A Comparison of Alternative Models for Estimating School Performance in Mathematics and Reading/Language Arts in Four State Accountability Systems: North Carolina Results

NCAASE Technical Report
December, 2017

Ann C. Schulte[b]
Joseph F. T. Nese[a]
Joseph J. Stevens[a]
Nedim Yel[c]
Gerald Tindal[a]
Daniel Anderson[a]
Stephen N. Elliott[b]

[a] University of Oregon
[b] Arizona State University
[c] Indiana University

Address all correspondence to Ann C. Schulte, Arizona State University, Sanford School of Social and Family Policy, Arizona State University, PO Box 873701, Tempe, AZ 85287-3701, ann.schulte@asu.edu

Table of Contents

List of Tables

A Comparison of Alternative Models for Estimating School Performance in Mathematics and Reading/Language Arts in Four State Accountability Systems: North Carolina Results

Background and Introduction

This technical report is one of a series of four technical reports that describe the results of a study comparing eight alternative models for estimating school academic achievement using data from the Arizona, North Carolina, Oregon, and Pennsylvania accountability systems. Our purpose was not to evaluate or examine the accountability systems in use by these states, but to evaluate a broader range of models commonly used for estimating school performance that are applied in many states and frequently reported in the school effectiveness research literature. This introduction briefly describes the study background and details the methods and procedures we used to estimate the eight school performance models and compare model results in all four states. The individual state technical reports including details on each state's accountability data, assessment instruments, and results are provided at: http://www.ncaase.com/publications/tech-reports.

Despite the central importance of analytic models used in evaluating teacher and school effects in modern accountability systems, there are relatively few studies of the reliability and validity of these high-stakes systems (see, for example, Goldschmidt, Choi, & Beaudoin, 2012). The results reported here examine eight models using operational state accountability data in mathematics and reading/language arts from the four participating states. We addressed four questions surrounding the use of analytic models for the evaluation of school performance:

1. Are estimates of school performance stable across successive cohorts of students?
2. How well do estimates of school performance correlate among models?
3. How do estimates of school performance correlate with variables describing the student composition of the school?
4. Do estimates of school performance vary from one model to another based on the school composition of students with disabilities (SWD)?

General Method Description

Sample

The sample from each state is described in each individual state technical report. In three of the four states, the sample consisted of all students who took the state's mathematics or reading/language arts general assessment in any one school year from 2007-08 through 2011-12, and whose records in each year were included in the state's calculation of Adequate Yearly Progress (AYP). Samples were separated into two grade level bands: a longitudinal elementary school sample (Grades 3 through 5) and a longitudinal middle school sample (Grades 6 through 8), each consisting of three cohorts (a) 2007/08 through 2009/2010; (b) 2008/09 through 2010/11; and (c) 2009/10 through 2011/12 (see research design schematic below). In Arizona, only one elementary and middle school cohort was used (2006/07 through 2008/09) due to changes in the Arizona testing program in 2010.

Instruments

The outcome measures for all analyses were the standardized mathematics and reading/language arts tests used for accountability in each state. In three of the states, the instruments used vertically linked developmental scales created using item response theory (IRT)

methods. In Pennsylvania, the test was not vertically linked over grades preventing the estimation of certain school performance models described in the next section. More detail about the North Carolina test is contained in a later section of this report.

Research design indicating academic years and longitudinal cohorts studied:

	Academic Year				
Grade	2007/08	2008/09	2009/10	2010/11	2011/12
3	**E1**	**E2**	**E3**		
4		**E1**	**E2**	**E3**	
5			**E1**	**E2**	**E3**
6	**M1**	**M2**	**M3**		
7		**M1**	**M2**	**M3**	
8			**M1**	**M2**	**M3**

Note. E denotes an elementary school cohort, M denotes a middle school cohort. Only one elementary and one middle school cohort were available in the Arizona data.

School Performance Models

For all models, we estimated school performance in the last focal year (Grade 5 or 8) of the two grade level bands, as well as using prior years of achievement data as dictated by the particular model. We applied eight alternative analytic models of school performance to the mathematics and reading/language arts achievement data in elementary and middle school for each state. The eight school performance models were: Percent Proficient (PP), gain score (Gain), transition matrix (TM), student growth percentile (SGP), value-added model (VAM), and three Multilevel Linear Model (MLM) estimates: focal year intercept or status (MLM0), focal year growth rate (Grate), and average MLM growth rate across the three years (AvGrate).

Percent Proficient (PP). PP was the NCLB required metric used by the state that calculated the percentage of students in each school that met or exceeded state benchmarks for proficiency in either mathematics or reading/language arts in each grade.

Average Gain Score. Gain scores were calculated as the prior academic year (Grade 4 or Grade 7) scale score in mathematics or reading/language arts subtracted from the focal year scale score (Grade 5 or Grade 8):

$$\text{Gain}_i = \Delta_i = Y_{it} - Y_{i(t-1)} \tag{1}$$

where Y_{it} was the assessment outcome for student i at time t. Student gain scores were averaged for each school (labeled "Gain" below).

Transition Matrix (TM). School performance estimates were computed from a table of the state's proficiency categories in the prior year crossed with the proficiency categories in the focal year (Grade 5 or Grade 8) which, in the case of five proficiency categories, created a transition matrix table of 25 cells. The percentage of students occurring in each of the cells was entered and then a weighting scheme was applied to each cell and the products were summed to create a TM school performance index. The weighting scheme awarded one of three scores: (a) -1 was recorded if the student moved down one or more categories from the previous year, (b) 0 was recorded if the student stayed in the same category, and (c) +1 was recorded if the student moved up one or more categories from the previous year (see Tindal, Nese, & Stevens, 2017). The weighted values were averaged across all cells to create an overall school TM index.

Student Growth Percentiles (SGP). Student growth percentiles were computed at the student level using the approach described by Betebenner (2009). A student's SGP was calculated by taking the current year test score and regressing it on the two prior years of test scores. Betebenner's (2009) approach uses ordinal methods (quantile regression) as well as B-spline, cubic polynomial smoothing of the resulting normative distribution of conditional regression estimates. The analysis results in a relative rank for each student in a conditional distribution of those who had similar scores in previous years. We used the R package *SGP* (Betebenner, & Iwaarden, 2011) to compute student estimates based on the regression of the two prior years of test scores on the current year's test score and then we aggregated student SGP for each school to create a median SGP as each school's SGP performance estimate.

Value-added Models (VAM). This mixed effects approach examined performance gains over years and included indicators for student membership in a particular school. This model is known generally as the "layered model" because layers of equations are added with each year of schooling (Ballou, Sanders, and Wright, 2004). For example, the model for our case with students with three years of data would be specified as follows:

$$Y_{0ij} = b_0 + u_0 + e_0 \tag{2a}$$
$$Y_{1ij} = b_1 + u_0 + u_1 + e_1 \tag{2b}$$
$$Y_{2ij} = b_2 + u_0 + u_1 + u_2 + e_2 \, , \tag{2c}$$

where Y_{tij} represents an assessment for student i at time t (grade) attending school j. The fixed mean for all students in the combination of grades and schools was μ_{tij}, while e_{tij} was the random deviation for student n from the mean, μ_{tij}. The layered model we used was limited to a maximum of three years and was applied separately to mathematics and reading/language arts.

Multilevel Linear Growth Model Initial Status, Focal Year Growth, and Average Growth (MLM0, MLM Growth Rate and MLM Average Growth Rate). We modeled student growth over the three elementary or three middle school grades with multilevel longitudinal analyses (Raudenbush & Bryk, 2002) using HLM 7.1 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011) and full maximum likelihood estimation. The conditional models included a level-1 model that specified student mathematics or reading/language arts scores predicted by a quadratic function of time of measurement, a level-2 model composed of the prediction of level-1 model parameters as a function of student mean values, and a level-3 model composed of the prediction of level-2 parameters as a function of school mean parameter values. Time was

centered on the focal year (Grade 5 or 8) for computation of MLM0 and MLM growth rate but was centered on the middle year (Grade 4 or 7) for computation of MLM average growth rate. We used a quadratic model based on previous findings (Bloom, Hill, Black, & Lipsey, 2008) as well as inspection of the data and statistical testing of alternative growth functions. Because only three time points were present, the model intercept and linear slope were random parameters but the variance of the quadratic parameter was fixed (note the omission of a residual term in equation 4c below) to obtain a model solution. We used two different centering definitions to take into account the curvilinear nature of growth. Although centering in the last, focal year is most consistent with the definition of other models we examined, it likely underestimates the amount of growth that occurs over the three year period because of deceleration. We therefore also centered on the middle grade in the three year span to produce an average growth rate over the three years. The resulting MLM model equations were:

Level 1 (Time):
$$(Y_{tij}) = \pi_{0ij} + \pi_{1ij}(\text{time}_{tij}) + \pi_{2ij}(\text{time squared}_{tij}) + e_{tij} \tag{3}$$

Level 2 (Students):
$$\pi_{0ij} = \beta_{00j} + r_{0ij} \tag{4a}$$
$$\pi_{1ij} = \beta_{10j} + r_{1ij} \tag{4b}$$
$$\pi_{2ij} = \beta_{20j} \tag{4c}$$

Level 3 (Schools):
$$\beta_{00j} = \gamma_{000} + u_{00j} \tag{5a}$$
$$\beta_{10j} = \gamma_{100} + u_{10j} \tag{5b}$$
$$\beta_{20j} = \gamma_{200} + u_{20j} \tag{5c}$$

where Y_{tij} was the mathematics or reading/language arts scale score for student i at time t in school j, π_{0ij} was the initial status or intercept for student i at time 0 in school j, π_{1ij} was the linear rate of change, π_{2ij} was the quadratic curvature representing the acceleration or deceleration in each student's growth trajectory and e_{tij} was the residual for each student. At level-2, the level-1 parameters were modeled using mean parameter values across students (β_{k0j}) and at level-3, the level-2 parameters were modeled using mean parameter values across schools (γ_{k0j}).

Comparison of Model Estimates

We used several comparison criteria to evaluate the comparability and stability of school estimates across school performance models and across cohorts. In each state technical report we describe the results of our evaluation of school performance estimates. We examined: (a) correlations of model estimates for each school across the three cohorts, (b) correlations among school estimates from one model to another, (c) correlations among the school estimates and school composition variables (e.g., percent free/reduced lunch in the school, percent minority students in the school), and (d) correlations of each model with the percentage of students with disabilities in the school.

Comparison of School Ranks Based on Model Estimates

Many states and districts create school ranks based on their accountability system results. To compare the alternative school performance models using this metric, we created school

percentile ranks (from 1 to 99, with 99 being the highest performance) based on each of the school performance model estimates described above. In one of the only studies evaluating school performance models, Goldschmidt, Choi, and Beaudoin (2012) compared models using quintiles. They examined the percentage of times schools remained in the same quintile band based on one school performance model versus another. Similarly, Castellano and Ho (2013) compared SGP and conditional regression models by examining the percentage of times schools remained within 1, 5 or 10 percentile ranks for each model. To maintain some comparability with each of these studies, we used three levels of similarity in school ranks, computing the percentage of schools within 5, 10, or 20 ranks of each other. We also computed the Spearman's correlation of school ranks from one cohort to another or from one school performance model to another. As a final comparison metric, we computed the root mean squared difference (RMSD) between school ranks based on each pair of cohorts or each pair of school performance models (see Castellano & Ho, 2013):

$$RMSD_{c,c} = \sqrt{\frac{\sum_{j=1}^{j}(Rank_{jc}-Rank_{jc})^2}{n}} \tag{6}$$

In equation 6, for a particular school performance model, the RMSD computes the difference ($Rank_{it}$) between each school's rank in one cohort (jt) versus the school's rank in a second cohort (ju), squaring the difference, summing across all schools, dividing by the number of schools, n, and taking the square root of the result.

$$RMSD_{mn} = \sqrt{\frac{\sum(Rank_{jm}-Rank_{jn})^2}{n}} \tag{7}$$

Similarly, in equation 7, the school ranks arising from alternative school performance models are compared in which $Rank_{jm}$ and $Rank_{jn}$ represent the rank of school j using school performance model m compared to that school's rank using school performance model n. As in equation 6, differences in ranks are then summed, squared, divided by the number of schools and taken to the ½ power. The RMSD was a measure of similarity in school performance models where a lower value indicates a pair of models that rank schools most similarly.

Summary

We evaluated eight models for estimating school academic performance in mathematics and reading/language arts using operational state accountability data. In NC, OR, and PA, we examined stability in model estimates across three successive student cohorts in mathematics and reading/language arts in both elementary and middle school grades. In all four states, we also compared the estimates of school performance from one model to another to determine whether the models provided similar or different depictions of school performance, although several models could not be estimated in Pennsylvania because their test did not have a vertically linked score scale. We then compared the degree to which model estimates correlated with variables that described the student composition of the school, a likely indication of construct irrelevant variance. Ideally estimates of school performance should not be related to the student composition of the school. Last, we evaluated the school performance models in terms of the way they ranked schools, the stability of school ranks across cohorts, and the degree of

agreement in school rankings from one school performance model to another. Detailed results of these analyses and comparisons follow for the state of North Carolina.

North Carolina Study

Method

Sample

The North Carolina sample was separated into an elementary school sample (Grades 3 through 5) and a middle school sample (Grades 6 through 8), each consisting of three successive cohorts of students enrolled in school years: (a) 2007/08 through 2009/2010; (b) 2008/09 through 2010/11; and (c) 2009/10 through 2011/12. The initial sample included students across the three cohorts whose Grade 5 (elementary school sample) or Grade 8 (middle school sample) North Carolina End-of-Grade reading comprehension or mathematics scores on the general or alternate assessment were included in the state calculation of Adequate Yearly Progress (AYP). There was a small number of cases where a unique student identifier appeared to have been associated with more than one student in a year. When conflicting reading or mathematics scores were associated with a student identifier, all records were removed for that student identifier in that year. The initial elementary school sample for the mathematics test was 335,071 students. The initial middle school sample for the mathematics test was 317,015 students. The initial elementary school sample for the reading comprehension test was 334,684 students. The initial middle school sample for the reading comprehension test was 316,669 students.

To create an analytic sample that was appropriate for our research questions, we only included students with valid test scores in all three years in schools that served all three grades (Grades 3 through 5 or 6 through 8). Students who did not follow the typical grade level sequence due to grade retention, acceleration, or dubious progressions were excluded from the sample; this included the transition from 2006/07 to 2007/08, so that no students present in 2007/08 had been retained or accelerated from the previous year. We included only schools that served all three grades for a cohort, and schools with $N \geq 10$ students in each of the three cohorts in the final reference year of the three-year grade level band (i.e., Grade 5 for elementary grades 3 to 5 and Grade 8 for middle grades 6 to 8). Students and schools that did not meet these criteria were excluded from analyses. As is the case in most operational and research applications of these models, we made no attempt to account for student mobility in years prior to the focal year or to make any attributions of "school effects" based on how many years the student had been in the focal year school. Our concern in creating the analytic sample was to maximize the interpretation of comparisons of the models rather than to ensure complete representativeness of the samples. These inclusion rules were applied to ensure that there were no differences in the analytic samples for different school models so that comparisons of school models were a function only of differences in the models and not the composition of the sample analyzed. The final elementary school analytic sample for the mathematics test was 230,492 students (68.79% of the initial sample). The final middle school analytic sample for the mathematics test was 224,492 students (70.81%). The final elementary school analytic sample for the reading comprehension test was 228,492 students (68.27%). The final middle school analytic sample for the reading comprehension test was 223,530 students (70.59%).

