Los Angeles Unified School District

Today's Learners, Tomorrow's Leaders

Academic Growth over Time: Technical Report on the LAUSD School-Level AGT Model Academic Year 2010-2011

Written by

Value-Added Research Center

Wisconsin Center for Education Research

University of Wisconsin-Madison

INTRODUCTION

This technical report describes Academic Growth over Time (a value-added model) used by the Los Angeles Unified School District (LAUSD) and developed in association with the Value-Added Research Center (VARC) of the Wisconsin Center for Education Research at the University of Wisconsin. The report is in four parts. The first part describes the data set used to produce the Academic Growth over Time (AGT) estimates. The second part describes the model used to estimate AGT for teachers and schools in LA. The third part describes the reporting of AGT. Finally, the fourth part presents the results of analyses of the properties of the AGT results.

Conceptually, AGT analysis is the use of statistical technique to isolate the component of measured student knowledge that is attributable to schools, teachers, or classrooms from other factors such as prior knowledge, student and classroom characteristics. In practice, AGT models focus on the improvement students make on annual assessments from one year to the next. AGT models often control for measurable student characteristics using available data such as race, income, and disability, and measurable classroom characteristics, to help isolate the impact of schooling. The model used in Los Angeles uses a large set of student and classroom characteristics to identify the extent to which schools contribute to the improvement of student achievement outcomes in their classrooms.

This document explains the technical details of the school level component of the AGT system that LAUSD is developing with VARC. A companion document will explain the classroom/teacher level component of the AGT system in the coming months.

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 AGT estimates. A separate analysis data set is produced for each grade, subject, and year. In total, eighty-one analysis data sets are produced, covering nine grades for English language arts (ELA) (third, fourth, fifth, sixth, seventh, eighth, ninth, tenth, and eleventh), six grades in Mathematics (third, fourth, fifth, sixth, seventh, and eighth) and twelve subjects added Fall 2011 (Algebra I, Algebra II, Geometry, Biology, Chemistry, Physics, Integrated Science I, Science Grade 5, Science Grade 8, History and Social Science Grade 8, US History, World History), over four years (2008-09, 2009-10, 2010-11). The analysis data sets include students with a posttest and pretest in consecutive grades in the same subject who could be assigned to a school, classroom, and teacher for that subject.

The analysis data set on which the AGT model is run includes both student-level and classroom-level variables. Variables at the student level provide information about individual students, while variables at the classroom level provide information about the classrooms students are in (including the average characteristics of the students in the classroom).

Student-level variables

Posttest and pretest variables

The test scores used in the data set are scores from the following California Standards Tests (CST) examinations:

- ELA in grades 2-11
- General Mathematics in grades 2-8
- Algebra I
- Algebra II
- Geometry
- Science Grade 5
- Science Grade 8
- Integrated Science I
- Biology
- Chemistry
- Physics
- History and Social Science Grade 8
- US History
- World History

For the AGT analysis, scale scores were converted into z-scores, which have a mean of 0 and a standard deviation of 1 across the district. Scale scores in math, ELA, science, and social studies are normalized within grade and year into z-scores by subtracting from scale scores the within-subject, within-grade, within-year mean and dividing that result by the within-subject, within-grade, within-year standard deviation. The normalization takes place across all students in the city with test scores. After z-scores are computed, duplicate observations are handled by dropping all observations with duplicate student IDs except that with the highest z-score.

In this year's analysis, the data included students who were continuously enrolled in the same school from the statewide school census date in October (CBEDS) through the date of testing in the spring (typically, May). These are the students who are also part of a school's Academic Performance Index (API) calculation. The results also only included students who could be associated with a particular classroom. Please see the Frequently Asked Questions for more information on these and other matters at http://agt.lausd.net.

The AGT system produces school measures for grades 3-8 in ELA using the prior year's CST in ELA and Mathematics. 9th ELA is produced using only 8th grade ELA as a pretest since students may have taken Algebra I or General Mathematics in grade 8. In Math results are produced for grades 3-7 including the prior year Mathematics and ELA scores. For grade 8 the posttest is either Algebra I or General Mathematics. These subjects are broken up into separate analyses as there is no current way to meaningfully combine separate subjects into one measure. Both of these analyses use prior Mathematics and ELA CST scores. Non-NCLB subject AGT results are produced using prior math and ELA CST scores, as well as other in-subject prior tests. Please see Appendix One for the pre-tests used for non-NCLB subjects.

Standard errors of measurement of pretest variables

The standard errors of measurement (SEM) of math and ELA z-scores are set to 1 minus the square root of Cronbach's alpha. Cronbach's alpha is available in the technical reports of the California Standards Tests Technical Report produced by the California Department of Education, Assessment and Accountability Division. Given the use of an unconditional measurement error measure, every student in the same year and grade has the same SEM for a given test. The standard errors of measurement are used for a correction for measurement error in the pretest. It is presumed that the covariance between the measurement errors of math and ELA pretests is zero.

