NATIONAL
CENTER *for* ANALYSIS *of* LONGITUDINAL DATA *in* EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

Chronically Low-performing Schools and Turnaround: Evidence from Three States

MICHAEL HANSEN AND KILCHAN CHOI

WORKING PAPER 60 • AUGUST 2012

Chronically Low-performing Schools and Turnaround: Evidence from Three States

Michael Hansen
American Institutes for Research

Kilchan Choi
American Institutes for Research

Contents

Acknowledgements

The research presented here was originally performed under contract with the Institute of Education Sciences (ED-04-CO-0025/0020). The study team was led by American Institutes for Research, with the Urban Institute, Decision Information Resources, and Policy Studies Associates. The preparation of this manuscript was not supported by this contract. Tommy Gonzalez provided superior research assistance, and Mike Garet, Jane Hannaway, and Rebecca Hermann provided critical guidance on the issues addressed here. We also acknowledge the North Carolina Education Research Data Center, the Florida Education Data Warehouse, and the Texas Schools Project - Education Research Center for providing access to the data utilized in the study and presented here.

 Author contact: mhansen@air.org and kchoi@air.org.

CALDER • American Institutes for Research
1000 Thomas Jefferson Street N.W., Washington, D.C. 20007
202-403-5796 • www.caldercenter.org

Chronically Low-performing Schools and Turnaround: Evidence from Three States

Michael Hansen and Kilchan Choi

CALDER Working Paper No. 60
August 2012

Abstract

This paper presents a new method of empirically identifying CLP schools and binning them into performance categories based on their trajectories. The method as it is applied here will be useful to state or local education agencies charged with similar tasks of identifying their lowest performers, and monitoring signs of improvement. Our findings show that low-performing schools turned around their performance more frequently than one might have presumed based on prior literature. In Florida, we identified approximately 15% of chronically low-performing elementary and 14% of chronically low-performing middle schools as turnarounds. Similar rates were observed in North Carolina—13% and 16%, respectively; and even higher in Texas—29% and 31%, respectively.

Introduction

Under Secretary Arne Duncan's leadership, the U.S. Department of Education has focused efforts and resources on turning around the nation's lowest performing schools. Turnaround strategies are integral elements of the administration's School Improvement Grant (SIG) program and the recent Race to the Top (RTT) initiative. Yet, in spite of the interest in intervening in the bottom five percent of the nation's chronically low-performing schools, the working definitions of "chronically low performing" and "turnaround" schools have been largely *ad hoc* in research and practice. No standard definition exists on what qualifies a school to be eligible for these designations, though various approaches abound.

This unsystematic approach to school turnaround is potentially risky for two reasons. First, intervention efforts may be inadvertently misallocated to schools that do not need them (false positives) while passing over struggling schools that do (false negatives). Second, the Department of Education's bold turnaround strategies may disrupt emerging improvements in some low-performing schools already engaged in their own (undetected) turnaround, which may potentially do more harm than good in such situations.

This paper presents an original methodology developed to identify chronically low-performing (CLP) and turnaround (TA) schools based on student performance. This method is uniformly applied to longitudinal data on student test scores from Florida, North Carolina, and Texas to address the study's research primary question—what proportion of CLP schools display specific performance trajectories over time? In particular, we seek to understand what proportion of CLP schools display: quick, dramatic improvement (TA schools); weak or mixed improvement (MI schools); or no signs of improvement (NI schools)?[1]

[1] The mixed improvement (MI) and non-improvement (NI) designations are original to this study to separately classify other CLP schools that fail to show large enough improvements to earn the TA label.

In summary, we find low-performing schools turned around their performance more frequently than one might have presumed based on prior research. In Florida, we identified approximately 15% of chronically low-performing elementary and 14% of chronically low-performing middle schools as turnarounds in at least one subject. Similar rates were observed in North Carolina—13% and 16%, respectively; and even higher in Texas—29% and 31%, respectively. We present supporting evidence to suggest that the improvements in performance identified with this model represent real gains, and are not due to demographic shifts in the student bodies of the schools.

Background and Context

The criteria for determining the student outcomes that define a school as having "turned around" are not well defined (Kutash, Nico, Gorin, Rahmatullah, & Tallant, 2010), and the definition of turnaround performance varies across studies (Aladjem, Birman, Orland, Harr-Robins, Heredia et al., 2010; Herman, Dawson, Dee, Greene, Maynard et al., 2008). Although current policy initiatives offer guidelines for identifying CLP schools, there is no standard definition or methodology in common usage. SIG and RTT, for example, require states to identify the lowest five percent of schools but allow states flexibility to set critical parameters such as the period over which performance should be measured.[2] Similarly, federal policy sets expectations that states will measure school improvement progress against benchmarks but allows states to establish the specific benchmarks (within parameters) and methods of measuring progress. Consequently, methods abound for identifying low-performing and turnaround schools.

All known attempts to identify CLP and TA schools in research and practice, up to this point, have relied on school-based performance measures—either mean test scores or, more commonly,

---

[2] The current emphasis on turnaround in SIG and RTT are variations on school restructuring policies under No Child Left Behind. See Malen and Rice (2009) for a review of school responses to such reforms.

percent proficient (for example, Brady, 2003; Meyers, Lindsay, Condon, Wan, 2012; Scott, 2008; Smarick, 2010; Stuit, 2010). Yet, school-based measures are problematic on three points (Kane and Staiger, 2002): first, percentage measures (e.g., percent proficiency, graduation rates) ignore important variations in outcomes that occur on either side of the cutpoint; second, school-based measures ignore the implied measurement error when calculating across schools of differing sizes; and third, school-based measures ignore compositional changes in the student body from cohort to cohort.

One may reasonably expect these problems with school-based performance data to be particularly detrimental in the case of identifying low-performing and turnaround schools. Attempting to identify low performers by definition focuses on schools with performance at the extreme low end of the distribution—these are the same schools that are most likely to be subject to large corrections over time if the observed performance was due to measurement error. Thus, distinguishing authentic low performance and authentic improvements from corrections due to random fluctuation in error-prone measures is the primary challenge of attempts to empirically identify this phenomenon. Consequently, studies in the turnaround literature that rely exclusively on school-level data (particularly data on measures such as percent proficiency) leave themselves vulnerable to such criticism.

Moreover, school-based accountability measures are generally status-based measures, which confound pre-existing differences among students with differences in school quality. Growth measures, conversely, are generally regarded as better measures of isolating schooling inputs (Meyer, 1997; Raudenbush, 2004). Status measures are much more stable over time (slower to show signs of improvement), while growth measures for a given student are likely to vary in different time periods and contexts in accordance with educational inputs (Kane & Staiger, 2002; Linn & Haug, 2002). This implies a focus on status-based performance measures would likely underidentify TA schools or identify improvements only several years after the turnaround actually occurred.[3]

This paper's contribution to the literature is its development of a method to identify low performance and turnaround in schools using student-based longitudinal data that uses both status- and growth-based performance measures. We apply this model to administrative data from Florida, North Carolina, and Texas to investigate the trajectories of low-performing schools. By utilizing student-level growth-based measures, this model overcomes the weaknesses inherent in methods that use school-level status-based measures raised above. In the model presented here, growth measures are calculated using student-level data along continuous measures, can be adjusted according to the imprecision in the estimate, and capture within-student improvements rather than differences between cohorts; as a result, we argue that we can better capture authentic performance and changes in performance that prior models have not been able to satisfactorily distinguish.

We wish to be clear this paper is not intended as an evaluation of any particular turnaround interventions in any of the three states. Rather, its purpose is to develop a credible empirical method of identifying low performance and turnaround that incorporates student-level outcome data and apply across several data sources to document the frequency of turnaround using this approach. All three states, however, had their own policies and consequences for restructuring low-performing schools that consistently failed to make adequate yearly progress under No Child Left Behind, which we show below is correlated with our identification method for low performers (though is intentionally different for reasons described above). Thus, many of the schools we identify as low performers were likely targeted with some level of improvement efforts, and all were exposed to accountability pressure.

[3] Ushomirsky and Hall (2010) propose a method for identifying low-performing schools (what they term "stuck" schools) using status and growth measures simultaneously. Meyers, et al. (2012) use a method that requires monotomic serial improvements in school performance to be labeled turnaround schools.

Modeling School Performance and Turnaround

Conceptually, the process of classifying schools' performance and identifying those that turn around appears straightforward—one simply isolates the lowest performers based on past performance and then observes current performance for improvements. Yet, this simplistic formulation ignores the complex and important decisions about measurement, data, and modeling that will ultimately influence the group of schools identified as CLPs and TAs. This section describes the specifics of our modeling approach employed here to measure school performance, and then describes the method applied to identify CLP and TA schools in the data.[4]

Measuring School Performance

Measuring school performance requires answering four questions: what data to use, in what subjects to measure performance, whether to measure performance as achievement levels or gains, and of which students?

