Differential-Effects Value-Added Growth by Student Characteristics by School

Michael S. Christian, *Education Analytics*
Robert H. Meyer, *Education Analytics*

March 2019

DRAFT: Please do not cite.

Abstract

We estimate a value-added growth model that allows for schools to have differential effects on students in different subgroups, using data from a large U.S. state. This model measures the extent to which differences in student growth outcomes across subgroups are school-specific. We find substantial differential effects by prior achievement and disability in both math and English language arts (ELA). These effects are robust to being modeled one at a time (i.e., measuring differential effects by prior achievement without controlling for differential effects by disability, or vice versa) or jointly (i.e., measuring coexisting differential effects in both). Differentials in growth by prior achievement are substantively correlated over consecutive years, while differentials by disability are only modestly correlated over time in math and only slightly correlated in ELA. We also measure differential effects by English language learner status, finding modest differential effects in math and virtually no differential effects in ELA. We find only small differential effects by race/ethnicity and economic disadvantage.

I. Introduction

The impacts of schools on specific subgroups, such as students with disabilities or English language learners (ELL), has been a historic focus of federal education accountability initiatives, the most recent of which is the Every Student Succeeds Act. We extend traditional value-added models to allow for the possibility that schools may be differentially effective across students in different subgroups, such as disability, ELL status, race/ethnicity, economic disadvantage, and prior achievement. This model produces measures of growth within schools by subgroup, which will better capture the causal impact of schools by subgroup than attainment measures such as percent proficient. The model is estimated using a multivariate shrinkage method that produces subgroup value-added measures that do not overstate differences across subgroups as a result of sampling error.

School differential effects on growth may be present for at least two reasons. First, teachers working in the school may, on average, have differential effects on students by subgroup. Second, students may be differentially sorted within schools by subgroup into classrooms with more effective and less effective teachers.

School value-added measures by subgroup address the extent to which differences in growth between subgroups are school-specific, relative to differences in growth in a broader state-wide or district-wide context. We can think of equity in student growth outcomes as having three aspects: across-school, average within-school, and school-specific. The first of these, across-school equity, is concerned with the extent to which students in different subgroups attend schools with different overall student growth outcomes. The second, average within-school equity, is concerned with the extent to which students in different subgroups experience different student growth outcomes within the average school. The third, school-specific equity, is concerned with the extent to which students in different subgroups experience different student growth outcomes within each specific school, relative to average within-school differences. The differential-effects value-added measures discussed here address this third aspect. After accounting for statewide differences across subgroups in student growth, what differences remain in individual schools? To what extent do schools differ from each other in the extent to which student subgroups within them experience differences in growth outcomes?

Prior research on differential effects in growth, which has primarily focused on differential effects at the teacher level, has found relatively small differences in the effects of teachers across subgroups (Aaronson et al., 2007; Lockwood and McCaffrey, 2009; Loeb et al., 2014; Fox, 2016), although some studies have found interactions between specific teacher and student characteristics (see Dee, 2004, for example, on student assignment to teacher by race).

We estimate value-added measures of school effects on student achievement by subgroup. The method employed involves estimating the joint distribution of these effects within schools across years and subgroups. The means of this distribution are the average effects across schools by year and subgroup. The variance-covariance matrix of this distribution describes the extent to which school effects are shared across years and subgroups or are specific to years and subgroups. Since differential-effects value-added measures for individual schools are likely to be noisy when the number of students in a subgroup is small, we employ a multivariate shrinkage estimator that uses the covariances of these effects across years and subgroups to minimize expected mean squared error and produce a best linear unbiased prediction.

We use data from a large U.S. state to measure value-added growth for subgroups within schools in English language arts (ELA) and mathematics in grades 4 through 8 over two years of growth. We find substantive differential effects within schools by prior achievement and by disability status in both math and ELA. This implies that schools differ from each other in the relative growth of students with and without disabilities and in relative student growth from higher and lower levels of achievement. We also find a modest differential effect by English language learner (ELL) status in math, virtually no differential effect by ELL status in ELA, and very small differential effects by race/ethnicity and economically disadvantaged status in math and ELA.

II. Estimating Differential-Effects School Growth Measures

Our approach to estimating differential-effects value-added begins with estimating a value-added model using a standard "constant-effects" regression specification that assumes that schools have the same impact on all students within a school. This model is a regression of current assessment score in math or ELA on lagged assessment scores, student demographics, and a set of school fixed effects. This specification, which has been referred to as a covariate adjustment model (McCaffrey et al., 2004) and a dynamic ordinary least squares model (Guarino et al., 2015), is widely employed in research in school-level value-added growth (Deming, 2014; Meyer and Dokumaci, 2015; Chiang et al., 2016; Ehlert et al., 2016; Angrist et al., 2016, 2017).

After estimating the constant-effects value-added regression, we compute student growth residuals equal to the sum of the student residual estimate and the school fixed effect estimate from the value-added regression. The mean of these growth residuals is equal to the school fixed effect estimate, which is a measure of the school's average value-added impact on student achievement; measures built from residuals like these are used as a teacher-level value-added measure in Chetty et al. (2014). We use these growth residuals to produce value-added measures by school that are specific to student characteristics, such as prior achievement, disability, English language learner, race/ethnicity, and low-income status. Using the growth residuals from an estimated constant-effects value-added model has the advantage of producing differential-effects value-added measures that are a "decomposition" of the constant-effects value-added measures.