Table 1 provides summary statistics describing the school-level analytical samples of North Carolina elementary and middle school students in the three cohorts for mathematics and reading comprehension. School composition variables reported in the table include the percent of English Language Learners (ELL), females, economically disadvantaged students (EDS), ethnic minorities, and students with disabilities (SWD). Although variation existed from cohort to cohort in sample demographic characteristics, generally the composition of the samples was quite similar across the three cohorts. One exception was a small, but consistent increase in the proportions of students who were EDS or racial/ethnic minorities across the three cohorts for both grade level bands. Also, a slightly greater percentage of SWD and English Language Learners (ELL) participated in the mathematics than the reading comprehension assessment across all cohorts and grade bands. There was much greater school level variation–as indicated by the values of the standard deviations in parentheses–in EDS and racial/ethnic minority student school composition than other student characteristics. It should be noted that when we refer to "school" composition, it references variables representing a particular cohort in each school in our analytic samples. Because we excluded students and schools to create our analytic samples, "total school" characteristics may differ slightly from the variables reported here.

Table 1

Proportion and Standard Deviation (in parentheses) of Student Subgroups for the North Carolina Analytical Samples by Content Area and Grade Level Band

		Cohort		
		1	2	3
Mathematics Elementary	ELL	0.054 (0.074)	0.053 (0.073)	0.054 (0.073)
	Female	0.502 (0.072)	0.497 (0.074)	0.499 (0.074)
	EDS	0.529 (0.248)	0.543 (0.243)	0.566 (0.248)
	Ethnic Minority	0.447 (0.304)	0.458 (0.302)	0.463 (0.300)
	SWD	0.091 (0.055)	0.090 (0.055)	0.090 (0.053)
Reading Comprehension Elementary	ELL	0.053 (0.073)	0.051 (0.072)	0.052 (0.071)
	Female	0.504 (0.072)	0.500 (0.074)	0.502 (0.074)
	EDS	0.528 (0.249)	0.541 (0.243)	0.565 (0.249)
	Ethnic Minority	0.447 (0.304)	0.457 (0.302)	0.462 (0.301)

	SWD	0.084	0.082	0.082
		(0.055)	(0.055)	(0.053)
Mathematics Middle	ELL	0.046	0.042	0.040
		(0.061)	(0.057)	(0.055)
	Female	0.511	0.509	0.506
		(0.065)	(0.065)	(0.062)
	EDS	0.480	0.500	0.526
		(0.225)	(0.221)	(0.222)
	Ethnic Minority	0.433	0.443	0.445
		(0.284)	(0.286)	(0.285)
	SWD	0.075	0.076	0.075
		(0.050)	(0.041)	(0.043)
Reading Comprehension Middle	ELL	0.046	0.041	0.038
		(0.060)	(0.056)	(0.053)
	Female	0.512	0.510	0.508
		(0.065)	(0.065)	(0.062)
	EDS	0.479	0.499	0.525
		(0.225)	(0.221)	(0.222)
	Ethnic Minority	0.433	0.442	0.444
		(0.285)	(0.286)	(0.285)
	SWD	0.072	0.072	0.070
		(0.050)	(0.041)	(0.044)

Instrument

The outcome measures for all analyses were the third editions of the North Carolina End-of-Grade Tests in Mathematics (EOG-M) and Reading Comprehension (EOG-RC), which are standardized, vertically scaled tests designed to measure the core content standards in the state curriculum (NC Department of Public Instruction, 2008, 2009). Both tests were in a multiple-choice format and were the primary assessments used in the state's school accountability model. EOG raw scores were converted to developmental scale scores based on the number of items answered correctly, taking item difficulty into account using the three-parameter logistic model of item response theory (IRT) methods and a vertical linking design over grades (NC Department of Public Instruction, 2008, 2009).

Results and Discussion

This technical report is organized in three sections: Section A describes school performance model estimates, Section B describes school ranks, and Appendices provide additional detailed results.

Section A: School Performance Estimates

Cohort stability. We first considered the stability of model estimates by computing the correlations among estimates across the three successive cohorts of students. It should be noted that cohort comparisons are both an indication of changes in the composition of students in the

school from one academic year to another as well as any other temporal changes that occur from one year to another including changes in policy, practice, instruction, or other factors that impact student test scores. Table 2 shows the correlation of model estimates across cohorts for mathematics and reading comprehension in the elementary school and middle school samples. As can be seen in Table 2, correlations generally ranged only from small to moderate for the model estimates (with the exception of the MLM0 and PP estimates) indicating some substantial instability in school performance estimates across cohorts. Correlations between adjacent years in the first two columns (cohort 1 with 2 or 2 with 3) are generally somewhat higher than the comparisons across two years (cohort 1 with 3). Although there is also some variation from elementary to middle school or from mathematics to reading/language arts, trends in cohort stability were fairly similar across content area and grade level band. To facilitate interpretation

Table 2

Correlations of School Performance Model Estimates Across Cohorts by Content Area and Grade Level Band

Elementary Schools

	Mathematics			Reading Comprehension		
Model	1 with 2	2 with 3	1 with 3	1 with 2	2 with 3	1 with 3
PP	0.724	0.675	0.641	0.721	0.648	0.648
MLM0	0.833	0.806	0.771	0.870	0.853	0.847
Gain	0.429	0.444	0.299	0.363	0.343	0.256
TM	0.378	0.378	0.264	0.325	0.340	0.253
SGP	0.486	0.460	0.332	0.316	0.335	0.268
VAM	0.523	0.498	0.354	0.416	0.405	0.325
Grate	0.397	0.425	0.278	0.288	0.253	0.220
AvGrate	0.551	0.523	0.355	0.632	0.618	0.539

Middle Schools

	Mathematics			Reading Comprehension		
Model	1 with 2	2 with 3	1 with 3	1 with 2	2 with 3	1 with 3
PP	0.794	0.745	0.703	0.875	0.871	0.842
MLM0	0.889	0.872	0.846	0.921	0.910	0.902
Gain	0.522	0.553	0.440	0.205	0.149	0.275
TM	0.439	0.076	0.003	0.409	0.443	0.440
SGP	0.594	0.616	0.489	0.355	0.338	0.355
VAM	0.642	0.637	0.503	0.501	0.484	0.465
Grate	0.503	0.508	0.405	0.215	0.179	0.218
AvGrate	0.631	0.624	0.469	0.561	0.442	0.423

of the cohort results, we also averaged correlations across the two content areas and grade levels (see Table 3). It can be seen that the correlations across cohorts were largest for the two status-based school performance measures (PP and MLM0) and noticeably lower for all other models that used two or three years of data to estimate school performance. The two rightmost columns of Table 3 show the overall mean and standard deviation across the cohort comparisons for each school performance model. The greatest agreement over cohorts, content, and grade level was for the MLM0 estimates (MLM focal year intercepts), closely followed by the PP model estimates. All remaining multiyear performance models had much greater instability. The standard deviation of correlations across cohort comparisons shown in the rightmost column of Table 3 also show the least variability over cohorts for the two status models and the greatest variability across cohort correlations for the Transition model followed by the AvGrate model.

Table 3

Average Correlations Across Content Area and Grade Level Band and Overall Mean and Standard Deviation (SD) Across the Three Cohort Comparisons

Model	1 with 2	2 with 3	1 with 3	Mean	SD
PP	0.778	0.735	0.708	0.740	0.037
MLM0	0.878	0.860	0.842	0.860	0.019
Gain	0.380	0.372	0.318	0.357	0.065
TM	0.388	0.309	0.240	0.312	0.091
SGP	0.438	0.437	0.361	0.412	0.049
VAM	0.520	0.506	0.412	0.479	0.059
Grate	0.351	0.341	0.280	0.324	0.048
AvGrate	0.594	0.552	0.446	0.531	0.081
Mean	0.541	0.514	0.451	--	--

Comparison of models. We computed the correlations of school performance estimates from one model to another within each of the three cohorts and then took the mean correlation across cohorts. Correlations of model estimates within each individual cohort are presented in Appendix A. Table 4 shows model correlations for mathematics and reading/language arts in the elementary school and middle school samples averaged over the three cohorts.

Table 4

Correlations of School Performance Estimates Across Models by Content Area and Grade Level Band

Elementary School Mathematics

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	0.914	0.246	0.272	0.441	0.488	0.208	0.257

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
MLM0		0.256	0.258	0.473	0.520	0.225	0.259
Gain			0.928	0.878	0.892	0.964	0.609
TM				0.822	0.833	0.895	0.564
SGP					0.966	0.764	0.809
VAM						0.777	0.838
Grate							0.394

Elementary School Reading Comprehension

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	0.849	-0.123	-0.158	0.382	0.457	-0.069	-0.426
MLM0		-0.198	-0.232	0.392	0.483	-0.109	-0.561
Gain			0.870	0.651	0.660	0.909	0.511
TM				0.551	0.547	0.785	0.489
SGP					0.905	0.514	0.343
VAM						0.550	0.333
Grate							0.216

Middle School Mathematics

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	0.897	0.178	0.233	0.523	0.559	0.139	0.174
MLM0		0.117	0.199	0.508	0.552	0.091	0.113
Gain			0.600	0.829	0.833	0.966	0.689
TM				0.559	0.568	0.564	0.438
SGP					0.971	0.730	0.797
VAM						0.734	0.810
Grate							0.508

Middle School Reading Comprehension

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	0.960	-0.072	-0.519	0.536	0.581	-0.020	-0.368

Model						
MLM0	-0.105	-0.570	0.510	0.580	-0.035	-0.420
Gain		0.709	0.588	0.598	0.876	0.446
TM			0.165	0.124	0.608	0.530
SGP				0.880	0.468	0.345
VAM					0.527	0.382
Grate						0.209

Average Over Content Area and Grade Level Band

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	0.905	0.057	-0.043	0.471	0.521	0.065	-0.091
MLM0		0.018	-0.086	0.471	0.534	0.043	-0.012
Gain			0.777	0.737	0.746	0.929	0.564
TM				0.525	0.518	0.713	0.505
SGP					0.931	0.619	0.573
VAM						0.647	0.591
Grate							0.332

As evident in Table 4, substantial variability was present in the degree to which school performance estimates for one model were related to other models and the correlations among models varied by content area and grade level band. For example, the correlation between the MLM0 model and the Transition model ranged from -.570 to +.258 and between PP and MLM AvGrate ranged from -.426 to +.257. The least variation in model correlations across content area and grade level band was for the MLM and VAM models from +.483 to +.580.

As shown in the last panel of Table 4, on average across content area and grade level band, the highest correlations were among the SGP and VAM models (+.931), the Gain and Grate models (+.929), and the MLM intercept (MLM0) with the PP model (+.905). The lowest correlations were between the PP and MLM average growth rate (AvGrate) models (-.091) and the MLM0 model and the Transition model (-.086). The average correlation of the two status models (PP, MLM0) with the remaining six multiyear models was only +.162. Average correlations among the six multiple year models ranged from +.505 to +.931 with one exception, the correlation of the Grate and AvGrate models was only +.332, with an average correlation among all six multiyear models of +.607.

We also examined the degree to which school performance model estimates were consistent from one content area to the other. Table 5 shows model estimate agreement across content areas in each cohort as well as the average across the three cohorts. As presented in Table 5, correlations were generally higher between content areas in elementary than middle school. On average, correlations for the two status models (PP and MLM0) were greater than +.750 and higher than average correlations for the other models that ranged from +.135 to +.532.

Table 5

Correlations of School Performance Model Estimates between Mathematics and Reading Comprehension by Grade Level Band in each Cohort and Averaged over Cohorts

| | Elementary Schools | | | | Middle Schools | | | |
| | Cohort | | | Mean | Cohort | | | |
Model	1	2	3		1	2	3	Mean
PP	0.998	0.997	0.781	0.925	0.780	0.776	0.778	0.778
MLM0	0.890	0.871	0.869	0.877	0.882	0.872	0.874	0.876
Gain	0.444	0.452	0.471	0.456	0.320	0.302	0.288	0.303
TM	0.352	0.375	0.365	0.364	0.225	0.167	0.013	0.135
SGP	0.457	0.477	0.449	0.461	0.320	0.284	0.375	0.326
VAM	0.538	0.522	0.536	0.532	0.452	0.359	0.448	0.420
Grate	0.424	0.411	0.440	0.425	0.297	0.309	0.262	0.289
AvGrate	0.441	0.506	0.512	0.486	0.483	0.400	0.505	0.463

Relation with school composition variables. We computed the correlation of model estimates with school composition variables to determine whether estimates were related to the aggregated student characteristics in each school. Table 6 shows the correlations of model estimates with school composition variables for mathematics and reading comprehension in the elementary school and middle school samples. Correlations of model estimates with school composition variables within each individual cohort are presented in Appendix B.

The rightmost column of Table 6 shows the average correlation of each school performance model with the school composition variables across all school composition variables. As can be seen, correlations of the status models, PP and MLM0, were negative and noticeably stronger than the correlations of the other school performance models with school composition variables. On average across content and grade level band, the correlation of the school composition variables was -0.253 for the PP model and -0.260 for the MLM0 model. In contrast, the average correlations of the school composition variables with the remaining models were quite low ranging from -0.064 to +0.130. Thus there was relatively little relation of the multiyear models with school composition, but for the status models performance estimates were higher when fewer students from protected subgroups were present in the school. No clear pattern was present for the relation between school size and model estimates.

