	56,319	55,591
Number of Public School Students	56,424	57,709	58,067	56,871	56,465	54,401	53,982	53,467
Number of Voucher Students	2,454	2,418	2,462	2,313	1,983	1,729	1,753	1,476
Number of Non-Voucher Private School Students	236	265	245	250	271	150	350	358
Total Number of Private School Students	2,690	2,683	2,707	2,563	2,254	1,879	2,103	1,834
Number of Public Schools	1,089	1,038	688	651	650	523	541	552
Number of Private Schools	132	133	131	131	121	58	60	56
Number of Public School District Codes	425	426	426	425	425	384	385	384
Posttest Mean	583.161	596.619	608.209	629.197	630.707	425.866	427.393	18.404
Posttest Standard Deviation	50.713	48.363	49.609	54.233	59.190	7.446	7.643	5.493
ELA Pretest Mean	557.701	582.034	601.879	611.078	629.185	635.474	426.891	428.048
ELA Pretest Standard Deviation	46.293	51.267	47.862	49.615	55.661	57.584	7.117	7.369
Math Pretest Mean	557.293	578.148	600.415	613.930	624.925	648.804	427.256	429.098
Math Pretest Standard Deviation	50.134	52.324	55.738	56.645	64.591	58.203	8.878	9.396
Proportion in ELP Level 1	0.005	0.002	0.002	0.003	0.003	0.002	0.001	0.001
Proportion in ELP Level 2	0.015	0.006	0.005	0.008	0.008	0.006	0.004	0.004
Proportion in ELP Level 3	0.042	0.029	0.020	0.026	0.020	0.016	0.015	0.015
Proportion in ELP Level 4	0.026	0.038	0.031	0.017	0.013	0.013	0.013	0.011
Proportion in ELP Level 5	0.001	0.002	0.002	0.001	0.001	0.001	0.001	0.001
Proportion in ELP Level 6 (former English learners)	0.014	0.027	0.042	0.046	0.048	0.049	0.049	0.049
Proportion Female	0.488	0.493	0.488	0.487	0.488	0.490	0.497	0.498
Proportion Asian	0.042	0.039	0.039	0.040	0.038	0.039	0.038	0.038
Proportion African American	0.104	0.103	0.099	0.097	0.094	0.078	0.061	0.063
Proportion Hispanic	0.136	0.137	0.135	0.135	0.127	0.119	0.114	0.109
Proportion Native American	0.010	0.010	0.011	0.011	0.011	0.010	0.009	0.010
Proportion Native Hawaiian or Other Pacific Islander	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Proportion Two or More Races	0.044	0.042	0.041	0.038	0.035	0.032	0.029	0.028
Proportion Special Education: Emotional Behavioral	0.015	0.016	0.017	0.016	0.017	0.013	0.011	0.010
Proportion Special Education: Learning/Intellectual	0.038	0.042	0.046	0.048	0.048	0.047	0.044	0.043
Proportion Special Education Autism	0.013	0.013	0.012	0.012	0.012	0.011	0.010	0.009
Proportion Special Education: Speech/Language	0.032	0.021	0.014	0.01	0.006	0.003	0.003	0.002

Variable	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
Proportion Special Education: Other	0.036	0.034	0.037	0.036	0.035	0.034	0.029	0.027
Proportion with Economic Disadvantage	0.457	0.456	0.444	0.429	0.404	0.369	0.329	0.32
Proportion Migrant	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

VALUE-ADDED MODEL

For the Wisconsin school-level model, 2018-19 value-added is measured in mathematics and English language arts (ELA) in grades four through eleven using the Forward assessment (4-8), the Aspire assessment (9-10), and the ACT (11). Schools are assigned single-year value-added measures that reflect student growth from Spring 2018 to Spring 2019. Once the schools get a growth value, these values are averaged to obtain the district's score, using the number of students attributed to each school as weights.[2] The single-year value-added measures for 2018-19 are averaged with value-added measures in previous years to smooth year-to-year variance in value-added measures.

The model, in brief

The value-added model is defined by six equations: a "best linear predictor" value-added model defined in terms of true student posttest and pretest achievement (i.e., student achievement in the absence of test measurement error) and five measurement error models for observed post and prior achievement:

Student achievement: $y_{2i} = \zeta + \lambda_1 y_{1i} + \lambda_1^{alt} y_{1i}^{alt} + \lambda_0 y_{0i} + \lambda_0^{alt} y_{0i}^{alt} + \beta' X_i + \alpha' S_i + e_i$ (1)

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

Same-subject, once-lagged pretest measurement error: $Y_{1i} = y_{1i} + v_{1i}$ (3)

Other-subject, once-lagged pretest measurement error: $Y_{1i}^{alt} = y_{1i}^{alt} + v_{1i}^{alt}$ (4)

Same-subject, twice-lagged pretest measurement error: $Y_{0i} = y_{0i} + v_{0i}$ (5)

Other-subject, twice-lagged pretest measurement error: $Y_{0i}^{alt} = y_{0i}^{alt} + v_{0i}^{alt}$ (6)

where:

- the subscript i denotes each individual student;
- y_{2i} is true post achievement;

[2] Note that students who changed schools within the districts are included in the district's score but not in a school score (see School Enrollment section).