Gender, race, and free- and reduced-price lunch

Gender, race, and free- and reduced-price lunch are drawn from the student biographical dataset. In the analysis data set, students are assigned the gender, race, and low-income status reported in the posttest year. Gender categories are male and female. Race categories are Asian, African American, Hispanic and White. Those students that do not have demographic data are rare, but are accounted for through a missing category. The states of free and reduced lunch status are concatenated in the data that LAUSD maintains and as such the analysis only considers free/reduced lunch status or not.

English language learner (ELL)

There are four categories of ELL status in the data: English Origin (EO), reclassified English language learner (RFEP), English language learner (LEP), and English as a second language (IFEP). IFEP are students that enter into the LAUSD system as proficient in English and have parents that speak a language other than English at home. RFEP is the designation for those ELL students that have been proficient on the CEDLT test three years in a row and are now considered proficient in English.

Disability

Students are categorized into two types of disability if they are listed in the special education data file. Specific learning disability or speech-language impaired were considered mild disabilities. All others (including autism, mental retardation, and traumatic brain injury) were considered severe. These students are only considered in the AGT framework if they took the CST two years in a row, so this is a strict subset of all disabled students in the district.

Homelessness

LAUSD tracks homelessness in its student information system. This variable is tracked by school staff and reported to the central office.

Classroom-level variables

Classroom means of variables in the student-level model

The student-level variables (including the pretests, though not including the posttests or the standard errors of measurement of the pretests) are averaged by classroom attended in the posttest year. The average pretest scores by classroom only include students for whom pretest scores are available. It does not include students for whom data is missing.

Teacher, classroom and school

Student-teacher-classroom-school link

Students were assigned to teachers and to schools using the marks data maintained by LAUSD. The data set linked each student to a math classroom and to an ELA classroom, linked each math classroom to a math teacher, and linked each ELA classroom to an ELA teacher. As a result, classroom is nested within teacher. Schools are given the opportunity to verify these data annually. The marks data also links students and teachers to schools. Teacher is not nested within school; it is possible in the input data set for teachers to teach in multiple schools.

Students are assigned to the school to which they are assigned in the marks data set. Students are assigned to the classroom within the school to which they are assigned in the student-teacher link data set with some exceptions. Students with blank course names are considered not assigned to a classroom. Students who are not assigned to a school are not assigned to a classroom.

Students are only included in analysis if they are successfully assigned to a school, classroom, and teacher. Students not assigned to a school, classroom, or teacher are not included in the AGT analysis. The following tables describe the sample used for the 2011 year:

English Language Arts Demographics Makeup 2011

CST	Asian	African American	Hispanic	White	FRL	Female	ELL-EO	ELL-IFEP	ELL-LEP	ELL-RFEP	SPED-Mild	SPED-Severe
ELA (GR 3)	7%	8%	75%	10%	83%	50%	35%	11%	32%	22%	5%	2%
ELA (GR 4)	7%	8%	75%	10%	83%	50%	34%	10%	27%	29%	4%	2%
ELA (GR 5)	7%	9%	75%	10%	83%	50%	34%	15%	18%	33%	5%	2%
ELA (GR 6)	7%	8%	74%	10%	82%	51%	34%	17%	10%	39%	5%	2%
ELA (GR 7)	6%	9%	77%	9%	82%	50%	31%	16%	11%	42%	6%	2%
ELA (GR 8)	8%	9%	71%	12%	76%	48%	32%	12%	19%	37%	10%	3%
ELA (GR 9)	7%	8%	75%	10%	71%	49%	27%	10%	16%	45%	7%	2%
ELA (GR 10)	5%	7%	80%	8%	73%	46%	26%	9%	22%	42%	11%	3%
ELA (GR 11)	10%	9%	71%	10%	73%	51%	28%	10%	12%	50%	6%	1%

Mathematics Demographics Makeup 2011

CST	Asian	African American	Hispanic	White	FRL	Female	ELL-EO	ELL-IFEP	ELL-LEP	ELL-RFEP	SPED-Mild	SPED-Severe
MATHEMATICS (GRADE 3)	5%	9%	76%	10%	92%	35%	34%	5%	53%	6%	63%	22%
MATHEMATICS (GRADE 4)	4%	10%	76%	9%	83%	29%	35%	5%	56%	3%	66%	23%
MATHEMATICS (GRADE 5)	3%	6%	86%	5%	94%	40%	22%	5%	59%	14%	59%	12%
MATHEMATICS (GRADE 6)	3%	1%	93%	3%	94%	43%	3%	1%	91%	4%	12%	2%
MATHEMATICS (GRADE 7)	3%	2%	92%	3%	93%	43%	4%	1%	92%	3%	17%	2%
MATHEMATICS (GR 8)	6%	4%	85%	5%	85%	43%	8%	3%	82%	5%	18%	4%
ALGEBRA I	7%	8%	78%	8%	78%	49%	25%	11%	19%	44%	6%	2%
ALGEBRA II	10%	7%	74%	9%	73%	52%	24%	11%	9%	55%	3%	1%
GEOMETRY	6%	7%	84%	4%	79%	40%	14%	4%	61%	20%	9%	2%