To produce differential-effects value-added measures by categorical variables such as disability or English language learner, we regress the growth residuals on interactions between a full set of school indicator variables and a full set of indicators for the categorical variable. The coefficients on these interactions measure the differential impacts of schools on students by category. To produce differential-effects value-added measures by prior achievement, we regress the growth residuals on a full set of school indicator variables and interactions between the school indicator variables and the pretest. The coefficients from this regression can be used to measure the impacts of schools on students at any level of prior achievement, assuming that the differential effect by pretest by school is linear. We employ a multivariate shrinkage approach to produce best linear unbiased predictors of the differential effects for individual schools. In many ways, this approach is similar to a hierarchical linear model (HLM) approach that includes random intercepts for subgroups within schools (in the case of a categorical subgroup variable) or random intercepts and prior achievement slopes for each school (in the case of the continuous prior achievement variable). A difference is that, in the approach used here, all regression coefficients are estimated as fixed, and shrinkage estimators are applied to the relevant coefficient estimates afterward. The steps involved in producing the differential-effects value-added measures are described in further detail below.

Constant-effects value-added

We begin by estimating the following posttest-on-pretest value-added model with school fixed effects:

$$y_{it} = y_{it-1}\lambda + y_{it-1}^{alt}\lambda^{alt} + X_{it}\beta + \alpha_{kt} + \varepsilon_{it} \qquad (1)$$

where student i attends school k at time t; y_{it} and y_{it-1} are student i's academic achievement in math or ELA at times t and t-1; y_{it-1}^{alt} is student i's academic achievement at time t-1 in the opposite subject from y_{it} (i.e., y_{it-1}^{alt} is achievement in math if y_{it} is achievement in ELA, and vice versa); X_{it} is a vector of student i's characteristics, which in this case is comprised of gender, ethnicity, disability, migrant, English language learner, and economic disadvantage; α_{kt} is a fixed effect for the school k attended by student i at time t; and ε_{it} is a residual term that includes all non-school impacts on student achievement that are uncorrelated with pretests y_{it-1} and y_{it-1}^{alt} and student characteristics X_{it}.

In practice, we do not measure y_{it}, y_{it-1}, and y_{it-1}^{alt}. Instead, we measure standardized assessment scores Y_{it}, Y_{it-1}, and Y_{it-1}^{alt}, which, being based on assessments of finite length, measure academic achievement with error. Simply replacing y_{it}, y_{it-1}, and y_{it-1}^{alt} in (1) with Y_{it}, Y_{it-1}, and Y_{it-1}^{alt} and estimating by ordinary least squares will yield biased estimates of the model coefficients due to the right-hand-side variables Y_{it-1} and Y_{it-1}^{alt} being measured with error. Instead, we estimate (1) using errors-in-variables regression (Fuller, 1987), which incorporates measures of the reliability of the pretest measures Y_{it-1} and Y_{it-1}^{alt} to produce consistent coefficient estimates.

We estimate the regression in (1) separately by grade, subject, and year. Over two subjects (math and ELA), five grades (four through eight), and two years of growth (2016-17 and 2017-18), this yields a total of twenty regressions. The assessment variables Y_{it}, Y_{it-1}, and Y_{it-1}^{alt} are transformed to rank-based z-scores; this step, which is common but not universally applied in value-added models, transforms measured scores to normally distributed variables with a statewide mean of zero and a standard deviation of one by grade, subject, and year. In addition, we demean the vector of demographic variables X_{it} in each regression sample. This yields a set of school fixed effect estimates $\hat{\alpha}_{kt}$ that sum to zero when weighted by the number of students associated with each school. As a result, these school fixed effects estimate the impact of individual schools relative to the average school impact experienced by students across the state. We use Cronbach's alpha to measure the reliability of the pretest assessments when implementing the errors-in-variables regression used to estimate (1).

In many value-added applications, a shrinkage estimator is applied to the estimates $\hat{\alpha}_{kt}$ to minimize expected mean squared error. This can be a simple univariate approach that multiplies the estimates $\hat{\alpha}_{kt}$ by their reliabilities, or a more complex multivariate approach that exploits the correlations in school value-added over time, within districts, or other aspects of the school. Given the focus of this study on differential effects, we do not produce a shrinkage estimate of α_{kt}; however, we do produce shrinkage estimates of the differential-effects value-added measures, which are described in the following sections.

Differential-effects value-added by a categorical variable

The school fixed-effect estimates $\hat{\alpha}_{kt}$ described in the previous section measure the impacts of schools under the assumption that a school's impact on any given student is the same across all students in the school by grade, subject, and year. However, we are interested in differences in school impact across students with different characteristics within a school. We can adapt equation (1) to allow for differences across a categorical variable that divides students into subgroups s = 1, 2, ..., S as follows:

$$y_{it} = y_{it-1}\lambda^* + y_{it-1}^{alt}\lambda^{alt*} + X_{it}\beta^* + \sum_{s=1}^{S} I_{ist}\alpha_{kst} + \varepsilon_{it}^* \qquad (2)$$

where I_{ist} is an indicator variable that equals 1 if student i is included in subgroup s at time t and α_{kst} is the impact of school k at time t on students in subgroup s.

We assume that $\lambda = \lambda^*$, $\lambda^{alt} = \lambda^{alt*}$, and $\beta = \beta^*$. This assumes that the relationships among y_{it}, y_{it-1}, y_{it-1}^{alt}, and X_{it} (with the exception of any indicator for the subgroup variable) are the same both within-school and within-subgroup as they are within-school and across-subgroup. When this is the case, the within-school, within-year mean of α_{kst}, weighted by the number of students associated with subgroup *s*, is equal to α_{kt} in equation (1). This has the appealing property of the differential effects α_{kst} across *s* within a given school *k* and year *t* being a "decomposition" of the constant or average effect α_{kt}.