- y_{1i} and y_{1i}^{alt} are true prior achievement, one year before post achievement, in the same subject and in the other subject (math in the ELA model, ELA in the math model), with slope parameters λ_1 and λ_1^{alt};
- y_{0i} and y_{0i}^{alt} are true prior achievement, two years before post achievement, in the same subject and in the other subject (math in the ELA model, ELA in the math model), with slope parameters λ_0 and λ_0^{alt};
- X_i is a vector of characteristics of student i, with slope parameter vector β;
- S_i is a vector of indicators for school;
- α is a vector of school effects;
- e_i is the error in predicting post achievement given the explanatory variables included in the model;
- Y_{2i} is measured post achievement;
- v_{2i} is measurement error in post achievement;
- Y_{1i} and Y_{1i}^{alt} are measured prior achievement, one year before post achievement, for the same subject and alternate subject, respectively;
- v_{1i} and v_{1i}^{alt} are measurement error in prior achievement, one year before post achievement, for the same subject and alternate subject, respectively;
- Y_{0i} and Y_{0i}^{alt} are measured prior achievement, two years before post achievement, for the same subject and alternate subject, respectively; and
- v_{0i} and v_{0i}^{alt} are measurement error in prior achievement, two years before post achievement, for the same subject and alternate subject, respectively.

Substituting the measurement error equations (2) through (6) into the student achievement equation (1) yields an equation defined in terms of measured student achievement:

Measured achievement: $Y_{2i} = \zeta + \lambda_1 Y_{1i} + \lambda_1^{alt} Y_{1i}^{alt} + \lambda_0 Y_{0i} + \lambda_0^{alt} Y_{0i}^{alt} + \beta' X_i + \alpha' S_i + \varepsilon_i$ (7)

where the error term ε_i includes both the original error component and the measurement error components:

Error in measured achievement: $\varepsilon_i = e_i + v_{2i} - \lambda_1 v_{1i} - \lambda_1^{alt} v_{1i}^{alt} - \lambda_0 v_{0i} - \lambda_0^{alt} v_{0i}^{alt}$ (8)

Estimating the measured student achievement equation (7) without controlling for pretest measurement error yields biased estimates of all parameters, including the value-added effects. This bias stems from the fact that measurement error in prior achievement causes the error term (8), which includes the measurement error components v_{1i}, v_{1i}^{alt}, v_{0i}, and v_{0i}^{alt}, to be correlated with measured prior achievement. The desired parameters, as defined in equation (1), can be estimated consistently if external information is available on the variance of measurement error for prior achievement; approaches for consistent estimation in the presence of measurement error are described in detail in Fuller (1987). Information about the variance of test measurement error is obtained from the reliability estimates reported in the technical manuals for the Forward and Aspire assessments.

In contrast to measurement error in the pretest variables, measurement error in the posttest does not cause any distortions in commonly used regression approaches and can safely be overlooked. This is because we do not expect posttest measurement error v_{2i} to be correlated with measured prior achievement or any of the other right-hand-side variables in the regression equation (7). We do not expect any such correlation because there is no reason to think that a student's good or bad luck on the posttest administration should have anything to do with their measured performance in the past, their demographic characteristics, or their school assignment. Given the absence of such a correlation, the presence of posttest measurement error v_{2i} in the regression error term in (8) will not bias coefficient estimates if it is overlooked. In fact, from the perspective of estimation technique, we can think of posttest measurement error v_{2i} as operating no differently from the structural error e_i.

Value-added regression

As mentioned, the value-added model is estimated using a least-squares regression approach that corrects for measurement error in the pretest variables. It estimates the coefficients λ, β, and α by regressing posttest on the pretests, other student-level variables, and a full set of school fixed effects. This regression is estimated using an approach that accounts for measurement error in the pretests Y_{1i}, Y_{1i}^{alt}, Y_{0i}, and Y_{0i}^{alt}. Recall from equation (8) above that v_{1i}, v_{1i}^{alt}, v_{0i}, and v_{0i}^{alt}, the measurement error components of the pretests, are part of the error term ε_i. As a result, estimating the regression using ordinary least squares (without controlling for pretest measurement error) will lead to biased estimates. The regression approach employed accounts for measurement error by removing the variance in the pretests that is attributable to measurement error. To illustrate the measurement error corrected regression, re-cast the above value-added regression equation into vector form:

$$Y_t = Y_{t-\ell}\lambda + W\delta + \varepsilon$$

where Y_t is an N × 1 vector of post-test scores, $Y_{t-\ell}$ is an N × 4 vector of same-subject and other-subject pre-test scores Y_{1i}, Y_{1i}^{alt}, Y_{0i}, and Y_{0i}^{alt}; λ is a 4 × 1 vector made up of λ_1, λ_1^{alt}, λ_0, and λ_0^{alt}; W is an N × K vector of the X demographic variables and S school indicators, δ is a K × 1 vector of the β and α coefficients, and ε is an N × 1 vector of error terms. The biased ordinary-least-squares estimates of the coefficients in λ and δ are equal to:

$$\begin{bmatrix} \hat{\lambda}_{OLS} \\ \hat{\delta}_{OLS} \end{bmatrix} = \begin{bmatrix} Y'_{t-\ell}Y_{t-\ell} & Y'_{t-\ell}W \\ W'Y_{t-\ell} & W'W \end{bmatrix}^{-1} \begin{bmatrix} Y'_{t-\ell}Y_t \\ W'Y_t \end{bmatrix}$$

The measurement-error-corrected estimates of the coefficients in λ and δ are equal to:

$$\begin{bmatrix} \hat{\lambda}_{EIV} \\ \hat{\delta}_{EIV} \end{bmatrix} = \begin{bmatrix} Y'_{t-\ell}Y_{t-\ell} - \left(\frac{N-K-4}{N}\right)\sum_{i=1}^{N} V_{it-\ell} & Y'_{t-\ell}W \\ W'Y_{t-\ell} & W'W \end{bmatrix}^{-1} \begin{bmatrix} Y'_{t-\ell}Y_t \\ W'Y_t \end{bmatrix}$$

where $V_{it-\ell}$ is a 4×4 variance-covariance matrix of the errors of measurement of the variables in $Y_{t-\ell}$ for student i. This model is described in section 2.2 of Fuller (1987).

To minimize the influence of test scores at the extreme of the distribution on the estimates of the pretest coefficients in λ, we estimated the value-added model in two steps in models of student growth in mathematics in which the most recent pretest was a Forward assessment. This method was found to be useful for the mathematics model because in some grades an appreciable percentage of students received the lowest observable scale score (LOSS) in mathematics on the Forward exam (see Table 3). In step one, model parameters are estimated using all students other than those at the LOSS on the mathematics pretests. In step two, the estimated parameters for the pretest variables in both math and reading are treated as known and the model is re-estimated using all students.

Table 3. Percentage of Students at Test Floor (Lowest Observable Scale Score, LOSS) for Pre- and Posttests

	Grade	Test Subject	Percent at Posttest Floor	Percent at Math Pretest Floor	Percent at ELA Pretest Floor
Included in Growth Analysis Data Set	4	ELA	0.0%	1.1%	0.0%
		Mathematics	1.6%	1.1%	0.0%
	5	ELA	0.0%	2.0%	0.0%
		Mathematics	3.1%	2.0%	0.0%
	6	ELA	0.0%	4.3%	0.0%
		Mathematics	3.0%	4.3%	0.0%
	7	ELA	0.0%	2.8%	0.0%
		Mathematics	3.1%	2.8%	0.0%
	8	ELA	0.0%	5.3%	0.0%
		Mathematics	2.3%	5.3%	0.0%
	9	ELA	0.0%	3.7%	0.0%
		Mathematics	0.0%	3.7%	0.0%

The variables in the model

In addition to posttest and pretest scores, the student-level variables included in the model (the X variables in equation 1) are gender, race/ethnicity, ELP category, economic disadvantage, disability status, and migrancy. No higher order terms or interactions of terms are used in the model. Refer to the section "Analysis Data Set: Student-Level Variables" for a more complete description of the categories that make up each student-level variable.

Incorporating Students with Only Two Years of Scores

The estimation approach above produces school growth measures based on the growth of students with measured scores in all three years (2018-19, 2017-18, and 2016-17). To include students with measured scores in 2018-19 and 2017-18 but not in 2016-17, we estimate a model that is identical to that described above except that it does not include the twice-lagged pretest variables y_{0i} and y_{0i}^{alt}. We then produce, for each student, a growth residual equal to an estimate of $\alpha' S_i + \varepsilon_i$, using the coefficients from the complete model that includes y_{0i} and y_{0i}^{alt} when the twice-lagged pretest measures Y_{0i} and Y_{0i}^{alt} are available, and using the coefficients from the single-lag model that does not include y_{0i} and y_{0i}^{alt} when the twice-lagged pretest measures Y_{0i} and Y_{0i}^{alt} are not available. This growth residual is demeaned to have a mean of zero by grade and subject and regressed on a full set of school indicators S_i using ordinary least squares. This produces unshrunk school value-added measures for each school by grade and subject.