<u>1. Using what data?</u> This study utilizes student-level data from state administrative data warehouses. These data come from three states: Florida, North Carolina, and Texas. The availability of high-quality, student-level data spanning multiple years was the primary driver in selecting these three states for the study. Each of the states has a large and diverse student population, with schools situated in urban, suburban, and rural settings. The datasets are uniformly constructed in order to apply a uniform identification strategy to all three states. This dataset spans grades 3 to 8 for the six school years between 2002–03 and 2007–08, and includes the universe of students and schools for these

[4] Further discussion on the issues inherent in attempts to empirically identify CLP and TA schools is presented in Hansen (2012).

grades. By using student-level observations, we capture variation in performance within and between students, not simply aggregated at the school level.[5]

2. Performance in what subjects? To measure school performance, we focus on student scores on state accountability assessments in reading/language arts and mathematics. Schools with poor performance (as defined below) in either subject are designated CLP, and schools are designated with subject-specific CLP status (either CLP-Reading or CLP-Math). Because school performance on both subjects is positively correlated, many schools are designated CLP in both subjects, and we pay particular, but not exclusive, attention to this subset of schools.

3. Performance levels or student growth? Achievement levels show where schools are at a given grade level. Student growth indicates the progress students make as they move from one grade to the next. Both measures are informative in assessing school performance—status measures represent accumulated learning (which may be from several sources, including schools), while growth measures show year-to-year performance differentials that would be most closely related to school inputs into learning. To represent a school's achievement levels, we used the performance of students in the school's terminal grade (i.e., grade 5 for elementary school or grade 8 for middle school); we use the term "status" to refer to this outcome measurement throughout the remainder of this study. To represent a school's achievement gains, we examined the between-grade growth in student achievement (i.e., the growth between grades 3 and 5 for elementary school and between grades 5 and 8 for middle school); we use the term "growth" for this outcome.[6]

4. Performance of which students? Student mobility may influence our identification of chronically low-performing schools. Student mobility is particularly relevant for this study because the

[5] Please see Section I of the Appendix for more documentation on the data and the sample selection methods used for the study, as well as a table of the descriptive statistics for all six samples.

[6] As shown in Section II of the Appendix, correlations of estimated status and growth parameters within schools during the baseline observation period for the study indicate a small positive correlation between these outcome measures within schools. This pattern is consistent across all of the data samples used in the study.

membership of a given cohort of students constantly changes across time periods, and mobility tends to be higher among schools serving disadvantaged populations (Raudenbush & Wilms, 1995; Robinson, Levin, Thomas, Pituch, & Vaughn, 2007). We could deal with student mobility using one of two alternatives: restricting analyses to stable students (i.e., those who exhibit standard grade progression and do not change schools) or including all students (i.e., including both stable and mobile students). Because low-performing schools are most vulnerable to misrepresentation through the omission of mobile students, we choose to include all students in the analyses.

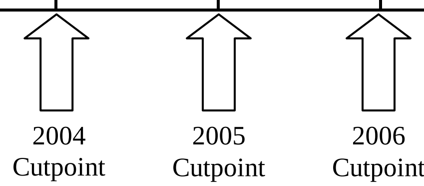
Classifying Schools as Chronically Low Performing

To classify schools as CLP, we address how we operationalize the meaning of "chronic" low performance and how low performance must be to warrant the CLP label.

<u>1. How long must a school be low performing?</u> The first component of chronic low performance deals with selecting an appropriate time span over which to evaluate school performance. Because we are dealing with data that have a limited time span for the current analysis, this issue is directly related to the issue of how long a chronically low-performing school must sustain improvements to be considered a turnaround school; we thus use the data structure in determining our identification strategy here. Figure 1 below presents the structure of the elementary-level data in the time span of the analysis. The columns represent years and the rows represent cohorts of students that progress through the school grades (labeled "Gr3" for grade 3, etc.). Our analysis is confined to the six-year period from the 2002–03 through 2007–08 school years. We observe eight unique cohorts of students progress through each school over this period: four cohorts (C3, C4, C5, and C6) are observed in each of the three elementary school grades for which we have data, two cohorts (C2 and C7) are observed in two of the three grades, and two cohorts (C1 and C8) are observed in only one grade each.

Figure 1: All Combinations of Base Period and Outcome Period Used in the Analysis

	2002–03	2003–04	2004–05	2005–06	2006–07	2007–08
C1	Gr5					
C2	Gr4	Gr5				
C3	Gr3	Gr4	Gr5			
C4		Gr3	Gr4	Gr5		
C5			Gr3	Gr4	Gr5	
C6				Gr3	Gr4	Gr5
C7					Gr3	Gr4
C8						Gr3



The task is to partition these data into two periods: a pre-period to determine CLP status and a post-period to categorize CLP schools as TA, NI, or MI. We refer to the break between the two periods as the cutpoint. CLP schools (and further subdivisions into TA, MI, and NI groupings) are categorized based on school performance over time; thus, we wish to aggregate school performance over multiple years to limit misclassification due to random error in a single year, since year to year changes in the performance of groups of students can be volatile (Linn & Haug, 2002). We employed an estimation strategy in which we inserted a cutpoint separating the pre- and post-periods in three different places, labeled as the 2004, 2005, and 2006 cutpoints in Figure 1 above. To clarify, the 2004 cutpoint is placed between the 2003-04 and 2004–05 school years, which implies the first two years of school performance are considered the pre-period and the final four years in the data are considered the post-period. Similarly, the 2005 cutpoint has a three-year pre-period and a three-year post-period, while the 2006 cutpoint has a four-year pre-period and a two-year post-period. We used three different cutpoints in the analysis because we have no way to determine *a priori* the actual time that a school changed its performance to warrant designation as a turnaround. Imposing several different cutpoints on the data and reevaluating the model under each allows us to flexibly identify improvements in school performance that may have happened anytime in any of these periods.

2. How low must performance be? Current federal policy prioritizes the bottom five percent of schools as the primary target for turnaround efforts. This study used this five-percent target in determining appropriate thresholds for classifying CLP schools; in particular, we iteratively searched through combinations of low-status and low-growth thresholds that resulted in approximately five percent of the total schools in each of the state samples jointly classified as CLP in both reading and math. Based on this exercise, we adopted a strategy that defines CLP schools as those that score in the lowest 15 percent of all schools in status and in the lowest 40 percent in growth within a given subject.[7] We preferred a lower status threshold to focus intervention efforts on schools that perform at the lowest levels overall. Yet, we additionally wanted to focus on low-status schools that demonstrated below-average growth (i.e., students are losing ground over time relative to peers). Low-status schools that have higher growth are already on a generally upward trajectory (which would presumably result in the emergence from the low-status designation in the near future if the trajectory were sustained), thus we omitted them from our classification as chronically low-performing schools. Low-growth, low-status schools are the most likely and appropriate targets for intervention efforts, given their low performance in both dimensions.

Note that the status threshold used here (those in the lowest 15 percent) is markedly higher than the bottom five percent target, which may be initially confusing. Because this method requires a school to show both low status *and* low growth to be labeled CLP, we must revise the status threshold upwards (as not all schools in the bottom five percent in status demonstrate low growth). Given the specified target of five percent of schools labeled as CLP, the status and growth thresholds are inversely

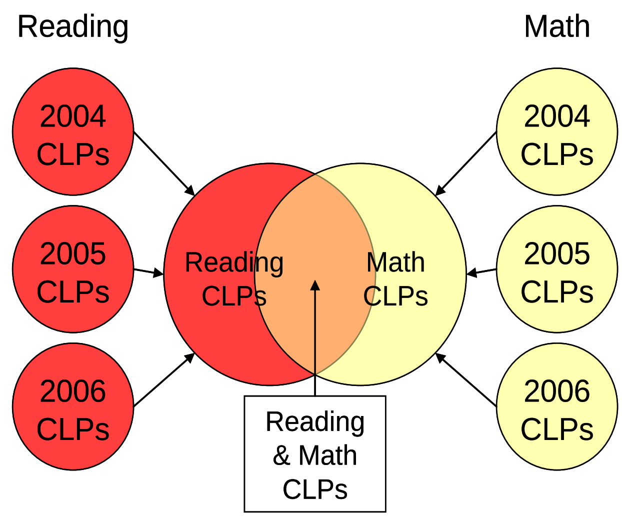
[7] For ease of interpretation, we chose to use two separate thresholds for status and growth in specifying the classification rules; however, other analysts may choose to specify different rules in identifying CLPs. For instance, combining weighted status and growth scores in some way will could be a meaningful approach.

related.[8] If we desired a low status threshold, the growth threshold would need to go up. Using both of these status and growth thresholds together results in getting as near the 5% target as we can in all of the data samples used in the analysis.

Our identification method essentially combines these two decisions. First, we observe student performance in the pre-period (using a given cutpoint, within a given subject) to estimate status and growth parameters; those schools scoring at or below both the 15th percentile in status and the 40th percentile in growth are labeled CLP in that subject-cutpoint combination. As there are six different subject-cutpoint combination, there are six different CLP classifications within each data sample (illustrated with Figure 2). All of schools that are considered CLP in any of these designations are considered CLPs in our reported results. Note, however, that many CLPs are common across cutpoints within subjects, and are again common across subjects, but that commonality is not a necessary condition to being considered a CLP for the study. As we report the results throughout, we present findings by subject (based on CLP status in either math or reading); and, we focus particular attention on the set of schools that are CLP in both reading and math (the intersection between reading and math CLPs in Figure 2) as these schools represent approximately five percent of all schools in each sample and we take these to be prime targets for turnaround intervention efforts.

