10-23-2008

Measuring Middle School Achievement Growth With Student Growth Percentile Methodology

Rachel Slaughter
Public Consulting Group's Center for Resource Management (PCG-CRM), rslaughter@pcgus.com

Measuring Middle School Achievement Growth With Student Growth Percentile Methodology

Rationale

A serious challenge for many schools and districts across the nation is that significant numbers of students enter the next grade level with performance levels well below proficiency. Traditional status and improvement indices used in accountability systems established by states to meet the requirements of the No Child Left Behind Act (NCLB) typically do not reflect achievement growth among those students (Choi, Seltzer, Herman, & Yamashiro, 2007; Goldschmidt, Roschewski, Choi, Auty, Hebbler, Blank, & Williams, 2005) even when some growth has been made.

The goal of improving programs and instruction so that all students are challenged to achieve high learning standards is certainly a critical one for education. If accountability indices fail to reflect the progress of students at the lowest performance levels, however, not only will that progress not be recognized but indices will fail to identify effective programs and instructional interventions for those students. Furthermore, educators in schools who are successfully promoting growth among the lowest achieving students may lose their sense of efficacy—a factor that has been empirically shown to be critical to improving student learning in schools with high populations of at-risk students (Northeast and Islands Regional Educational Laboratory, 2000; The Carnegie Corporation, 2002).

Thus, failure to identify achievement progress among the lowest achieving students and the schools serving those students can lead to loss of resources and opportunities to identify and capitalize on effective practices. Supplementing traditional status and accountability indicators of school performance with indicators of student achievement *growth* keeps learning standards high while providing important information about the achievement of students across the entire achievement spectrum.

The purpose of the study reported here was to analyze the achievement growth of a same-student cohort of middle school students in a large school district over three years. The study focused on

the growth of the majority (68 %) of students in an 8th grade same-student cohort in a large district who scored below the proficiency level on the statewide test in 6th grade, with almost one-third of them (32 %) having achievement at the lowest performance level. There was additional interest in whether the achievement growth of 6th grade students at the lowest performance levels varied substantially across different schools. The study also investigated the relationship of the growth metric to proficiency metrics.

Selection of the Methodology

There are many different types of growth models that are currently being implemented around the country. As of June 2008, the U. S. Department of Education had approved proposals from a total of eleven (11) states for incorporating some kind of growth metric into state-level AYP plans under NCLB. Growth models are also being used operationally or on a pilot basis at the district-level. Growth models are being rapidly and widely adopted because of their great promise for demonstrating gains throughout the achievement spectrum over time and thus may function as better measures than traditional status and improvement indices for holding teachers and schools accountable for student learning (Betebenner, 2008; Ho, 2008).

Gain or growth metrics have a long-standing tradition in the educational research literature but their widespread application within current accountability systems is a recent phenomenon (CCSSO, January 2008; Willett, 1994). This factor, along with the complexity of educational accountability systems; psychometric issues related to characteristics of test scores; and statistical issues related to methodologies for deriving growth metrics from test scores; has lead to the development of many different approaches to measuring student achievement growth (Goldsmith, et al., 2005).

Growth models share many common features but diverge widely with respect to other features. Commonalities include the analysis of the achievement of same-student cohort groups over two or more years; the requirement that test scores allow meaningful comparisons across at least adjacent grades; the implementation of statistical procedures for calculating growth metrics at the individual student and/or school level; and, typically, an explicit link with traditional status or improvement indices (e.g., Are students “on track” to achieve proficiency within a specified

number of years?). Key differences include whether performance levels or scaled scores are used to calculate the growth metric; the extent to which the models require test scores derived from vertically-scaled tests; the statistical sophistication (lack of transparency) of the methodology; and whether the progress of students is judged against growth targets set by policy and/or in relation to observed growth ("normative" growth). While student background characteristics and other variables related to the school context may be and often are factors in research studies about achievement growth, such variables cannot be included in growth models under the NCLB accountability system wherein all students are held to the same proficiency targets.

Table 1 provides an overview of the types of growth models currently in use for accountability pilots under NCLB as well as in other state-level and district-level accountability systems. A brief description of each follows as background for the selection of a growth model for this study.

Table 1. Overview of Growth Model

Growth Model	Description	Requirements	Outputs
Value Table	Measures growth according to the weighted value stakeholders place on movement and maintenance within and across performance levels.	1. Performance levels have a consistent meaning from year to year. 2. Table representing stakeholder agreement on the value of year-to-year achievement outcomes.	1. School/district growth scores. 2. Sublevel growth scores that permit analysis of growth at lowest performance levels.
Categorical	Measures growth according to whether students below mastery move into higher performance sublevel categories. Does not reward growth in students who decline in achievement and regain previously attained levels or move to higher levels below proficiency.	Performance levels have a consistent meaning from year to year.	Results in a "Proficiency Index" that reflects both the students who are at proficiency and who are progressing toward proficiency in terms of movement into higher proficiency levels (and no backward shifts).
Growth Trajectory	Measures progress against growth targets that are based on an individual student's unique trajectory determined by baseline score, gap between actual score and proficiency, and prescribed time within which the student must reach proficiency.	Test scores must be vertically scaled for results to be interpreted meaningfully.	The identification of which students are "on target" for reaching proficiency each year in addition to those students that have reached proficiency targets.
Projection (Value-Added)	Predicts an individual student's future test scores based on prior test scores.	1. Development of a statistical model that "fits" the test data. 2. Complex statistical procedures (sometimes proprietary) to estimate likely future student achievement.	Estimates of whether students are likely to reach proficiency within a specific timeframe.
Student Growth Percentile (SGP)	Describes a student's growth by examining current achievement relative to academic peers—those students with identical prior achievement.	1. Quantile regression procedures used to derive estimates. 2. Does not require vertically-equated tests for meaningful interpretation of outcomes.	1. Individual SGP scores. 2. SGP scores can be aggregated/disaggregated to examine growth for individual students and across schools, student subgroups, programs, etc. 3. Provides empirical ("normative") basis for achievement growth to inform decisions about adequate ("criterion-referenced") growth.

Growth Models Using Performance Level Categories

A value table is the central component of the value table growth model approach (Hill, 2006; *Delaware's Proposal for a Growth Model*, Submitted to U. S. Department of Education, February 17, 2006.) The value table represents how states or districts want to see individual students within schools progress across performance levels from year to year, (i.e., the "values" placed on different achievement outcomes over time). The points assigned within the value table are then multiplied by the number of students in a school who demonstrate each achievement outcome across two years. A school's growth score, based on the value table, will be higher if most of the students within the school demonstrate the achievement outcomes that are most highly valued by the state or district.

The categorical growth model approach is somewhat similar to the value table approach in that movement into performance level categories provides the growth metric (*No Child Left Behind Growth Model Pilot Proposal*, Submitted by the Iowa Department of Education to U. S. Department of Education, April 30, 2007). In this approach, non-proficient performance levels are further sub-divided to provide greater sensitivity to growth among low-achieving students). It clearly communicates state goals for student achievement growth in terms of performance levels. It does not use a value table and is based on achievement outcomes for students *below proficiency only*.