Aggregation to multiple-grade value-added

The value-added regression to obtain unshrunk school value-added is performed separately for each grade and subject combination. For schools that have results for more than one grade level, these estimates are averaged across grades, using the number of students attributed to the school and grade as weights, to produce unshrunk multiple-grade value-added estimates. Before aggregation, value-added measures are normalized by subject and grade, so they are on a similar scale (i.e. with a mean of 0 and a true standard deviation of 1). This normalization is done by dividing the measures by an estimate of the standard deviation of within-grade value-added. This aggregation is made separately at the elementary/middle (grades 4-8) and high school (grades 9-11) levels.

Shrinkage of value-added

At all levels, the unshrunk value-added estimates are shrunk using an Empirical Bayes multivariate shrinkage technique described in Longford (1999). This procedure is employed to bring value-added estimates based on smaller sample sizes closer to the state average, so that schools with fewer students are not overrepresented among the highest- and lowest-value-added cases simply due to randomness. It is also employed to reduce year-by-year variation in value-added scores within schools.

To use this multivariate shrinkage approach, we begin with single-year value-added measures for the 2018-19 and 2017-18 school years. Let $\hat{\alpha}_{kt}$ be the estimated value-added for school *k* in year *t*. We can group the value-added estimates for a given school *k* into a *T* x 1 column vector $\hat{\alpha}_k$, where *T* is the number of years in which value-added is measured for school *k*.(In this application, *T* will usually be 2, although it will equal 1 in schools in which value-added is measured in 2018-19 but not 2017-18 or vice versa.) Also let α_{kt} be the true value-added (which is unmeasured, and equal to what estimated value-added would be in the absence of sampling error) for school *k* in year *t*, which can be grouped by school into a *T* x 1 column vector α_k. Let the variance of α_k be the *T* x *T* matrix $Var[\alpha_k] = \Omega$, which reflects the within-year variance and across-year covariance of true value-added across schools. Also let the variance of $\hat{\alpha}_k$ conditional on α_k be the *T* x *T* matrix $Var[\hat{\alpha}_k|\alpha_k] = \Sigma_{kk}$, which reflects the within-year variance and across-year covariance of sampling error in $\hat{\alpha}_k$. We produce shrunk estimates of value-added using the following equation:

$$\alpha_k^* = \Omega[\Omega + \Sigma_{kk}]^{-1}\hat{\alpha}_k$$

where α_k^* is a *T* x 1 column vector of shrunk value-added measures for school *k* over the *T* years in which value-added is measured for school *k*. The expected mean squared error of the shrunk value-added estimates α_k^* is equal to:

$$EMSE_k = \Omega - \Omega[\Omega + \Sigma_{kk}]^{-1}\Omega$$

In practice, we use estimates of Ω and Σ_{kk} to estimate α_k^* and its expected mean squared error. The estimate of the matrix Σ_{kk} is the estimated variance-covariance matrix of the value-added estimates in $\hat{\alpha}_k$. Let $\hat{\sigma}_{t\tau kk}$ be the entry of this matrix in the row corresponding to $\hat{\alpha}_{kt}$ and the column corresponding to $\hat{\alpha}_{k\tau}$. The diagonal entries of this matrix are the squares of the estimated standard errors of the value-added estimates in $\hat{\alpha}_k$.

The diagonal entries of Ω, which are equal to the variance of α_{kt} across schools in a given year *t* and which we denote ω_{tt}, are estimated by computing the variance across schools *k* within year *t* of the unshrunk value-added estimates $\hat{\alpha}_{kt}$, then subtracting from that the average across schools *k* within year *t* of $\hat{\sigma}_{ttkk}$, the estimated squared standard error of $\hat{\alpha}_{kt}$. This estimates the variance of the true school value-added for each year *t*, excluding variance due to randomness in the value-added estimates. The square root of this variance measure is also used for normalizing value-added measures by grade before aggregation to multiple-grade measures. The off-diagonal entries of Ω, which we denote $\omega_{t\tau}$ and are equal to the covariance of α_{kt} and $\alpha_{k\tau}$ across schools between years *t* and τ, is estimated by computing the covariance of the unshrunk value-added estimates $\hat{\alpha}_{kt}$ and $\hat{\alpha}_{k\tau}$, and then subtracting from that the average error covariance estimate $\hat{\sigma}_{t\tau kk}$.