[8] Consider the function {Target ≈ (status threshold) x (growth threshold)}, which provides an approximation of the actual amount of schools captured by the thresholds (equality holds if the status and growth measures are uncorrelated, in practice there is a small positive correlation). If we desired a low status threshold, the growth threshold would need to go up to meet a fixed target amount. For instance, having a threshold on status of 5 percent would imply a threshold of 1 on growth in order to achieve the target (i.e. all schools in the bottom 5 percent of status, regardless of their actual growth). Since we want to require that schools have low growth (below the 40th percentile), this function requires the status threshold to increase to hit the 5 percent target. The thresholds we set resulted in identifying slightly more than the lowest 5% of schools in each sample as CLP in both subjects.

Figure 2: CLP Status Across Cutpoints and Subjects



Categorizing Performance Trajectories and Identifying Turnaround, Mixed-Improvement, and Not-Improving Schools

Last, we define our strategy in categorizing CLP schools into turnaround (TA), mixed-improvement (MI), and not-improving (NI) schools based on their demonstrated performance in the post period of the study.

<u>1. Measuring post-period status and growth.</u> All CLPs identified during the pre-period are categorized as TA, MI, or NI schools based on their observed performance in the post-period. Because we are interested in changes in two performance outcomes (status and growth) within schools over time, we made comparisons in two different ways. First, the change in a school's status over time is estimated by comparing the terminal-grade performance of student cohorts before the cutpoint relative to those after the cutpoint. For instance, using the 2005 cutpoint in elementary grades, the 5th grade performance of cohorts 1, 2, and 3 is compared with that of cohorts 4, 5, and 6 (referring back to Figure 1).

The second comparison (determining changes in growth outcomes) is slightly more complex because growth represents the marginal improvement observed in a cohort of students over time and, regardless of the cutpoint used, some cohort of students will straddle the cutpoint. This presents us with an attribution problem for these straddling cohorts. Attributing all of the growth in those cohorts to either the pre- or post-period could attenuate the resulting estimate of the difference in performance between the two periods. To counter this problem, we divided a cohort's growth into one-year blocks (e.g., the marginal growth from grade 3 to grade 4) and assigned that block of growth to either the pre- or post-period depending on when the growth took place relative to the cutpoint. For instance, using the 2005 cutpoint compares all the growth observed in cohorts 2 and 3 with cohorts 6 and 7 (again, referring to Figure 1). Note, however, that cohorts 4 and 5 both straddle the cutpoint. In the case of cohort 4, the grade 3 to grade 4 growth is attributed to the pre-period, while the grade 4 to grade 5 growth is attributed to the post-period. In the case of cohort 5, all of the cohort's growth is attributed to the post-period, since only grade 3 occurred prior to the cutpoint (grade 4 occurred after the cutpoint, implying the growth between these points is attributable to the latter period). The data structure presented in Exhibit 2 represents elementary schools, but a similar coding strategy is applied to middle schools, using grade 5 as a pre-test score. For middle schools, we divided growth into three segments (grade 5–6, 6–7, and 7–8), and assigned them to the pre- or post-period depending on when the growth occurred relative to the cutpoint.

2. How large must improvements be? Based on the school's demonstrated performance in the post-period, we then categorized these CLP schools into three categories. Previously we described these three categories as

- a policy-relevant improvement in performance (TA schools),
- mixed, or weak evidence of improvement (MI schools), and
- no signs of improvement or declining performance (NI schools).

We implemented these three categories as follows, depicted in Figure 3. First, TA schools showed improvements in both status and growth outcomes in a given subject. Specifically, these schools had an average estimated performance in the post-period that was both a statistically significant increase over average performance in the base period (using a one-tailed test with an alpha level of 0.05) and higher by an improvement threshold that we considered meaningful. This improvement threshold varied for status and growth outcomes: for status, an improvement of at least 5 percentile points (relative to the pre-period percentile ranking) had to be obtained; for growth, the school's post-period estimate must have met or exceeded the 65th percentile of all schools in the post-period to ensure the school is on an upward trajectory.[9] Next, MI schools demonstrate an average estimated performance in the post-period (in either status or growth, or in both) that was a statistically significant increase over pre-period performance but did not obtain statistical significance and exceed the meaningful targets in both status and growth simultaneously. Finally, NI schools exhibited either a nonsignificant increase in performance between the pre- and post-periods, or a decline in performance (significance not tested for decreases) for both status and growth.

Figure 3: Classification of Chronically Low-Performing Schools

		Post-period status improvement		
		Not significantly positive	Significantly positive / < 5 percentile points	Significantly positive / ≥ 5 percentile points
Post-period growth estimate	Not significantly positive	NI	MI	MI
	Significantly positive / < 65th percentile	MI	MI	MI
	Significantly positive / ≥ 65th percentile	MI	MI	TA

[9] The threshold we implemented on meaningful improvement in status measures is notably lower than that in other studies in the turnaround literature. This is intentional as status measures by construction slowly evolve in response to changes in inputs over time. Note, however, that we also implement a high bar on improving growth measures, which other studies generally do not employ. A lengthier discussion of improvement in status and growth measures in low-performing schools is presented in Hansen (2012).

Note that CLP schools are subdivided into TA, MI, and NI separately in each subject; thus, schools that are CLP in both reading and math can be considered in TA in reading but not math (or vice versa). As depicted in Figure 3, the subgroups for TA and NI schools are fairly selective, the MI subgroup is a catch-all category; the results we report below also show the TA and NI groups are considerably smaller than the MI group in the data. Later, we compare our method of identifying turnaround in schools with changes in performance using a more familiar policy measure, percent proficiency, within a school. As will be shown, the TA schools we identify under this method appear to show consequential gains in performance using this metric.

Statistical Model

The model we employ to estimate pre- and post-period performance in schools is a three-level hierarchical linear model, in which student's test scores over time are nested within students, which are in turn nested in schools. The model for the elementary school sample is described below.

Level-1 (within-student):

$$Y_{tik} = \pi_{0ik} + \pi_{1ik}\text{Time1}_{tik} + \pi_{2ik}\text{Time2}_{tik} + \varepsilon_{tik} \qquad \varepsilon_{tik} \sim N(0, \sigma^2) \qquad (1)$$

Level-2 (between-student; within-school):

$$\pi_{0ik} = \beta_{00k} + \beta_{01k}\,\text{Post\_status}_{jk} + r_{0ik} \qquad r_{0ik} \sim N(0, \tau_{\pi 0}) \qquad (2a)$$

$$\pi_{1ik} = \beta_{10k} + \beta_{11k}\text{Post\_g34}_{jk} \qquad (2b)$$

$$\pi_{2ik} = \beta_{20k} + \beta_{21k}\text{Post\_g45}_{jk} \qquad (2c)$$

Level-3 (between-school):

$$\beta_{00k} = \theta_{000} + V_{00k} \qquad V_{00k} \sim N(0, \tau_{00}) \qquad (3a)$$

$$\beta_{01k} = \theta_{010} + V_{01k} \qquad V_{01k} \sim N(0, \tau_{01}) \qquad (3b)$$

$$\beta_{10k} = \theta_{100} + V_{10k} \qquad V_{10k} \sim N(0, \tau_{10}) \qquad (3c)$$

$$\beta_{11k} = \theta_{110} + V_{11k} \qquad V_{11k} \sim N(0, \tau_{11}) \qquad (3d)$$

$$\beta_{20k} = \theta_{200} + V_{20k} \qquad V_{20k} \sim N(0, \tau_{20}) \qquad (3e)$$

$$\beta_{21k} = \theta_{210} + V_{21k} \qquad V_{21k} \sim N(0, \tau_{21}) \qquad (3f)$$

This model allows both status and growth parameters for both pre- and post-periods in all schools to be estimated simultaneously. Specifically, in Equation 1, Time1_{tik} took values -1 for grade 3, 0 for grade 4, and 0 for grade 5. Time2_{tik} was coded -1 for grade 3, -1 for grade 4, and 0 for grade 5. In this coding scheme, π_{0ijk}, π_{1ijk}, and π_{2ijk} represent, respectively, status at grade 5, growth between grades 3 and 4, and growth between grades 4 and 5. In Equations 2a, b, and c, each predictor was coded 0 for outcomes in the pre-period and 1 for those in the post-period. Note that Post_status_{jk} is 1 if the cohort was in 5th grade in the post-turnaround period. Post_g34_{jk} is 1 if the cohort was in 4th grade in the post-turnaround period. Post_g45_{jk} is identical to Post_status_{jk} within cohort, because 5th grade achievement and the grade 4–5 gain are realized simultaneously. Note that we constrain growth between grades to be constant for all students within a school (during either the pre- or post-period), to make the model tractable; though we allow the intercept to vary across students within schools. This model was applied separately for each subject-cutpoint combination in the data.