We estimate α_{kst} by estimating a "value-added residual" that sums the value-added effect estimate $\hat{\alpha}_{kt}$ and the residual estimate $\hat{e}_{it}$ from the errors-in-variables regression estimate of equation (1):

$$\hat{q}_{it} = \hat{\alpha}_{kt} + \hat{e}_{it} = Y_{it} - Y_{it-1}\hat{\lambda} - Y_{it-1}^{alt}\hat{\lambda}^{alt} - X_{it}\hat{\beta} \quad (3)$$

where $\hat{\lambda}$, $\hat{\lambda}^{alt}$, and $\hat{\beta}$ are estimates of λ, λ^{alt}, and β from the errors-in-variables regression estimate of equation (1). We then regress $\hat{q}_{it}$ on a full set of school-subgroup interaction variables to produce estimates $\hat{\alpha}_{kst}$. We normalize the estimates $\hat{\alpha}_{kst}$ using an estimate of the standard deviation of the constant effects α_{kt} described in the previous section. This expresses the differential effects in units that can be understood as relative to the variance of the more commonly reported constant-effects value-added measures.

After the estimates $\hat{\alpha}_{kst}$ and their estimated standard errors $\hat{\sigma}_{kst}$ are produced, we apply a multivariate shrinkage estimator (Longford, 1999) to minimize expected mean squared error. This shrinkage estimator is motivated by understanding the school-subgroup-year effects α_{kst} as jointly normally distributed across schools *k*:

$$\alpha_{k(sub)} \sim N\left(\mu_{(sub)}, \Omega_{(sub)}\right) \quad (4)$$

where $\alpha_{k(sub)}$ is a column vector that stacks α_{kst} over all combinations of subgroup *s* and year *t* in school *k*. The multivariate shrinkage estimate of $\alpha_{k(sub)}$ is equal to:

$$\tilde{\alpha}_{k(sub)} = \mu_{(sub)} + \Omega_{(sub)}\left[\Omega_{(sub)} + \Sigma_{k(sub)}\right]^{-1}\left(\hat{\alpha}_{k(sub)} - \mu_{(sub)}\right) \quad (5)$$

where $\hat{\alpha}_{k(sub)}$ is a column vector of the $\hat{\alpha}_{kst}$ in the same order by school, subgroup, and year as $\alpha_{k(sub)}$ and $\Sigma_{k(sub)} = E[(\hat{\alpha}_{k(sub)} - \alpha_{k(sub)})(\hat{\alpha}_{k(sub)} - \alpha_{k(sub)})']$. The prediction standard errors of the entries $\tilde{\alpha}_{kst}$ in $\tilde{\alpha}_{k(sub)}$ are the square roots of the diagonal entries of $\tilde{\Sigma}_{k(sub)}$, where

$$\tilde{\Sigma}_{k(sub)} = \Omega_{(sub)}\left[\Omega_{(sub)} + \Sigma_{k(sub)}\right]^{-1}\Sigma_{k(sub)} \qquad (6)$$

In practice, we use estimates $\hat{\Omega}_{(sub)}$ and $\hat{\Sigma}_{k(sub)}$ in place of $\Omega_{(sub)}$ and $\Sigma_{k(sub)}$ when applying shrinkage in equation (5). When estimating $\Sigma_{k(sub)}$, we assume that the sampling errors of $\hat{\alpha}_{kst}$ are uncorrelated across school and subgroup. Given that students are nested within school, subgroup, and year, this assumption should be sufficiently close to correct; any correlations of estimation errors will only be a side effect of sampling error in $\hat{\lambda}$, $\hat{\lambda}^{alt}$, and $\hat{\beta}$. Under this assumption, we set $\hat{\Sigma}_{k(sub)}$ to a diagonal matrix whose diagonal entries are the squares of the standard error estimates $\hat{\sigma}_{kst}$.

When estimating $\Omega_{(sub)}$ for each grade and subject, we use a data set that includes all pairs of $\hat{\alpha}_{kst}$ within schools across subgroups and years. We use this data set to estimate the elements of $\Omega_{(sub)}$ using method of moments, assuming no restrictions other than that the within-year, within-subgroup variance is the same across years for any specific subgroup and that the across-year, across-subgroup covariance is the same across consecutive years for any specific pair of subgroups. The diagonal components of $\Omega_{(sub)}$, which measure variances of α_{kst} by subgroup s, are estimated:

$$\hat{\omega}_s^2 = Var[\hat{\alpha}_{kst}] - Mean[\hat{\sigma}_{kst}^2] \qquad (7)$$

while the off-diagonal components of $\Omega_{(sub)}$, which measure covariances between α_{kst} and $\alpha_{ks^*t^*}$, are estimated:

$$\hat{\omega}_{s,t,s^*,t^*} = Cov[\hat{\alpha}_{kst}, \hat{\alpha}_{ks^*t^*}] \qquad (8)$$

No adjustment for correlated sampling error is needed when estimating the covariance between α_{kst} and $\alpha_{ks^*t^*}$ in equation (8) because the sampling errors of $\hat{\alpha}_{kst}$ and $\hat{\alpha}_{ks^*t^*}$ are uncorrelated.

If this approach produces an estimate of $\Omega_{(sub)}$ that is not positive semidefinite, we impose a stricter set of assumptions that limits the number of parameters in $\Omega_{(sub)}$ to four: within-year within-subgroup variance; across-year within-subgroup covariance; within-year across-subgroup covariance; and across-year across-subgroup covariance. When this is the case, we restrict the values of these four parameters so that, if we expressed α_{kst} as the sum of four orthogonal, normally distributed components as below:

$$\alpha_{kst} = u_k + u_{ks} + u_{kt} + u_{kst} \tag{9}$$

the variance of none of the u components are implicitly less than zero. An exception to the above rule is the case of differential effects by race/ethnicity, in which we always used the more limited approach to estimating $\Omega_{(sub)}$ because of the number of race/ethnicity categories; this also has the advantage of producing for the tables in the results section a single summary statistic that covers all race/ethnicity categories. In practice, this restriction was employed in differential effects by English language learner status in fifth-grade through eighth-grade ELA, in differential effects by low-income status in eighth-grade ELA, and (as noted above) in differential effects by race/ethnicity in all grades and subjects.