Science Demographics Makeup 2011

CST	Asian	African American	Hispanic	White	FRL	Female	ELL-EO	ELL-IFEP	ELL-LEP	ELL-RFEP	SPED-Mild	SPED-Severe
SCIENCE (GR 5)	4%	14%	74%	8%	87%	42%	40%	9%	40%	11%	49%	13%
SCIENCE (GR 8)	3%	17%	78%	3%	92%	36%	24%	14%	43%	16%	11%	5%
INTEGRATED SCIENCE 1	3%	20%	73%	4%	77%	37%	30%	3%	55%	12%	52%	14%
BIOLOGY	9%	8%	75%	8%	73%	49%	26%	10%	17%	47%	7%	2%
CHEMISTRY	10%	9%	72%	9%	72%	51%	27%	11%	13%	49%	4%	1%
PHYSICS	13%	6%	70%	11%	81%	51%	24%	12%	8%	56%	4%	1%

Social Studies Demographics Makeup 2011

CST	Asian	African American	Hispanic	White	FRL	Female	ELL-EO	ELL-IFEP	ELL-LEP	ELL-RFEP	SPED-Mild	SPED-Severe
HISTORY & SOCIAL SCIENCE (GR 8)	3%	11%	81%	5%	87%	37%	29%	4%	57%	10%	51%	11%
US HISTORY	5%	10%	81%	4%	75%	45%	25%	5%	39%	31%	31%	6%
WORLD HISTORY	6%	13%	78%	4%	78%	37%	24%	4%	53%	14%	32%	6%

ACADEMIC GROWTH OVER TIME MODEL

For the LAUSD school level model, Academic Growth over Time (AGT) is measured in math in grades three through eight, English Language Arts (ELA) in grades three through eleven, and the secondary level subjects. Schools receive single-year AGT measures that reflect student growth in 2010-11 as well as multiple-year AGT measures that reflect student growth over as many as three years. AGT results were also computed for student subgroups within the school such as students with disabilities, English language learners, gender, free/reduced lunch status and students in certain proficiency categories of the CST based on prior achievement. The model measures average achievement among a teacher's students, controlling for prior achievement in both math and ELA and a large number of student and classroom characteristics.

The elementary-level model, in brief

The AGT model for elementary subjects (math and ELA grades three through eight, science grades five and eight, history and social science grade eight) is defined by four equations: a "best linear predictor" AGT model defined in terms of true student post and prior achievement and three measurement error models for observed post and prior achievement:

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

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

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

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

where:

- y_{1i} is true post achievement;
- y_{0i} and y_{0i}^{alt} are true prior achievement in the same subject and in the other subject (math in the ELA model, ELA in the math model), with slope parameters λ and λ^{alt};
- X_i is a vector of characteristics of student *i*, with slope parameter vector β;
- Z_i is a vector of characteristics of student *i*'s classroom, with slope parameter vector γ;
- S_i is a vector of school indicators;
- α is a vector of school value-added effects (where α_k is the value-added effect for school *k*);
- e_i is the error in predicting post achievement given the explanatory variables included in the model;
- Y_{1i} is measured post achievement;
- v_{1i} is measurement error in post achievement;
- Y_{0i} and Y_{0i}^{alt} are measured prior achievement; and
- v_{0i} and v_{0i}^{alt} are measurement error in prior achievement.

Substituting the measurement error equations (2), (3), and (4) into the student achievement equation (1) yields an equation defined in terms of measured student achievement:

$$\text{Measured achievement: } Y_{1i} = \zeta + \lambda Y_{0i} + \lambda^{alt} Y_{0i}^{\ alt} + \beta' X_i + \gamma' Z_i + \alpha' S_i + \varepsilon_i \qquad (5)$$

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

$$\text{Error in measured achievement: } \varepsilon_i = e_i + v_{1i} - \lambda v_{0i} - \lambda^{alt} v_{0i}^{\ alt} \qquad (6)$$

Estimating the measured student achievement equation (5) without controlling for pretest measurement error yields biased estimates of all parameters, including the value-added teacher effects. This bias stems from the fact that measurement error in prior achievement causes the error term (6), which includes the measurement error components 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 Wayne Fuller, *Measurement Error Models* (Wiley, 1987). Information about the variance of test measurement error is reported in the technical manuals for the CST.

When estimating the teacher effects, a shrinkage approach is employed to ensure that schools with fewer students are not overrepresented among the highest- and lowest-value-added teachers due to randomness. The approach, Empirical Bayes shrinkage, is described in J. N. K. Rao, *Small Area Estimation* (Wiley, 2003).

Not only are overall school effects estimated, but so are school effects for student subgroups. These effects are produced by extending the above model to allow for schools to have different effects for students with different characteristics. These extensions make it possible to produce school wide AGT by pretest score, gender, ethnicity, English language learner, and disability.