Value table and categorical growth model approaches have many advantages, including being easy for educators to understand; generating growth scores through simple calculations; and using performance levels which have meaning for educators and clearly communicate state goals for growth over time. Drawbacks include the requirement that performance levels have a consistent meaning across at least adjacent grade levels; the formidable task of creating value tables or performance level subcategories that appropriately reflect stakeholder values for growth outcomes; reliance on policy decisions about expected growth that may not be realistic; the lack of an obvious, intuitive relationship between the growth metric and a focus on school-level metrics rather than individual student metrics.

Growth Trajectory Models Using Scaled Scores

Several states are piloting growth approaches whereby trajectories are determined for individual students using baseline test scores; time to proficiency; and proficiency cut scores so that interim progress toward meeting proficiency targets can be rewarded (*North Carolina's Proposal to Pilot the Use of a Growth Model for AYP Purposes in 2005-2006*, 4/16/06; *Proposal for a Growth Model to Evaluate Adequate Yearly Progress for Schools and Districts*, Arizona

Department of Education, July 2, 2007; *Florida's Application for the NCLB Growth Model,* September 15, 2006).

In these approaches, students below proficiency in any year may be classified as "proficient" if their performance indicates that they are on a growth trajectory that is like to result in proficiency within the specified time span. Some states scale tests so that it is possible to determine linear growth trajectories. Other states take into account non-linear growth trajectories across different grade spans and content areas.

These approaches, like those that use movement across performance levels, permit schools to be rewarded for students who are making progress toward reaching proficiency as well as for students who are proficient. In comparison with the performance level approaches, however, these approaches permit more precise measurement of growth if assumptions about the score scale are met. Drawbacks include the fact that many statewide assessment systems do not have vertically scaled tests; and the lack of an empirical basis for determining whether growth targets set by policy are reasonable.

Projection Growth Models Using Predicted Scores

Several states and districts are piloting or using projection or projection to proficiency models (*NCLB Growth Model Pilot Program: Proposal to the U. S. Department of Education,* Tennessee Department of Education, February 16, 2006; *Proposal to the US Department of Education for Participation in the No Child Left Behind (NCLB) Growth Model Pilot Program,* The Pennsylvania Department of Education, October 31, 2006). These models employ a "mixed-model, longitudinal methodology" to predict future test performance for a student using all prior test score information. For example, to determine a 6^{th} grade student's projected score for an 8^{th} grade reading test, all of the student's prior test data is analyzed including the scores of students who have the same historical pattern of test scores (to adjust for missing data). Thus if the student has 3^{rd}, 5^{th}, and 6^{th} grade scores but is missing a score for 4^{th} grade, the methodology estimates regression coefficients for the missing score based on the subset of other 6^{th} grade students who also have scores at grades 3, 5, and 6 only. The regression estimates are then used to calculate the student's projected score on the 8^{th} grade reading test. The results are used to

indicate whether the progress the student has made from grades 3-6 indicates that the student will reach proficiency by 8th grade.

The projection approach permits more precise growth information to be generated, if the assumptions of the statistical methodology are met, including taking into account missing data. There are many challenges in implementing these approaches, however, including the need for multiple years of test data for each student (more than two); the use of complex statistical procedures which may be proprietary and which teachers and administrators are unlikely to understand; costly computer software and hardware requirements; and the need for high-level statistical expertise to perform the analyses (Hibpshman, 2004; *Next Generation of Value-Added Models and Indicators*, 2008).

Student Growth Percentile Model Using Conditional Percentile Ranks

The student growth percentile model is a newly emerging approach that has been adopted by Colorado and is being considered at the state-wide level by other states (*Colorado's Academic Growth Model*, 2008; Betebenner, 2008). This approach focuses on estimating the observed growth of a student in relation to students with the same prior academic achievement in order to establish a normative baseline for growth in order to better inform decisions about adequate growth. This is in contrast to most other growth models where growth targets are set by policy, not in relation to empirical information about typical growth.

In this approach, student growth percentile (SGP) scores describe a students' growth by locating the student's current score within the distribution of students who had identical prior achievement. For example, if a student's SGP score is determined to be 70 %, only 30 % of the students who had the same prior achievement had the same or higher achievement—that student's growth was substantially above average in relation to the student's academic peers. If, on the other hand, a student's SGP score is determined to be 20 %, 80 % of the students who had the same prior achievement had higher achievement—that student's growth was substantially below average in relation to the student's academic peers.

Quantile regression is used to determine the relationship between the score distributions over time and to derive the estimated SGP scores. Applying quantile regression permits the conditional density associated with the student's score at time *t* to be estimated using the student's prior scores at times 1,2,…,*t*-1 as the conditioning variable (Betebenner, 2008). Given the conditional density for the student's score at time *t*, the student's growth percentile is defined as the percentile of the score within the time *t* conditional density.

Quantile regression is a natural extension of least-squares regression. In quantile regression, one or more conditional quantile functions are estimated for the response variable—a year 2 test score distribution, for instance, compared with the year 1 test score distribution—instead of the conditional mean function (Koenker & Hallock, 2001). The power of quantile regression for distributions of achievement scores is that several conditional quantiles can be estimated which takes into account differences in the bivariate distribution of scores across years at different points of the achievement continuum. For instance, there is generally more score dispersion at the lowest and highest achievement levels. *Quantile* is a general term for dividing a distribution of scores into parts. The median, for instance, represents the score in the distribution that divides the observations exactly in half. The 1st quartile is the point that divides the distribution such that 25 % of the observations are below and 75 % are above. The 3rd quartile divides the distribution at the 75th percentile. Distributions can also be divided into quintiles—10 %, 20 %, 30 %, 90 %. 90 %. Quantile regression can also be done with different levels of precision throughout the distribution--5 %, 10 %, 15 %, 20 %, 30 % , 50 %, 60 %, 65 %, 70 %,… etc.—permitting a great deal of flexibility in the analysis according to how specific sets of achievement scores may be distributed.

Quantile regression has traditionally been used by economists to describe relationships among variables that are not expected to be normally distributed and are like to have different levels of dispersion (variance) at different points in the score distribution. For example, income levels and amount of income spent on food. While income level may be highly predictive of percentage of income spent on food at lower income levels, at upper income levels there is likely to be more variability and, thus, simply predicting average food expenditure based on income level would misrepresent the relationship. The same properties of achievement score distributions make

quantile regression analysis appropriate for examining the relationship of test scores over time. Achievement test scores generally are not normally distributed and scores at both the lower and higher ends of the achievement continuum tend to be more variable than scores toward the middle.

Advantages of the SGP approach using quantile regression analysis procedures to derive the SGP scores include being able to provide a solid estimate of observed growth (a growth *norm*) that provides better information for decisions about "adequate" growth (criterion-referenced decision) and whether policies about interim growth trajectories toward proficiency represent reasonable expectations for growth. The SGP approach also yields growth metrics that can be used at the individual student level; provides metrics that are familiar to educators (percentiles) and employs statistical methodology that, while it is complex, is suitable for the score distributions used in education in that it is robust to outliers, uncorrelated with prior achievement, and does not require vertically-scaled test scores.