Table 6

Correlations of Model Estimates with School Composition Variables by Content Area and Grade Level Band

Elementary School Mathematics

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.591	-0.279	-0.113	0.001	-0.534	0.216	-0.216
MLM0	-0.688	-0.303	-0.089	-0.017	-0.532	0.291	-0.223
Gain	0.068	0.105	-0.043	0.000	0.123	-0.002	0.042
TM	0.043	0.093	-0.037	0.004	0.105	0.001	0.035
SGP	-0.056	0.043	-0.039	0.003	0.033	0.033	0.003
VAM	-0.086	0.031	-0.042	0.005	0.013	0.039	-0.007
Grate	0.026	0.085	-0.050	-0.008	0.076	0.032	0.027
AvGrate	0.138	0.119	0.009	0.022	0.197	-0.090	0.066

Elementary School Reading Comprehension

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.648	-0.349	-0.104	-0.002	-0.571	0.230	-0.241
MLM0	-0.818	-0.458	-0.084	0.007	-0.638	0.278	-0.285
Gain	0.261	0.213	-0.011	-0.021	0.311	-0.054	0.116
TM	0.280	0.218	-0.004	-0.021	0.339	-0.067	0.124
SGP	-0.197	-0.039	-0.027	0.003	-0.055	0.072	-0.041
VAM	-0.273	-0.067	-0.036	0.005	-0.099	0.097	-0.062
Grate	0.139	0.130	-0.029	-0.029	0.197	0.013	0.070
AvGrate	0.584	0.411	0.076	0.018	0.542	-0.234	0.233

Middle School Mathematics

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.578	-0.317	-0.190	0.057	-0.547	0.121	-0.242
MLM0	-0.675	-0.310	-0.151	0.048	-0.527	0.212	-0.234
Gain	0.203	0.131	-0.023	0.021	0.153	-0.047	0.073

TM	0.048	0.037	0.003	-0.008	0.040	0.008	0.021
SGP	-0.071	-0.002	-0.076	0.058	-0.060	0.004	-0.024
VAM	-0.097	-0.009	-0.090	0.046	-0.071	0.010	-0.035
Grate	0.174	0.119	-0.022	0.011	0.127	-0.013	0.066
AvGrate	0.246	0.148	-0.024	0.038	0.210	-0.148	0.078

Middle School Reading Comprehension

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.784	-0.455	-0.146	0.052	-0.690	0.153	-0.312
MLM0	-0.813	-0.438	-0.136	0.050	-0.647	0.198	-0.298
Gain	0.100	0.060	0.055	-0.036	0.127	0.008	0.052
TM	0.507	0.274	0.075	-0.008	0.435	-0.126	0.193
SGP	-0.388	-0.190	-0.051	0.018	-0.280	0.048	-0.140
VAM	-0.427	-0.221	-0.058	-0.004	-0.291	0.102	-0.150
Grate	0.037	0.011	0.017	-0.031	0.067	0.056	0.026
AvGrate	0.401	0.235	0.064	-0.037	0.358	-0.173	0.141

Relation of model estimates to SWD school composition. Because of the NCAASE emphasis on the performance and academic growth of SWD, we also focused more specifically on the relations between the percentage of SWD students served by a school and the school performance model estimates. Correlations of model estimates with SWD school composition within each individual cohort are presented in Appendix C. Table 7 shows the correlation of model estimates with the percentage of SWD in each school for mathematics and reading/language arts in the elementary school and middle school samples averaged over cohorts. As can be seen in the bottom row of Table 7, average school performance estimates based on the single-year, status models (PP and MLM0) had correlations with school SWD composition that were substantially stronger in magnitude than the other school performance models. With the PP and MLM0 models, school performance estimates were higher the smaller the percentage of SWD students in the school and lower to the extent that the school served larger proportions of SWD.

Table 7

Average School Performance Model Estimates as a Function of the Percentage of SWD in the School by Content and Grade Level Band

Content Area and Grade Level Band	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
Math Elementary	-0.113	-0.089	-0.043	-0.037	-0.039	-0.042	-0.050	0.009

Reading Comprehension Elementary	-0.104	-0.084	-0.011	-0.004	-0.027	-0.036	-0.029	0.076
Math Middle	-0.190	-0.151	-0.023	0.003	-0.076	-0.090	-0.022	-0.024
Reading Comprehension Middle	-0.146	-0.136	0.055	0.075	-0.051	-0.058	0.017	0.064
Mean	-0.138	-0.115	-0.005	0.009	-0.048	-0.056	-0.021	0.031

Summary of Section A. We evaluated eight alternative models for estimating school academic performance in mathematics and reading/language arts using operational North Carolina state accountability data. We observed substantial variability in model estimates across three successive student cohorts in mathematics and reading/language arts in both elementary and middle school grades. Variability across cohorts was considerably less for the two status models (PP and MLM0) than for the models that used more than one year of data. We also compared the estimates of school performance from one model to another and found substantial disagreement across models. In general, correlations within model type (i.e., single year or multiyear) were stronger than correlations where a status model was paired with a model using multiple years of data.

We also compared school performance estimates in mathematics with those in reading/language arts. Again, agreement was greater across content areas for the status models than for the multiple year models. The correlations of the status models (PP and MLM0) with student composition were stronger than the correlations of the multiple year models with student composition. Larger proportions of protected student subgroups were associated with lower school performance. Finally, we correlated school performance estimates with the percentage of SWD in each school. Ideally, estimates of school performance should be unrelated to the student composition of the school, but as with the other school composition variables, we found that the status models were more strongly correlated with SWD school composition than the multiyear model estimates.

Section B: School Ranks Based on School Performance Estimates

In this section, we focus on the examination of school ranks based on the school performance estimates reported in the previous section. It is a common practice for states and other jurisdictions is to rank schools as a method for evaluating academic performance. Therefore, using the estimates of school performance generated by the eight models described previously, we computed percentile ranks for each school. We then compared school ranks within each school performance model across the three cohorts used in the study. Next, we compared the school ranks for each model to the ranks obtained from each of the other models. Finally, we examined the relation between school ranks from each model with variables describing the student composition of each school. Three criteria were used to evaluate the comparisons of school ranks: (a) the Spearman's correlation between school ranks, (b) the proximity of absolute school ranks, and (b) the root mean square difference (RMSD) in school ranks.

Comparison of cohorts. We first consider the stability of school ranks within each school performance model across the three successive cohorts of students in mathematics and reading/language arts in the elementary and middle school grades. We computed the Spearman's correlation of the school ranks from one cohort to the school ranks from each of the other two cohorts within each of the eight school performance models to determine the stability of school ranks. As mentioned in Section A, cohort comparisons are both an indication of changes in the composition of students in the school from one academic year to another as well as any other temporal changes that occur from one year to another including changes in policy, practice, instruction, or other factors that impact student test scores. Table 8 shows the correlation of school ranks across cohorts for mathematics and reading/language arts in the elementary school and middle school samples. As can be seen in Table 8, the majority of the correlations fell in the moderate range indicating considerable variability in school ranks from one cohort to another. As would be expected, correlations between adjacent years in the first two columns (cohort 1 with 2 or 2 with 3) were generally somewhat higher than the comparison across two years (cohort 1 with 3). Results for mathematics tended to be more similar to those for reading in the elementary grades than the middle grades, and correlations in school ranks were more similar in the middle grades than the elementary grades.

Table 8

Spearman's Correlations of Model School Ranks for Each Pair of Cohorts by Content Area and Grade Level Band

	Elementary Schools					
	Mathematics			Reading Comprehension		
Model	1 with 2	2 with 3	1 with 3	1 with 2	2 with 3	1 with 3
PP	0.726	0.693	0.653	0.724	0.650	0.655
MLM0	0.816	0.782	0.748	0.861	0.837	0.837
Gain	0.414	0.436	0.307	0.325	0.316	0.237
TM	0.365	0.374	0.273	0.279	0.303	0.229
SGP	0.486	0.443	0.329	0.309	0.318	0.260
VAM	0.518	0.482	0.355	0.400	0.379	0.309
Grate	0.380	0.417	0.290	0.269	0.248	0.212
AvGrate	0.539	0.498	0.348	0.612	0.603	0.526

<div align="center">Middle Schools</div>

	Mathematics			Reading Comprehension		
Model	1 with 2	2 with 3	1 with 3	1 with 2	2 with 3	1 with 3
PP	0.805	0.753	0.733	0.873	0.857	0.838
MLM0	0.880	0.854	0.827	0.911	0.898	0.893
Gain	0.488	0.492	0.405	0.212	0.168	0.245
TM	0.444	0.059	0.003	0.440	0.474	0.436
SGP	0.565	0.574	0.448	0.360	0.368	0.349
VAM	0.618	0.607	0.477	0.476	0.478	0.468
Grate	0.469	0.455	0.379	0.205	0.164	0.215
AvGrate	0.592	0.593	0.437	0.543	0.444	0.435

To facilitate further interpretation, we averaged the results shown in Table 8 across content area and grade level band. As can be seen in Table 9, on average the greatest stability was for the two status models, PP and MLM0. Noticeably lower correlations occurred for the remaining school performance models, all of which were based on more than one year of data, with the least stability for the TM, Grate, and Gain models.

Table 9

Spearman's Correlations of Model School Ranks Averaged Across Content Area and Grade Level Band and Overall Mean and Standard Deviation (SD) Across the Three Cohort Comparisons

Model	1 with 2	2 with 3	1 with 3	Mean	SD
PP	0.782	0.738	0.720	0.747	0.033
MLM0	0.867	0.843	0.826	0.845	0.021
Gain	0.360	0.353	0.298	0.337	0.051
TM	0.382	0.302	0.235	0.306	0.089
SGP	0.430	0.426	0.347	0.401	0.048
VAM	0.503	0.486	0.402	0.464	0.054
Grate	0.331	0.321	0.274	0.309	0.042
AvGrate	0.572	0.534	0.436	0.514	0.074

Our second criterion for comparing school ranks was to determine how much a school's rank changed from one cohort to another. Table 10 shows the proportion of schools that were within 5, 10, or 20 ranks in one cohort versus another for each school performance model in mathematics and reading/language arts at each grade level band. The last table entry for each school performance model shows the average differences in school ranks averaged over content area and grade level band. It can be seen that on average for the PP model, about one quarter to one third of the schools differed by only 5 percentile ranks or less, about half of schools differed by 10 ranks or less, and about 70-75% differed by 20 ranks or less. This also indicates that about 25% of schools differed by more than 20 ranks from one cohort to another. The results for the MLM0 model showed somewhat greater agreement in school ranks across cohorts. However, the level of agreement in school ranks across cohorts was noticeably lower for all of the remaining models that were based on two or more years of achievement data. For example, school ranks based on the remaining models (Gain, TM, SGP, VAM, Grate, and AvGrate) differed by more than 20 ranks for about 50% of the schools.

Table 10

Proportion of Elementary or Middle Schools Within 5, 10, or 20 Ranks of Each Other for Each School Performance Model for Each Pair of Cohorts in Mathematics and Reading Comprehension

PP

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.271	0.464	0.699
	2 vs. 3	0.247	0.432	0.676
	1 vs. 3	0.253	0.422	0.666
Reading Comprehension Elementary	1 vs. 2	0.262	0.460	0.698
	2 vs. 3	0.247	0.408	0.644
	1 vs. 3	0.242	0.417	0.643
Mathematics Middle	1 vs. 2	0.367	0.556	0.779
	2 vs. 3	0.320	0.501	0.763
	1 vs. 3	0.274	0.479	0.728
Reading Comprehension Middle	1 vs. 2	0.410	0.625	0.856
	2 vs. 3	0.367	0.578	0.860
	1 vs. 3	0.343	0.562	0.813
Mean	1 vs. 2	0.328	0.526	0.758
	2 vs. 3	0.295	0.480	0.736
	1 vs. 3	0.278	0.470	0.712

MLM0

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.347	0.522	0.775
	2 vs. 3	0.299	0.488	0.744
	1 vs. 3	0.306	0.474	0.705
Reading Comprehension Elementary	1 vs. 2	0.383	0.583	0.842
	2 vs. 3	0.330	0.543	0.806
	1 vs. 3	0.330	0.531	0.794
Mathematics Middle	1 vs. 2	0.418	0.629	0.850
	2 vs. 3	0.391	0.611	0.846
	1 vs. 3	0.345	0.544	0.813
Reading Comprehension Middle	1 vs. 2	0.462	0.730	0.909
	2 vs. 3	0.444	0.663	0.880
	1 vs. 3	0.402	0.647	0.872
Mean	1 vs. 2	0.402	0.616	0.844
	2 vs. 3	0.366	0.576	0.819
	1 vs. 3	0.346	0.549	0.796

Gain

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.186	0.326	0.528
	2 vs. 3	0.180	0.319	0.532
	1 vs. 3	0.160	0.298	0.493
Reading Comprehension Elementary	1 vs. 2	0.164	0.285	0.495
	2 vs. 3	0.156	0.272	0.470
	1 vs. 3	0.133	0.241	0.439

	Cohort	r = 5	r = 10	r = 20
Mathematics Middle	1 vs. 2	0.209	0.343	0.554
	2 vs. 3	0.207	0.357	0.568
	1 vs. 3	0.170	0.302	0.519
Reading Comprehension Middle	1 vs. 2	0.158	0.294	0.481
	2 vs. 3	0.124	0.262	0.440
	1 vs. 3	0.146	0.268	0.450
Mean	1 vs. 2	0.179	0.312	0.514
	2 vs. 3	0.167	0.302	0.502
	1 vs. 3	0.152	0.277	0.475

TM

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.171	0.307	0.511
	2 vs. 3	0.183	0.313	0.502
	1 vs. 3	0.153	0.262	0.471
Reading Comprehension Elementary	1 vs. 2	0.142	0.263	0.467
	2 vs. 3	0.148	0.279	0.459
	1 vs. 3	0.148	0.256	0.440
Mathematics Middle	1 vs. 2	0.185	0.353	0.570
	2 vs. 3	0.108	0.185	0.369
	1 vs. 3	0.120	0.245	0.393
Reading Comprehension Middle	1 vs. 2	0.183	0.325	0.564
	2 vs. 3	0.168	0.304	0.572
	1 vs. 3	0.181	0.302	0.536
Mean	1 vs. 2	0.170	0.312	0.528
	2 vs. 3	0.152	0.270	0.476
	1 vs. 3	0.150	0.266	0.460

SGP

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.174	0.340	0.561
	2 vs. 3	0.165	0.293	0.551
	1 vs. 3	0.173	0.282	0.485
Reading Comprehension Elementary	1 vs. 2	0.151	0.263	0.487
	2 vs. 3	0.150	0.281	0.478
	1 vs. 3	0.160	0.257	0.459
Mathematics Middle	1 vs. 2	0.209	0.349	0.592
	2 vs. 3	0.229	0.359	0.604
	1 vs. 3	0.205	0.357	0.550
Reading Comprehension Middle	1 vs. 2	0.148	0.264	0.495
	2 vs. 3	0.174	0.298	0.501
	1 vs. 3	0.160	0.294	0.471
Mean	1 vs. 2	0.170	0.304	0.534
	2 vs. 3	0.180	0.308	0.534
	1 vs. 3	0.174	0.298	0.491

VAM

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.211	0.352	0.579
	2 vs. 3	0.187	0.325	0.557
	1 vs. 3	0.174	0.286	0.498
Reading Comprehension Elementary	1 vs. 2	0.177	0.282	0.498
	2 vs. 3	0.178	0.285	0.500
	1 vs. 3	0.148	0.274	0.453

Mathematics Middle	Cohort	r = 5	r = 10	r = 20
Mathematics Middle	1 vs. 2	0.219	0.365	0.635
	2 vs. 3	0.211	0.349	0.647
	1 vs. 3	0.209	0.331	0.560
Reading Comprehension Middle	1 vs. 2	0.170	0.325	0.533
	2 vs. 3	0.205	0.310	0.542
	1 vs. 3	0.191	0.310	0.548
Mean	1 vs. 2	0.194	0.331	0.561
	2 vs. 3	0.195	0.317	0.562
	1 vs. 3	0.180	0.300	0.515

Grate

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.184	0.315	0.531
	2 vs. 3	0.179	0.323	0.541
	1 vs. 3	0.149	0.279	0.485
Reading Comprehension Elementary	1 vs. 2	0.137	0.254	0.460
	2 vs. 3	0.142	0.240	0.448
	1 vs. 3	0.121	0.230	0.435
Mathematics Middle	1 vs. 2	0.195	0.329	0.544
	2 vs. 3	0.183	0.310	0.554
	1 vs. 3	0.144	0.290	0.479
Reading Comprehension Middle	1 vs. 2	0.166	0.278	0.452
	2 vs. 3	0.140	0.264	0.430
	1 vs. 3	0.122	0.243	0.432
Mean	1 vs. 2	0.170	0.294	0.497
	2 vs. 3	0.161	0.284	0.493
	1 vs. 3	0.134	0.260	0.458

AvGrate

	Cohort	r = 5	r = 10	r = 20
Mathematics Elementary	1 vs. 2	0.205	0.356	0.609
	2 vs. 3	0.193	0.336	0.565
	1 vs. 3	0.159	0.293	0.496
Reading Comprehension Elementary	1 vs. 2	0.226	0.365	0.601
	2 vs. 3	0.208	0.365	0.604
	1 vs. 3	0.211	0.351	0.573
Mathematics Middle	1 vs. 2	0.207	0.377	0.611
	2 vs. 3	0.262	0.398	0.629
	1 vs. 3	0.201	0.325	0.538
Reading Comprehension Middle	1 vs. 2	0.164	0.316	0.560
	2 vs. 3	0.144	0.310	0.529
	1 vs. 3	0.166	0.320	0.548
Mean	1 vs. 2	0.200	0.354	0.595
	2 vs. 3	0.202	0.352	0.582
	1 vs. 3	0.184	0.322	0.539

Our third criterion for comparing school ranks was to calculate the root mean square difference (RMSD) between cohorts or models as defined in the report introduction. Table 11 shows the RMSD across pairs of cohorts by content area and grade level band for each of the eight school performance models and in the last two columns the mean and standard deviation (SD) across cohort comparisons. As can be seen in the table, the smallest differences in rank were for the MLM0 model, about 12-18 ranks on average, followed by the PP model. Average differences in school rank across cohorts for the remaining models ranged from about 26 to 36.