Student group value-added

Value-added is also measured by student groups defined by certain student characteristics. Specifically, we calculated differential value-added effects for:

- the seven race/ethnicity groups;
- students with and without disabilities;
- economically disadvantaged and non-economically disadvantaged students;
- English-language learners[3] and non-English-language learners;
- students who were proficient and not proficient in the same subject in the previous year; and
- students who are in (and not in) a target group made up of students who scored below the 25th percentile within their school in the same subject in the previous year.

To produce the group results by school for all subgroups other than the proficiency and target group subgroups, we produce unshrunk value-added effects for both 2017-18 and 2018-19 for each subgroup for each school. These are produced by computing the sum of the school effects and the residual, $\alpha' S_i + \varepsilon_i$, for each student, and then computing the average of this variable by year, school, and subgroup. We then shrink these measures using a multivariate shrinkage approach that considers correlations in school- and subgroup-level value-added across subgroups and across years. After shrinkage, the subgroup measures are re-centered for consistency so the average of school growth across the subgroups, weighted by the number of students in each subgroup, is equal to the school's overall value-added.

To produce the group results by school for the proficiency subgroups, we regress the sum of the school effects and residual, $\alpha' S_i + \varepsilon_i$, on same-subject, once-lagged prior achievement within each school. This regression is estimated in a way that accounts for measurement error in prior achievement, using approaches described in section 2.5 of Fuller (1987), and is estimated separately for growth in 2017-18 and in 2018-19. This regression produces a separate intercept and slope for each school for each year, with the intercept measuring the school's effect on a student with average prior achievement and the slope measuring the school-specific relationship between student growth and prior achievement within the school. We then shrink these intercepts and slopes using a multivariate shrinkage approach that considers correlations among the intercepts and slopes both with each other and over time. After

[3] English-language learners includes students who reached English language proficiency in the last four years.

shrinkage, the intercepts are re-centered for consistency so that school growth at average prior achievement within the school is equal to the school's overall value-added. We then use the shrunk intercepts and slopes to produce school growth measures for each year for a representative non-proficient student, evaluated at a z-statistic of prior achievement of -0.67, and for a representative proficient student, evaluated at a z-statistic of prior achievement of +0.86. These scores corresponded to the average z-statistic scores, across grades and subjects, of non-proficient and proficient students in 2017-18.

To produce the group results by school for the target group subgroups, we estimate unshrunk value-added effects for 2016-17, 2017-18, and 2018-19 in the same way as they are produced for the non-proficiency subgroups (English-language learner, etc.) described earlier. These unshrunk value-added effects will generally be biased upward in the lower-scoring target group and biased downward in the higher-scoring target group. This is because the pretest assessment used to determine whether students are in the target group is inevitably measured with some degree of error. Some of the students assigned to the target group will have been assigned to the target group simply as a result of pretest measurement error with negative sign. Since we do not expect pretest measurement error to have any effect on the posttest, we expect these students to have higher measured growth, even if their actual growth in knowledge of the content being assessed is itself not higher. Similarly, some of the students who were not assigned to the target group will have been so assigned as a result of pretest measurement error with positive sign, which in turn will lead to lower measured growth given that pretest measurement error should have no effect on the posttest.

We adjust for this bias by subtracting from the unshrunk value-added effects an estimate of this bias, based on the standard error of measurement of the pretest assessment and an assumption that pretest assessment error is normally distributed. The adjustments are equal to:

$$adj\_target_k = -\lambda \frac{\sigma^2_{v(k)}}{\sqrt{\sigma^2_{y*(k)} + \sigma^2_{v(k)}}} \frac{\phi(z_k)}{\Phi(z_k)}$$

$$adj\_nontarget_k = +\lambda \frac{\sigma^2_{v(k)}}{\sqrt{\sigma^2_{y*(k)} + \sigma^2_{v(k)}}} \frac{\phi(z_k)}{\left(1 - \Phi(z_k)\right)}$$

where adj_target_k and $adj\_nontarget_k$ are added to the target and non-target group measures for school *k;* λ is the coefficient on the same-subject pretest in the previous year; $\sigma^2_{y*(k)}$ is an estimate of the variance in school *k* of same-subject pretest in the previous year adjusted for measurement error; $\sigma^2_{v(k)}$ is an estimate of the variance in school k of measurement error in the same-subject pretest in the previous year; z_k is the cutoff score in school k for inclusion in the target group given a normalized pretest; and $\phi(.)$ and $\Phi(.)$ are the standard normal probability density and cumulative distribution functions.

After making these adjustments, it is still not necessarily the case that the average of the unshrunk growth measures across schools within the target or non-target group was equal to zero. We made a further adjustment that subtracted the mean across schools by target or non-target group from the target and non-target group measures to ensure that this was the case. The unshrunk growth measures by target and non-target group were shrunk using a bivariate shrinkage approach that takes into account the correlation of growth within schools between the target and non-target group. This step was implemented to control for noise in the estimation of target/non-target group effects. The shrunk growth measures were then re-centered within school to ensure that the average of school growth across the target and non-target groups, weighted by the number of students in the two groups, averaged to the school's overall growth measure. This latter adjustment ensured that the growth estimates for the target and non-target group estimates were consistent with the reported overall growth measures.