While quantile regression methodology has not been widely used in the educational field to date, information about how to apply it is becoming more widely available (Hao & Naiman, 2007). There is also software available in the public domain for analysis—the R language—which includes extensive documentation (Koenker, 2006). R language can be installed within Windows or Mac applications and also as an "add-on" program in SPSS. Statistical software for conducting quantile regression analysis is also available within other commercial statistical software packages such as STATA and SAS.

Research Questions

This study was undertaken to help the district examine achievement growth in the district's middle schools. The district has undertaken multiple initiatives to improve reading and mathematics achievement for middle school students over the last several years and is interested in knowing what impact those interventions are having. Status and improvement (trend) indicators—even the results from following the movement of same-student cohorts across performance levels—while important to the larger picture and while they do reflect some improvement, are not sufficient. The overwhelming majority of 6th grade students in the district

enter the middle schools with below proficient performance levels. A critical issue for the district is whether the educational experiences these students have in the district's middle schools lead to growth in reading and mathematics achievement.

The study reported for this paper addressed the following research questions:

1. What progress did low-achieving students make in reading in the district's middle schools from 6th grade to 8th grade?
2. Were there differences among middle schools in the growth of students in reading achievement from grade 6 to grade 8?
3. What was the relationship between the percent of students at reading proficiency across the middle schools in grade 8 and measures of growth for those schools?

Sample

Recent reading test scores from a criterion-referenced, standards-based statewide assessment for a cohort of 8th grade students were analyzed in order to address the research questions. Students in the longitudinal, same-student cohort had complete sets of reading test scores for consecutive years from grade 6 to grade 8. There were a total of 1256 students. Table 2 shows the distribution of students in the cohort by program assignment and school.

Table 2. Distribution of Student Cohort By Program Assignment and School (Middle Schools)

	1		2		3		Total	
Program Assignment	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Regular Education (RegEd)	227	76%	390	76%	366	82%	983	78%
Special Education (SpEd)	70	24%	95	18%	72	16%	237	19%
English Language Learner (ELL)	0	0%	23	4%	6	1%	29	2%
ELL & SpEd	0	0%	7	1%	0	0%	7	1%
Total	297	100%	515	100%	444	100%	1256	100%

Methods and Procedures

The student growth percentile growth model was chosen to address the research questions for this study for several reasons. First of all, it permits a more fine-grained analysis of student growth at the lowest achievement levels than is possible with other approaches. Second, it provides the district with a normative base line for student growth at the middle school level to better inform discussions about defining adequate growth; to ensure that reasonable growth targets are set; and to indicate the level of intervention that may be needed to help students meet proficiency targets. Finally, the methodology is appropriate for the test score distributions: it does not rely on tests being vertically-scaled; it takes into account the non-normal distribution of achievement test scores; and it enables growth to be measured across the achievement spectrum.

A student growth percentile was calculated for each student using quantile regression whereby the conditional density associated with the student's score at time *t* is estimated using the student's prior scores at times 1,2,…,*t*-1 as the conditioning variable (Betebenner, 2008). The regression analysis was done at the quintile level, that is, the probability of a student's score in the second year given the conditional distribution of students with the same academic achievement in the first year was estimated at the 10^{th}, 20^{th}, 30^{th}, 40^{th}, …90^{th} percentile rank in the bivariate distribution. Figure 1 shows the quantile regression plot for the analyses for the grade 6 – grade 7 reading scores.

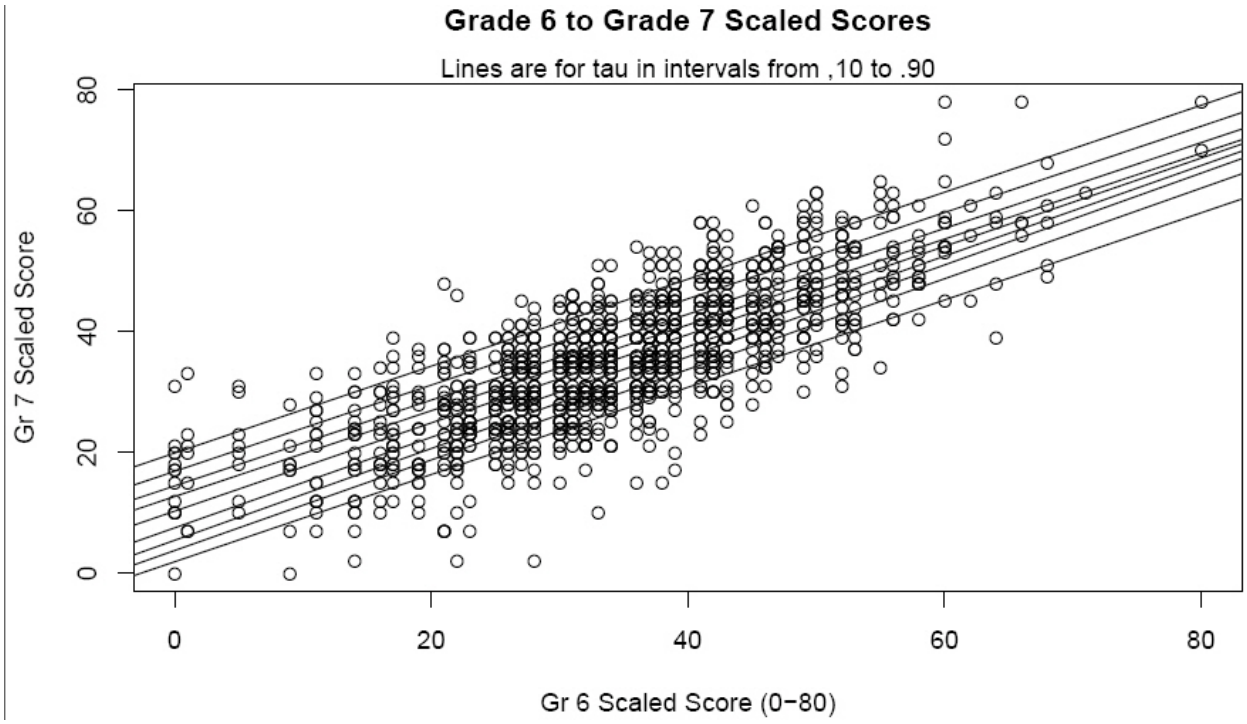


Figure 1. Quantile Regression Plot for Grade 6 To Grade 7 Scores.

A student growth percentile (SGP) score was assigned to each student based on achievement over two time intervals: 6^{th} – 7^{th} grade and 7^{th} – 8^{th} grade. The SGP scores were then aggregated for comparisons among schools and among subgroups within the cohort according to grade 6 achievement levels.

Steps in the quantile regression analysis and the assignment of student growth percentile scores are described in Appendix A.

Results

Research Question #1

What progress did low-achieving students make in reading in the district's middle schools from 6th grade to 8th grade?