Table 11

RMSD in School Ranks for Each Student Cohort for Each School Performance Model by Content Area and Grade Level Band

Elementary School Mathematics

Model	1 with 2	2 with 3	1 with 3	Mean	SD
PP	21.114	22.383	23.798	22.432	1.343
MLM0	17.313	18.844	20.256	18.804	1.472
Gain	30.906	30.330	33.624	31.620	1.759
TM	32.189	31.939	34.439	32.856	1.377
SGP	28.965	30.143	33.078	30.729	2.118
VAM	28.043	29.060	32.445	29.849	2.305
Grate	31.802	30.845	34.019	32.222	1.628
AvGrate	27.413	28.608	32.613	29.545	2.724
Mean	27.218	27.769	30.534	--	--

Elementary School Reading Comprehension

Model	1 with 2	2 with 3	1 with 3	Mean	SD
PP	21.196	23.864	23.733	22.931	1.504
MLM0	15.058	16.292	16.323	15.891	0.722
Gain	33.185	33.415	35.276	33.959	1.147
TM	34.284	33.708	35.468	34.487	0.897
SGP	33.581	33.361	34.739	33.894	0.740
VAM	31.280	31.834	33.571	32.228	1.195
Grate	34.531	35.017	35.850	35.133	0.667
AvGrate	25.138	25.459	27.799	26.132	1.453
Mean	28.532	29.119	30.345	--	--

Middle School Mathematics

Model	1 with 2	2 with 3	1 with 3	Mean	SD
PP	17.828	20.038	20.838	19.568	1.559
MLM0	13.986	15.407	16.755	15.383	1.385
Gain	28.851	28.764	31.128	29.581	1.340
TM	30.075	39.131	40.287	36.498	5.592
SGP	26.597	26.306	29.981	27.628	2.043
VAM	24.938	25.282	29.171	26.464	2.351
Grate	29.384	29.792	31.795	30.324	1.290

AvGrate	25.769	25.724	30.26	27.251	2.606
Mean	24.678	26.306	28.777	--	--

Middle School Reading Comprehension

Model	1 with 2	2 with 3	1 with 3	Mean	SD
PP	14.368	15.237	16.249	15.285	0.941
MLM0	12.001	12.892	13.203	12.699	0.624
Gain	35.801	36.795	35.058	35.885	0.872
TM	30.201	29.251	30.307	29.920	0.582
SGP	32.260	32.054	32.531	32.282	0.239
VAM	29.199	29.137	29.414	29.250	0.145
Grate	35.964	36.890	35.746	36.200	0.607
AvGrate	27.258	30.092	30.322	29.224	1.706
Mean	27.131	27.794	27.854	--	--

Comparison of models. We next compared school ranks from one model to another within each of the three cohorts. We first computed the Spearman's correlations among school ranks for the different models. These values were quite similar to the Spearman's correlations among school model estimates (see Table 4 and Appendix A) and for this reason they are not included in this report. Our second criterion for comparing school ranks was to determine how much a school's rank changed from one model to another. For each pair of school performance models, Table 12 shows the average percentage of schools that were within 5, 10, or 20 percentile ranks in one model versus the other. As can be seen in the table, three pairs of models produced results that were quite similar: (a) SGP vs. VAM, (b) Gain vs. Grate, and (c) PP vs. MLM0. In each of these three pairings, over 74% of schools were within 10 ranks of each other and over 92% were within 20 ranks of each other.

When a single year model (PP or MLM) was paired with a model that made use of multiyear results, the level of agreement in school ranks was much lower than when a single year model was paired with another single year or status model (PP and MLM) or a multiyear model with a multiyear model.

Table 12

Proportion of Elementary or Middle Schools Within 5, 10, or 20 Ranks of Each Other for Each Pair of School Performance Models in Mathematics and Reading Comprehension Averaged Over Cohorts

Model Comparison:	r = 5	r = 10	r = 20
PP vs. MLM0			
Math Elementary	0.494	0.750	0.939

Reading Comprehension Elementary	0.393	0.607	0.837
Math Middle	0.493	0.743	0.936
Reading Comprehension Middle	0.659	0.871	0.982
Mean	0.510	0.743	0.924

PP vs. Gain

Math Elementary	0.136	0.250	0.447
Reading Comprehension Elementary	0.098	0.183	0.344
Math Middle	0.124	0.237	0.417
Reading Comprehension Middle	0.100	0.193	0.368
Mean	0.114	0.216	0.394

PP vs. TM

Math Elementary	0.140	0.257	0.460
Reading Comprehension Elementary	0.095	0.169	0.325
Math Middle	0.143	0.231	0.429
Reading Comprehension Middle	0.066	0.115	0.255
Mean	0.111	0.193	0.367

PP vs. SGP

Math Elementary	0.170	0.293	0.516
Reading Comprehension Elementary	0.167	0.303	0.509
Math Middle	0.191	0.345	0.560
Reading Comprehension Middle	0.204	0.350	0.565
Mean	0.183	0.323	0.538

PP vs. VAM

Math Elementary	0.167	0.314	0.538
Reading Comprehension Elementary	0.181	0.324	0.536
Math Middle	0.211	0.366	0.591
Reading Comprehension Middle	0.203	0.368	0.608
Mean	0.190	0.343	0.568

PP vs. Grate

Math Elementary	0.127	0.243	0.437

Reading Comprehension Elementary	0.094	0.182	0.356
Math Middle	0.126	0.215	0.396
Reading Comprehension Middle	0.095	0.189	0.366
Mean	0.110	0.207	0.389

PP vs. AvGrate

Math Elementary	0.125	0.241	0.439
Reading Comprehension Elementary	0.076	0.143	0.284
Math Middle	0.128	0.239	0.436
Reading Comprehension Middle	0.085	0.161	0.295
Mean	0.104	0.196	0.364

MLM0 vs. Gain

Math Elementary	0.129	0.239	0.446
Reading Comprehension Elementary	0.101	0.188	0.336
Math Middle	0.113	0.219	0.398
Reading Comprehension Middle	0.093	0.191	0.340
Mean	0.109	0.209	0.380

MLM0 vs. TM

Math Elementary	0.130	0.245	0.449
Reading Comprehension Elementary	0.093	0.174	0.316
Math Middle	0.133	0.234	0.417
Reading Comprehension Middle	0.063	0.119	0.231
Mean	0.105	0.193	0.353

MLM0 vs. SGP

Math Elementary	0.166	0.299	0.520
Reading Comprehension Elementary	0.173	0.296	0.500
Math Middle	0.178	0.316	0.548
Reading Comprehension Middle	0.183	0.331	0.552
Mean	0.175	0.310	0.530

MLM0 vs. VAM

Math Elementary	0.182	0.315	0.549

Reading Comprehension Elementary	0.179	0.318	0.541
Math Middle	0.199	0.346	0.565
Reading Comprehension Middle	0.201	0.359	0.586
Mean	0.190	0.334	0.560

MLM0 vs. Grate

Math Elementary	0.127	0.234	0.432
Reading Comprehension Elementary	0.104	0.196	0.351
Math Middle	0.125	0.204	0.386
Reading Comprehension Middle	0.100	0.189	0.364
Mean	0.114	0.206	0.383

MLM0 vs. AvGrate

Math Elementary	0.128	0.238	0.434
Reading Comprehension Elementary	0.078	0.138	0.259
Math Middle	0.119	0.218	0.416
Reading Comprehension Middle	0.076	0.147	0.288
Mean	0.100	0.185	0.349

Gain vs. TM

Math Elementary	0.465	0.709	0.921
Reading Comprehension Elementary	0.377	0.592	0.834
Math Middle	0.347	0.535	0.733
Reading Comprehension Middle	0.268	0.444	0.663
Mean	0.364	0.570	0.788

Gain vs. SGP

Math Elementary	0.379	0.598	0.847
Reading Comprehension Elementary	0.226	0.388	0.625
Math Middle	0.318	0.510	0.764
Reading Comprehension Middle	0.219	0.379	0.609
Mean	0.286	0.469	0.711

Gain vs. VAM

Math Elementary	0.398	0.605	0.867

Reading Comprehension Elementary	0.228	0.387	0.635
Math Middle	0.321	0.507	0.744
Reading Comprehension Middle	0.204	0.351	0.607
Mean	0.288	0.462	0.713

Gain vs. Grate

Math Elementary	0.589	0.839	0.981
Reading Comprehension Elementary	0.446	0.692	0.921
Math Middle	0.595	0.851	0.988
Reading Comprehension Middle	0.410	0.675	0.913
Mean	0.510	0.764	0.951

Gain vs. AvGrate

Math Elementary	0.224	0.372	0.605
Reading Comprehension Elementary	0.191	0.332	0.553
Math Middle	0.226	0.401	0.635
Reading Comprehension Middle	0.178	0.315	0.532
Mean	0.205	0.355	0.581

TM vs. SGP

Math Elementary	0.321	0.519	0.785
Reading Comprehension Elementary	0.206	0.346	0.569
Math Middle	0.258	0.409	0.659
Reading Comprehension Middle	0.133	0.232	0.401
Mean	0.230	0.376	0.604

TM vs. VAM

Math Elementary	0.327	0.530	0.791
Reading Comprehension Elementary	0.190	0.335	0.561
Math Middle	0.260	0.421	0.650
Reading Comprehension Middle	0.125	0.222	0.392
Mean	0.226	0.377	0.598

TM vs. Grate

Math Elementary	0.396	0.627	0.877

Reading Comprehension Elementary	0.300	0.490	0.732
Math Middle	0.291	0.471	0.692
Reading Comprehension Middle	0.204	0.356	0.584
Mean	0.298	0.486	0.721

TM vs. AvGrate

Math Elementary	0.208	0.355	0.584
Reading Comprehension Elementary	0.183	0.321	0.550
Math Middle	0.176	0.309	0.540
Reading Comprehension Middle	0.203	0.347	0.571
Mean	0.192	0.333	0.561

SGP vs. VAM

Math Elementary	0.609	0.854	0.987
Reading Comprehension Elementary	0.419	0.672	0.898
Math Middle	0.668	0.878	0.989
Reading Comprehension Middle	0.424	0.657	0.893
Mean	0.530	0.765	0.942

SGP vs. Grate

Math Elementary	0.275	0.460	0.716
Reading Comprehension Elementary	0.185	0.328	0.549
Math Middle	0.265	0.422	0.671
Reading Comprehension Middle	0.165	0.308	0.529
Mean	0.222	0.380	0.616

SGP vs. AvGrate

Math Elementary	0.317	0.515	0.758
Reading Comprehension Elementary	0.150	0.270	0.472
Math Middle	0.301	0.477	0.730
Reading Comprehension Middle	0.164	0.283	0.481
Mean	0.233	0.386	0.610

VAM vs. Grate

Math Elementary	0.288	0.471	0.715
Reading Comprehension Elementary	0.194	0.329	0.561

Math Middle	0.258	0.423	0.663
Reading Comprehension Middle	0.182	0.321	0.560
Mean	0.230	0.386	0.625

Grate vs. AvGrate

Math Elementary	0.172	0.292	0.501
Reading Comprehension Elementary	0.139	0.252	0.429
Math Middle	0.179	0.308	0.533
Reading Comprehension Middle	0.132	0.236	0.433
Mean	0.156	0.272	0.474

Our last criterion for comparing school ranks across cohorts was the RMSD between pairs of school performance model rankings. Appendix E shows the RMSD between pairs of school performance model rankings for each individual cohort. Table 13 shows the RMSD averaged over the three cohorts by content area and grade level band. The RMSD values reflect the same patterns of results for models as described previously. The SGP vs. VAM, Gain vs. Grate, and PP vs. MLM0 pairings produced school rankings that were quite similar.

When a single year model (PP or MLM) was paired with a model that made use of multiyear results, the level of agreement in school ranks was much lower (difference of about 37 ranks on average across all model pairings of this type) than when the two single year models were paired (MLM and PP pairs differed by 11 ranks on average), or a multiyear model was paired with another multiyear model (difference of about 22 ranks, on average).