The secondary-level model, in brief

The AGT model for secondary level subjects (Algebra I, Algebra II, Geometry, Biology, Chemistry, Physics, Integrated Science I, US History, and World History) allows us to produce estimates for teachers whose students took the same posttest (for example, Physics), but different pretests (for example, Biology or Chemistry).

$$y_{1i} = \lambda_A \mathrm{I}_i^A y_{0Ai} + \lambda_B \mathrm{I}_i^B y_{0Bi} + \gamma \mathrm{I}_i^A + \beta X_i + \alpha' C_i + \varepsilon_i$$

where:

- y_{1i} is achievement on the posttest;
- y_{0Ai} is achievement on pretest A;
- y_{0Bi} is achievement on pretest B;
- I_i^A is an indicator variable that equals 1 if student *i* took pretest A and 0 if student *i* did not take pretest A;
- I_i^B is an indicator variable that equals 1 if student *i* took pretest B and 0 if student *i* did not take pretest B;
- γ represents the average difference in posttest score between students who took pretest A and students who took pretest B
- X_i is a vector of student characteristics of student *I*, with slope parameter vector β;
- C_i is a vector of classroom indicators;
- α is a vector of classroom AGT effects (where α_j is the AGT effect for classroom *j*);
- ε_i is the error in predicting post achievement given the explanatory variables included in the model.

The above model is the most basic example of a value added model for students who share a posttest but took different pretests for a single subject. Many models not only include multiple possible pretests for one subject but also pretests from other subjects. For example, a model for students taking the Algebra 2 posttest could include pretests from Algebra 1 and Geometry as well as the ELA8 and ELA9 pretests to account for students that take different combinations of pretests.

In the above model, both pretest variables, y_{0Ai} and y_{0Bi}, are multiplied by indicator variables, I_i^A and I_i^B, respectively, that are one if student *i* took the relevant posttest and zero if not. λ_A and λ_B represent how each posttest varies with pretest A and pretest B, respectively. These coefficients do not affect the predicted score when a student did not take the pretest. For example, if student *i* took pretest B, $\mathrm{I_i}^{\mathrm{A}}$ would equal zero. Therefore, λ_A would be multiplied by zero and not affect the predicted score.

Students who take different pretests may be different from each other. For students taking the Algebra 2 posttest, we might expect students who took Geometry in the previous grade to differ from those who took Algebra 1. To control for average differences in posttest scores between students who took different pretests, the model includes the term $\gamma\,\mathrm{I_i}^{\mathrm{A}}$.

Similarly, students in different grades taking the same posttest may be different from each other, necessitating additional control variables for grade. In these models, however, grade is collinear or near collinear with the ELA pretest. For example, most tenth graders only have pretest scores for ELA9 and almost never have pretest scores for ELA8. Therefore, the ELA pretest variables soak up average differences between students of different grades

The variables in the model

The student-level variables included in the model (the X variables in equation 1) include gender, race, English language learner (English Origin (EO), English as a second language (IFEP), English language learner (LEP), reclassified (RFEP)), free-and reduced-price lunch, disability (severe and mild) and homelessness. The classroom-level variables included in the model (the Z variables in equation 1) include classroom averages of pretests and the student-level variables in X.

Stage one regression (student-level regression)

The value-added regression is run in three stages. The first stage estimates the coefficients λ on the pretests after correcting for test measurement error. It regresses posttest on same-subject pretest, other-subject pretest, student-level variables, and a full set of classroom fixed effects. This can be expressed mathematically as:

$$Y_{1i} = \lambda Y_{0i} + \lambda^{alt} Y_{0i}^{alt} + \beta' X_i + \alpha^{*\prime} C_i + \varepsilon_i$$

where C_i is a vector of classroom dummies that affect posttest with parameters α^*. For a given classroom c, α_c^* is equal to $\zeta + \gamma' Z_c + \alpha_j$, where Z_c is the characteristics of classroom c and α_j is the value added of school j in classroom c.

This regression is estimated using an approach that accounts for measurement error in the pretests Y_{0i} and Y_{0i}^{alt}. Recall from equation (6) above that the measurement error components of Y_{0i} and Y_{0i}^{alt}, v_{0i} and v_{0i}^{alt}, are part of the error term ε_i. As a result, estimating the regression using ordinary least squares 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:

$$Y_t = Y_{t-1}\lambda + W\delta + \varepsilon$$

where Y_t is an N × 1 vector of post-test scores, Y_{t-1} is an N × 2 vector of same-subject and other-subject pre-test scores Y_{t-1} and Y_{t-1}^{alt}, λ is a 2 × 1 vector made up of λ and λ^{alt}, W is an N × K vector of the X demographic variables, 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:

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

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

$$\begin{bmatrix} \hat{\lambda}_{CORR} \\ \hat{\delta}_{CORR} \end{bmatrix} = \begin{bmatrix} Y'_{t-1}Y_{t-1} - \sum_{i}^{N} sem_{it-1} & Y'_{t-1}W \\ W'Y_{t-1} & W'W \end{bmatrix}^{-1} \begin{bmatrix} Y'_{t-1}Y_t \\ W'Y_t \end{bmatrix}$$

where $sem_{it\text{-}1}$ is a 2 × 2 variance-covariance matrix of the errors of measurement of $Y_{it\text{-}1}$ and $Y_{it\text{-}1}{}^{alt}$ for student i. This model is described in section 2.2 of Wayne Fuller, *Measurement Error Models* (Wiley, 1987).