After the grade 7 and grade 8 SGP scores were assigned to each student in the cohort, the distributions were examined for the total group and for subgroups formed according to 6th grade performance levels in order to examine the growth low-achieving students made from grade 6 to grade 8. The number and percent of students at each of several student growth percentile ranges for both grade 7 and grade 8 and by grade 6 performance levels are shown in Tables 3 and 4, respectively. Level 1 is the lowest grade 6 performance level; Level 2 is the next highest but were still below proficient; Level 3 and Level 4 represent the proficient and advanced levels.

Examining these tables, the most important result of this study in terms of the primary research question is that, in grade 7, well over one-third (38 % - 44 %) of students at the lowest 6th grade proficiency levels demonstrated above average or substantially above average growth in reading in relation to their 6th grade academic peers. In fact, the growth demonstrated by students at the lowest 6th grade achievement levels in grade 7 was not dissimilar to the growth profile demonstrated by the group of students scoring at or above proficiency. Forty-two percent of students with Level 3 or 4 proficiency levels in grade 6 demonstrated above average or substantially above average growth in grade 7 in relation to students who had the same level of achievement in grade 6.

Table 3. Grade7 Reading Achievement Growth by Grade 6 Proficiency Levels

		Level 1		Level 2		Levels 3 / 4		Total	
Growth	**SGP**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**
Substantially Above Average	80 - 90	20.2%	80	18.2%	83	24.9%	100	20.9%	263
Above Average	60 - 70	18.1%	72	25.6%	117	16.9%	68	20.5%	257
Average	40 - 50	22.7%	90	23.4%	107	21.1%	85	22.5%	282
Below Average	20 - 30	20.4%	81	19.3%	88	18.2%	73	19.3%	242
Substantially Below Average	0 - 10	18.6%	74	13.6%	62	18.9%	76	16.9%	212
Total		100.0%	397	100.0%	457	100.0%	402	100.0%	1,256

From grade 7 to grade 8, SGP scores declined somewhat for students in the Level 1 and Level 2 proficiency level groups with 27 % of the Level 1 group demonstrating above average growth or substantially above average growth in grade 8 compared to 38 % in grade 7. Thirty-nine percent of the Level 2 group demonstrated above average or substantially above average growth in grade 8 compared to 44 % in grade 7. On the other hand, the students in the Level 3 / 4 group demonstrated increased growth with 51 % demonstrating above average to substantially above average growth compared to 42 % in grade 7.

Table 4. Grade 8 Growth by Grade 6 Proficiency Levels

		Level 1		Level 2		Levels 3 / 4		Total	
Growth	**SGP**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**
Substantially Above Average	80 - 90	10.8%	43	18.8%	86	31.1%	125	20.2%	254
Above Average	60 - 70	16.4%	65	20.4%	93	19.9%	80	18.9%	238
Average	40 - 50	21.7%	86	19.9%	91	20.9%	84	20.8%	261
Below Average	20 - 30	23.2%	92	21.0%	96	14.9%	60	19.7%	248
Substantially Below Average	0 - 10	28.0%	111	19.9%	91	13.2%	53	20.3%	255
Total		100.0%	397	100.0%	457	100.0%	402	100.0%	1,256

The *Wilcoxon signed-rank test* was used to test for the differences across the two years as the non-parametric equivalent of the dependent samples *t-test* (Hayslett, 1968). As shown in Tables 5-7, the results of the tests indicated that:

- the trend of lower growth from grade 7 to grade 8 was statistically significant only for students at the lowest 6th grade performance level;
- the upward trend in growth from grade 7 to grade 8 for students with achievement levels; and
- the differences were not statistically significant for students at performance Level 2 in grade 6.

Table 5. Results of the Wilcoxon Signed Ranks Tests for Students at Level 1 in Grade 6

		N	Mean Rank	Sum of Ranks	Test Statistics	
Gr 7-8 SGP - Gr 6-7 SGP a. Gr 7-8 < Gr 6-7 b. Gr 7-8 > Gr 6-7 c. Gr 7-8 = Gr 6-7	Negative Ranks	219a	191.92	42031.00	Z	-3.862a
	Positive Ranks	150b	174.89	26234.00	Asymp. Sig. (2-tailed)	.000
	Ties	28c				a. Based on positive ranks.
	Total	397				b. Wilcoxon Signed Ranks Test

Table 6. Results of the Wilcoxon Signed Ranks Tests for Students at Level 2 in Grade 6

		N	Mean Rank	Sum of Ranks	Test Statistics	
Gr 7-8 SGP - Gr 6-7 SGP a. Gr 7-8 < Gr 6-7 b. Gr 7-8 > Gr 6-7 c. Gr 7-8 = Gr 6-7	Negative Ranks	227a	210.97	47891.00	Z	-1.575a
	Positive Ranks	192b	208.85	40099.00	Asymp. Sig. (2-tailed)	.115
	Ties	38c				a. Based on positive ranks.
	Total	457				b. Wilcoxon Signed Ranks Test

Table 7. Results of the Wilcoxon Signed Ranks Tests for Students at Level 3 / 4 in Grade 6

		N	Mean Rank	Sum of Ranks	Test Statistics	
Gr 7-8 SGP - Gr 6-7 SGP a. Gr 7-8 < Gr 6-7 b. Gr 7-8 > Gr 6-7 c. Gr 7-8 = Gr 6-7	Negative Ranks	155a	187.93	29128.50	Z	-2.525a
	Positive Ranks	215b	183.75	39506.50	Asymp. Sig. (2-tailed)	.012
	Ties	32c				a. Based on positive ranks.
	Total	402				b. Wilcoxon Signed Ranks Test

In spite of some declines from grade 7 to grade 8, over one-quarter of students who had the lowest level of achievement in grade 6 demonstrated growth that was substantially above average in relation to students who had the same level of achievement in grade 6. Well over one-third of students in the next highest (but still below proficient) 6^{th} grade performance level demonstrated substantially above average achievement. It is also important to note that considerable growth from grade 6 to grade 8 occurred across the achievement spectrum, with

51 % of students with 6th grade performance levels at or above proficient demonstrating substantially above average growth.

Research Question #2

Were there differences among middle schools in the growth of students from grade 6 to grade 8?

The first step in addressing this research question was to examine the distribution of SGP scores across the schools. Table 8 shows the summary statistics for the grade 8 SGP score distribution across schools. Schools 2 and 3 had median SGP scores of 50 % and 3rd quartile SGP scores of 75 % and 80 %, respectively, in comparison with School 1 which had a median SGP score of 30 % and a 3rd quartile SGP score of 60 %. Overall, there were fairly substantial differences among schools in the growth of students from grade 6 to grade 8 especially for School 1 in comparison with Schools 2 and 3.