We compute district-level measures for the target and non-target groups by averaging the analogous school-level measures across schools within the district. We do not include in district-level measures for the target and non-target groups students who were not enrolled in a school for the full academic year. This is because the target group is defined by students' prior achievement level relative to other students within their school.

Final stage for estimation of school and district value-added results

MULTI-YEAR AGGREGATION

Final estimates of school value-added effects are measured as a weighted moving three-year average of estimates for 2016-17, 2017-18, and 2018-19. The weights used are equal to the number of students in the school's value-added measure, multiplied by 1.5 for 2018-19, 1.0 for 2017-18, and 0.5 for 2016-17. The averaged value-added measure includes the 2016-17 and/or 2017-18 value-added measures only if there are at least twenty (in the case of subgroup measures, ten) students associated with that specific year's value-added measure. The multi-year average value-added measures are rescaled, based on the number of years included, to have a variance similar to that of a single-year value-added measure.

CALCULATING DISTRICT-LEVEL SCORES

Final estimates of district value-added effects are obtained by averaging the shrunk combined value-added estimates (as described above) for all the schools in each district, with weights determined by the number of students in each school in 2018-19. As mentioned earlier, the district results include students if they were FAY at the district even if they were not FAY at any of the district's schools. Thus, students who moved from one school in a district to another school in the district are included. These students are incorporated into the estimation of the model using a fixed effect estimate for a placeholder school for each district for students who were FAY in the district but not FAY in any school in the district.

PROPERTIES OF THE VALUE-ADDED RESULTS

Coefficients on student-level variables in the model

The coefficients estimated in the value-added model are presented in Tables 4 and 5. To interpret these coefficients, note that both pretest and posttest are measured using standardized scores; therefore, all coefficients are measured in the posttest standard deviation scale. For example, note that the coefficient on female gender is -0.055 in grade 4 Math. The posttest standard deviation for grade 4 Math is 51.114 (see Table 1). This implies that male students improved about 2.811 scale score points more on the grade 4 Math test from spring to spring than otherwise similar female students.

It is important to keep in mind the standard errors of the coefficients when interpreting them. A span of 1.96 standard errors in both the positive and negative directions provides a 95 percent confidence range for a coefficient. Continuing with the example of the coefficient on female gender in grade 4 Math, note that the standard error of this coefficient estimate is 0.004 in posttest SD units or 0.204 in scale score points. This means that, while our best estimate of the difference in growth between female and male students is -2.811 scale score points, a 95 percent confidence interval for the difference ranges from -3.211 to -2.411 scale score points.

Table 4. Coefficients on Student-Level Variables, 2018-19 Math

	Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 9		Grade 10		Grade 11	
Variable	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**
Math Pretest (lag 1)	0.785	0.007	0.479	0.011	0.619	0.010	0.550	0.011	0.321	0.012	0.490	0.011	0.604	0.008	0.633	0.013
ELA Pretest (lag 1)	0.092	0.006	0.148	0.010	0.116	0.009	0.182	0.012	0.266	0.012	0.104	0.010	0.133	0.009	-0.050	0.013
Math Pretest (lag 2)	-	-	0.359	0.010	0.300	0.009	0.253	0.010	0.542	0.013	0.446	0.010	0.194	0.007	0.280	0.012
ELA Pretest (lag 2)	-	-	-0.073	0.010	-0.003	0.009	-0.006	0.012	-0.145	0.013	-0.017	0.010	0.017	0.009	0.094	0.012
ELP Level 1	-0.133	0.032	-0.130	0.067	0.110	0.060	0.252	0.055	0.182	0.055	0.313	0.054	-0.067	0.104	0.215	0.082
ELP Level 2	-0.052	0.020	0.069	0.033	0.103	0.035	0.078	0.028	0.282	0.028	0.241	0.032	0.118	0.038	0.133	0.037
ELP Level 3	0.026	0.013	0.018	0.015	0.117	0.018	0.008	0.016	0.152	0.019	0.185	0.020	0.050	0.019	0.032	0.019
ELP Level 4	0.076	0.016	0.081	0.014	0.054	0.015	0.002	0.019	0.086	0.022	0.042	0.022	0.047	0.020	0.032	0.021
ELP Level 5	0.182	0.087	0.015	0.058	-0.054	0.054	0.078	0.066	0.121	0.069	-0.007	0.098	-0.097	0.091	-0.038	0.090
ELP Level 6	0.060	0.020	0.040	0.015	0.025	0.013	0.026	0.013	0.000	0.013	-0.024	0.013	-0.004	0.012	-0.004	0.012
Female	-0.055	0.004	0.005	0.005	0.042	0.005	-0.034	0.005	0.056	0.005	0.025	0.005	-0.021	0.005	-0.104	0.005
Asian	0.038	0.013	0.092	0.014	0.054	0.014	0.002	0.014	0.130	0.015	0.043	0.014	0.009	0.013	0.025	0.013
African-American	-0.098	0.011	-0.021	0.011	0.005	0.011	-0.071	0.011	0.026	0.012	0.043	0.012	-0.040	0.012	-0.019	0.011
Hispanic	-0.019	0.009	0.002	0.009	0.010	0.009	-0.009	0.009	-0.019	0.010	-0.013	0.010	-0.021	0.009	-0.009	0.009
American Indian or Alaskan Native	-0.029	0.026	-0.035	0.025	0.035	0.025	-0.050	0.024	-0.034	0.026	-0.028	0.025	-0.048	0.024	-0.067	0.023
Native Hawaiian or Other Pacific Islander	0.023	0.083	0.093	0.089	-0.109	0.097	-0.092	0.083	0.164	0.083	0.106	0.095	-0.134	0.094	0.092	0.086
Two or More Races	-0.005	0.011	-0.008	0.011	-0.010	0.012	-0.026	0.012	-0.033	0.013	-0.001	0.013	-0.026	0.013	-0.006	0.013
Special Education EBD	-0.107	0.018	-0.101	0.018	-0.073	0.018	0.030	0.018	-0.055	0.019	0.033	0.020	-0.031	0.020	0.060	0.021

	Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 9		Grade 10		Grade 11	
Special Education LD/ID	-0.115	0.012	-0.122	0.011	0.043	0.011	-0.007	0.011	0.148	0.011	0.130	0.011	0.067	0.011	0.090	0.011
Special Education A	-0.093	0.020	-0.128	0.020	-0.059	0.021	0.094	0.021	0.142	0.022	0.044	0.022	-0.014	0.022	0.015	0.022
Special Education SL	-0.027	0.013	-0.034	0.015	0.032	0.019	0.020	0.023	0.065	0.031	0.070	0.040	0.071	0.038	-0.061	0.041
Special Education Other	-0.127	0.012	-0.154	0.012	-0.026	0.012	0.032	0.012	0.086	0.013	0.076	0.013	0.023	0.013	0.014	0.013
Economic Disadvantage	-0.034	0.005	-0.024	0.005	-0.004	0.005	0.008	0.005	-0.010	0.006	-0.032	0.006	-0.039	0.005	-0.041	0.005
Migrancy Status	-0.081	0.156	0.345	0.203	-0.112	0.241	0.162	0.267	-0.087	0.214	-0.133	0.272	0.260	0.203	-0.236	0.158

Table 5. Coefficients on Student-Level Variables, 2018-19 ELA

	Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 9		Grade 10		Grade 11	
Variable	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**	**Coeff.**	**SE**
Math Pretest (lag 1)	0.076	0.007	0.038	0.008	0.034	0.006	0.055	0.009	0.043	0.006	0.031	0.006	0.045	0.007	0.062	0.012
ELA Pretest (lag 1)	0.815	0.007	0.585	0.010	0.614	0.010	0.676	0.013	0.650	0.012	0.576	0.010	0.714	0.009	0.526	0.012
Math Pretest (lag 2)	-	-	0.022	0.008	0.045	0.006	0.001	0.008	-0.010	0.009	0.053	0.006	0.013	0.006	0.043	0.011
ELA Pretest (lag 2)	-	-	0.284	0.010	0.212	0.010	0.235	0.013	0.273	0.014	0.227	0.010	0.166	0.008	0.309	0.011
ELP Level 1	0.022	0.034	-0.038	0.062	-0.057	0.057	0.101	0.052	0.101	0.050	0.180	0.060	0.297	0.126	0.156	0.087
ELP Level 2	0.006	0.021	0.052	0.030	0.085	0.034	0.098	0.027	0.162	0.026	0.110	0.031	0.073	0.036	0.097	0.037
ELP Level 3	0.026	0.014	-0.011	0.014	0.062	0.017	0.014	0.015	0.076	0.017	0.041	0.019	0.076	0.018	0.073	0.018
ELP Level 4	0.052	0.016	0.018	0.013	0.047	0.014	0.026	0.018	0.093	0.020	-0.038	0.020	0.053	0.019	0.028	0.020
ELP Level 5	0.187	0.092	0.017	0.054	-0.004	0.052	0.093	0.064	-0.017	0.062	0.017	0.090	0.137	0.082	0.000	0.084
ELP Level 6	0.057	0.021	0.057	0.014	0.083	0.012	0.039	0.012	0.045	0.012	-0.017	0.012	0.035	0.011	-0.019	0.011
Female	0.063	0.005	0.091	0.004	0.114	0.005	0.034	0.005	0.049	0.005	0.161	0.005	0.075	0.004	-0.042	0.004
Asian	-0.004	0.014	0.029	0.013	0.075	0.013	0.083	0.013	0.077	0.013	0.088	0.013	0.058	0.012	-0.007	0.012
African-American	-0.033	0.012	-0.007	0.011	-0.030	0.011	-0.010	0.011	-0.002	0.011	-0.042	0.011	-0.029	0.011	-0.041	0.011