Stage two regression (classroom-level regression)

The second stage regression re-estimates the coefficients β on the student level variables and estimates the coefficients γ on the classroom-level variables. Let $q_{1i} = Y_{1i} - \lambda Y_{0i} - \lambda^{alt} Y_{0i}{}^{alt}$. Then we can express the second-stage regression mathematically as:

$$q_{1i} = \zeta + \beta' X_i + \gamma' Z_i + w_i$$

where w_i is equal to $\alpha' S_i + \varepsilon_i$. When estimating this regression, we have to use as our left-hand-side variable an estimate of q_{1i}, which is computed using the estimates of λ and λ^{alt} from the first-stage regression. When this regression is run, it takes into account that the errors w_i are correlated within classrooms via $\alpha' S_i$ by specifying a classroom random effect.

Stage three regression (school-level AGT)

Now that all the other variables have been controlled for, the third-stage regression estimates the value added measures α_j. This can be expressed using the equation:

$$w_i = \alpha' S_i + \varepsilon_i.$$

where $w_i = Y_{1i} - \zeta - \lambda Y_{0i} - \lambda^{alt} Y_{0i}{}^{alt} - \beta' X_i - \gamma' Z_i$. When we estimate this regression, it is necessary to use an estimate of w_i, which is drawn from the residuals of the second-stage regression.

This is a very easy regression to estimate. All one needs to do is compute the average of w_i within school k to produce estimates $\hat{\alpha}_k$. Once this is done, compute estimates of the error term ε_i by subtracting $\hat{\alpha}_k$ from the estimate of w_i. The standard errors of the estimates $\hat{\alpha}_k$ are equal to the square root of the ratio of the sample variance of the estimates of ε_i to the number of observations for school k. The variance-covariance matrix of $\hat{\alpha}$ is diagonal, and the n-weighted mean of $\hat{\alpha}_k$ across schools is zero. It is important to note that the standard errors computed under this approach ignore error that comes from having used estimates of λ, β, and γ to control for pretests, student-level variables, and classroom-level variables rather than the true values of λ, β, and γ instead.

Single-year and multiple-year measures of AGT

The three-stage regression described above is run separately for each combination of grade, subject, and year over four years of data. This produces unshrunk single-year school-level value added estimates: $\hat{\alpha}_k$. When we wish to measure a multiple-year measure of value added, we run the first two stages of the regression separately by year. When we come to the third stage, we pool our estimates of w_i over multiple years and compute the multiple-year versions of $\hat{\alpha}_k$ over the pooled data using the same technique as if it were a single-year estimate.

Shrinkage of teacher-level AGT

The unshrunk value-added estimates $\hat{\alpha}_k$ are shrunk using a Empirical Bayes univariate shrinkage technique described in J. N. K. Rao, *Small Area Estimation* (Wiley, 2003).

The first step in shrinking the estimates $\hat{\alpha}_k$ is to estimate the variance of the true (rather than the estimated) school effects α_k. This is relatively straightforward. Let $\hat{\sigma}_k^2$ be the squared standard error of $\hat{\alpha}_k$. Also let $\hat{\omega}_{est}^2$ be the variance of $\hat{\alpha}_k$ across schools. We estimate the variance of α_k as $\hat{\omega}^2 = \hat{\omega}_{est}^2 - \bar{\hat{\sigma}}_k^2$, where $\bar{\hat{\sigma}}_k^2$ is the mean of $\hat{\sigma}_k^2$ across schools.

The second step in shrinking the estimates is to compute shrunk value-added estimates using simple Empirical Bayes shrinkage. This is accomplished by multiplying the unshrunk values added by their reliabilities. The estimated reliability of value added of school k is equal to $r_k = \hat{\omega}^2 / (\hat{\omega}^2 + \hat{\sigma}_k^2)$. Shrunk value added for school k is equal to $\hat{\alpha}_k^{EB} = r_k \hat{\alpha}_k$, and the standard error of shrunk value added is equal to $\hat{\sigma}_k^{EB} = r_k^{1/2} \hat{\sigma}_k$.