By extension, the middle-school version of this model was expanded to include 5th grade test scores as a pre-test score. That model included a Time0_{tik} variable in the first level coded as -1 in grade 5, and 0 for all middle school grades, which represents the first year of growth in the middle school; the model also included a Post_g56_{jk} indicator variable in Level 3 to identify when this growth took place relative to the turnaround point.

Schools are classified as CLP based on the estimated random-effect parameters V_{00k}, V_{10k}, and V_{20k}, which represent a given school's average pre-period status, growth from grade 3 to 4, and growth from grade 4 to 5, respectively. If a school's estimate of status, V_{00k}, fell in the lowest 15 percent of all school estimates and its estimated growth over the two periods, ($V_{10k} + V_{20k}$), fell in the lowest 40 percent, the school was considered a CLP school in that particular subject-cutpoint. For a school to be considered a TA, the post-period status (V_{01k}) and growth parameters ($V_{11k} + V_{21k}$) had to show statistically significant increases during the post-period, in addition to meeting the meaningful thresholds of both a 5 percentile-point improvement in the school's percentile in the post-period for status and ranking at or above the 65th percentile of all schools' growth in the post-period. If either of the parameters (status or growth) are statistically significant but do not meet the meaningful thresholds, the school was categorized as an MI school. Schools that exhibited no statistically significant improvement in either parameter estimate were labeled NI schools.

Results

We used the strategy outlined above to estimate school-level status and growth parameters in both the pre- and post-periods for each of the study's six datasets (3 states x (elementary + middle)).[10] Table 1 presents the total counts for the sample size, and chronically low performing, TA, MI, and NI schools in each of the samples. Totaling across all samples, we count 1,102 CLP schools (in any subject-cutpoint combination), which are subdivided into 249 TA schools, 574 MI schools, and 279 NI schools.

[10] Section II of the Appendix presents some evidence on the correlations of the estimated status and growth parameters for each school within each data sample (across subject-cutpoint combinations). Additional evidence is presented on the demographic characteristics of schools selected based on low status versus low performance measures.

Table 1: Count of Schools Identified as CLPs and Subdividing into TA, MI, and NI

Final Count of CLP, TA, MI, and NI Schools by State					
	Total Schools	**CLP Schools**	**TA**	**MI**	**NI**
Florida					
Elementary	1,599	224	34	91	99
Middle	535	57	8	28	21
Total	2,134	281	42	119	120
North Carolina					
Elementary	1,095	154	20	48	86
Middle	504	80	13	29	38
Total	1,599	234	33	77	124
Texas					
Elementary	2,662	466	136	220	110
Middle	1,023	121	38	65	18
Total	3,685	587	174	285	128
Three-state Total					
Elementary	5,356	844	190	359	295
Middle	2,062	258	59	122	77
Total	7,418	1,102	249	481	372

Note: Schools that are CLP in both reading and math are separately assigned to TA, MI, and NI status for each subject (e.g., a school could be TA in reading and MI in math). In this table, a TA school is TA in either reading or mathematics, an MI school is neither TA nor NI in reading or mathematics, and an NI school is NI in either reading or mathematics and never TA in reading or mathematics.

We are particularly interested in focusing on the TA classification in the groups of identified CLP schools across states and wish to better understand where turnaround occurs most frequently in the data (rather than looking at aggregated numbers as in Table 1). Table 2 below presents the incidence of identified TA schools when looking at subject-specific designations in each of the six data samples used in the study.

Table 2. Turnaround among CLP Schools by State, School Level, and Subject

	Florida		North Carolina		Texas	
	Elem.	Middle	Elem.	Middle	Elem.	Middle
CLP in Reading only						
Total CLPs across cutpoints	66	11	46	21	145	20
Total TAs	5	0	2	1	22	3
Turnaround rate	8%	0%	4%	5%	15%	15%
CLP in Math only						
Total CLPs across cutpoints	77	23	44	27	131	66
Total TAs	15	4	3	5	40	23
Turnaround rate	19%	17%	7%	19%	31%	35%
CLP in Reading & Math						
Total CLPs across cutpoints	81	23	64	32	190	35
Total TAs in Reading only	0	1	0	1	15	6
Total TAs in Math only	13	3	13	6	41	3
Total TAs in Reading & Math	1	0	2	0	18	3
Turnaround rate	17%	17%	23%	22%	39%	34%

Four particular findings from Table 2 are noteworthy. First, all three states have a much higher incidence of turnaround in math than in reading, whether among CLP schools designated as CLP in one subject only or in both subjects. Second, the group of CLPs in both subjects is of particular interest because of their low performance across all dimensions (status and growth in both reading and math scores)—we find turnaround rates among this group ranging from 17 to 39 percent, which is considerably larger than what we had expected based on prior studies in the literature (e.g., Smarick, 2010; Stuit, 2012). Third, the surprisingly high turnaround rate among these CLP schools in both subjects appears to be due mostly to improvements in math; the incidence of turnaround in both subjects was very rare. Fourth, the turnaround rates in Florida and North Carolina are fairly consistent across the data samples used; Texas is markedly larger in comparison for the given subject-specific CLP designation.[11]

[11] While we cannot know for certain, the reason for the higher turnaround rate in Texas may relate to the state's documented ceiling effect on the standardized exams (Koedel & Betts, 2008). This ceiling effect means that growth is primarily identified among students scoring at the low end of the test distribution and underestimated in high-

Descriptive View of TA, MI, and NI School Trajectories

If our categorization of schools as TA, MI, and NI is successful, we should observe differences in the achievement trajectories of the three groups of schools. We illustrate the differences in these trajectories by taking an in-depth view of the CLP schools identified in the North Carolina elementary school dataset.[12]

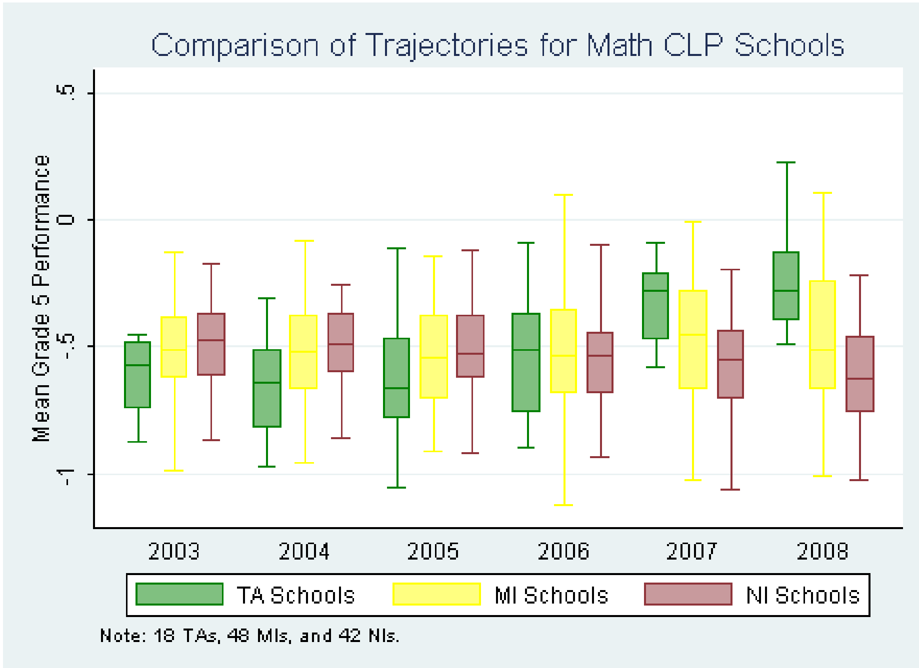
First, we plotted the trajectories of grade-5 performance among the CLP schools to ensure that the three subdivisions of these CLP schools are really capturing schools on different performance trajectories. This graphic is depicted as a box whisker plot in Figure 4. The box whisker shows all TA, MI, and NI schools had somewhat similar distributions for the first several years in the period; however, over the last few years the TA schools increased quite dramatically relative to the NI schools. By the end of the six-year period, the 25th percentile of TA schools was higher than the 75th percentile of the NI school distribution. The MI schools show some improvement over the period, taking a position between the TA and NI distributions; the NI schools show little improvement in the range of performance for these schools.

---

scoring students. This implicitly lowers the bar for showing significant improvements in growth among Texas CLP schools.

[12] Results from other five datasets are qualitatively similar to the tables and figures presented below.

Figure 4: Trajectories of Grade 5 Performance in Math in NC Elementary Chronically Low-Performing Schools



The box-whisker diagram clearly shows different trajectories in TA, MI, and NI schools, but recall that grade-5 performance is our status parameter—and we wish to verify our TA schools are improving in both status and growth. To look at this, we plot average performance in TA schools in math separately by cohort, as presented in Figure 5.