Differential-effects value-added by a continuous variable measured with error

We can also adapt equation (1) to measure differential effects by a continuous pretest variable as follows:

$$y_{it} = y_{it-1}\lambda^{\dagger} + y_{it-1}^{alt}\lambda^{alt\dagger} + X_{it}\beta^{\dagger} + \theta_{0kt} + y_{it-1}\theta_{1kt} + \varepsilon_{it}^{\dagger} \tag{10}$$

This version includes not just a separate intercept for each school θ_{0kt}, but also a separate slope θ_{1kt} for the same-subject pretest y_{it-1} for each school. Since y_{it-1} is normalized to a distribution with a mean of zero and a standard deviation of one, the school intercept θ_{0kt} is equal to the impact of school k in year t on students whose same-subject pretest achievement is at the mean. In addition, the impact of school k in year t is $\theta_{0kt} + \theta_{1kt}$ for students whose pretest achievement is one standard deviation above the sample-wide mean and $\theta_{0kt} - \theta_{1kt}$ for students whose pretest achievement is one standard deviation below the sample-wide mean.

We assume, similarly to the case of differential effects by a categorical variable, that $\lambda = \lambda^{\dagger}$, $\lambda^{alt} = \lambda^{alt\dagger}$, and $\beta = \beta^{\dagger}$. This implies that the within-school relationships among y_{it}, y_{it-1}, and X_{it} are, on average, the same regardless of whether or not you control for interactions between school assignment and same-subject prior achievement. Given this assumption, we estimate θ_{0k} and θ_{1k} by regressing $\hat{q}_{it}$, as defined in (3), on measured prior achievement Y_{it-1} separately for each school k. Given that $\hat{q}_{it}$ and Y_{it-1} both include the error in measurement in Y_{it-1} as a component, we estimate these regressions using errors-in-variables regression approaches described in Fuller (1987). To avoid imprecise estimates of the variance-covariance matrix of the estimates $\hat{\theta}_{0k}$ and $\hat{\theta}_{1k}$, we assume, when estimating the variance-covariance matrix, that the error $\varepsilon_{it}^{\dagger}$ is homoskedastic across schools. As was the case of the categorical differential-effects measures, we normalize the estimates $\hat{\theta}_{0k}$ and $\hat{\theta}_{1k}$ using an estimate of the standard deviation of the constant effects α_{kt}.

We shrink the estimates $\hat{\theta}_{0kt}$ and $\hat{\theta}_{1kt}$ using a multivariate shrinkage approach very similar to that employed in the case of differential effects by a categorical variable. We assume that

$$\theta_k \sim N(\mu_\theta, \Omega_\theta) \quad (11)$$

where θ_k is a column vector that stacks θ_{0kt} and θ_{1kt} across years. The shrinkage estimate of θ_k is equal to:

$$\tilde{\theta}_k = \mu_\theta + \Omega_\theta[\Omega_\theta + \Sigma_\theta]^{-1}(\hat{\theta}_k - \mu_\theta) \quad (12)$$

where $\hat{\theta}_k$ stacks $\hat{\theta}_{0kt}$ and $\hat{\theta}_{1kt}$ across years in the same order as θ_k, and Σ_θ, the variance-covariance matrix of $\hat{\theta}_k$, is equal to $E[(\hat{\theta}_k - \theta_k)(\hat{\theta}_k - \theta_k)']$.

We estimate Σ_θ by assuming that Σ_θ is block-diagonal by year and setting the blocks equal to the estimated variance-covariance matrices of $\hat{\theta}_{0kt}$ and $\hat{\theta}_{1kt}$ within each year. We use a method-of-moments approach similar to that used in the case of differential effects by a categorical variable to estimate Ω_θ; the most substantive difference is that an adjustment for correlated sampling error is needed when computing the covariance between θ_{0kt} and θ_{1kt}. We create a data set that includes every within-school pairing of $\hat{\theta}_{0kt}$ and $\hat{\theta}_{1kt}$, both within and across year, by grade and subject. Using this data set, we estimate the elements in Ω_θ using method of moments, with the only restrictions being that the within-year variances of the intercept θ_{0kt} and the slope θ_{1kt} are the same across years and that the across-year covariances of θ_{0kt} with θ_{1kt} are the same for any pair of consecutive years.

III. Evidence for School Differentials in Academic Growth

We estimate the value-added models described above using assessment data from a large U.S. state. As mentioned above, we estimate the value-added models separately by subject, grade, and year, which produces school effect estimates that are specific to subject, grade, and year. There are twenty combinations of subject, grade, and year, covering two subjects (math and English language arts), five grades (four through eight), and two years (2016-17 and 2017-18). We present descriptive statistics for the sample in Table 1.

The differential-effects model described in the previous section measures separate value-added measures by school by subgroup. For example, for each school with students both with and without disabilities, we produce one value-added measure using the growth of students with disabilities, and another value-added measure using the growth of students without disabilities.

This model takes into account the possibility that schools may have a different impact on students in particular subgroups relative to students in other subgroups. We estimate this model in four different ways: one that differentiates students by disability (with disability/without disability); another that differentiates students by English language learner (English language learner/not English language learner); a third that differentiates students by race/ethnicity (using seven categories: Asian, black, Hispanic, Pacific Islander, multiracial, Native American, and white); and a fourth that differentiates students by economic disadvantage (economically disadvantaged/not economically disadvantaged).