Subgroups

AGT is also estimated by subgroup. In a subgroup model, we assume that teachers have different effects for students with different characteristics. ELL is used as an example here, but the results generalize to special education, race, pretest category and gender. In the case of ELL, we replicate the student achievement model (1) with the following model:

$$y_{1i} = \zeta + \lambda y_{0i} + \lambda^{alt} y_{0i}^{alt} + \beta X_i + \gamma Z_i + \theta_0 S_i + \theta_1 [S_i \times (ELL_i - \mu_{ELL(k)})] + e_i \quad (1')$$

where θ_0 is a vector of S intercepts, θ_1 is a vector of S slopes, ELL_i is an indicator variable for student i being ELL and $\mu_{ELL(k)}$ is equal to the mean of ELL_i within school k.

When this is estimated, we impute the estimated ζ, λ, β, and γ from the first- and second-stage regressions in overall value added, leaving us with the estimated residual terms w_i previously used in computing overall value added. This residual term is regressed on (ELL_i - $\overline{ELL}_k$) within schools, where $\overline{ELL}_k$ is the sample mean of ELL_i among school k's students. This yields estimates of the intercept $\hat{\theta}_{0k}$ and slope $\hat{\theta}_{1k}$ for each school k. Because the subgroup variable has been interpreted as a deviation from a mean, the estimate of the intercept $\hat{\theta}_{0k}$ is equal to the unshrunk estimate of overall value added $\hat{\alpha}_k$. The measurement error in the slope term, $\hat{\theta}_{1k}$, will be uncorrelated with the measurement error in the intercept term $\hat{\theta}_{0k}$, except for a component that derives from the substitution of $\overline{ELL}_k$ for $\mu_{ELL(k)}$ that is ignored.

The slope terms $\hat{\theta}_{1k}$ are shrunk using a Emprical Bayes approach that is the same as that described above for overall value added. When the variance of θ_{1k} is estimated for shrinkage, schools for whom the standard error of $\hat{\theta}_{1k}$ is 0.5 or greater are excluded from the computation. These are badly measured estimates of $\ddot{\theta}_{1j}$ that in some cases lead to negative estimates of the variance of θ_{1k}. The slope terms $\hat{\theta}_{1k}$ are demeaned before shrinkage to have a mean of zero across schools within the group with a standard error small enough to be included in the variance computation.

From the shrunk overall value added estimate $\hat{\alpha}_k^{EB}$ and the shrunk slope $\hat{\theta}_{1k}^{EB}$ (both shrunk using Empirical Bayes), we compute value added among students both in and not in the subgroup. In the case of ELL, value added among ELL students for school k is equal to:

$$\hat{\alpha}_k^{EB} + \hat{\theta}_{1k}^{EB}\left(1 - \overline{ELL}_k\right)$$

with a squared standard error equal to the squared standard error of $\hat{\alpha}_k^{EB}$ plus $\left(1 - \overline{ELL}_k\right)^2$ times the squared standard error of $\hat{\theta}_{1k}^{EB}$. This presumes that, across schools, overall value added α_k and slope θ_{1k} are uncorrelated. Value added for non-ELL students for school k is equal to:

$$\hat{\alpha}_k^{EB} - \hat{\theta}_{1k}^{EB}\overline{ELL}_j$$

with a squared standard error equal to the squared standard error of $\hat{\alpha}_k^{EB}$ plus $\overline{ELL}_k^{\;2}$ times the squared standard error of $\hat{\theta}_{1k}^{EB}$.

REPORTING AGT

After the AGT analysis is completed, each school has a large number of results about the improvement of its students. Each school in the covered grades has a single-year overall AGT that covers 2010-11; a multiple-year overall AGT that covers 2008-09, 2009-10, and 2010-11; single-year AGT measures for 2010-11 specific to students with disabilities, ELL students, male students, female students, and students in the top, middle, and bottom thirds of the pretest distribution (by proficiency category); and multiple-year AGT measures that cover the same subgroups for all three periods. If a school has not been open long enough to have all three AGT measures, the average score is based off of the maximal amount of data. In the extreme case, a school only has one year of AGT. In that case the average estimate will not be reported.

These results are produced for each grade in the school and each subject where applicable. There is also an aggregate school wide measure that encompasses all of the grades in that school. This is meant to be a descriptive summary of the school since the tests in LAUSD are not vertically equated. In other words since the tests in different grades do not have a comparable scale, the school wide average is only an index of grade level AGT rather than a direct AGT measure. This is simply the sample size weighted average of the AGT measures across all of the grades in a particular subject in a school. Note that since most of the variation in AGT is within schools (across grades) the variance of the school level estimates will be smaller than the grade level estimates.

Confidence intervals

The AGT measure is our best estimate of the school's effects on its students given the data, and is often referred to as a *point estimate* given that it is a single number. However, every AGT measure is based on a finite number of students and, consequently, includes some error from randomness in the students a school has.

In reports, AGT is presented as a point estimate surrounded by a 95 percent confidence range. The maximum point within this range is equal to the AGT point estimate plus 1.96 times the standard error of AGT. The minimum point is equal to the point estimate minus 1.96 times the standard error of AGT. Values outside of this range can be rejected with 95 percent confidence as the teacher's AGT score.