Figure 5: Trajectory of Average TA Schools in Math Using NC Elementary Sample

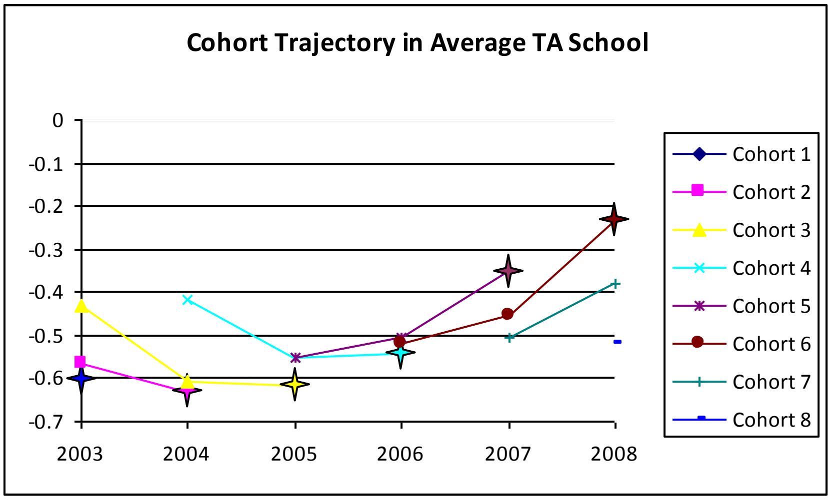
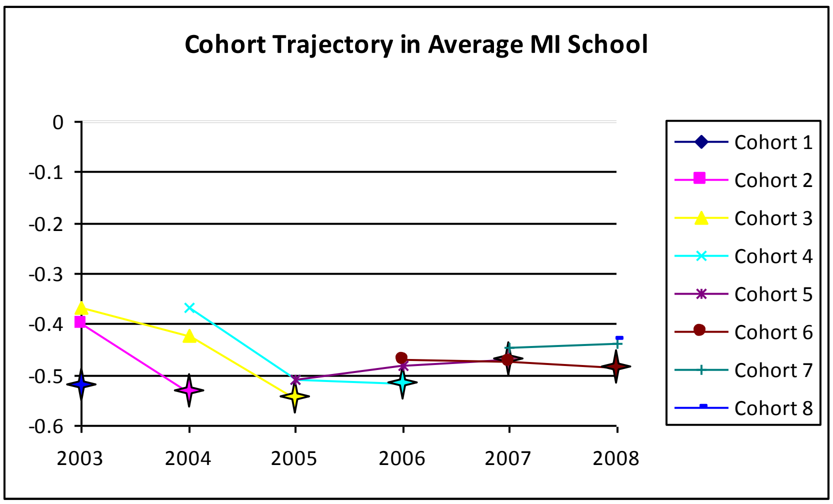


Figure 5 depicts the test-score outcomes for each cohort of students for each grade-year combination over the six-year analysis period. The x-axis represents the school year and the y-axis represents the normalized equivalent score. Each series of points represents the trajectory of each cohort that progressed through that school over time. Recall that we model status as the terminal-grade achievement level (grade 5 in this elementary school case); each cohort's 5th grade performance is accentuated with a star icon. Growth is depicted in this figure by the slope of each of the cohort lines; a positive slope indicates that the cohort is gaining ground in comparison with other students in the state, a negative slope represents losing ground. This figure shows the early cohorts of the TA schools were on a clear downward trajectory. Starting with Cohort 5, however, students began showing positive growth that stayed consistent (on average) in each of the succeeding cohorts. From this figure, there is evidence of both improved status and growth among the TA schools.

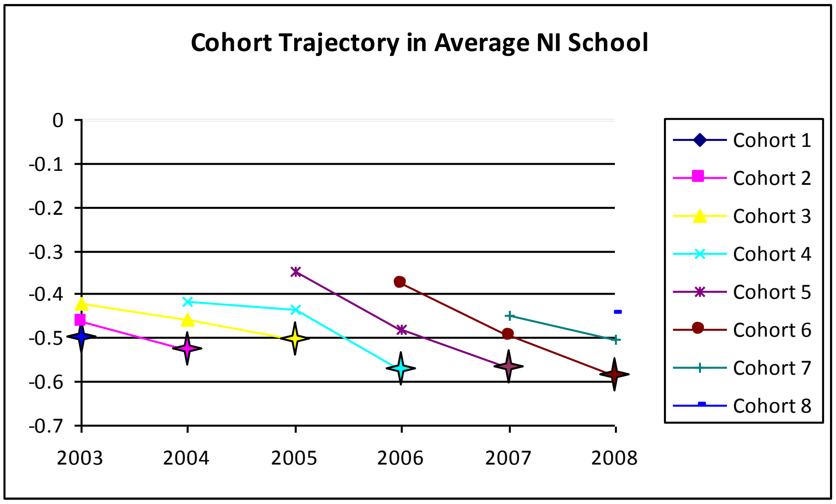
Contrast the performance of these TA schools with the trajectory of the schools we identify as MIs in Figure 6. Similar to Figure 5 above, this illustration represents the cohort average performance as they progress through schools. Beginning with Cohort 5, we observe a conspicuous break in the trajectory of these schools—to be considered an MI school, status or growth must significantly improve in the statistical sense in the post-period, but fail to meet the threshold improvement targets for both status and growth. These schools were clearly performing poorly on a consistent basis during the early cohorts and show some stabilization (on average) in status and growth outcomes in the post-period. Overall, the net improvement in the latter period is an improvement over the post-period, but certainly no net gains in absolute performance appear to be evident.

Figure 6: Trajectory of Average MI Schools in Math Using NC Elementary Sample



Finally, Figure 7 illustrates the achievement trajectories of the schools identified as NIs. These schools show no sign of improvement in either status or growth; rather, performance over the period begins low and declining and continues the trend throughout the period.

Figure 7: Trajectory of Average NI Schools in Math Using NC Elementary Sample



Validation of Results

The analyses above indicate that TA, MI, and NI schools differ in achievement trajectories, as anticipated. One question that remains to be addressed is whether the observed change in performance for TA schools might have been driven by changes in the underlying student demographics. To judge whether demographics played a role in performance improvements, we compared the demographics of TA schools before and after the 2005 cutpoint and compared these side-by-side with those of the MI and NI schools, which is presented in Table 3 for the NC elementary school sample. These tables show the variation in student body demographics from the pre- to the post-period was generally small. Moreover, changes in the TA schools were commonly mirrored in the MI and NI schools as well; for example, the percentage of students who are English language learners (ELL) appears to be uniformly increasing across all chronically low-performing schools, whether identified as a TA or not.

Table 3: Comparison of School Demographics Before and After Cutpoint in NC Elementary School Sample

Elementary Schools							
		Math TA		Reading TA		CLP Non-TA	
		Pre	Post	Pre	Post	Pre	Post
School Size	CLP Math	205	205			223	211
	CLP Reading			226	191	206	192
Percentage White	CLP Math	27.4%	27.7%			24.9%	22.3%
	CLP Reading			20.4%	21.8%	16.2%	13.9%
Percentage ELL	CLP Math	4.9%	8.9%			7.2%	15.0%
	CLP Reading			3.0%	8.6%	5.5%	11.9%
Percentage Free/Reduced Price Lunch	CLP Math	82.0%	80.4%			81.2%	82.1%
	CLP Reading			82.3%	82.3%	85.3%	86.7%

Additionally, we examined the time trend in the percent proficient statistic reported in the North Carolina Report Card system under its accountability program. The state's time trend of percent proficient for each of the TA, MI, and NI subgroups are presented in Table 4.

Table 4: Comparison of School Percent Proficient by CLP Subgroup in NC Elementary School Sample

Elementary Schools (Grade 5)	2003	2004	2005	2006	2007	2008
Reading Percent Proficient Measures						
All Sample Mean	86.5%	88.7%	89.0%	88.8%	88.8%	54.7%
TA-Reading Schools	68.3%	74.5%	77.7%	71.2%	81.8%	35.9%
MI-Reading Schools	76.3%	78.9%	77.1%	76.0%	80.1%	34.3%
NI-Reading Schools	75.3%	78.7%	78.7%	77.6%	77.4%	31.9%
Math Percent Proficient Measures						
All Sample Mean	90.5%	92.7%	90.1%	62.7%	66.1%	69.3%
TA-Math Schools	81.0%	82.6%	78.4%	41.2%	54.2%	62.8%
MI-Math Schools	84.2%	86.7%	79.3%	42.3%	49.6%	51.7%
NI-Math Schools	82.7%	85.0%	82.9%	39.9%	45.2%	47.1%

Note that the state's percent proficiency measure took a discrete jump downwards in 2008 for reading and in 2006 for math, from changes in how the state calculates the statistic. In reading, TAs showed a slight improvement in their percent proficiency measures over this period, relative to the MI and NI schools. In math, the difference in trajectories among the TA, MI, and NI groups had a considerably larger magnitude than that observed in reading proficiency scores. For instance, the TA schools perform at approximately similar levels in mathematics with MI and NI schools for the first four years, then emerge from the pack in the final two years of the sample. In the 2007–08 school year, TA schools are near the school mean of the sampled schools than the mean of either the MI or NI groups.

Conclusion and Discussion

This paper develops a new methodology to identify CLP and TA schools using longitudinal, student-level data on standardized test scores. By using student-level data, this model is immune to the criticisms of models that use school-level data to track school performance, particularly ignoring measurement error and demographic shifts in the underlying student population (Kane and Staiger, 2002). In addition, this model's use of both status- and growth-based performance measures in classifying CLP and TA schools ensures that CLP schools are genuinely performing poorly and TA schools are showing authentic improvements, rather than identifying schools based primarily on the socioeconomic inputs of the student body.