Table 8. Grade 8 SGP Summary Statistics by School

SGP Distribution	School 1	School 2	School 3	Total
Min	0	0	0	0
1st Quartile	10	25	30	20
Median	30	50	50	40
3rd Quartile	60	75	80	70
Max	90	90	90	90
Total n	297	515	444	1256

Because of the focus on the growth of low-achieving 6th graders, the 8th grade SGP distribution was examined for differences among growth scores across schools for students with different grade 6 proficiency levels. Table 9 provides the frequencies for SGP scores in the above average range (60 % to 90 %) by grade 6 performance level and school. As can be seen, the percentage of students with grade 6 performance levels with grade 8 SGP scores above average was similar across schools, ranging from 27 % to 33 %. The percentage of students with grade 6 performance at Level 2 who had above average grade 8 SGP scores, however, was substantially higher for Schools 2 and 3. School 3 also had a substantially higher percentage of students in the above average growth category who had 6th grade performance at Level 3 than did the other schools.

The differences in the 8th grade SGP score distributions by 6th grade performance levels are also shown graphically by the box plots in Figures 2–4.

Table 9. Percent of Students with Above Average Grade 8 SGP Scores by Grade 6 Performance Levels

Grade 6 Performance Level	School		
	1	2	3
1	27%	33%	27%
2	25%	41%	49%
3	32%	42%	63%

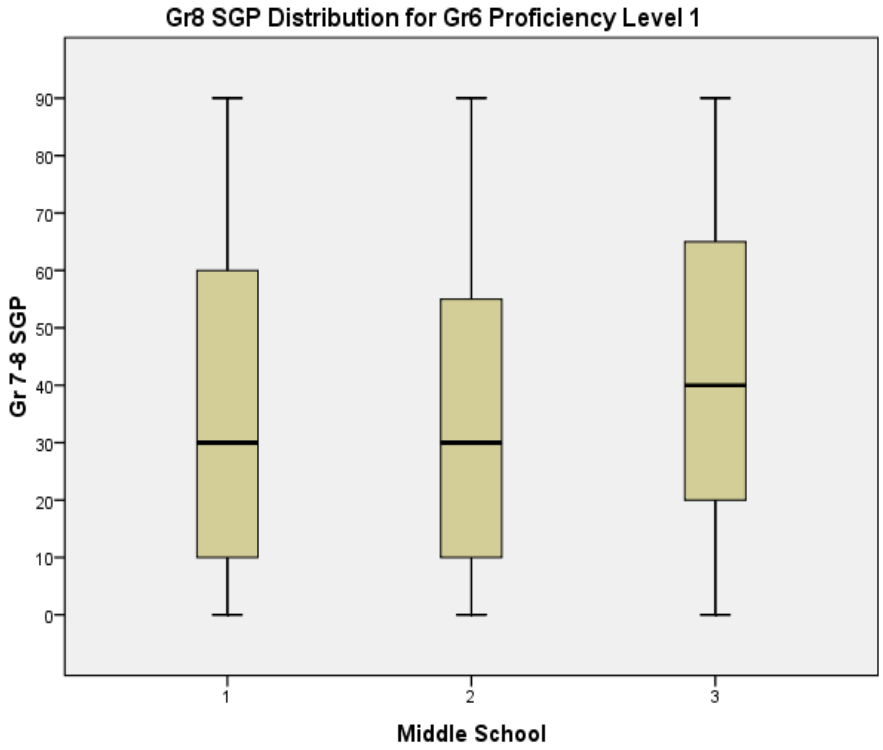


Figure 2. Grade 8 SGP Score Distribution by School and Grade 6 Performance Level 1.

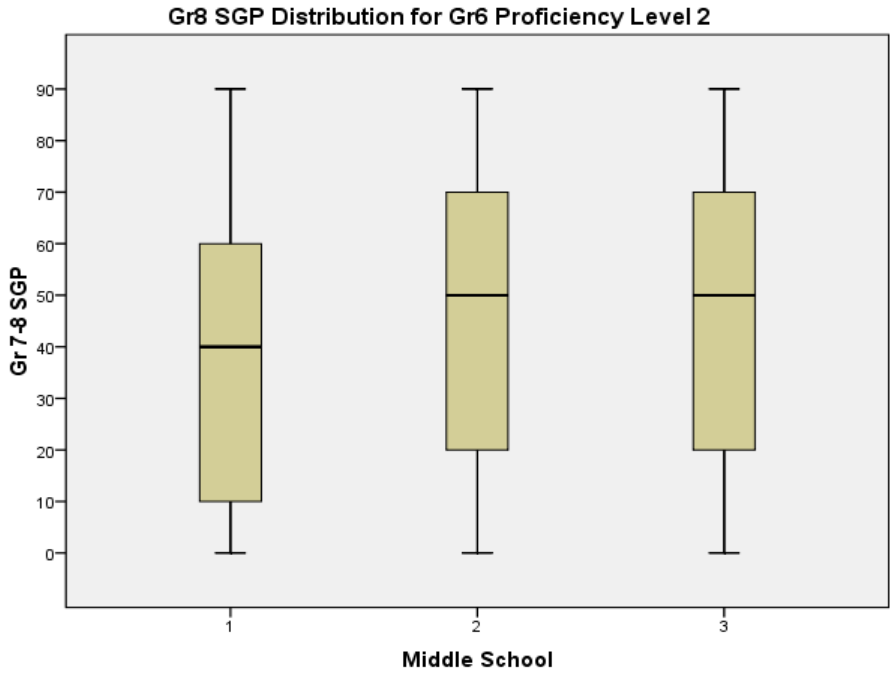


Figure 3. Grade 8 SGP Score Distribution by School and Grade 6 Performance Level 2.

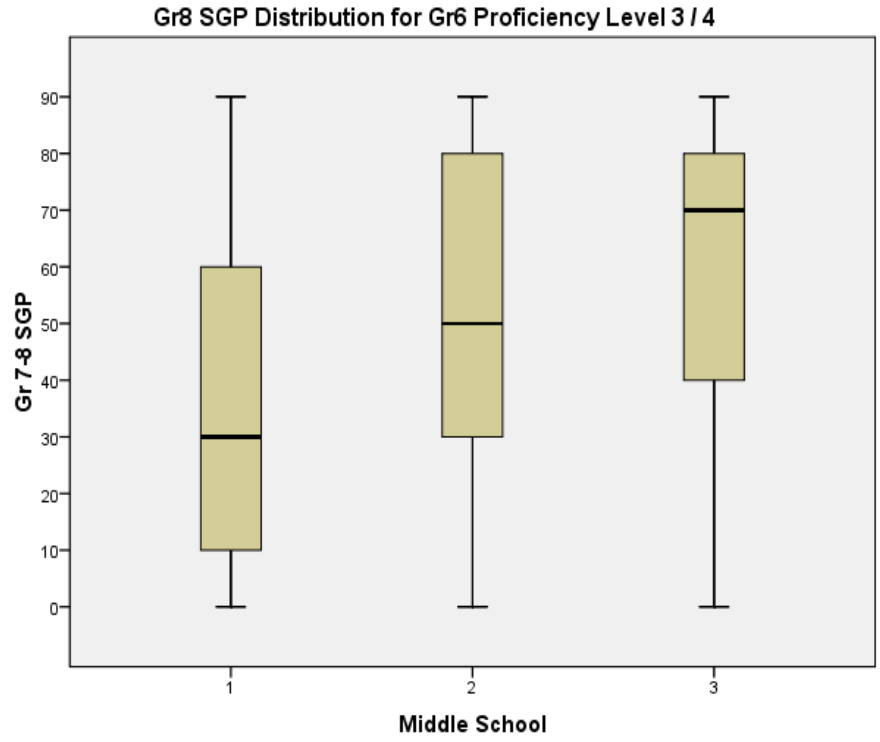


Figure 4. Grade 8 SGP Score Distribution by School and Grade 6 Performance Level 3 / 4.