Table 13

Average Across Cohorts of RMSD in School Ranks Between School Performance Models by Content Area and Grade Level Band

Elementary School Mathematics

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	10.317	35.007	34.261	30.217	28.974	35.803	34.992
MLM0		34.638	34.473	29.573	28.223	35.398	34.910
Gain			11.419	14.483	13.666	7.707	25.939
TM				17.399	16.904	13.457	27.311
SGP					7.336	20.113	18.121
VAM						19.676	17.040
Grate							31.878

Elementary School Reading Comprehension

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	14.667	42.688	43.433	31.689	29.756	41.692	48.221
MLM0		44.072	44.839	31.692	29.518	42.475	50.335
Gain			15.364	24.795	23.920	11.170	28.608
TM				27.812	27.556	18.882	29.074
SGP					12.333	28.500	33.385
VAM						27.643	33.646
Grate							35.937

Middle School Mathematics

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	10.465	37.160	35.679	28.054	26.800	37.870	36.955
MLM0		38.133	36.310	28.508	27.143	38.701	38.052
Gain			21.546	17.498	17.679	7.378	24.034
TM				24.823	24.593	23.207	30.307
SGP					6.821	21.790	19.593
VAM						21.938	19.300
Grate							29.676

Middle School Reading Comprehension

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	7.274	41.472	49.986	27.609	26.122	40.646	46.909
MLM0		42.184	50.868	28.386	26.597	41.071	47.742
Gain			22.793	26.075	25.223	11.416	30.097
TM				37.581	38.282	25.424	27.719
SGP					12.685	29.350	32.910
VAM						28.327	32.247
Grate							36.362

We also evaluated the extent to which school ranks agreed from one content area to the other. Table 14 shows the Spearman's correlation of school ranks in mathematics with school ranks in reading comprehension by cohort and grade level band. The table also shows the mean correlation across cohorts at the two grade level bands. As can be seen in Table 14, on average correlations of school ranks across mathematics and reading comprehension in elementary schools ranged from about .33 to .92 for the different school performance models. For middle schools, the average correlations ranged from about .14 to .77. Correlations were higher for the two status models, and lower for the multiyear models at both grade level bands. Average correlations at the middle school level were lower than for the elementary level for all models.

Table 14

Spearman's Correlations of School Performance Model Estimates Across Mathematics and Reading Comprehension by Cohort

	Elementary Schools				Middle Schools			
Model	Cohort 1	Cohort 2	Cohort 3	Mean	Cohort 1	Cohort 2	Cohort 3	Mean
PP	0.997	0.997	0.777	0.924	0.781	0.762	0.763	0.769
MLM0	0.882	0.855	0.852	0.863	0.865	0.852	0.856	0.858
Gain	0.425	0.402	0.425	0.417	0.328	0.323	0.273	0.308
TM	0.331	0.329	0.342	0.334	0.214	0.175	0.018	0.136
SGP	0.453	0.459	0.432	0.448	0.316	0.290	0.354	0.320
VAM	0.533	0.505	0.524	0.521	0.420	0.352	0.431	0.401
Grate	0.395	0.388	0.404	0.396	0.280	0.306	0.235	0.274
AvGrate	0.428	0.481	0.470	0.460	0.489	0.379	0.477	0.448

Table 15 shows the proportion of schools that shared similar ranks in mathematics as in reading comprehension for each school performance model by school level and averaged over grade level band. Similar to results previously described, Table 15 shows greater agreement for the PP and MLM0 models than the other school performance models with about 82% or more of the schools having ranks within 20 places across grade level bands. In contrast, there was substantially less agreement across the two content areas for the remaining, multiyear models with only approximately 50% of schools agreeing within 20 ranks for most models in either grade level band.

Table 15

Proportion of Elementary or Middle Schools Within 5, 10, or 20 Ranks of Each Other in Mathematics versus Reading Comprehension for Each School Performance Model Averaged Over Cohorts

Model Comparison	r = 5	r = 10	r = 20
PP			
Elementary	0.751	0.833	0.917
Middle	0.310	0.509	0.765
Mean	0.530	0.671	0.841
MLM0			
Elementary	0.376	0.596	0.840
Middle	0.380	0.573	0.825
Mean	0.378	0.584	0.832
Gain			
Elementary	0.171	0.306	0.517

Middle	0.146	0.268	0.478
Mean	0.158	0.287	0.498

TM

Elementary	0.156	0.280	0.493
Middle	0.126	0.241	0.419
Mean	0.141	0.260	0.456

SGP

Elementary	0.176	0.318	0.529
Middle	0.167	0.287	0.477
Mean	0.172	0.302	0.503

VAM

Elementary	0.188	0.332	0.557
Middle	0.181	0.314	0.500
Mean	0.184	0.323	0.528

Grate

Elementary	0.171	0.301	0.506
Middle	0.147	0.264	0.440
Mean	0.159	0.282	0.473

AvGrate

Elementary	0.181	0.307	0.523
Middle	0.167	0.307	0.521
Mean	0.174	0.307	0.522

Calculation of the RMSD in school ranks for mathematics versus reading comprehension by cohort and grade level band and averaged over cohorts showed similar results (see Table 16). The difference in school ranks averaged over cohorts for the PP and MLM0 models ranged from about 8 to 19. Average differences in rank across the two content areas were substantially greater for the remaining models ranging from 28 to 37 depending on model and grade level band.

Table 16

RMSD in School Ranks for Mathematics and Reading Comprehension by Cohort and Grade Level Band and Overall Means

	Elementary Schools				Middle Schools			
Model	Cohort 1	Cohort 2	Cohort 3	Mean	Cohort 1	Cohort 2	Cohort 3	Mean
PP	2.105	2.269	19.058	7.811	18.868	19.667	19.643	19.393
MLM0	13.899	15.364	15.511	14.925	14.845	15.526	15.281	15.217
Gain	30.610	31.222	30.615	30.816	33.080	33.196	34.395	33.557

TM	33.040	33.079	32.759	32.959	35.757	36.653	39.973	37.461
SGP	29.884	29.700	30.438	30.007	33.385	33.985	32.424	33.265
VAM	27.607	28.410	27.876	27.964	30.710	32.480	30.422	31.204
Grate	31.415	31.587	31.174	31.392	34.230	33.607	35.283	34.373
AvGrate	30.546	29.108	29.412	29.689	28.849	31.789	29.173	29.937

Relation with school composition variables. We computed the correlation of school ranks based on each school performance model with school composition variables to determine whether estimates were related to the aggregated student characteristics in each school. Table 17 shows these correlations for mathematics and reading comprehension in the elementary school and middle school samples. Correlations of model estimates with school composition variables within each individual cohort are presented in Appendix F. The rightmost column of Table 17 shows the correlation of each school performance model averaged over all of the school composition variables. As can be seen, correlations of the status models, PP and MLM0, ranged from -.218 to -.303 depending on content and grade level band and were noticeably stronger in magnitude than the correlations of the other school performance models with school composition variables, which ranged from -.148 to +.218 depending on content and grade level band.

Table 17

Spearman's Correlations of School Ranks With School Composition Variables by Content and Grade Level Band

Elementary School Mathematics

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.602	-0.299	-0.102	0.004	-0.511	0.203	-0.218
MLM0	-0.685	-0.293	-0.071	-0.015	-0.527	0.272	-0.220
Gain	0.045	0.081	-0.036	-0.006	0.100	0.011	0.032
TM	0.022	0.073	-0.037	-0.001	0.083	0.013	0.026
SGP	-0.070	0.032	-0.041	0.004	0.023	0.036	-0.003
VAM	-0.101	0.015	-0.039	0.005	0.000	0.044	-0.013
Grate	0.009	0.064	-0.039	-0.013	0.059	0.038	0.020
AvGrate	0.123	0.100	0.000	0.028	0.174	-0.085	0.057

Elementary School Reading Comprehension

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.656	-0.355	-0.088	0.002	-0.546	0.216	-0.238
MLM0	-0.816	-0.433	-0.067	0.009	-0.624	0.262	-0.278

Gain	0.251	0.198	-0.023	-0.021	0.288	-0.049	0.107
TM	0.275	0.206	-0.017	-0.017	0.319	-0.073	0.116
SGP	-0.207	-0.049	-0.031	0.002	-0.056	0.069	-0.045
VAM	-0.270	-0.069	-0.047	0.007	-0.092	0.093	-0.063
Grate	0.131	0.118	-0.036	-0.032	0.182	0.005	0.061
AvGrate	0.576	0.364	0.060	0.017	0.513	-0.223	0.218

Middle School Mathematics

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.584	-0.286	-0.195	0.036	-0.525	0.060	-0.249
MLM0	-0.668	-0.270	-0.144	0.023	-0.534	0.157	-0.239
Gain	0.182	0.113	-0.020	0.011	0.127	-0.029	0.064
TM	0.038	0.046	0.004	-0.006	0.035	0.020	0.023
SGP	-0.087	-0.020	-0.083	0.044	-0.070	0.010	-0.034
VAM	-0.115	-0.029	-0.096	0.037	-0.083	0.010	-0.046
Grate	0.158	0.112	-0.011	0.003	0.102	-0.002	0.060
AvGrate	0.228	0.088	-0.050	0.044	0.176	-0.127	0.060

Middle School Reading Comprehension

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.780	-0.408	-0.112	0.031	-0.660	0.111	-0.303
MLM0	-0.813	-0.388	-0.111	0.030	-0.640	0.154	-0.295
Gain	0.094	0.052	0.014	-0.024	0.104	0.000	0.040
TM	0.526	0.248	0.023	-0.005	0.432	-0.129	0.182
SGP	-0.386	-0.177	-0.068	-0.002	-0.266	0.044	-0.143
VAM	-0.423	-0.192	-0.070	-0.007	-0.278	0.080	-0.148
Grate	0.036	0.015	-0.007	-0.027	0.058	0.045	0.020
AvGrate	0.398	0.200	0.038	-0.026	0.328	-0.154	0.131

Relation of school ranks with SWD school composition. We specifically examined the relations between the percentage of SWD students served by a school and the school ranks based on the school performance model. Table 18 shows these correlations for mathematics and reading comprehension in the elementary school and middle school samples averaged over cohorts. Correlations of model estimates with SWD school composition within each individual cohort are presented in Appendix G. As can be seen in the bottom row of Table 18, on average, correlations of the status models (PP and MLM0) with school SWD composition were

substantially stronger in magnitude than the correlations for the other school performance models. With the PP and MLM0 models, school ranks were higher with lower percentages of SWD students in the school and school ranks were lower as schools served larger proportions of SWD. Little relation was present between school ranks based on the other models and SWD school composition.

Table 18

Average School Rank as a Function of the Percentage of SWD in the School by Model, Content Area, and Grade Level Band

Content Area and Grade Level Band	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
Math Elementary	-0.102	-0.071	-0.036	-0.037	-0.041	-0.039	-0.039	0.000
Reading Comprehension Elementary	-0.195	-0.144	-0.020	0.004	-0.083	-0.096	-0.011	-0.050
Math Middle	-0.088	-0.067	-0.023	-0.017	-0.031	-0.047	-0.036	0.060
Reading Comprehension Middle	-0.112	-0.111	0.014	0.023	-0.068	-0.070	-0.007	0.038
Mean	-0.124	-0.098	-0.016	-0.007	-0.056	-0.063	-0.023	0.012

Summary of Section B. We evaluated the school ranks arising from eight alternative models for estimating school academic performance in mathematics and reading comprehension across three sequential cohorts of students. As with the school performance estimates described in Section A, substantial variability in school ranks was present across the three student cohorts regardless of content area or grade level band. When we compared school ranks arising from one model to school ranks from other models, we found three pairs of models produced similar results across the members of a pair. Those models were (a) SGP and VAM, (b) Gain and Grate, and (c) MLM intercept (MLM0) and PP model. In general, pairs of models that combined a status model with a model making use of multiple years of test data showed the most discrepant results.

Comparison of model estimates to school composition variables showed that the status models (PP and MLM0) had correlations stronger in magnitude than the remaining school performance models. Finally, we correlated school ranks arising from the eight performance models with the percentage of SWD in each school. As with the school performance model estimates, we found the status models were more strongly correlated with SWD school composition but there was little relation of the other model estimates with the percentage of SWD students in the school.

Conclusion

This report described the North Carolina results of a large study examining eight alternative methods of estimating school performance. The eight alternative methods were representative of types of models often used in state accountability models, although none were the actual model used in North Carolina at the time. We represented school performance in two

ways, the actual model estimates and school ranks based on model estimates. In addition to this North Carolina report, there are reports describing results for the three other states (AZ, OR, PA) included in the study. Our primary interest in these comparisons was estimating the impact of cohort and student composition (including the percent of SWD) on school performance estimates, as well examining the extent to which different estimates of school performance correlated with each other.

A number of general conclusions can be drawn from the results of the North Carolina analyses. First, model representations of school performance over successive cohorts of students were very unstable, irrespective of whether representations were based on school performance model estimates or on school ranks. There was somewhat greater stability for status models (PP, MLM0) than for the multiyear models. Nonetheless, even with the most stable model, MLM0, Spearman's correlations showed that less than 75% of the variance was common across cohorts, and over all the models, there was substantial instability over cohorts. These results were also reflected in the examination of differences in absolute or average (RMSD) differences in ranks over cohorts.

Our examination of the relations of the school performance models with each other produced similar results. Generally, there was agreement between the two status model estimates (PP and MLM0) that were based on a single year of data, but these two models did not agree with the remaining multiyear models. However, there was some substantial agreement of the multiyear models with each other with some variations. In general, the AvGrate model showed the least agreement with the other multiyear models.

We also examined the relation of school performance model estimates with variables describing the student composition of the schools. These results showed a pattern of results that differed between the status and the multiyear models. The two status models had substantially higher correlations with school composition variables than the multiyear models. This was also true in terms of the percentage of SWD students served by a school. The larger the percentage of SWD in the school, the lower the status model estimates of school performance.

Thus, the North Carolina results showed consistent patterns of instability of estimates of school performance over successive cohorts of students, different estimates of school performance depending on the model chosen, especially for status versus multiyear models, and stronger relations of status models with the student composition of the school than multiyear models. Taken together, these results suggest the need for substantial caution in the way that school performance models are used and interpreted. Cohort instability suggests that rolling averages or some other mechanism is needed to provide more dependable depictions of school performance that are more stable over time. The substantial disagreement among the school performance models suggests that the choice of model matters a great deal. This choice should be made very carefully. A single model estimate of school performance may not be trustworthy and may need to be augmented by the results from additional models or metrics of school performance.

References

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics,* 29(1), 37-65.

Betebenner, D. (2009). Norm-and criterion-referenced student growth. *Educational Measurement: Issues and Practice*, *28*(4), 42-51.

Betebenner, D. W., & Iwaarden, A. V. (2011). *SGP: An textupR package for the calculation and visualization of student growth percentiles* [Computer software manual]. (R package version 0.4-0.0 available at http://cran.r-project.org/web/packages/SGP/)

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*, 289‑328.

Castellano, K. E., & Ho, A. D. (2013). Contrasting OLS and quantile regression approaches to student "growth" percentiles. *Journal of Educational and Behavioral Statistics, 38*, 190–215.

Goldschmidt, P., Choi, K., & Beaudoin, J. P. (2012, February). *Growth model comparison study: Practical implications of alternative models for evaluating school performance.* Council of Chief State School Officers: Washington, DC.

North Carolina Department of Public Instruction. (2009). *North Carolina Reading Comprehension Tests: Technical report*. Raleigh, NC: Author. Downloaded from http://www.ncpublicschools.org/docs/accountability/testing/reports/ eogreadingtechman3.pdf

North Carolina Department of Public Instruction. (2008). *North Carolina Mathematics Test edition 3 technical report (Citable draft)*. Raleigh, NC: Author. Downloaded from http://www.dpi.state.nc.us/docs/accountability/reports/mathtechmanualdrafted2.pdf

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *MLM 7: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Tindal, G., Nese, J. F. T., and Stevens, J. J. (2017). Estimating School Effects with a State Testing Program Using Transition Matrices. *Educational Assessment, 22*, 189-204.