Applying this method to longitudinal data from three states, we find the incidence of turnaround overall to range from 13 to 31 percent across the six data samples we utilize here. Across all three states, turning around reading performance was much less common using our definition of turnaround. Schools that were identified as CLPs in both subjects (the weakest of all schools overall, and what we take to be our bottom five percent of schools) showed turnaround rates that were similar or better than the turnaround rates of CLP schools in only one subject; yet, the incidence of turnaround in both subjects in a school was very uncommon. Notably, schools in Texas were identified as turning around much more frequently than in either Florida or North Carolina, though we cannot rule out that this differential may be due to testing regimes.

The issue of measuring low performance and turnaround is important and timely, given the U.S. Department of Education's recent interest in making heavy investment in such efforts. The methodology presented here provides states and policymakers with a framework to model school performance and identify the target schools in a systematic way without all of the confounding errors implicit in school-level performance measures.

A natural direction for future research based on our analysis is to next investigate the differences in policies, practices, and programs in low-performing schools that may be associated with turnaround across these schools. These questions are addressed in related work from this project, which conduct principal surveys, analyze administrative data, and conduct site visits in both TA and non-TA CLP schools identified in the current study.[13] The collective findings of these studies suggest there are many unanswered questions about how we can most efficiently replicate turnaround in low-performing schools, and this area is ripe for future research.

[13] See Herman and Huberman (2012), Hansen (2012b), and Turnbull and Arcaira (2012).

References

Aladjem, D. K., Birman, B. F., Orland, M., Harr-Robins, J., Heredia, A., Parrish, T. B., et al. (2010). *Achieving dramatic school improvement: An exploratory study.* Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. Retrieved December 5, 2011, from http://www.wested.org/online_pubs/dramatic-improvement-report.pdf

Hansen, M. (2012a). Key Issues in Empirically Identifying Chronically Low-Performing and Turnaround Schools. *Journal of Education for Students Placed at Risk, 17* (1), 55-69.

Hansen, M. (2012b). Investigating the Role of Human Resources in School Turnaround: Evidence from Two States. Paper Presented at the 2012 Fall Research Conference of the Society for Research on Educational Effectiveness.

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., et al. (2008). *Turning around chronically low-performing schools: A practice guide* (NCEE 2008-4020). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved December 5, 2011, from http://ies.ed.gov/ncee/wwc/pdf/practice_guides/Turnaround_pg_04181.pdf

Herman, R. & Huberman, M. (2012). Differences in the Policies, Programs, and Practices (PPPs) and Combination of PPPs across Turnaround, Moderately Improving, and Not Improving Schools. Paper Presented at the 2012 Fall Research Conference of the Society for Research on Educational Effectiveness.

Kane, T. J., & Staiger, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *The Journal of Economic Perspectives, 16*(4), 91–114. doi:10.1257/089533002320950993

Koedel, C., & Betts, J. R. (2008). *Test score ceiling effects and value-added measures of school quality.* Paper presented at the 2008 Annual Meetings of the American Statistical Association.

Kutash, J., Nico, E., Gorin, E., Rahmatullah, S., & Tallant, K. (2010). *The school turnaround field guide.* Boston: FSG Social Impact Advisors. Retrieved December 5, 2011, from http://www.wallacefoundation.org/KnowledgeCenter/KnowledgeTopics/CurrentAreasofFocus/EducationLeadership/Documents/school-turnaround-field-guide.pdf

Linn, R. L., & Haug, C. (2002). Stability of school-building accountability scores and gains. *Educational Evaluation and Policy Analysis*, *24*(1), 29–36.

Malen, B. & Rice, J. K. (2009). "School Reconstitution and School Improvement: Theory and Evidence." in *Handbook of Education Policy Research*, ed. G. Sykes, B. Schneider, & D. N. Plank. (New York, NY: Routledge).

Meyer, R. H. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review 16*(3), 283–301. doi:10.1016/S0272-7757(96)00081-7

Meyers, C., Lindsay, J., Condon, C., & Wan,Y. (2012). A Statistical Approach to Identifying Schools Demonstrating Substantial Improvement in Student Learning. *Journal of Education for Students Placed at Risk, 17* (1), 70-91.

Raudenbush, S. W. (2004). *Schooling, statistics, and poverty: Can we measure school improvement?* (William H. Angoff Memorial Lecture Series). Princeton, NJ: Educational Testing Service. Retrieved March 14, 2012, from www.ets.org/Media/Education_Topics/pdf/angoff9.pdf

Turnbull, B. & Arcaira, E. (2012). Implementation of Turnaround Strategies in Chronically Low-Performing Schools. Paper Presented at the 2012 Fall Research Conference of the Society for Research on Educational Effectiveness.

Ushomirsky, N., & Hall, D. (2010). *Stuck Schools: A framework for identifying schools where students need change—now!* Washington, DC: The Education Trust.

Appendix

I. Data and Sample Selection

This analysis draws on longitudinal data from three states (Florida, North Carolina, and Texas) to identify CLP schools and then sub-divide these into TA, MI, and NI groups. When constructing the data, we ensured uniformity of the three samples to facilitate comparison of the results across states. Specifically, the data samples are similar in three primary ways: they span the same time frame, they span the same grade levels, and they use student achievement on the state's standardized test scores as the primary outcome measure. Though each state uses different tests and vary in their scaling, we transform all scaled test scores into Normal Curve Equivalents (NCEs) to standardize outcomes, providing a measure of student performance relative to other students at the same grade level and year.[14] Below, we describe some specifics of the data in each state further, and then describe the sample selection process.

Florida

The administrative data from Florida is provided from the Florida Department of Education K-20 Education Data Warehouse. Students in Florida are again linked to schools and teachers over time. The Florida Comprehensive Assessment Test (FCAT) was first administered in 1998 and, beginning in the 2001-02 school year, Florida has administered the FCAT test to all students in grades 3-10 for accountability purposes. The FCAT is a criterion-referenced, vertically aligned test. The Florida data contain approximately 200,000 unique students per grade over the six-year period in the study.

North Carolina

The North Carolina Department of Instruction collects detailed data on students and schools that are compiled annually by the North Carolina Education Research Data Center at Duke University. Each student enrolled in North Carolina public schools is assigned a unique randomized identifier that allows researchers to track individual students over time. The standardized tests in North Carolina are criterion-referenced and vertically aligned so that the scaling remains constant across grades. [15] On average, North Carolina has slightly more than 100,000 students per grade per year.

Texas

The administrative data from Texas is warehoused at three separate education research centers in the state; our access to the data was provided through the Texas Schools Project housed at the University of Texas at Dallas. Students in Texas are likewise linked to schools and teachers over time. Beginning in the 2002-03 school year, Texas has administered the Texas Assessment of Knowledge and Skills (TAKS) to all students in grades 3-8 for accountability purposes, which is a criterion-referenced, vertically aligned test.[16] The Texas data observed approximately 300,000 unique students per grade in the 2002-03 school year, and this number grew to almost 350,000 by the 2007-08 school year.

---

[14] Results based on NCEs have been shown to be very highly consistent with results based on vertically equated scale scores (Goldschmidt, Choi, Martinez, & Novak, under review; Choi, 2005). Note that our use of NCEs implies that all performance improvements are not absolute improvements but are relative to the sample of all students (and by extension, all schools).

[15] Each time a new version of an end-of-grade subject test was introduced in North Carolina, the scaled scores were all increased by a constant. During the time span of the sample used in the study, North Carolina introduced new forms of the reading test in 2004 and 2008 and introduced a new version of the mathematics test in 2006.

[16] While the TAKS documentation claims that the test is vertically aligned across grades, we found the empirical distribution of test scores in each grade and year showed a strong negative skew, indicating consequential ceiling effects on the test. Koedel and Betts (2008) analyze the effects of a test ceiling on school value-added measures and

Sample Selection

Using the data from these three states, we isolate two samples of schools in each state: one for elementary schools and one for middle schools. However, early in the sample selection process for this study, we confronted a common misconception concerning the identity of schools themselves. Commonly, schools are construed as static structures that do not change over time, but this construct does not reflect the reality of the dynamic schools we observe in the data. For example, within the time span of the data, we observe schools both opening and closing, as well as beginning or discontinuing service to a particular grade. Because of these dynamic changes, we cannot simply use all data from any school for this analysis.

We make three restrictions to the universe of schools for inclusion in the study. First, to ensure that the school structure is consistent with our estimation approach, we require that a school report test outcomes for all the grades of interest (grades 3–5 for elementary schools, grades 6–8 for middle schools). Second, to ensure that we have sufficient data over a long period on which to judge improvements in performance, we require that a school serve all these grades for all six school years in the analysis time span. Finally, to ensure that schools are internally stable over time, we require that each school have at least 50 percent of its students re-enroll in the following year for each of the five observed year-to-year transitions in the data. This final restriction is intended to filter out schools that may have undergone a high level of internal change (e.g., changes to school catchment) that could potentially coincide with changes in the school's overall performance; such changes are beyond the scope of the policies, programs, and practices analyzed in this study. Though not explicitly stated above, our restrictions require a school to maintain the same organizational identification and affiliation over time, omitting schools that have undergone any institutional change, such as a redistricting.[17]

Appendix Tables 1 and 2 present descriptive means from the Florida datasets. The first two columns of the tables are from the 2002–03 and 2006–07 school years, respectively, reporting on the universe of schools in the data that serve any of the grades of interest during that school year (grades 3-5 for the elementary sample, 6-8 for the middle sample). In the third column we report the same statistics (from the 2006–07 school year) for the sample of schools that we used in the study. The first four rows report demographic means of the students in the data, followed by the next section that reports descriptive variables on the schools themselves.