	Grade 4		Grade 5		Grade 6		Grade 7		Grade 8		Grade 9		Grade 10		Grade 11	
Hispanic	0.005	0.010	0.003	0.009	-0.014	0.009	0.005	0.009	0.001	0.009	-0.038	0.009	-0.003	0.008	-0.028	0.008
American Indian or Alaskan Native	-0.047	0.027	0.012	0.024	-0.027	0.023	-0.028	0.023	-0.035	0.024	0.003	0.024	0.011	0.022	-0.070	0.021
Native Hawaiian or Other Pacific Islander	-0.054	0.087	0.145	0.082	-0.093	0.092	-0.056	0.080	0.171	0.075	0.036	0.088	-0.029	0.089	0.002	0.081
Two or More Races	0.000	0.012	-0.008	0.011	-0.021	0.011	0.022	0.012	-0.013	0.012	-0.017	0.012	0.004	0.012	-0.019	0.012
Special Education EBD	-0.129	0.019	-0.134	0.017	-0.195	0.017	-0.016	0.018	-0.010	0.017	-0.079	0.019	0.013	0.019	0.060	0.021
Special Education LD/ID	-0.073	0.013	-0.115	0.011	-0.097	0.011	0.020	0.011	0.034	0.011	-0.149	0.011	0.047	0.011	0.142	0.010
Special Education A	-0.104	0.021	-0.119	0.019	-0.050	0.020	0.150	0.020	0.127	0.020	-0.027	0.021	0.085	0.020	0.043	0.021
Special Education SL	-0.013	0.013	-0.018	0.014	-0.008	0.018	0.057	0.022	0.045	0.028	-0.056	0.037	0.037	0.036	0.044	0.039
Special Education Other	-0.123	0.013	-0.092	0.012	-0.118	0.012	0.043	0.012	0.026	0.012	-0.100	0.012	0.048	0.012	0.064	0.013
Economic Disadvantage	-0.047	0.006	-0.037	0.005	-0.032	0.005	-0.016	0.005	-0.025	0.005	-0.049	0.005	-0.025	0.005	-0.065	0.005
Migrancy Status	0.206	0.163	0.314	0.188	0.180	0.228	0.174	0.256	0.048	0.194	0.271	0.250	0.276	0.184	-0.130	0.157

Test of model neutrality: Correlation with average prior attainment

In this test, we calculate correlations between growth estimates and school-level prior attainment. This is a method for validating whether the variables included on the right-hand side of our regression adequately control for school-level factors influencing growth estimates. The higher the correlation magnitude, the higher the level of "non-neutrality".

Our results show a low correlation at the school-and-grade level and a modest correlation at the overall school level between average prior attainment--a measure of average performance in the previous year--and value-added. In general, schools were somewhat more likely to have a high value-added score than a low score if their students began the year with high pretest scores rather than low scores.

Table 6. Correlations between Prior Attainment and Value-Added

Subject	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Overall
ELA	0.059	0.016	0.009	-0.123	-0.032	0.122	0.32	-0.061	0.199
Math	0.256	-0.029	-0.028	-0.18	0.112	0.098	0.362	0.001	0.266

Correlation between Math and ELA value-added

There were substantive positive correlations between math and ELA value-added within each school. Schools that were high value-added in math were also more often than not high value-added in ELA. This implies that schools with a higher-than-average impact in mathematics also had a higher-than-average impact in English language arts.

Table 7. Correlations between Subjects

Subject	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11	Overall
2018-19 Math and ELA	0.543	0.509	0.624	0.451	0.373	0.644	0.65	0.624	0.58

CONTACT

For more information, contact the Principal Investigator for this project, Dr. Robert Meyer, at rhmeyer@edanalytics.org.

REFERENCES

Fuller, W. (1987). *Measurement Error Models*, John Wiley and Sons.

Longford, N. T. (1999). Multivariate shrinkage estimation of small area means and proportions. *Journal of the Royal Statistical Society* 162 (Part 2), 227-245.