Appendix A

Correlations among School Performance Model Estimates for Each Individual Cohort by Content Area and Grade Level Band

Mathematics Elementary Schools

Cohort 1

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.922	0.269	0.299	0.477	0.522	0.224	0.287
MLM0			0.263	0.268	0.491	0.538	0.224	0.277
Gain				0.922	0.865	0.885	0.965	0.599
TM					0.812	0.824	0.888	0.556
SGP						0.962	0.749	0.809
VAM							0.767	0.837
Grate								0.384

Cohort 2

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.914	0.233	0.260	0.425	0.472	0.196	0.236
MLM0			0.237	0.237	0.455	0.501	0.205	0.241
Gain				0.932	0.876	0.889	0.962	0.584
TM					0.818	0.830	0.897	0.539
SGP						0.966	0.757	0.799
VAM							0.767	0.827
Grate								0.358

Cohort 3

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.908	0.235	0.257	0.421	0.468	0.205	0.247
MLM0			0.270	0.268	0.473	0.519	0.247	0.259

	TM	SGP	VAM	Grate	AvGrate
Gain	0.931	0.892	0.902	0.966	0.642
TM		0.835	0.844	0.901	0.597
SGP			0.969	0.787	0.821
VAM				0.796	0.849
Grate					0.438
AvGrate					

Mathematics Middle Schools

Cohort 1

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.898	0.155	0.318	0.535	0.568	0.112	0.155
MLM0			0.096	0.218	0.523	0.569	0.068	0.098
Gain				0.922	0.811	0.812	0.962	0.677
TM					0.804	0.810	0.876	0.626
SGP						0.966	0.702	0.776
VAM							0.708	0.792
Grate								0.485

Cohort 2

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.900	0.153	0.318	0.486	0.539	0.121	0.151
MLM0			0.103	0.250	0.477	0.533	0.087	0.086
Gain				0.929	0.828	0.835	0.966	0.678
TM					0.835	0.849	0.887	0.646
SGP						0.972	0.729	0.805
VAM							0.736	0.805
Grate								0.491

Cohort 3

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.894	0.227	0.062	0.548	0.572	0.185	0.218
MLM0			0.153	0.128	0.524	0.554	0.117	0.154
Gain				-0.049	0.848	0.853	0.971	0.713
TM					0.036	0.046	-0.072	0.041
SGP						0.974	0.760	0.809
VAM							0.760	0.834
Grate								0.548

Reading Comprehension Elementary Schools

Cohort 1

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.812	-0.136	-0.175	0.379	0.464	-0.109	-0.395
MLM0			-0.215	-0.241	0.417	0.509	-0.158	-0.560
Gain				0.876	0.619	0.629	0.910	0.508
TM					0.530	0.528	0.799	0.485
SGP						0.907	0.464	0.324
VAM							0.500	0.305
Grate								0.220

Cohort 2

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.789	-0.105	-0.148	0.343	0.412	-0.048	-0.406
MLM0			-0.200	-0.225	0.362	0.454	-0.091	-0.584
Gain				0.869	0.676	0.681	0.913	0.536
TM					0.567	0.562	0.796	0.489
SGP						0.907	0.561	0.362

VAM						0.595	0.351
Grate							0.255

Cohort 3

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.946	-0.128	-0.152	0.424	0.495	-0.050	-0.477
MLM0			-0.178	-0.229	0.398	0.487	-0.078	-0.539
Gain				0.865	0.657	0.669	0.903	0.490
TM					0.557	0.550	0.761	0.492
SGP						0.902	0.516	0.343
VAM							0.555	0.344
Grate								0.171

Reading Comprehension Middle Schools

Cohort 1

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.956	-0.059	-0.462	0.539	0.572	-0.012	-0.383
MLM0			-0.124	-0.539	0.509	0.558	-0.046	-0.444
Gain				0.726	0.582	0.606	0.878	0.478
TM					0.202	0.181	0.634	0.557
SGP						0.879	0.501	0.333
VAM							0.566	0.397
Grate								0.284

Cohort 2

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.963	-0.139	-0.582	0.500	0.567	-0.059	-0.468
MLM0			-0.156	-0.613	0.477	0.570	-0.050	-0.525
Gain				0.700	0.559	0.555	0.847	0.403
TM					0.106	0.054	0.554	0.531
SGP						0.862	0.394	0.272

						0.460	0.282
VAM						0.460	0.282
Grate							0.092

Cohort 3

Model	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP		0.961	-0.018	-0.513	0.568	0.604	0.012	-0.255
MLM0			-0.037	-0.558	0.545	0.613	-0.007	-0.290
Gain				0.701	0.622	0.634	0.902	0.456
TM					0.188	0.136	0.635	0.503
SGP						0.899	0.508	0.430
VAM							0.555	0.467
Grate								0.252

Appendix B

Correlations of School Performance Model Estimates with School Composition Variables for Each Individual Cohort by Content Area and Grade Level Band

Mathematics Elementary Schools

Cohort 1

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.595	-0.290	-0.125	-0.006	-0.568	0.211
MLM0	-0.670	-0.312	-0.094	-0.032	-0.557	0.274
Gain	0.069	0.080	-0.028	-0.016	0.083	-0.022
TM	0.033	0.077	-0.047	0.000	0.071	-0.016
SGP	-0.070	0.012	-0.019	-0.012	-0.027	0.015
VAM	-0.096	-0.002	-0.030	-0.014	-0.043	0.016
Grate	0.037	0.070	-0.037	-0.024	0.047	0.013
AvGrate	0.116	0.084	0.024	0.011	0.146	-0.112

Cohort 2

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.587	-0.283	-0.108	0.027	-0.543	0.216
MLM0	-0.687	-0.304	-0.084	0.010	-0.537	0.294
Gain	0.076	0.120	-0.045	-0.008	0.139	0.031
TM	0.051	0.099	-0.029	-0.008	0.110	0.033
SGP	-0.025	0.064	-0.032	-0.004	0.064	0.034
VAM	-0.061	0.049	-0.041	0.006	0.039	0.053
Grate	0.028	0.097	-0.057	-0.008	0.084	0.070
AvGrate	0.175	0.135	0.016	0.013	0.230	-0.099

Cohort 3

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.590	-0.263	-0.105	-0.016	-0.491	0.221
MLM0	-0.708	-0.294	-0.088	-0.028	-0.501	0.307
Gain	0.058	0.115	-0.057	0.023	0.149	-0.014
TM	0.045	0.104	-0.035	0.020	0.135	-0.012
SGP	-0.074	0.053	-0.066	0.024	0.063	0.048
VAM	-0.101	0.047	-0.056	0.022	0.044	0.048

Grate	0.014	0.088	-0.055	0.008	0.097	0.013
AvGrate	0.122	0.137	-0.011	0.042	0.215	-0.061

Mathematics Middle Schools

Cohort 1

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.567	-0.295	-0.152	0.076	-0.560	0.133
MLM0	-0.662	-0.283	-0.108	0.056	-0.541	0.202
Gain	0.246	0.159	0.043	-0.050	0.194	-0.053
TM	0.157	0.075	0.031	-0.075	0.067	-0.007
SGP	-0.069	0.043	0.004	0.059	-0.073	-0.006
VAM	-0.093	0.025	-0.025	0.019	-0.085	0.004
Grate	0.205	0.137	0.030	-0.071	0.164	-0.016
AvGrate	0.263	0.199	0.028	0.043	0.202	-0.165

Cohort 2

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.562	-0.310	-0.179	0.029	-0.538	0.111
MLM0	-0.661	-0.309	-0.145	0.041	-0.510	0.192
Gain	0.194	0.149	-0.025	0.081	0.166	-0.032
TM	0.059	0.078	-0.077	0.071	0.014	-0.010
SGP	-0.045	-0.001	-0.068	0.095	-0.013	0.004
VAM	-0.077	0.004	-0.079	0.083	-0.031	0.007
Grate	0.162	0.142	-0.036	0.087	0.139	0.004
AvGrate	0.272	0.149	0.001	0.052	0.251	-0.145

Cohort 3

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.606	-0.346	-0.240	0.066	-0.544	0.119
MLM0	-0.702	-0.336	-0.200	0.045	-0.531	0.241
Gain	0.169	0.085	-0.085	0.030	0.100	-0.057
TM	-0.072	-0.040	0.056	-0.021	0.040	0.055
SGP	-0.097	-0.050	-0.164	0.019	-0.092	0.014
VAM	-0.122	-0.056	-0.164	0.038	-0.096	0.018
Grate	0.154	0.079	-0.059	0.016	0.079	-0.029

AvGrate	0.203	0.098	-0.102	0.020	0.176	-0.133

Reading Comprehension Elementary Schools

Cohort 1

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.589	-0.286	-0.118	-0.011	-0.564	0.211
MLM0	-0.804	-0.429	-0.094	-0.022	-0.650	0.277
Gain	0.245	0.223	0.045	0.019	0.278	-0.066
TM	0.256	0.216	0.043	0.011	0.294	-0.064
SGP	-0.227	-0.037	0.000	0.012	-0.121	0.061
VAM	-0.299	-0.054	-0.012	0.004	-0.159	0.081
Grate	0.144	0.143	0.020	-0.013	0.179	-0.001
AvGrate	0.583	0.411	0.091	0.042	0.531	-0.267

Cohort 2

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.588	-0.282	-0.103	0.024	-0.544	0.216
MLM0	-0.819	-0.470	-0.077	0.060	-0.649	0.265
Gain	0.289	0.238	-0.040	-0.042	0.325	-0.052
TM	0.297	0.239	-0.033	-0.037	0.357	-0.063
SGP	-0.153	-0.002	-0.023	0.006	-0.024	0.054
VAM	-0.228	-0.034	-0.035	0.015	-0.078	0.073
Grate	0.152	0.150	-0.054	-0.024	0.198	0.022
AvGrate	0.611	0.443	0.084	-0.033	0.561	-0.246

Cohort 3

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.766	-0.479	-0.092	-0.021	-0.605	0.264
MLM0	-0.831	-0.475	-0.082	-0.016	-0.615	0.293
Gain	0.247	0.179	-0.039	-0.039	0.331	-0.045
TM	0.286	0.199	-0.023	-0.038	0.366	-0.051
SGP	-0.211	-0.078	-0.059	-0.008	-0.021	0.097
VAM	-0.291	-0.111	-0.063	-0.005	-0.059	0.138
Grate	0.121	0.097	-0.052	-0.049	0.212	0.017
AvGrate	0.559	0.380	0.054	0.044	0.533	-0.190

Reading Comprehension Middle Schools

Cohort 1

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.779	-0.410	-0.131	0.058	-0.699	0.159
MLM0	-0.805	-0.399	-0.115	0.051	-0.647	0.196
Gain	0.140	0.068	0.005	-0.047	0.106	-0.075
TM	0.513	0.261	0.041	0.021	0.399	-0.164
SGP	-0.397	-0.144	-0.071	0.070	-0.326	0.010
VAM	-0.377	-0.200	-0.065	0.001	-0.312	0.036
Grate	0.068	0.026	-0.002	0.000	0.053	0.011
AvGrate	0.437	0.205	0.047	-0.042	0.328	-0.229

Cohort 2

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.787	-0.454	-0.137	0.070	-0.702	0.137
MLM0	-0.809	-0.439	-0.145	0.049	-0.659	0.190
Gain	0.109	0.086	0.058	-0.022	0.154	0.065
TM	0.501	0.299	0.067	0.066	0.474	-0.122
SGP	-0.367	-0.175	-0.007	-0.006	-0.285	0.032
VAM	-0.437	-0.207	-0.061	-0.016	-0.312	0.135
Grate	0.030	0.025	-0.003	-0.023	0.096	0.129
AvGrate	0.466	0.290	0.089	-0.063	0.381	-0.179

Cohort 3

Model	EDS	EL	SWD	Female	Minority	School Size
PP	-0.786	-0.499	-0.169	0.027	-0.671	0.162
MLM0	-0.826	-0.478	-0.148	0.049	-0.634	0.207
Gain	0.050	0.026	0.101	-0.041	0.123	0.033
TM	0.507	0.263	0.119	-0.112	0.431	-0.093
SGP	-0.399	-0.252	-0.074	-0.009	-0.228	0.103
VAM	-0.469	-0.258	-0.049	0.003	-0.248	0.134
Grate	0.014	-0.018	0.057	-0.070	0.053	0.029
AvGrate	0.300	0.210	0.056	-0.006	0.367	-0.110

Appendix C

Correlations of School Performance Model Estimates with School Percentage SWD for Each Individual Cohort by Content Area and Grade Level Band

Mathematics Elementary Schools

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.125	-0.094	-0.028	-0.047	-0.019	-0.030	-0.037	0.024
2	-0.108	-0.084	-0.045	-0.029	-0.032	-0.041	-0.057	0.016
3	-0.105	-0.088	-0.057	-0.035	-0.066	-0.056	-0.055	-0.011

Mathematics Middle Schools

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.152	-0.108	0.043	0.031	0.004	-0.025	0.030	0.028
2	-0.179	-0.145	-0.025	-0.077	-0.068	-0.079	-0.036	0.001
3	-0.240	-0.200	-0.085	0.056	-0.164	-0.164	-0.059	-0.102

Reading Comprehension Elementary Schools

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.118	-0.094	0.045	0.043	0.000	-0.012	0.020	0.091
2	-0.103	-0.077	-0.040	-0.033	-0.023	-0.035	-0.054	0.084
3	-0.092	-0.082	-0.039	-0.023	-0.059	-0.063	-0.052	0.054

Reading Comprehension Middle Schools

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.131	-0.115	0.005	0.041	-0.071	-0.065	-0.002	0.047
2	-0.137	-0.145	0.058	0.067	-0.007	-0.061	-0.003	0.089
3	-0.169	-0.148	0.101	0.119	-0.074	-0.049	0.057	0.056

Appendix D

Proportion of Elementary or Middle Schools within 5, 10, or 20 Ranks of Each Other for Each Pair of School Performance Models in Mathematics and Reading Comprehension by Cohort