Comparing the demographic variables across the universe of data with those in the sample of schools, we see the sample is reasonably consistent with the characteristics of the larger student population generally. Moving to the variables on the schools' organization, however, we see differences between the sample and the universe of data. For instance, 3.8 percent of the schools serving any of grades 3–5 are charter schools in the 2007 universe of data, but these account for 2.4 percent of our

---

find even relatively severe test ceilings have a reasonably small effect on value-added measures. Particularly because we are interested in school performance at the low end of the distribution, where few students were subject to the test ceiling, our analysis is likely minimally influenced by this ceiling.

[17] In the NC elementary data, for example, two schools underwent a change in organizational identification (e.g., school closing and re-opening under different name, two districts merging all schools into a single district, etc.) within the time period and retained a sufficient level of their student body to pass the threshold on internal stability. To avoid the possibility that these schools may have turned around as a result of the institutional change (e.g., new staff, new district oversight, etc.), they were excluded. A particular case of schools undergoing an institutional change may be of particular policy relevance: when failing schools convert to charter schools. In the North Carolina data, we observed a handful of elementary schools that made such a change, but this conversion to a charter was also coincident with a high level of internal change in the student body, precluding their inclusion in the sample.

sample.[18] We also see that the K–5 structure is the most common for elementary grades in Florida, comprising over 85 percent of the schools in the universe of data and over 88 percent of our final sample.[19, 20]

The next two measures reflect school stability. The first reports the number of schools where at least 50 percent of their students (grades 3 and 4) return in the subsequent year, reflecting internal stability of the student body. Schools with such high student turnover are explicitly removed from the sample. The second measure is the count of schools that are not observed in the subsequent year of data (presumably indicating that the school is either temporarily or permanently closing). Such external changes also exclude schools from the original sample. The final line reports the number of unique students in these schools, showing over 86 percent of the unique students in the elementary data are retained in the sample, while approximately 78 percent of the observations are retained in the middle school sample.

Appendix Tables 3 and 4 present descriptive means from the North Carolina datasets. As with the Florida data above, there are no consequential differences between the sample and the population of students. The K-5 and 6-8 organizational structures were not as frequent as that observed in Florida, and we accordingly lose more schools from the sample as a result. Approximately two-thirds of all students in 2007 are retained in the sample in both elementary and middle schools.

Appendix Tables 5 and 6 present descriptive means from the Texas datasets. Here, approximately two-thirds of all students in 2007 are retained in the sample in the elementary sample and roughly half of all students are retained in the middle school sample (grade 7-8 middle schools are quite common in TX, but are omitted from the sample).

[18] As is apparent by comparing Column 1 with Column 2 in all tables, charter schools grew considerably in all three states over this period. Charter schools commonly increase their grade range as cohorts of students progress through these grades (e.g. in year 1, the school serves grades K-2, in year 2 it serves grades K-3, etc.). Because we required schools to serve all grades for all years, still-expanding charters were excluded from the analysis.

[19] This number includes schools that also serve pre-kindergarten.

[20] Note that in the first two columns, the sum of these K–5, K–8, and "3–5 with alternative structure" percentages do not equal 100%, whereas the final column does sum to 100%; this is because the first two columns report summary statistics on schools that serve *any* (but not necessarily *all*) of the elementary grades of interest. Approximately 10 percent of the schools in columns 1 or 2 serve some but not all of the grades (for instance, a school serving K–4 only); these schools are excluded from the final sample and are not counted in the three categories shown in the table.

Appendix Table 1: Descriptive Statistics of Florida Elementary School Sample

Descriptive Statistics of FL Elementary Schools			
	2003 Universe of Data	2007 Universe of Data	2007 Sample
Demographic Variables			
Percentage female	48.7%	48.5%	48.6%
Percentage White	49.5%	44.5%	44.4%
Percentage LEP	8.0%	7.9%	7.8%
Percentage ever eligible for free/reduced price lunch	62.0%	63.5%	64.0%
Institutional Variables			
Percentage charter	2.1%	3.8%	2.4%
Percentage K-5	86.8%	85.5%	88.1%
Percentage K-8	2.4%	3.4%	2.5%
Percentage serving 3-5 with alternative structure	9.7%	9.9%	9.4%
Number of schools with forward percentage <50%	41*	37**	0
Number of schools that close in the following year	43	80	0
Number of unique students in grades 3, 4, or 5	574,125	565,596	490,237
Total number of schools	1,906	2,148	1,599
*out of 1750 schools (serving all of grades 3-5 at minimum in 2003 and 2004)			
**out of 1938 schools (serving all of grades 3-5 at minimum in 2007 and 2008)			
Note: A school with an "alternative structure" serves all grades 3-5, but is not K-5 or K-8, and may potentially serve grades outside the 3-5 range. The forward percentage is calculated as the number of students in grades 4 and 5 in school S in year t who were in grades 3 and 4 in S in year t-1 out of the total number of students in grades 3 and 4 in S in year t-1.			

Appendix Table 2: Descriptive Statistics of Florida Middle School Sample

Descriptive Statistics of FL Middle Schools			
	2003 Universe of Data	2007 Universe of Data	2007 Sample
Demographic Variables			
Percentage female	48.9%	48.2%	48.8%
Percentage White	50.3%	46.7%	47.1%
Percentage LEP	5.7%	5.0%	5.2%
Percentage ever eligible for free/reduced price lunch	51.5%	65.1%	65.2%
Institutional Variables			
Percentage charter	2.3%	4.0%	2.3%
Percentage K-8	2.5%	4.6%	3.5%
Percentage 6-8	82.0%	80.6%	89.9%
Percentage serving 6-8 with alternative structure	10.0%	8.4%	6.5%
Number of schools with forward percentage <50%	179*	152**	0
Number of schools that close in the following year	59	126	0
Number of unique students in grades 6, 7, or 8	583,465	565,813	443,913
Total number of schools	1,182	1,426	535
*out of 782 schools (serving all of grades 6-8 at minimum in 2003 and 2004)			
**out of 963 schools (serving all of grades 6-8 at minimum in 2007 and 2008)			
Note: A school with an "alternative structure" serves all grades 6-8, but is not 6-8 or K-8, and serve grades outside the 6-8 range. The forward percentage is calculated as the number of students in grades 7 and 8 in school S in year t who were in grades 7 and 8 in school S in year t who were in grades 6 and 7 in S in year t-1 out of the total number of students in grades 6 and 7 in S in year t-1.			

Appendix Table 3: Descriptive Statistics of North Carolina Elementary School Sample

Descriptive Statistics of NC Elementary Schools			
	2003 Universe of Data	2007 Universe of Data	2007 Sample
Demographic Variables			
Percentage female	49.1%	49.4%	49.4%
Percentage White	59.2%	56.3%	56.8%
Percentage LEP	4.1%	9.4%	9.0%
Percentage ever eligible for free/reduced price lunch	49.6%	53.7%	53.8%
Institutional Variables			
Percentage charter	4.9%	7.1%	2.1%
Percentage K-5	71.4%	70.2%	82.8%
Percentage K-8	7.3%	9.1%	8.5%
Percentage serving 3-5 with alternative structure	10.8%	10.9%	8.7%
Number of schools with forward percentage <50%	30*	67**	0
Number of schools that close in the following year	17	18	0
Number of unique students in grades 3, 4, or 5	369,647	387,346	262,653
Total number of schools	1,413	1,529	1, 095

*out of 1,252 schools (serving all of grades 3-5 at minimum in 2003 and 2004)

Note: A school with an "alternative structure" serves all grades 3-5, but is not K-5 or K-8, and may potentially serve grades outside the 3-5 range. The forward percentage is calculated as the number of students in grades 4 and 5 in school S in year t who were in grades 3 and 4 in S in year t-1 out of the total number of students in grades 3 and 4 in S in year t-1.