Reporting AGT

School level AGT results are most useful when they are based on enough students to draw conclusions about the growth in a school. Consequently, AGT is only reported if a grade level had a sufficient number of students. In most cases the grade level AGT is suppressed if there are less than 11 students in the grade. For the differential effects estimates, the result is suppressed for a number of reasons:

- The subgroup has less than 11 students (if it is binary the opposing category will be suppressed as well)
- The subgroup is defined by the state as not having a significant number of students
- The subgroup does not have enough variance across the district to calculate the differentials as described above
- The subgroup does not have sufficient "balance" to calculate an effect. For instance if a school consists of 100% Hispanic students, there can be no comparison between other ethnicities since there is no data on how that school performs with other ethnicities.

AGT is normalized at the grade level by dividing each measure by the estimated variance of school AGT (as calculated in the empirical bayes step) and adding 3. This centers the AGT results around 3 with a variance that is biased downwards due to shrinkage. Since the reliability of the estimates is so high, this variance is not shrunk much and it was decided through the model co-build to not employ more advanced shrinkage techniques to inflate the variance.

Specific colors are used to indicate significance levels. If a final measure is significantly above 3, it is colored green; if it is significantly above 4 it is colored blue; if it is significantly below 3 it is colored yellow; if it is significantly below 2 it is colored red; and if it is not significantly different from 3 it is colored gray.

In a given school and year, correlations between grade level AGT is low. Since this is the case, when the AGT results across grades are averaged to the school level the variation of school level AGT across the district shrinks considerably. This was not re-normalized to reflect the fact that in a given school, some grades may be high and some may be low and that there is lower variation between schools than there is within a school. The next technical report (teacher level AGT) will explore the variance of AGT within schools as opposed to across schools.

PROPERTIES OF THE AGT RESULTS

Coefficients on student- and classroom-level variables in the model

The coefficients estimated in the AGT model for a single grade, subject, and year (grade 4 ELA for 2010-11) are presented below. To interpret the below coefficients, note that both pretest (3rd grade tests in math and ELA) and posttest (4th grade test in ELA) are measured using z-scores with a mean of 0 and a standard deviation of 1 across students in LAUSD. Consequently, all coefficients are measured in student-level standard deviations. For example, note that the coefficient on female gender is 0.093. This implies that female students improved 0.093 standard deviations more on the CST ELA test from 2009-10 to 2010-11 than otherwise similar male students. On the fourth grade ELA test in 2011, a standard deviation is equal to about 58 scale score points.

Coefficients on student-level variables, 4th grade ELA, 2010-11

Variable	***Coeff.***	***Std. Err.***
ELA pretest	0.568	0.005
Math pretest	0.171	0.005
Homeless	0.033	0.018
IEP Severe	-0.140	0.021
IEP Mild	-0.186	0.014
ELL- RFEP	0.062	0.062
ELL- LEP	-0.164	0.009
ELL-IFEP	0.058	0.010
Free/Reduced Lunch	-0.076	0.010
White	0.062	0.012
African American	-0.072	0.012
Asian	0.072	0.012
Female	0.093	0.005

The coefficients below are coefficients on the classroom variables in the model. These measure the relationship between classroom characteristics and student improvement on the test. For example, the coefficient on proportion free lunch is -0.097. This means that a 10 percentage point increase in the share of free-lunch students in the classroom (say, from 50 percent to 60 percent) is associated with a 0.0097 (10 percent of 0.097) standard deviation decrease in the scores of students in that classroom, regardless of whether the students are free-lunch or not.

Coefficients on classroom-level variables, 4th grade ELA, 2010-11

Variable	***Coeff.***	***Std. Err.***
Average classroom Math pretest	-0.021	0.027
Average classroom ELA pretest	0.062	0.029
Proportion homeless	-0.106	0.1
Proportion IEP severe	-0.02	0.062
Proportion IEP mild	0.004	0.048
Proportion ELL-REFP	0.089	0.047
Proportion ELL-LEP	-0.021	0.041
Proportion ELL-IFEP	0.129	0.071
Proportion free/reduced lunch	-0.097	0.052
Proportion white	0.008	0.06
Proportion female	0.046	0.048
Proportion African American	-0.142	0.046
Proportion Asian	0.208	0.06

It is important to keep in mind the standard errors of the coefficients in both the student- and classroom-level models when interpreting them. A span of two standard deviations in both the positive and negative directions provides a 95 percent confidence range for a coefficient. For example, note that the coefficient on proportion free lunch is -0.097. The standard error on this coefficient is 0.052. This means that, while our best estimate of the effect of proportion free lunch on classroom-wide growth is -0.097 standard deviations, a 95 percent confidence range for the effect estimate would range from -0.044 to +0.149 standard deviations. Since this range includes zero, we cannot reject with 95 percent confidence the hypothesis that proportion free lunch has no effect on student improvement in the classroom.

Correlation with average prior proficiency

AGT results show a very low correlation between average prior proficiency--a measure of average performance in the previous year among the teacher's students--and AGT. In general, teachers were not more or less likely to have a low AGT than a high one if their students came in with low pretest scores rather than high ones.