Model Comparison	Cohort 1			Cohort 2			Cohort 3		
	r = 5	r = 10	r = 20	r = 5	r = 10	r = 20	r = 5	r = 10	r = 20
PP vs. MLM0									
Math Elementary	0.517	0.776	0.946	0.495	0.743	0.945	0.469	0.731	0.928
Reading Comprehension Elementary	0.313	0.523	0.785	0.305	0.498	0.763	0.561	0.801	0.962
Math Middle	0.503	0.753	0.935	0.503	0.765	0.957	0.473	0.710	0.917
Reading Comprehension Middle	0.647	0.854	0.974	0.659	0.878	0.980	0.673	0.882	0.990
Mean	0.495	0.726	0.910	0.490	0.721	0.911	0.544	0.781	0.949
PP vs. Gain									
Math Elementary	0.137	0.246	0.452	0.145	0.263	0.446	0.124	0.240	0.443
Reading Comprehension Elementary	0.101	0.176	0.341	0.105	0.181	0.344	0.089	0.192	0.347
Math Middle	0.122	0.227	0.406	0.124	0.233	0.422	0.124	0.250	0.422
Reading Comprehension Middle	0.114	0.215	0.375	0.083	0.176	0.347	0.103	0.189	0.381
Mean	0.119	0.216	0.394	0.114	0.213	0.390	0.110	0.218	0.398
PP vs. TM									
Math Elementary	0.135	0.256	0.466	0.142	0.268	0.474	0.142	0.247	0.438
Reading Comprehension Elementary	0.088	0.163	0.334	0.094	0.164	0.320	0.103	0.180	0.321
Math Middle	0.150	0.252	0.462	0.154	0.249	0.467	0.126	0.193	0.359
Reading Comprehension Middle	0.061	0.120	0.260	0.059	0.105	0.243	0.077	0.120	0.262
Mean	0.108	0.198	0.380	0.112	0.196	0.376	0.112	0.185	0.345

PP vs. SGP

Math Elementary	0.179	0.290	0.531	0.169	0.294	0.517	0.162	0.294	0.500
Reading Comprehension Elementary	0.160	0.299	0.502	0.163	0.299	0.505	0.178	0.309	0.519
Math Middle	0.174	0.343	0.562	0.203	0.339	0.554	0.197	0.353	0.564
Reading Comprehension Middle	0.187	0.339	0.578	0.215	0.361	0.554	0.209	0.351	0.562
Mean	0.175	0.318	0.543	0.188	0.323	0.532	0.186	0.327	0.536

PP vs. VAM

Math Elementary	0.169	0.305	0.539	0.171	0.319	0.531	0.162	0.317	0.543
Reading Comprehension Elementary	0.179	0.314	0.536	0.169	0.314	0.526	0.194	0.343	0.546
Math Middle	0.197	0.363	0.602	0.207	0.367	0.598	0.229	0.369	0.574
Reading Comprehension Middle	0.187	0.361	0.596	0.221	0.371	0.604	0.201	0.373	0.625
Mean	0.183	0.336	0.568	0.192	0.343	0.565	0.196	0.351	0.572

PP vs. Grate

Math Elementary	0.126	0.241	0.439	0.132	0.243	0.445	0.124	0.246	0.426
Reading Comprehension Elementary	0.091	0.174	0.339	0.085	0.175	0.346	0.108	0.198	0.383
Math Middle	0.110	0.207	0.387	0.144	0.217	0.398	0.124	0.221	0.402
Reading Comprehension Middle	0.089	0.193	0.377	0.095	0.181	0.359	0.103	0.193	0.361
Mean	0.104	0.204	0.386	0.114	0.204	0.387	0.115	0.214	0.393

PP vs. AvGrate

Math Elementary	0.126	0.245	0.447	0.121	0.225	0.433	0.130	0.252	0.437
Reading Comprehension Elementary	0.082	0.143	0.297	0.076	0.153	0.288	0.070	0.133	0.267
Math Middle	0.116	0.225	0.418	0.136	0.239	0.426	0.132	0.252	0.464

Reading Comprehension Middle	0.065	0.138	0.268	0.087	0.166	0.290	0.103	0.179	0.327
Mean	0.097	0.188	0.358	0.105	0.196	0.359	0.109	0.204	0.374

MLM0 vs. Gain

Math Elementary	0.132	0.251	0.439	0.135	0.243	0.451	0.121	0.224	0.448
Reading Comprehension Elementary	0.101	0.191	0.332	0.103	0.188	0.337	0.100	0.186	0.339
Math Middle	0.103	0.227	0.412	0.128	0.209	0.404	0.108	0.221	0.379
Reading Comprehension Middle	0.110	0.215	0.361	0.085	0.176	0.314	0.083	0.183	0.345
Mean	0.112	0.221	0.386	0.113	0.204	0.376	0.103	0.204	0.378

MLM0 vs. TM

Math Elementary	0.129	0.241	0.441	0.140	0.252	0.466	0.121	0.242	0.440
Reading Comprehension Elementary	0.094	0.171	0.322	0.095	0.186	0.325	0.090	0.163	0.300
Math Middle	0.154	0.280	0.473	0.144	0.243	0.408	0.101	0.179	0.369
Reading Comprehension Middle	0.077	0.120	0.215	0.053	0.110	0.233	0.059	0.126	0.247
Mean	0.114	0.203	0.363	0.108	0.198	0.358	0.093	0.178	0.339

MLM0 vs. SGP

Math Elementary	0.174	0.303	0.522	0.159	0.294	0.510	0.164	0.300	0.529
Reading Comprehension Elementary	0.169	0.297	0.509	0.167	0.29	0.493	0.182	0.301	0.498
Math Middle	0.179	0.325	0.572	0.179	0.31	0.538	0.174	0.314	0.533
Reading Comprehension Middle	0.183	0.325	0.560	0.178	0.327	0.533	0.189	0.339	0.562
Mean	0.176	0.312	0.541	0.171	0.305	0.518	0.177	0.314	0.530

MLM0 vs. VAM

Math Elementary	0.183	0.313	0.547	0.173	0.311	0.540	0.191	0.323	0.559
Reading Comprehension Elementary	0.188	0.322	0.542	0.169	0.310	0.533	0.179	0.321	0.547
Math Middle	0.213	0.369	0.588	0.183	0.331	0.564	0.199	0.337	0.544
Reading Comprehension Middle	0.211	0.367	0.576	0.195	0.347	0.566	0.197	0.363	0.615
Mean	0.199	0.343	0.563	0.180	0.325	0.551	0.192	0.336	0.566

MLM0 vs. Grate

Math Elementary	0.128	0.243	0.429	0.128	0.231	0.440	0.124	0.228	0.427
Reading Comprehension Elementary	0.107	0.192	0.334	0.106	0.201	0.351	0.100	0.196	0.367
Math Middle	0.134	0.205	0.393	0.114	0.213	0.383	0.126	0.195	0.383
Reading Comprehension Middle	0.099	0.199	0.369	0.089	0.168	0.369	0.112	0.199	0.355
Mean	0.117	0.210	0.381	0.109	0.203	0.386	0.116	0.205	0.383

MLM0 vs. AvGrate

Math Elementary	0.136	0.236	0.433	0.120	0.231	0.429	0.129	0.247	0.442
Reading Comprehension Elementary	0.080	0.139	0.272	0.080	0.145	0.251	0.072	0.131	0.253
Math Middle	0.118	0.227	0.418	0.118	0.199	0.406	0.120	0.227	0.424
Reading Comprehension Middle	0.075	0.122	0.278	0.069	0.144	0.254	0.085	0.174	0.331
Mean	0.102	0.181	0.350	0.097	0.180	0.335	0.102	0.195	0.362

Gain vs. TM

Math Elementary	0.469	0.713	0.911	0.438	0.702	0.927	0.488	0.713	0.925
Reading Comprehension Elementary	0.377	0.605	0.850	0.378	0.587	0.832	0.376	0.585	0.821
Math Middle	0.467	0.698	0.907	0.477	0.720	0.931	0.097	0.187	0.361

Reading Comprehension									
Middle	0.276	0.477	0.684	0.256	0.420	0.645	0.270	0.434	0.661
Mean	0.397	0.623	0.838	0.387	0.607	0.834	0.308	0.48	0.692

Gain vs. SGP

Math Elementary	0.360	0.573	0.845	0.391	0.604	0.839	0.386	0.616	0.858
Reading Comprehension Elementary	0.215	0.36	0.596	0.242	0.419	0.643	0.220	0.384	0.637
Math Middle	0.294	0.479	0.751	0.341	0.515	0.775	0.318	0.535	0.765
Reading Comprehension Middle	0.225	0.406	0.619	0.203	0.363	0.592	0.229	0.367	0.617
Mean	0.273	0.454	0.703	0.294	0.475	0.712	0.288	0.476	0.719

Gain vs. VAM

Math Elementary	0.390	0.605	0.863	0.395	0.593	0.855	0.410	0.617	0.882
Reading Comprehension Elementary	0.214	0.369	0.600	0.246	0.405	0.650	0.225	0.388	0.653
Math Middle	0.298	0.483	0.738	0.323	0.517	0.748	0.343	0.521	0.748
Reading Comprehension Middle	0.223	0.361	0.631	0.178	0.337	0.576	0.211	0.355	0.613
Mean	0.281	0.454	0.708	0.286	0.463	0.707	0.297	0.470	0.724

Gain vs. Grate

Math Elementary	0.600	0.843	0.980	0.566	0.827	0.981	0.602	0.846	0.983
Reading Comprehension Elementary	0.468	0.708	0.924	0.452	0.711	0.933	0.419	0.658	0.908
Math Middle	0.602	0.824	0.990	0.588	0.850	0.986	0.596	0.880	0.986
Reading Comprehension Middle	0.444	0.696	0.945	0.337	0.606	0.864	0.450	0.724	0.931
Mean	0.528	0.768	0.960	0.486	0.748	0.941	0.517	0.777	0.952

Gain vs. AvGrate

Math Elementary	0.215	0.365	0.619	0.215	0.355	0.585	0.242	0.394	0.610
Reading Comprehension Elementary	0.202	0.326	0.541	0.202	0.364	0.577	0.170	0.308	0.542
Math Middle	0.252	0.422	0.631	0.215	0.385	0.629	0.211	0.396	0.645

Reading Comprehension Middle	0.178	0.320	0.544	0.172	0.316	0.507	0.185	0.310	0.544
Mean	0.212	0.358	0.584	0.201	0.355	0.574	0.202	0.352	0.585

TM vs. SGP

Math Elementary	0.317	0.516	0.783	0.307	0.506	0.775	0.339	0.533	0.796
Reading Comprehension Elementary	0.188	0.323	0.548	0.212	0.358	0.580	0.217	0.358	0.578
Math Middle	0.318	0.489	0.785	0.345	0.535	0.803	0.112	0.203	0.391
Reading Comprehension Middle	0.140	0.249	0.422	0.114	0.203	0.375	0.144	0.245	0.406
Mean	0.241	0.394	0.635	0.244	0.400	0.633	0.203	0.335	0.543

TM vs. VAM

Math Elementary	0.324	0.532	0.795	0.315	0.527	0.785	0.340	0.531	0.791
Reading Comprehension Elementary	0.190	0.327	0.540	0.192	0.340	0.578	0.189	0.339	0.564
Math Middle	0.335	0.499	0.769	0.323	0.550	0.801	0.122	0.213	0.379
Reading Comprehension Middle	0.120	0.243	0.414	0.128	0.197	0.361	0.126	0.227	0.400
Mean	0.242	0.400	0.630	0.240	0.404	0.631	0.194	0.328	0.534

TM vs. Grate

Math Elementary	0.386	0.625	0.876	0.393	0.629	0.875	0.411	0.627	0.879
Reading Comprehension Elementary	0.308	0.508	0.758	0.304	0.507	0.738	0.288	0.454	0.701
Math Middle	0.381	0.625	0.862	0.394	0.594	0.850	0.097	0.193	0.363
Reading Comprehension Middle	0.207	0.365	0.584	0.205	0.349	0.568	0.201	0.355	0.602
Mean	0.320	0.531	0.770	0.324	0.520	0.758	0.249	0.407	0.636

TM vs. AvGrate

Math Elementary	0.223	0.363	0.588	0.184	0.318	0.557	0.217	0.384	0.606

Reading Comprehension Elementary	0.184	0.333	0.550	0.189	0.312	0.556	0.177	0.318	0.544
Math Middle	0.229	0.389	0.633	0.219	0.371	0.635	0.081	0.168	0.353
Reading Comprehension Middle	0.193	0.349	0.584	0.221	0.337	0.566	0.195	0.355	0.562
Mean	0.207	0.358	0.589	0.203	0.334	0.578	0.168	0.306	0.516

SGP vs. VAM

Math Elementary	0.583	0.854	0.982	0.604	0.837	0.987	0.640	0.873	0.992
Reading Comprehension Elementary	0.410	0.670	0.898	0.424	0.669	0.904	0.423	0.678	0.892
Math Middle	0.625	0.858	0.982	0.704	0.892	0.988	0.675	0.886	0.996
Reading Comprehension Middle	0.410	0.655	0.882	0.418	0.649	0.890	0.444	0.667	0.907
Mean	0.507	0.759	0.936	0.538	0.762	0.942	0.546	0.776	0.947

SGP vs. Grate

Math Elementary	0.277	0.448	0.710	0.277	0.457	0.712	0.272	0.474	0.725
Reading Comprehension Elementary	0.162	0.304	0.515	0.208	0.355	0.583	0.184	0.325	0.548
Math Middle	0.262	0.416	0.655	0.266	0.424	0.649	0.266	0.426	0.708
Reading Comprehension Middle	0.162	0.316	0.542	0.158	0.306	0.519	0.176	0.304	0.527
Mean	0.216	0.371	0.606	0.227	0.386	0.616	0.224	0.382	0.627

SGP vs. AvGrate

Math Elementary	0.313	0.514	0.759	0.308	0.494	0.737	0.330	0.536	0.779
Reading Comprehension Elementary	0.156	0.277	0.474	0.141	0.256	0.477	0.153	0.277	0.466
Math Middle	0.282	0.456	0.708	0.306	0.497	0.753	0.316	0.479	0.730
Reading Comprehension Middle	0.166	0.292	0.467	0.152	0.256	0.471	0.174	0.302	0.503
Mean	0.229	0.385	0.602	0.227	0.376	0.609	0.243	0.398	0.620

VAM vs. Grate									
Math									
Elementary Reading Comprehension	0.284	0.469	0.723	0.274	0.464	0.701	0.306	0.480	0.722
Elementary Math Middle	0.182	0.305	0.523	0.217	0.369	0.594	0.184	0.314	0.566
	0.266	0.422	0.659	0.256	0.416	0.635	0.252	0.430	0.694
Reading Comprehension Middle	0.195	0.329	0.586	0.170	0.300	0.536	0.181	0.333	0.556
Mean	0.232	0.381	0.623	0.229	0.387	0.616	0.231	0.389	0.634
Grate vs. AvGrate									
Math									
Elementary Reading Comprehension	0.165	0.283	0.502	0.158	0.278	0.490	0.194	0.315	0.510
Elementary Math Middle	0.142	0.269	0.428	0.152	0.263	0.453	0.122	0.225	0.407
	0.185	0.325	0.542	0.179	0.296	0.509	0.172	0.302	0.546
Reading Comprehension Middle	0.144	0.241	0.471	0.103	0.209	0.389	0.150	0.258	0.440
Mean	0.159	0.279	0.486	0.148	0.262	0.460	0.160	0.275	0.476

Appendix E

RMSD in School Ranks for Pairs of School Performance Models for Each Individual Cohort by Content Area and Grade Level Band