Table 1. Descriptive statistics

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Economic disadvantage	0.45	0.44	0.42	0.40	0.39
Female	0.49	0.49	0.49	0.49	0.49
Black	0.10	0.10	0.10	0.09	0.09
Hispanic	0.13	0.13	0.13	0.12	0.12
English language learner	0.08	0.06	0.05	0.04	0.04
Disability	0.13	0.13	0.12	0.12	0.12
No. students					
Math, 2016-17	60,063	58,954	58,473	58,896	58,128
Math, 2017-18	59,998	60,519	59,077	58,709	58,926
ELA, 2016-17	60,064	58,940	58,480	58,909	58,162
ELA, 2017-18	59,999	60,529	59,093	58,708	58,940
No. schools					
Math, 2016-17	1,275	1,211	848	803	816
Math, 2017-18	1,265	1,209	843	797	797
ELA, 2016-17	1,275	1,211	848	803	815
ELA, 2017-18	1,265	1,209	843	797	797

Note: Demographic measures are proportions pooled across years and subjects.

Another version of the differential-effects value-added model described in the previous section measures school value-added in a way that allows for a school's impact on a given student to increase (or decrease) with that student's performance on the previous year's assessment in the same subject. This model takes into account the possibility that some schools may have a different impact on students who had been higher-achieving students in the past relative to students who had been lower-achieving students. We assume that the way in which a school's impact is different by prior achievement is linear: for every standard deviation higher that a student scores on the previous year's assessment, the school's impact on that student's growth is higher or lower by some constant, school-specific slope coefficient. This approach makes it possible to measure the school's impact on students at any given level of prior achievement, given the linearity assumption.

Magnitude of differences in school value-added effects by student characteristics

In Table 2 below, we present correlations in school value-added across subgroups within school and year. These correlations in Table 2 are estimated in a way that accounts for sampling error in the estimated value-added measures, so that they estimate correlations among the actual differential impacts of schools on different subgroups. They can be thought of as estimates of what the correlations among the value-added estimates would be if every school had an extremely large number of students. These correlations should be indicative of the extent to which schools have differential effects across students in different subgroups. If the correlation is high, then the effects of schools are typically very similar across students in different subgroups. In contrast, if the correlation is low, then individual schools are more likely to have different effects for students in different subgroups.

In the case of the categorical subgroup variables, such as disability and race/ethnicity, the correlations in Table 2 are computed from estimates of $\Omega_{(sub)}$ and are equal to:

$$Corr(\alpha_{kst}.\alpha_{ks^*t}) = \frac{Cov(\alpha_{kst}.\alpha_{ks^*t})}{\sqrt{Var(\alpha_{kst})Var(\alpha_{ks^*t})}} \tag{13}$$

where α_{kst} is the subgroup effect for school *k* for students in subgroup *s* at time *t* and subgroups *s* and *s** are different subgroups (for example, English language learner and not English language learner, or students with disabilities and students without disabilities). In the case of the continuous pretest variable, these correlations are derived from estimates of Ω_θ and are equal to:

$$Corr(\theta_{0kt} + \theta_{1kt}x, \theta_{0kt} - \theta_{1kt}x) = \frac{Var(\theta_{0kt}) - Var(\theta_{1kt})x^2}{\sqrt{Var(\theta_{0kt}+\theta_{1kt}x)Var(\theta_{0kt}-\theta_{1kt}x)}} \quad (14)$$

where

$$Var(\theta_{0kt} + \theta_{1kt}x) = Var(\theta_{0kt}) + Var(\theta_{1kt})x^2 + 2Cov(\theta_{0kt}, \theta_{1kt})x$$

and

$$Var(\theta_{0kt} - \theta_{1kt}x) = Var(\theta_{0kt}) + Var(\theta_{1kt})x^2 - 2Cov(\theta_{0kt}, \theta_{1kt})x$$

and θ_{0kt} is the impact of school k at time *t* on students with average prior student achievement and θ_{1kt} is the slope by which school *k*'s impact at time *t* increases or declines with a standard deviation of prior student achievement. We set *x* to 0.8, which measures the correlation between a school's impact on students who had previously scored 0.8 standard deviations above the statewide average and its impact on students who had previously scored 0.8 standard deviations below the statewide average. This corresponds to the average achievement of students in the upper and lower halves of an assessment with normally distributed scores.

Table 2. Across-subgroup, within-year correlation in school value-added

	Pretest	Disability	ELL	Race/ethnicity	Low-income
Math					
Grade 4	0.80	0.63	0.81	0.88	0.96
Grade 5	0.75	0.61	0.87	0.85	0.96
Grade 6	0.77	0.62	0.91	0.90	0.96
Grade 7	0.59	0.46	0.69	0.76	0.95
Grade 8	0.65	0.53	0.70	0.87	0.93
ELA					
Grade 4	0.72	0.56	0.85	0.86	0.91
Grade 5	0.74	0.54	0.99	0.95	0.99
Grade 6	0.81	0.64	1.00	0.97	0.98
Grade 7	0.81	0.58	1.00	1.00	0.95
Grade 8	0.78	0.63	0.97	0.93	0.99

The correlations in Table 2 suggest that there are substantive differential effects within schools by prior achievement and by disability. The correlations between schools' impacts by prior achievement range from 0.59 (seventh-grade math) to 0.81 (sixth- and seventh-grade ELA). This suggests that, while the impacts of schools on higher-achieving students is substantially correlated with their impacts on lower-achieving students, the correlation is not close to perfect, and many schools will have substantially different impacts by prior achievement.

Similarly, the correlations in impacts on students with and without disabilities, which range from 0.46 (seventh-grade math) and 0.64 (sixth-grade ELA), imply substantive differentials in the impacts of schools by disability status. One caution in interpreting this result is that some of this differential may be an effect of some schools serving students with more severe disabilities than others, which would appear as a differential value-added effect if the severity of disability is insufficiently controlled for in the value-added regression in equation (1). The estimated value-added regression included disability not simply as a binary variable indicating whether students had or did not have a disability, but broken out into five separate indicators by type of disability (learning or intellectual disability; emotional behavioral disability; speech or language impairment; autism; and all other disabilities). Consequently, for the differential-effects measures to be biased by sorting of students with more or less severe disabilities across schools, it would need to be the case that there were differences across schools in the severity of disability within (and not across) the five disability types described above.