While not large in magnitude, there do seem to be some statistically significant negative correlations in early grades. We cannot make causal statement about these correlations, but one plausible explanation could be that in lower grades, the district as a whole is doing slightly better with lower attaining students.

Correlations between Prior Attainment and AGT 2010-2011

Corr	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*
ELA	-0.09*	-0.12*	-0.09*	-0.07	0.05	-0.04
Math	-0.11*	-0.09*	-0.08	-0.22*	0.05	0.28*

* significant at the 5% level

Stability

Another property of the AGT results was stability over time. Schools that were high AGT in one year were, more often than not, also high AGT in the following year. In general, AGT in 2009-10 had a substantively positive correlation with AGT in 2008-09, particularly in math. In the companion piece to this we examine classroom level correlations which tend to be even higher. In the data analyzed roughly fifty percent of teachers are not seen teaching in the same grade and subject in consecutive years. Indeed a grade level team in a particular school in a particular year is unlikely to remain the same the next year so a correlation on the effectiveness of that team will be lower than the correlations at the classroom level. Grade 10 and 11 ELA are under observation and will be reconsidered for inclusion in next year's school level AGT results. They will not be included in the teacher level analysis.

Correlations between Years 2010-2011

	Grade 3	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*	*Grade 9*	*Grade 10*	*Grade 11*
ELA	0.49*	0.37*	0.44*	0.58*	0.36*	0.35*	0.56*	0.12	-0.03
Math	0.46*	0.46*	0.47*	0.66*	0.53*	0.20			

* significant at the 5% level

Correlations between Years 2010-2011

Algebra I	0.63*
Algebra II	0.65*
Geometry	0.54*
Biology	0.72*
Chemistry	0.68*
Physics	0.61*
Integrated Science I	0.67*
Science Grade 8	0.67*
Science Grade 5	0.61*
History and Social Science Grade 8	0.66*
US History	0.44*
World History	0.65*

*significance at the 5% level

Correlation between Math and ELA AGT

There were also substantive positive correlations between math and ELA AGT within each school. Schools that were high AGT in math were also more often than not also high AGT in ELA. This trend diminishes as grade increases, most likely because in secondary school departmentalization takes hold and the teaching of math and ELA becomes unlinked.

Correlations between Subjects

Correlation	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*
2011 Math and ELA	0.77*	0.72*	0.64*	0.64*	0.53*	0.45*

* significant at the 5% level

CONTINUOUS INPROVEMENT

LAUSD and VARC have agreed to make continuous improvements to the LAUSD AGT system as time progresses. In this first release an advanced system was put in place that will change over time as the data improve. The following features are planned to be implemented when the data can support them:

- Dealing with student mobility (both across schools and into and out of the LAUSD system)
- Controlling for the effects of varying levels of attendance
- Measuring AGT for schools and educators in grades K-2, expanded measurement to other subjects such as Science and Social Studies, expanded measurement in more complicated grade arrangements such as high school math and later high school grades.
- Studying the effect of staff mobility and creating new reporting structures for such events

CONCLUSION

This technical report described the AGT model used at LAUSD and developed in association with the Value-Added Research Center of the Wisconsin Center for Education Research at the University of Wisconsin.

For more information on the value-added research of the Value-Added Research Center of the Wisconsin Center for Education Research at the University of Wisconsin, visit VARC's website at:

http://varc.wceruw.org

For information on LAUSD guidance to schools for how to use this data, see the LAUSD AGT portal at:

http://agt.lausd.net

APPENDIX ONE: Pre-Tests for Non-NCLB Models

Subject	Prior Year Math CST	Prior Year ELA CST	Prior Year Science CST	Prior Year Social Studies CST
ELA Grade 3	X	X		
ELA Grade 4	X	X		
ELA Grade 5	X	X		
ELA Grade 6*	X	X		
ELA Grade 6*	X	X		
LEA Grade 7	X	X		
ELA Grade 8	X	X		
ELA Grade 9	X	X	X	X
ELA Grade 10	X	X	X	
ELA Grade 11	X	X	X	X
Math Grade 3	X	X		
Math Grade 4	X	X		
Math Grade 5	X	X		
Math Grade 6*	X	X		
Math Grade 6*	X	X		
Math Grade 7	X	X		
Math Grade 8	X	X		
Algebra I	X	X	X	X
Algebra II	X	X		
Geometry	X	X	X	X
Science Grade 5	X	X		
Science Grade 8	X	X		
Biology	X	X	X	X
Chemistry	X	X	X	
Integrated Science 1	X	X	X	X
Physics	X	X	X	
History and Social Studies Grade 8	X	X		
US History	X	X	X	X
World History	X	X	X	X
* Two separate models for sixth grade are run because some sixth grade classes are included in the elementary marks data set (for elementary schools that include grade 6) and some classes are in the secondary marks data set (for middle schools that include grade 6)				