Research Question #3

What was the relationship between the percent of students at proficiency across the middle schools in grade 8 and measures of growth for those schools?

As shown in Table 10, one-half of the students in School 3 were at or above proficiency in grade 8. Only a little more than one-third (35 % and 37 %) of students in School 1 and School 2 were at or above proficiency. Table 11 shows the number and percent of students at each performance level in grade 6 by school and overall. Comparing Tables 10 and 11, the changes in the percent of students proficient across years, while an important part of the big picture, may be disappointing to the districts in terms of the impact of intervention efforts reflecting only relatively modest increases in the percent of students proficient or above across years, even for longitudinal cohorts. In this case, there were only 8 % more students proficient in grade 8 compared to grade 6 overall; 12 % more for School 2; 8 % more for School 1; and 7 % more for School 3.

Table 10. Number and Percent of Students at each Grade 8 Performance Level by School

Grade 8	School							
Performance	1		2		3		Total	
Levels	n	%	n	%	n	%	n	%
1	89	30%	160	31%	97	22%	346	28%
2	105	35%	165	32%	124	28%	394	31%
3 / 4	103	35%	190	37%	223	50%	516	41%
Total	297	100%	515	100%	444	100%	1256	100%

Table 11. Number and Percent of Students at each Grade 6 Performance Level by School

Grade 6	School							
Performance	1		2		3		Total	
Levels	n	%	n	%	n	%	n	%
1	101	34%	184	36%	112	25%	397	32%
2	115	39%	202	39%	140	32%	457	36%
3 / 4	81	27%	129	25%	192	43%	402	32%
Total	297	100%	515	100%	444	100%	1256	100%

The SGP scores for 8th grade students, however, provide a much more encouraging picture of the growth students are making in these schools across the achievement continuum and suggest that intervention efforts may be having a positive impact although much more still needs to be done to help more students reach proficiency by the end of 8th grade. Table 12 shows the percent of students at each 8th grade SGP score range by school and 8th grade performance level. As can be seen there, almost half (48 %) of the students in School 3 who were still below proficient in grade 8 demonstrated above average growth and a substantial percent of students in Schools 1 and 2 who had achievement below proficient also demonstrated above average growth (30 % - 38 %) in relation to students.

Table 12. Percent of Students at 8th Grade SGP Score Ranges by School and Performance Level

School	Grade 8 Performance Level	0-10	20-30	40-50	60-70	80-90	Total
1	1	60%	26%	4%	7%	3%	100%
	2	24%	27%	30%	14%	6%	100%
	3	14%	12%	20%	28%	26%	100%
2	1	43%	25%	21%	8%	3%	100%
	2	19%	29%	23%	19%	9%	100%
	3	3%	11%	19%	28%	39%	100%
3	1	34%	30%	21%	11%	4%	100%
	2	15%	23%	29%	23%	10%	100%
	3	2%	8%	18%	22%	49%	100%

Table 13 shows the number of students at each 8th grade SGP score range by school and performance level. A total of 149 (12 %) students who had performance levels below proficient in grade 8 demonstrated above average growth in relation to their academic peers.

Table 13. Number of Students at 8th Grade SGP Score Ranges by School and Performance Level

School	Grade 8 Performance Level	0-10	20-30	40-50	60-70	80-90	Total
1	1	53	23	4	6	3	89
	2	25	28	31	15	6	105
	3	14	12	21	29	27	103
2	1	69	40	34	13	4	160
	2	32	48	38	32	15	165
	3	5	21	36	54	74	190
3	1	33	29	20	11	4	97
	2	19	29	36	28	12	124
	3	5	18	41	50	109	223
	Total	255	248	261	238	254	1256

Table 14 shows the number and percent of students at each 8th grade SGP score range for those students who had below proficient performance levels in grade 6 and achieved performance levels of proficient or above in grade 8. There were a total of 184 students in this category--15% of the cohort. Most of the students who had performance levels below proficient in grade 6 and who had performance levels at or above proficient in grade 8 demonstrated growth that was better than at least 60% of their academic peers. It is apparent that students who had performance below proficient in grade 6 had to make extraordinary achievement growth to reach proficiency by grade 8.

Table 14. SGP Scores for Below Proficient 6th Grade Students Achieving Proficiency in Grade 8

Grade 8 SGP	School							
	1		2		3		Total	
Score	%	n	%	n	%	%	%	n
0 -10	2%	1	0%	0	0%	0	1%	1
20-30	7%	3	5%	4	2%	1	4%	8
40-50	13%	6	16%	14	19%	10	16%	30
60-70	36%	16	35%	30	24%	13	32%	59
80-90	42%	19	44%	37	56%	30	47%	86
Total	100%	45	100%	85	100%	54	100%	184

As shown in Table 15, there were also students who scored at or above proficient in grade 6 who scored below proficient in grade 8. There were a total of 70 students in this category—6 % of the cohort. Most of these students demonstrated achievement growth well below average with 80 % or more of their 6th grade academic peers showing better growth. Clearly, attention must be paid to the continued growth of students who start middle school at or above proficient levels of achievement as well as to the growth of low-achieving students toward proficiency.

Table 15. SGP Scores for Proficient 6th Grade Students Declining to Below Proficient in Grade 8

	School							
Grade 8 SGP	1		2		3		Total	
Score	%	n	%	n	%	n	%	n
0 -10	61%	14	38%	9	30%	7	43%	30
20-30	26%	6	29%	7	17%	4	24%	17
40-50	9%	2	25%	6	35%	8	23%	16
60-70	4%	1	8%	2	13%	3	9%	6
80-90	0%	0	0%	0	4%	1	1%	1
Total	100%	23	100%	24	100%	23	100%	70

Summary and Conclusions

This study represents an application of an emerging methodology for examining student achievement growth patterns normatively (“what is”) in order to provide a more solid foundation for policies about expected or adequate growth (“what should be”) and, beyond that, what excellent growth would look like (“what could be”). As the results of this study show, the comparison of student growth percentile (SGP) scores across students grouped by initial achievement levels over all and among schools permits a more fine-grained analysis of student progress than is possible with traditional status and improvement indicators. In addition this methodology yields growth metrics that can be easily explained to educators, and provides data appropriate for grouping low-achieving students demonstrating “extraordinary” growth in relationship to their academic peers. This approach permits factors associated with those successes to be identified and replicated to improve the learning progress of other low-achieving students. For example, the data can lead to meaningful investigation of related educational

variables such as: What background characteristics and/or educational experiences might the 8th grade students who made above average growth share?