Elementary School Mathematics: Cohort 1

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	9.811	34.738	33.702	29.268	28.226	35.644	34.335
MLM0		34.765	34.303	29.231	27.983	35.614	34.658
Gain			11.622	14.815	13.781	7.572	25.798
TM				17.444	16.918	13.486	27.304
SGP					7.623	20.293	18.098
VAM						19.625	16.999
Grate							31.625

Elementary School Mathematics: Cohort 2

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	10.393	35.037	34.228	30.486	29.104	35.789	35.271
MLM0		34.861	34.623	30.033	28.630	35.539	35.441
Gain			11.434	14.749	14.063	7.919	27.078
TM				17.800	17.223	13.596	28.395
SGP					7.464	20.546	18.912
VAM						20.291	17.746
Grate							33.098

Elementary School Mathematics: Cohort 3

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	10.746	35.245	34.854	30.898	29.592	35.976	35.370
MLM0		34.287	34.493	29.454	28.056	35.042	34.632
Gain			11.200	13.885	13.154	7.630	24.940
TM				16.954	16.571	13.291	26.236
SGP					6.920	19.500	17.354
VAM						19.112	16.376
Grate							30.911

Elementary School Reading Comprehension: Cohort 1

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	16.834	43.224	43.915	31.670	29.676	42.858	47.744
MLM0		44.503	44.974	31.096	28.826	43.501	50.239
Gain			14.814	25.973	25.125	11.221	28.647
TM				28.477	28.131	18.265	28.974
SGP					12.413	30.026	33.759
VAM						29.155	34.254
Grate							35.760

Elementary School Reading Comprehension: Cohort 2

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	18.442	42.166	43.054	32.702	30.702	41.099	47.758
MLM0		44.063	44.714	32.490	30.260	42.130	50.637
Gain			15.406	23.799	23.112	10.618	27.845
TM				27.195	27.042	18.229	29.122
SGP					12.185	27.069	32.792
VAM						26.300	33.141
Grate							34.936

Elementary School Reading Comprehension: Cohort 3

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	8.724	42.674	43.329	30.695	28.889	41.118	49.161
MLM0		43.649	44.830	31.491	29.470	41.794	50.129
Gain			15.874	24.614	23.523	11.672	29.332
TM				27.763	27.494	20.152	29.125
SGP					12.400	28.407	33.602
VAM						27.473	33.544
Grate							37.116

Middle School Mathematics: Cohort 1

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	10.302	37.637	34.425	28.068	26.961	38.263	37.474
MLM0		38.589	36.045	28.557	27.325	39.058	38.621
Gain			11.425	18.041	18.097	7.470	24.211
TM				18.262	18.104	13.618	25.730
SGP					7.249	22.225	20.435
VAM						22.243	19.867
Grate							29.887

Middle School Mathematics: Cohort 2

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	10.002	37.412	33.666	28.467	26.823	38.169	37.024
MLM0		38.189	34.743	29.096	27.318	38.688	38.371
Gain			11.459	17.364	17.702	7.633	24.219
TM				16.470	16.100	14.059	24.952
SGP					6.565	21.925	18.922
VAM						22.055	19.376
Grate							29.976

Middle School Mathematics: Cohort 3

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	11.090	36.432	38.946	27.626	26.615	37.177	36.368
MLM0		37.620	38.142	27.871	26.787	38.358	37.163
Gain			41.754	17.088	17.238	7.032	23.671
TM				39.738	39.573	41.944	40.238
SGP					6.649	21.219	19.422
VAM						21.515	18.656
Grate							29.167

Middle School Reading Comprehension: Cohort 1

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	7.440	41.060	49.210	27.709	26.584	40.279	47.274
MLM0		42.069	50.306	28.488	27.281	40.948	48.213
Gain			22.142	25.851	24.554	10.435	29.034
TM				36.895	36.877	24.483	26.988
SGP					13.219	28.717	33.110
VAM						27.332	31.784
Grate							34.898

Middle School Reading Comprehension: Cohort 2

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	7.421	42.822	51.085	28.486	26.560	41.476	48.373
MLM0		43.370	51.784	29.141	26.899	41.591	49.262
Gain			23.011	26.899	26.662	13.378	31.370
TM				38.554	39.704	26.847	27.507
SGP					12.713	30.542	34.470
VAM						30.061	34.433
Grate							38.801

Middle School Reading Comprehension: Cohort 3

Model	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
PP	6.960	40.535	49.665	26.631	25.223	40.182	45.080
MLM0		41.114	50.515	27.531	25.611	40.673	45.751
Gain			23.228	25.476	24.453	10.434	29.888
TM				37.295	38.263	24.942	28.663
SGP					12.123	28.792	31.150
VAM						27.588	30.522
Grate							35.387

Appendix F

***Correlations of School Ranks with School Composition Variables by Content Area and Grade
Level Band for Each Individual Cohort***

Elementary School Mathematics: Cohort 1

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.614	-0.302	-0.093	-0.004	-0.554	0.192	-0.229
MLM0	-0.675	-0.292	-0.060	-0.029	-0.558	0.260	-0.226
Gain	0.056	0.046	-0.029	-0.020	0.067	0.001	0.020
TM	0.015	0.043	-0.036	-0.009	0.045	0.010	0.011
SGP	-0.074	-0.013	-0.023	-0.006	-0.035	0.029	-0.020
VAM	-0.100	-0.029	-0.026	-0.006	-0.055	0.030	-0.031
Grate	0.023	0.039	-0.037	-0.029	0.036	0.030	0.010
AvGrate	0.115	0.054	0.026	0.024	0.122	-0.103	0.040

Elementary School Mathematics: Cohort 2

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.588	-0.304	-0.101	0.019	-0.517	0.211	-0.213
MLM0	-0.680	-0.304	-0.071	0.004	-0.537	0.279	-0.218
Gain	0.052	0.105	-0.047	-0.004	0.111	0.042	0.043
TM	0.023	0.095	-0.045	-0.011	0.085	0.041	0.031
SGP	-0.038	0.058	-0.050	0.004	0.046	0.029	0.008
VAM	-0.077	0.043	-0.053	0.010	0.018	0.053	-0.001
Grate	0.007	0.082	-0.047	-0.008	0.064	0.076	0.029
AvGrate	0.161	0.130	-0.016	0.020	0.205	-0.098	0.067

Elementary School Mathematics: Cohort 3

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.604	-0.290	-0.111	-0.002	-0.461	0.206	-0.210
MLM0	-0.701	-0.282	-0.082	-0.018	-0.487	0.278	-0.215
Gain	0.027	0.093	-0.031	0.005	0.122	-0.009	0.034
TM	0.027	0.080	-0.029	0.018	0.118	-0.013	0.034
SGP	-0.098	0.051	-0.049	0.013	0.058	0.049	0.004
VAM	-0.127	0.032	-0.039	0.011	0.036	0.050	-0.006

Grate	-0.002	0.071	-0.032	-0.003	0.076	0.007	0.020
AvGrate	0.092	0.116	-0.009	0.039	0.195	-0.055	0.063

Elementary School Reading Comprehension: Cohort 1

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.608	-0.299	-0.092	-0.008	-0.551	0.191	-0.228
MLM0	-0.803	-0.393	-0.065	-0.030	-0.632	0.260	-0.277
Gain	0.244	0.187	0.010	-0.012	0.250	-0.067	0.102
TM	0.254	0.195	0.011	-0.029	0.261	-0.091	0.100
SGP	-0.240	-0.057	-0.003	-0.009	-0.124	0.068	-0.061
VAM	-0.295	-0.067	-0.029	-0.002	-0.155	0.088	-0.077
Grate	0.135	0.114	0.004	-0.041	0.164	0.003	0.063
AvGrate	0.579	0.346	0.065	0.041	0.494	-0.256	0.212

Elementary School Reading Comprehension: Cohort 2

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.589	-0.303	-0.088	0.016	-0.520	0.211	-0.212
MLM0	-0.815	-0.451	-0.061	0.059	-0.640	0.254	-0.276
Gain	0.273	0.232	-0.030	-0.015	0.302	-0.052	0.118
TM	0.286	0.230	-0.029	-0.002	0.343	-0.067	0.127
SGP	-0.158	-0.010	-0.033	0.019	-0.026	0.045	-0.027
VAM	-0.225	-0.050	-0.044	0.026	-0.078	0.057	-0.052
Grate	0.145	0.158	-0.046	-0.011	0.193	0.006	0.074
AvGrate	0.594	0.378	0.065	-0.030	0.530	-0.226	0.218

Elementary School Reading Comprehension: Cohort 3

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.769	-0.463	-0.086	-0.002	-0.568	0.247	-0.274
MLM0	-0.831	-0.456	-0.075	-0.002	-0.599	0.272	-0.282
Gain	0.236	0.176	-0.050	-0.035	0.312	-0.029	0.102
TM	0.286	0.194	-0.032	-0.021	0.352	-0.062	0.120
SGP	-0.222	-0.081	-0.057	-0.002	-0.019	0.093	-0.048
VAM	-0.289	-0.089	-0.069	-0.004	-0.042	0.133	-0.060
Grate	0.112	0.082	-0.067	-0.045	0.189	0.006	0.046
AvGrate	0.555	0.368	0.051	0.039	0.513	-0.187	0.223

Middle School Mathematics: Cohort 1

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.591	-0.279	-0.187	0.035	-0.552	0.078	-0.249
MLM0	-0.673	-0.258	-0.123	0.035	-0.545	0.159	-0.234
Gain	0.206	0.130	0.031	-0.067	0.154	-0.040	0.069
TM	0.125	0.060	-0.003	-0.077	0.045	-0.020	0.022
SGP	-0.094	0.001	-0.018	0.012	-0.088	-0.012	-0.033
VAM	-0.109	-0.014	-0.047	0.000	-0.094	-0.017	-0.047
Grate	0.178	0.128	0.038	-0.086	0.131	-0.006	0.064
AvGrate	0.243	0.109	-0.033	0.049	0.159	-0.153	0.062

Middle School Mathematics: Cohort 2

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.564	-0.290	-0.161	0.013	-0.505	0.049	-0.243
MLM0	-0.649	-0.267	-0.117	0.007	-0.516	0.144	-0.233
Gain	0.185	0.119	-0.044	0.091	0.147	-0.016	0.080
TM	0.061	0.061	-0.076	0.077	0.010	0.009	0.024
SGP	-0.050	-0.017	-0.085	0.111	-0.019	0.016	-0.007
VAM	-0.092	-0.021	-0.092	0.101	-0.042	0.010	-0.023
Grate	0.156	0.125	-0.038	0.092	0.120	0.013	0.078
AvGrate	0.260	0.080	-0.035	0.077	0.225	-0.130	0.080

Middle School Mathematics: Cohort 3

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.598	-0.290	-0.237	0.059	-0.520	0.052	-0.256
MLM0	-0.681	-0.285	-0.191	0.026	-0.541	0.169	-0.250
Gain	0.156	0.090	-0.046	0.008	0.078	-0.031	0.042
TM	-0.070	0.017	0.091	-0.017	0.052	0.073	0.024
SGP	-0.117	-0.043	-0.145	0.010	-0.103	0.026	-0.062
VAM	-0.145	-0.050	-0.147	0.010	-0.114	0.035	-0.068
Grate	0.139	0.084	-0.033	0.003	0.056	-0.014	0.039
AvGrate	0.182	0.076	-0.083	0.006	0.142	-0.099	0.037

Middle School Reading Comprehension: Cohort 1

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.780	-0.370	-0.106	0.031	-0.672	0.118	-0.296
MLM0	-0.814	-0.358	-0.105	0.026	-0.642	0.158	-0.289
Gain	0.116	0.031	-0.003	0.014	0.058	-0.072	0.024
TM	0.521	0.211	-0.006	0.020	0.375	-0.167	0.159
SGP	-0.395	-0.162	-0.075	0.043	-0.319	0.001	-0.151
VAM	-0.379	-0.199	-0.074	0.012	-0.304	0.010	-0.156
Grate	0.054	0.008	-0.018	0.020	0.018	-0.007	0.012
AvGrate	0.437	0.158	0.012	-0.023	0.297	-0.206	0.112

Middle School Reading Comprehension: Cohort 2

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.788	-0.420	-0.109	0.046	-0.676	0.102	-0.308
MLM0	-0.808	-0.388	-0.108	0.033	-0.653	0.149	-0.296
Gain	0.108	0.106	0.011	-0.031	0.146	0.056	0.066
TM	0.539	0.266	0.018	0.029	0.483	-0.141	0.199
SGP	-0.381	-0.166	-0.039	-0.024	-0.263	0.034	-0.140
VAM	-0.445	-0.164	-0.064	-0.015	-0.295	0.123	-0.143
Grate	0.027	0.060	-0.017	-0.026	0.104	0.118	0.044
AvGrate	0.450	0.236	0.053	-0.060	0.356	-0.153	0.147

Middle School Reading Comprehension: Cohort 3

Model	EDS	EL	SWD	Female	Ethnic Minority	School Size	Mean
PP	-0.771	-0.434	-0.120	0.015	-0.632	0.112	-0.305
MLM0	-0.816	-0.417	-0.120	0.032	-0.623	0.156	-0.298
Gain	0.057	0.019	0.035	-0.055	0.108	0.017	0.030
TM	0.518	0.266	0.056	-0.064	0.438	-0.078	0.189
SGP	-0.382	-0.202	-0.089	-0.026	-0.217	0.095	-0.137
VAM	-0.446	-0.214	-0.072	-0.017	-0.235	0.106	-0.146
Grate	0.027	-0.024	0.014	-0.076	0.051	0.025	0.003
AvGrate	0.305	0.205	0.049	0.006	0.332	-0.104	0.132

Appendix G

***Correlations of School Ranks with School Percentage SWD for Each Individual Cohort by
Content Area and Grade Level Band***

Elementary School Mathematics

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.093	-0.060	-0.029	-0.036	-0.023	-0.026	-0.037	0.026
2	-0.101	-0.071	-0.047	-0.045	-0.050	-0.053	-0.047	-0.016
3	-0.111	-0.082	-0.031	-0.029	-0.049	-0.039	-0.032	-0.009

Elementary School Reading Comprehension

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.092	-0.065	0.010	0.011	-0.003	-0.029	0.004	0.065
2	-0.088	-0.061	-0.030	-0.029	-0.033	-0.044	-0.046	0.065
3	-0.086	-0.075	-0.050	-0.032	-0.057	-0.069	-0.067	0.051

Middle School Mathematics

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.187	-0.123	0.031	-0.003	-0.018	-0.047	0.038	-0.033
2	-0.161	-0.117	-0.044	-0.076	-0.085	-0.092	-0.038	-0.035
3	-0.237	-0.191	-0.046	0.091	-0.145	-0.147	-0.033	-0.083

Middle School Reading Comprehension

Cohort	PP	MLM0	Gain	TM	SGP	VAM	Grate	AvGrate
1	-0.106	-0.105	-0.003	-0.006	-0.075	-0.074	-0.018	0.012
2	-0.109	-0.108	0.011	0.018	-0.039	-0.064	-0.017	0.053
3	-0.120	-0.120	0.035	0.056	-0.089	-0.072	0.014	0.049