The extent of differential value-added effects by English language learner differs by subject. We find virtually no differential effects by ELL status in ELA, while we find evidence of modestly-sized differential effects in math. This would be the case if schools with strong ELA instruction for ELL students also have very good instruction in ELA for non-ELL students, while the quality of instruction in math for ELL students and non-ELL students, while still quite related, is not as closely related as in ELA.

The correlations in the last two columns of Table 2 measure the correlation between school effects by race/ethnicity and economic disadvantage. In both cases, the correlations are very high, suggesting that differential effects by these characteristics are small. This suggests that schools that are effective with low-income students also tend to a very high degree to be effective with high-income students, and that the impacts of schools across race/ethnicity groups are similarly very closely related.

Persistence of differences in school value-added effects by student characteristics

The results in Table 2 speak to the magnitude of differential effects within schools within a given year. Given that we have two years of growth measures, we can also measure the extent to which differential effects persist from one year to the next. Table 3 presents the correlation from one year to the next in the difference in school effects across subgroups for cases where a substantive differential effect was measured in Table 2 (pretest and disability in both subjects, as well as ELL in math). If a school's effect is greater for one subgroup than for another in one year, these correlations measure the extent to which the same school's effect will be greater for the same subgroup relative to the same other subgroup in the following year. As is the case in Table 2, the correlations in Table 3 are measured in a way that adjusts for sampling error, so as to estimate the year-to-year correlation in the differences in the school's actual impacts on students in different subgroups, and can be thought of as what the correlations among measures would be were those measures based on an extremely large number of students. In the case of categorical subgroup variables, this correlation is computed from estimates of $\Omega_{(sub)}$ and is equal to

$$Corr(\alpha_{kst} - \alpha_{ks^*t}, \alpha_{kst^*} - \alpha_{ks^*t^*}) = \frac{Cov(\alpha_{kst}.\alpha_{kst^*}) + Cov(\alpha_{ks^*t}.\alpha_{ks^*t^*}) - 2Cov(\alpha_{kst}.\alpha_{ks^*t^*})}{Var(\alpha_{kst}) + Var(\alpha_{ks^*t}) - 2Cov(\alpha_{kst}.\alpha_{ks^*t})} \qquad (15)$$

In the case of the continuous pretest variables, this correlation is computed from estimates of Ω_θ and is equal to

$$Corr(2\theta_{1kt}x, 2\theta_{1kt^*}x) = \frac{Cov(\theta_{1kt}, \theta_{1kt^*})}{Var(\theta_{1kt})} \quad (16)$$

The results in Table 3 suggest that there is substantial year-to-year persistence in differences in school value-added effects by prior achievement. This is particularly the case in math, where the correlation from year to year in the difference between a school's impacts on high-achieving and low-achieving students ranges from 0.56 (sixth grade) to 0.69 (seventh grade). This similar to the correlation from year to year in overall, constant-effects value-added in math, which ranges from 0.52 (fifth grade) to 0.63 (sixth grade). The high persistence in the value-added differential by prior achievement in math implies that a school with a substantively greater impact on high-achieving students relative to low-achieving students in one year is likely to also have a substantively greater impact on high-achieving students relative to low-achieving students in the following year. The persistence of differences in school value-added by prior achievement is also high in ELA, where the average year-to-year correlation of the differential across grades is 0.29. This is comparable to the year-to-year correlation of overall, constant-effects value-added in ELA, which also has an average across grades of 0.29.

Table 3. Across-year correlation of across-subgroup differences in value-added

	Pretest		Disability		ELL
	Math	ELA	Math	ELA	Math
Grade 4	0.57	0.34	0.20	-0.02	0.21
Grade 5	0.56	0.36	0.17	0.01	0.19
Grade 6	0.56	0.35	0.21	0.32	0.91
Grade 7	0.69	0.04	0.30	0.06	0.50
Grade 8	0.65	0.37	0.20	0.11	0.15

In contrast, the persistence over time of differences in school value-added effects by disability status is modest in math, ranging from 0.17 to 0.30, and is very low in ELA, where it is often less than 0.10. That the persistence of differentials is so low in ELA might lend some support to the possibility that differential effects by disability are driven by differences across schools in severity of disability, although this would only be the case if the extent of severity varies substantially from year to year within grades within schools. We would also expect this possibility to affect the persistence of disability differentials in math, where the persistence of differences in value-added by disability is modest but also consistently positive.

We also find modest persistence over time in differences in school value-added effects in math between ELL and non-ELL students, with the correlation differing from grade to grade. Note that the grade with the most substantial persistence correlation in Table 3, grade six, is also the grade with the smallest differential effect in terms of magnitude in Table 2.

Shrinkage estimates of school value-added effects by student characteristics

In practice, when value-added is measured for individual schools, a shrinkage estimate is employed (Longford, 1999). Shrinkage takes into account that school value-added is generated by a distribution. It produces a best linear unbiased prediction of the individual school's value-added effect from that distribution and from the school fixed-effect value-added estimate. One effect of shrinkage is that it tempers value-added measures based on smaller numbers of students in the direction of the average value-added effect. If a multivariate shrinkage approach is employed, value-added measures are also adjusted to be closer to each other when the effects that they measure are correlated with each other.

We produce differential value-added effects for individual schools using a multivariate shrinkage approach that assumes that value-added effects are correlated within schools. This approach assumes that subgroup value-added effects are correlated within schools, both between years and between subgroups. The impact of this shrinkage approach is that value-added measures based on smaller numbers of students will be adjusted toward the statewide mean. In addition, the shrinkage approach will also adjust differential-effects value-added measures for the same school toward each other, both between subgroups and over time. The statistical formula for the multivariate shrinkage estimator is described in equations (5) and (12) above.