Specifically, this analysis showed that a substantial proportion of 6th graders in this district made above average growth from grade 6 to grade 7 and from grade 7 to grade 8 in reading achievement. Growth from grade 6 to grade 7 was greater for 6th grade students who were below proficient than was growth from grade 7 to 8. That is, 38% of the Level 1 group demonstrated above average growth in 7th grade compared to 27% in 8th grade and 44% of the Level 2 group demonstrated above average growth in 7th grade compared to 39% in 8th grade. More of the Level 3/4 group, however, demonstrated above average growth in 8th grade (51%) compared to 7th grade (42%).

Despite the downward trend of growth scores from 7th grade to 8th grade for the students who were below proficient in 6th grade, there was above average growth for more than one-quarter (27 %) of students at the lowest performance level in grade 6 and for well over one-third (39 %) of students at the next lowest performance level in grade 6. This is an encouraging finding especially in comparison with only very modest increases in the overall percent of students at the proficient level from grade 6 to grade 8. While the district certainly needs to continue and perhaps intensify intervention initiatives efforts toward helping more middle school students achieve reading proficiency by grade 8, it is apparent that current intervention efforts are having a positive impact on the progress of some 6th grade students at the lowest levels of achievement.

While the schools were about equally effective in supporting the growth of students at the lowest levels of 6th grade achievement, some schools had substantially higher percentages of students with above average growth for students with 6th grade achievement at Level 2 (the next to lowest performance level) and Level 3 / 4 (proficient or advanced). For one of the schools, over 40% of students in these categories demonstrated above average growth and for another, 49% - 63% had above average growth. One of the schools had a substantially higher percent of low-achieving 6th graders who achieved proficiency in grade 8 in comparison with the other two schools. Examination of what kind of support low-achieving students received in that school could yield valuable insights that support educational improvement planning.

Finally, limiting the measure of improvement at the percent of students who are proficient across schools in any one year and changes (improvement) in those percentages does not reflect growth that low-achieving students are making within the schools toward reaching proficiency. Nor does it indicate which schools may be more effective in supporting low-achieving students. Likewise, the percent of students proficient each year, or improvements in that index, provides little information about which students may be failing to maintain proficient status and which schools may be more effective in supporting students in continuing to demonstrate growth when they have achieved performance levels of proficiency or above.

The results of this study demonstrate levels of individual student achievement progress throughout the achievement continuum; differences among schools in student growth; and growth trajectories needed from grade 6 to grade 8 to reach proficiency that traditional status and improvement indices cannot capture. In addition, because each student has a SGP score, the analysis permits districts to identify specific students who demonstrated growth substantially above average and to examine what educational experiences may have contributed to the discrepancy between their growth patterns and those of their 6^{th} grade academic peers. This will help the district to target intervention efforts more effectively.

Implications for Further Research

Further analysis such as differences in SGP scores among student subgroups by program assignment (regular education, special education, ELL) as well as additional analyses over time, are needed both to see if growth results are similar among different student subgroups and in order to cross-validate the findings. More investigation of the downward trend for SGP scores from grade 7 to grade 8 is also needed to determine if this is a typical pattern which the district needs to investigate and correct or whether it might be something specific to this sample and these time periods. It is also possible that the pattern might related to factors unrelated to student achievement growth such as the content of the 7^{th} and 8^{th} grade tests or it might be an artifact of the analysis. For instance, in this study, students were assigned SGP scores of 80, for instance, if their residuals were positive, *above* the 80^{th} decile, and below the 90^{th}. If a student's residual was near the 80^{th} decile, the student was assigned a SGP score of 70. This procedure provided a

conservative estimate of SGP scores—something determined to be appropriate for the purposes of the study—but it may also have introduced opposing trends at the two ends of the achievement spectrum in terms of the relationship between SGP scores across grade 7 and grade 8.

References

10 Capacities for Initiating and Sustaining School Improvement at the Elementary Level (2000). The Northeast and Islands Regional Laboratory (The Lab), A Program of the Education Alliance at Brown University. Retrieved September 20, 2008 from http://www.alliance.brown.edu/db/ea_catalog.ph

Betebenner, D. W. (March, 20, 2008). *Norm- and Criterion-Referenced Student Growth*. The National Center for the Improvement of Educational Assessment. Retrieved May 27, 2008 from http://www.nceia.org

Choi, K., Seltzer, M., Herman, J., & Yamashiro, K. (Fall 2007). Children left behind in AYP and non-AYP schools: Using student progress and the distribution of student gains to validate AYP. *Educational Measurement: Issues and Practices*, pp. 21–32.

Colorado's Academic Growth Model. (February 13, 2008). Report of the Technical Advisory Panel for the Longitudinal Analysis of Student Assessment Convened Pursuant to Colorado HB 07-1048.

Delaware's Proposal for a Growth Model, Submitted to U. S. Department of Education, February 17, 2006.

Florida's Application for the NCLB Growth Model. (September 15, 2006). Florida Department of Education.

Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (October 2005). *Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ?* Council of Chief State School Officers (CSSO).

Hao, L. & Naiman, D. Q. (2007). *Quantile Regression.* Sage Publications, Inc. Thousand Oaks, CA.

Hayslett, H. T., Jr. (1968). *Statistics Made Simple*. Doubleday: New York.

Hibpshman, Terry. (September 2004). *A Review of Value-Added Models*. Kentucky Education Professional Standards Board.

Hill, R. (April, 2006). *Using value tables for a school-level accountability system*. Paper presented at the NCME Annual Conference. The National Center for the Improvement of Educational Assessment.

Ho, A. D. (2008). The problem with "Proficiency": Limitations of statistics and policy under No Child Left Behind. *Educational Researcher*, Vol. 37, No. 6, pp.351-360.

Implementer's Guide to Growth Models. (January 2008). A paper commissioned by the CCSSO Accountability Systems and Reporting State Collaborative Project. Council of Chief State School Officers, Washington, DC. Retrieved September 25, 2008 from http://www.ccsso.org/publications/details.cfm?PublicationID=360

Koenker, R. (May 20, 2006). *Quantile Regression in R: A Vignette.* Retrieved October 21, 2008 from *http://www.econ.uiuc.edu/~roger/research/rq/vig.pdf*.

Koenker, R., and Hallock, K. F., (Fall 2001). Quantile Regression. *Journal of Economic Perspectives*, Vol. 15 (4), pp. 143-156.

Next Generation of Value-Added Models and Indicators. Wisconsin Center for Education Research at the School of Education, University of Wisconsin-Madison. Retrieved February 25, 2008 from http://www.wcer.wisc.edu/projects/projects.php?project_num=77

NCLB Growth Model Pilot Program: Proposal to the U. S. Department of Education. (February 16, 2006). Tennessee Department of Education.

No Child Left Behind Growth Model Pilot Proposal, Submitted by the Iowa Department of Education to U. S. Department of Education, January 19, 2007, Revised April 30, 2007.