Table 4. Across-subgroup, within-year correlation in value-added, post-shrinkage

	Pretest	Disability	ELL	Race/ethnicity	Low-income
Math					
Grade 4	0.88	0.77	0.91	0.97	0.99
Grade 5	0.84	0.75	0.96	0.96	0.99
Grade 6	0.83	0.75	0.95	0.96	0.99
Grade 7	0.67	0.60	0.84	0.91	0.99
Grade 8	0.71	0.69	0.87	0.97	0.98
ELA					
Grade 4	0.82	0.72	0.95	0.97	0.98
Grade 5	0.84	0.72	1.00	1.00	1.00
Grade 6	0.90	0.79	1.00	0.99	1.00
Grade 7	0.89	0.74	1.00	1.00	0.99
Grade 8	0.85	0.79	0.99	0.99	1.00

In Table 4, we present the correlations across subgroups in differential-effects value-added after having applied the shrinkage estimator. The correlations are considerably higher among the shrinkage estimates than in the sampling-error-adjusted correlations presented in Table 2. This is a result of overshrinkage, in which the variance of shrunk parameter estimates is typically smaller than the estimates of the variance of the parameters (Louis, 1984). In particular, this is a side effect of the multivariate shrinkage approach, which takes into account that value-added is highly correlated across subgroups within schools and adjusts the value-added measures of different subgroups within the same school toward each other, particularly when those value-added measures are measured with substantial sampling error. This approach minimizes the expected mean squared error of the differential-effects value-added measures, and produces the best linear unbiased predictors of value-added by subgroup for each individual school. For this reason, the use of multivariate shrinkage when measuring differential-effects value-added is strongly recommended. However, it does have the side effect of measuring smaller differentials across subgroups on average.

Partial differential effects by prior achievement and disability

The results presented above suggest that schools have different impacts on students with different levels of prior achievement. They also suggest that schools have different impacts on students with and without disabilities. However, in both cases, the measured differential effects were estimated using completely separate models in which the *only* differential effect was by prior achievement or by disability; there was no attempt to control for differential effects by prior achievement when measuring differential effects by disability or vice versa.

We specify a model that allows for there to simultaneously be differential effects by prior achievement and by disability in equation (17) below:

$$y_{it} = y_{it-1}\lambda' + y_{it-1}^{alt}\lambda^{alt\prime} + X_{it}\beta' + \sum_{s=1}^{2} I_{st}\varphi_{0kst} + y_{it-1}\varphi_{1kt} + \varepsilon_{it}' \qquad (17)$$

where *s* = 1 indicates students with disabilities and *s* = 2 indicates students without disabilities; I_{st} is an indicator variable that equals 1 among students in disability subgroup *s* in year *t*; φ_{0kst} is the effect of school *k* at time *t* among students in disability subgroup *s* with an average score on the prior year's assessment; and φ_{1kt} is the extent to which the effect of school *k* increases or decreases at time *t* with student *i*'s prior-year assessment score y_{it-1}. We estimate (17) in the much the same way as the other differential-effects models are estimated. We assume that the coefficients λ', $\lambda^{alt\prime}$, and β' are the same as their analogues in the constant-effects value-added model. This allows us to estimate the parameters φ_{0kst} and φ_{1kt} by regressing, using errors-in-variables methods, the residual $\hat{q}_{it}$ as defined in equation (3) on interactions between school and disability status and interactions between school and prior achievement y_{it-1}. We estimate the variances and covariances among the φ_{0kst} and φ_{1kt} using a method-of-moments approach, which can be employed to produce shrinkage estimates of φ_{0kst} and φ_{1kt}.

Table 5. Effect of measuring differential effects by pretest and disability simultaneously on across-subgroup, within-year correlation in school value-added

	Pretest			Disability	
	Without disability diff. effect	W/ disability diff. effect, students w/o disabilities	W/ disability diff. effect, students w/ disabilities	Without pretest diff. effect	W/ pretest diff. effects, students at mean pretest
Math					
Grade 4	0.80	0.80	0.86	0.63	0.74
Grade 5	0.75	0.77	0.83	0.61	0.68
Grade 6	0.77	0.76	0.86	0.62	0.72
Grade 7	0.59	0.57	0.77	0.46	0.64
Grade 8	0.65	0.62	0.81	0.53	0.70
ELA					
Grade 4	0.72	0.73	0.83	0.56	0.67
Grade 5	0.74	0.75	0.84	0.54	0.61
Grade 6	0.81	0.82	0.89	0.64	0.72
Grade 7	0.81	0.81	0.86	0.58	0.71
Grade 8	0.78	0.76	0.85	0.63	0.67

In Table 5, we compare the across-subgroup, within-year correlations in school value-added from the models in equations (2) and (10) that assume differential effects by disability *only* or by prior achievement *only* with the same correlations from the model in equation (17) that allows for coexisting differential effects by both prior achievement and disability. As is the case in Tables 2 and 3, these correlations adjust for sampling error. In the model with coexisting effects, we measure the correlation between value-added effects among students with and without disabilities for a student with an average prior achievement score. We measure the correlation between value-added effects among higher-achieving and lower-achieving students for a student with either disability status; these correlations are not the same given we do not assume that the covariance between the disability-specific intercept φ_{0kts} and the slope φ_{1kt} is the same across disability status. The first and third columns of Table 5 repeat the correlations from Table 2, while the second and fourth columns present the analogous correlations from the coexisting-effects model.

In general, we find that the differential effects by disability and by prior achievement described in Table 2 are largely robust to specifying the model to include differential effects in both disability and prior achievement simultaneously. In other words, a differential effect in prior achievement remains even when controlling for a differential effect by disability, and vice versa.

The differential effect by disability is diminished in magnitude when pretest differential effects are included in the model. We can see this by comparing the correlations in the two rightmost columns of Table 5; including the pretest differential effect leads to a substantially higher correlation between value-added among students with and without disabilities, which narrows the differences between the modeled impacts of schools on students with and without disabilities. This suggests that, to some degree, the measured differential effect by disability presented in Table 2 is driven somewhat by an unmodeled differential effect by pretest. However, even after a pretest differential effect is included in the model, a substantial differential effect by disability remains, given that the correlations in the rightmost column of Table 5 are still substantively below one.

In the model that includes differential effects by both disability and prior achievement, there are two correlations between value-added for high-achieving and low-achieving students: one for students with disabilities, and one for students without disabilities. The two correlations are different because, conditional on prior achievement being at the mean, the variance of value-added for students with disabilities is greater than the variance of value-added for students without disabilities. This leads to a higher correlation in value-added between high-achieving and low-achieving students among students with disabilities. Regardless of disability status, a substantial differential effect by prior achievement remains after controlling for differential effects by disability. This is particularly the case for students without disabilities, for whom there is virtually no effect on the measured correlation in value-added between low- and high-achieving students.

IV. Conclusions

Using data from a large state, we estimate a value-added model that allows for schools to have different impacts on the academic achievement of students with different characteristics. Estimating this model measures school-specific differences in growth across subgroups, relative to across-subgroup differences in growth at the average school, and speaks to heterogeneity across schools in equity in growth outcomes by subgroup. We find evidence of differential school effects by prior achievement and by disability. These differential effects are largely robust to whether they are modeled separately (estimating a model with differential effects in prior achievement but not in disability, or vice versa) or whether they are modeled simultaneously (estimating a model with differential effects in *both* prior achievement and in disability). Differences in school-level growth by prior achievement are substantially correlated from one year to the next in both math and ELA; in contrast, differences by disability are only modestly correlated in math and slightly correlated--if at all--in ELA.

We also find evidence of modest differential school effects by English language learner status in mathematics, but not in English language arts. Finally, we measure only small differential school effects by race/ethnicity and by economically disadvantaged status.

References

Aaronson, Daniel, Lisa Barrow, and William Sander (2007), "Teachers and Student Achievement in Chicago Public Schools", *Journal of Labor Economics* 25(1), pp. 95-135.

Angrist, Joshua, Peter Hull, Parag Pathak, and Christopher Walters (2016), "Interpreting Tests of School VAM Validity", *American Economic Review: Papers and Proceedings* 106(5), pp. 388-392.

Angrist, Joshua, Peter Hull, Parag Pathak, and Christopher Walters (2017), "Leveraging Lotteries for School Value-Added: Testing and Estimation", *Quarterly Journal of Economics* 132(2), pp. 871-919.

Chaing, Hanley, Stephen Lipscomb, and Brian Gill (2016). "Is School Value Added Indicative of Principal Quality?", *Education Finance and Policy* 11(3), pp. 283-309.

Chetty, Raj, John Friedman, and Jonah Rockoff (2014), "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates", *American Economic Review* 104(9), pp. 2593-2932.

Dee, Thomas (2004), "Teachers, Race, and Student Achievement in a Randomized Experiment", *Review of Economics and Statistics* 86(1), pp. 195-210.

Deming, David (2014), "Using School Choice Lotteries to Test Measures of School Effectiveness", *American Economic Review: Papers and Proceedings* 104(5), pp. 406-411.

Ehlert, Mark, Cory Koedel, Eric Parsons, and Michael Podgursky (2016), "Selecting Growth Measures for Use in School Evaluation Systems: Should Proportionality Matter?", *Educational Policy* 30(3), pp. 465-500.

Fox, Lindsay (2016), "Playing to Teachers' Strengths: Using Multiple Measures of Teacher Effectiveness to Improve Teacher Assignments", *Education Finance and Policy* 11(1), pp. 0-96.

Fuller, Wayne (1987), *Measurement Error Models* (Hoboken, N.J.: John Wiley & Sons)

Guarino, Cassandra, Mark Reckase, and Jeffrey Wooldridge (2015). "Can Value-Added Measures of Teacher Performance Be Trusted?", *Education Finance and Policy* 10(1), pp. 117-156.

Lockwood, J.R., and Daniel McCaffrey (2009), "Exploring Student-Teacher Interactions in Longitudinal Achievement Data", *Education Finance and Policy* 4(4), p. 439-467.

Loeb, Susanna, James Soland, and Lindsay Fox (2014), "Is a Good Teacher a Good Teacher For All? Comparing Value-Added of Teachers With Their English Learners and non-English Learners", *Educational Evaluation and Policy Analysis* 36(4), pp. 457-475.

Longford, Nicholas (1999), "Multivariate Shrinkage Estimation of Small Area Means and Proportions", *Journal of the Royal Statistical Society* 162(2), pp. 227-245.
Louis, Thomas (1984), "Estimating a Population of Parameter Values Using Bayes and Empirical Bayes Methods", *Journal of the American Statistical Association* 79(386), pp. 393-398.

McCaffrey, Daniel, J. R. Lockwood, Daniel Koretz, Thomas Louis, and Laura Hamilton (2004). "Models for Value-Added Modeling of Teacher Effects", *Journal of Educational and Behavioral Statistics* 29(1), pp. 67-101.

Meyer, Robert, and Dokumaci, Emin (2015), "Value-Added Models and the Next Generation of Assessments", in R. W. Lissitz (Ed.), *Value added modeling and growth modeling with particular application to teacher and school effectiveness* (Charlotte, NC: Information Age Publishing), pp. 139-190.