North Carolina's Proposal to Pilot the Use of a Growth Model for AYP Purposes in 2005-2006. (April, 16, 2006) NCDPI/Accountability Services.

Proposal for a Growth Model to Evaluate Adequate Yearly Progress for Schools and Districts. (Revised July 2, 2007). Arizona Department of Education.

Proposal to the United States Department of Education for Employing a Growth Model for No Child Left Behind Accountability Purposes. (October 31, 2006). State of Ohio, Ohio Department of Education.

Proposal to the US Department of Education for Participation in the No Child Left Behind (NCLB) Growth Model Pilot Program. (October 31, 2006). The Pennsylvania Department of Education.

SAS® EVAAS® for K-12: Value-Added Assessment—A Powerful Diagnostic Tool. SAS. Cary, NC. Retrieved February 25, 2008 from http://www.sas.com/govedu/edu/evass.pdf

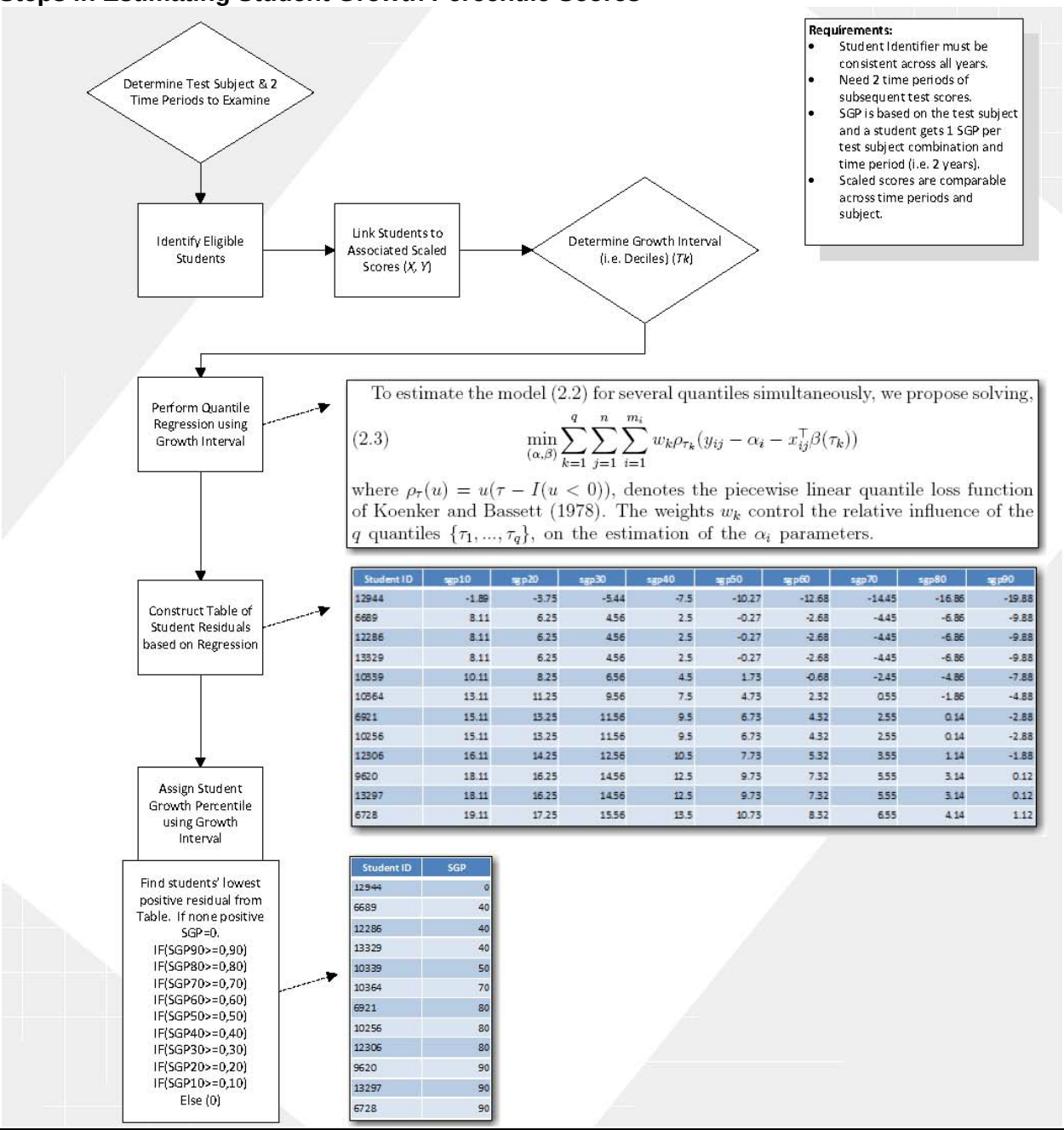
The Annenberg Challenge: Lessons and Reflections on Public School Reform (2002). Carnegie Corporation: New York. Retrieved September 20, 2008 from http://www.annenbergfoundation.org/usr_doc/Lessons_&Reflections_report.pdf

Willet, J. B. (1994). *Measurement of Change.* In T. Husen and T. N. Postlethwaite (Eds.). (2nd Ed). International Encyclopedia of Education, Oxford, UK: Elsevier Science Press, 671 – 678. Retrieved September 25, 2008 from http://gseacademic.harvard.edu/~willetjo/change.htm

Appendix A

Quantile Regression Analysis Steps

Steps in Estimating Student Growth Percentile Scores



Student ID	sgp10	sgp20	sgp30	sgp40	sgp50	sgp60	sgp70	sgp80	sgp90
12944	-1.89	-3.75	-5.44	-7.5	-10.27	-12.68	-14.45	-16.86	-19.88
6689	8.11	6.25	4.56	2.5	-0.27	-2.68	-4.45	-6.86	-9.88
12286	8.11	6.25	4.56	2.5	-0.27	-2.68	-4.45	-6.86	-9.88
13329	8.11	6.25	4.56	2.5	-0.27	-2.68	-4.45	-6.86	-9.88
10339	10.11	8.25	6.56	4.5	1.73	-0.68	-2.45	-4.86	-7.88
10364	13.11	11.25	9.56	7.5	4.73	2.32	0.55	-1.86	-4.88
6921	15.11	13.25	11.56	9.5	6.73	4.32	2.55	0.14	-2.88
10256	15.11	13.25	11.56	9.5	6.73	4.32	2.55	0.14	-2.88
12306	16.11	14.25	12.56	10.5	7.73	5.32	3.55	1.14	-1.88
9620	18.11	16.25	14.56	12.5	9.73	7.32	5.55	3.14	0.12
13297	18.11	16.25	14.56	12.5	9.73	7.32	5.55	3.14	0.12
6728	19.11	17.25	15.56	13.5	10.73	8.32	6.55	4.14	1.12

Student ID	SGP
12944	0
6689	40
12286	40
13329	40
10339	50
10364	70
6921	80
10256	80
12306	80
9620	90
13297	90
6728	90