Appendix Table 4: Descriptive Statistics of North Carolina Middle School Sample

Descriptive Statistics of NC Middle Schools			
	2003 Universe of Data	2007 Universe of Data	2007 Sample
Demographic Variables			
Percentage female	49.1%	49.0%	49.2%
Percentage White	60.3%	56.2%	56.8%
Percentage LEP	2.6%	6.9%	6.7%
Percentage ever eligible for free/reduced price lunch	45.5%	52.1%	51.7%
Institutional Variables			
Percentage charter	8.5%	13.3%	5.6%
Percentage K-8	13.7%	16.8%	17.1%
Percentage 6-8	48.4%	48.5%	71.0%
Percentage serving 6-8 with alternative structure	20.5%	17.6%	11.5%
Number of schools with forward percentage <50%	50*	58**	0
Number of schools that close in the following year	5	11	0
Number of unique students in grades 6, 7, or 8	358,438	389,649	273,411
Total number of schools	752	829	504

*out of 616 schools (serving all of grades 6-8 at minimum in 2003 and 2004)

Note: A school with an "alternative structure" serves all grades 6-8, but is not 6-8 or K-8, and serve grades outside the 6-8 range. The forward percentage is calculated as the number of students in grades 7 and 8 in school S in year t who were in grades 6 and 7 in S in year t-1 out of the total number of students in grades 6 and 7 in S in year t-1.

Appendix Table 5: Descriptive Statistics of Texas Elementary School Sample

Descriptive Statistics of TX Elementary Schools			
	2003 Universe of Data	2007 Universe of Data	2007 Sample
Demographic Variables			
Percentage female	49.9%	49.8%	49.8%
Percentage White	39.5%	35.2%	31.9%
Percentage LEP	4.1%	9.4%	9.0%
Percentage ever eligible for free/reduced price lunch	54.5%	57.0%	59.2%
Institutional Variables			
Percentage charter	3.1%	4.7%	1.8%
Percentage K-5	47.2%	51.6%	67.8%
Percentage K-8	2.5%	2.6%	2.9%
Percentage serving 3-5 with alternative structure	27.5%	22.8%	29.3%
Number of schools with forward percentage <50%	187*	144*	0
Number of schools that close in the following year	77	67	0
Number of unique students in grades 3, 4, or 5	874,904	967,619	604,783
Total number of schools	4,146	4,386	2,661
*Out of schools serving grades 3-5 in 2003 and 2004			
Note: A school with an "alternative structure" serves all grades 3-5, but is not K-5 or K-8, and may potentially serve grades outside the 3-5 range. The forward percentage is calculated as the number of students in grades 4 and 5 in school S in year t who were in grades 3 and 4 in S in year t-1 out of the total number of students in grades 3 and 4 in S in year t-1.			

Appendix Table 6: Descriptive Statistics of Texas Middle School Sample

Descriptive Statistics of TX Middle Schools			
	2003 Universe of Data	2007 Universe of Data	2007 Sample
Demographic Variables			
Percentage female	50.2%	49.8%	49.8%
Percentage White	43.1%	37.4%	35.5%
Percentage LEP	5.8%	13.5%	15.3%
Percentage ever eligible for free/reduced price lunch	47.9%	52.3%	53.6%
Institutional Variables			
Percentage charter	5.5%	7.9%	3.7%
Percentage K-8	3.4%	4.0%	5.9%
Percentage 6-8	28.9%	32.7%	71.4%
Percentage serving 6-8 with alternative structure	16.7%	14.8%	22.8%
Number of schools with forward percentage <50%	203*	128*	0
Number of schools that close in the following year	116	130	0
Number of unique students in grades 6, 7, or 8	855,386	918,936	492,962
Total number of schools	2,871	2,826	1,023
*Out of schools serving grades 6-8 in 2003 and 2004			
Note: A school with an "alternative structure" serves all grades 6-8, but is not 6-8 or K-8, and serve grades outside the 6-8 range. The forward percentage is calculated as the number of students in grades 7 and 8 in school S in year t who were in grades 6 and 7 in S in year t-1 out of the total number of students in grades 6 and 7 in S in year t-1.			

II. Status and Growth Parameter Estimates

We estimated the three-level model presented in the text to estimate school-level status and growth parameters in both the pre- and post-periods. Identification as a CLP school was based on these estimated status and growth parameter values in the pre-period only. Note that these values (hence, a school's designation as a CLP) varied depending on the subject (reading or mathematics) and the cutpoint (2004, 2005, or 2006) used in the model. While the estimated parameter values varied across models, the correlation between these models was in fact quite high. Additionally, other correlations between these different parameters are worth highlighting, which we do here.

Appendix Tables 7 and 8 present correlation matrices for the pre-period status and growth parameters in each of the 6 different models (2 subjects x 3 cutpoints) we estimated in the Florida data. To aid in interpretation, the highest correlations are identified with red fill; mid-range correlations are orange; low correlations are yellow; negative correlations are light blue.

Two particular patterns of these correlation coefficients are worth highlighting. First, there are patterns of very high correlation (near perfect in many cases) within subject across all cutpoints, but moderately lower levels of correlation when comparing the relative performance of schools across subjects (the southwest quadrant of the matrix).

A second noteworthy pattern in the correlation matrices is that high and low correlations fluctuate along each diagonal (moving away from the main diagonal), indicating high correlations within outcome measurement (status or growth) models and low correlations (in all cases below 0.3) when measuring the correlation across these outcome measures.

Appendix Tables 9 and 10 present similar correlation matrices using the North Carolina samples. Generally speaking, the patterns are very similar to that observed in the Florida sample, with the exception that the correlations between status and growth parameter estimates in NC are low positive values overall, whereas they had very low negative values in many instances in the FL samples.

Appendix Tables 11 and 12 presents the results using the Texas samples. The primary difference in these correlation matrices (compared to the Florida and North Carolina results above) is the correlation between status and growth measures is not as consistent. This different pattern may be an artifact of the prominent test ceiling in the TAKS tests (see footnote 4 in Section I above for more detail). Also, note the moderate negative correlations between status and growth measures in the middle school reading test results—this negative correlation is consistent with the test ceiling, but it appears that the ceiling is particularly strong among these tests.

Across all three states, the lowest correlations observed are those for status and growth, suggesting that the schools low in one outcome may be observably different from schools low in the other dimension. In an attempt to understand how low-status schools differ from low-growth schools, we compared the socioeconomic characteristics of these schools side by side. Appendix Tables 13 through 18 presents the means (and standard deviations) of various demographic measures in schools that rank in the lowest 15 percent of schools based on status measures (columns 1 and 2) and schools that rank in the lowest 40 percent of schools based on growth measures (columns 3 and 4). Note from the text that the intersection of these two groups will constitute the CLP schools. These schools presented in these tables are included based on their pre-parameter estimates using the 2005 cutpoint.

The primary point of interest in these tables is that schools at the bottom of the distribution using status outcomes have significantly lower socioeconomic indicators than those identified in the bottom 40 percent using growth as the outcome measure. In particular, the low performers under the status models are composed of significantly fewer white students, fewer students with parents holding at least a college degree, and more students eligible for the free or reduced-price lunch program.[21] These findings indicate a school's status and growth are very different measures: almost any school can have low growth, but those with low status are schools that are most commonly characterized as disadvantaged.[22]

Low-growth schools are those that contribute the smallest amount of learning to each student (i.e., students learn less and fall behind relative to students in other schools), so these are the schools where intervention is necessary. However, low growth in and of itself is insufficient to warrant categorizing a school as CLP—there is little public interest in turning around a low-growth school in an affluent district. The main target group for the study is disadvantaged schools, thus low status was also used as one of the criteria in attaching the chronically low-performing label to a school.

In other words, the intersection of low-performing schools in status and growth comprised the group labeled CLP schools. The strategy of targeting schools based on low observed performance in achievement levels (status) and gains (growth) is consistent with the 'stuck schools' approach detailed in Ushomirsky and Hall (2010). While there are some key advantages to our data and modeling approach (i.e., the work of these authors relies on rankings of school-level proficiency rates whereas our approach uses an HLM model on student-level data and standardized test scores), combining low-status measures with low-growth measures provides some advantages in identifying low-performing schools.

[21] T-tests on the difference in these means between low-growth and low-status schools were uniformly significant (rejecting the null hypothesis in virtually every case). T-tests on differences in the percentage of students with limited English proficiency were split between marginally rejecting the hypothesis and marginally failing to reject.

[22] We isolated the bottom 30 percent of low-status schools and compared them against the bottom 30 percent of low-growth schools and saw results mirroring those presented above. Hence, the significant differences between the groups are not attributable to the size difference between the groups.

Appendix Table 7: Correlation Matrix for FL Elementary School Sample

Correlation Matrix on Pre-period Parameter Estimates (FL Elementary Schools)

		Reading 2004 Cutpoint		Reading 2005 Cutpoint		Reading 2006 Cutpoint		Math 2004 Cutpoint		Math 2005 Cutpoint		Math 2006 Cutpoint	
		Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth
Reading 2004	Status	1.00									$.75<\rho<1$		
Cutpoint	Growth	0.17	1.00								$.30<\rho<.75$		
Reading 2005	Status	1.00	0.20	1.00							$0<\rho<.30$		
Cutpoint	Growth	0.01	0.84	0.04	1.00						$\rho<0$		
Reading 2006	Status	0.99	0.22	1.00	0.06	1.00							
Cutpoint	Growth	0.00	0.74	0.03	0.94	0.06	1.00						
Math 2004	Status	0.92	0.23	0.92	0.09	0.92	0.09	1.00					
Cutpoint	Growth	-0.07	0.50	-0.05	0.49	-0.04	0.44	0.13	1.00				
Math 2005	Status	0.91	0.26	0.92	0.12	0.92	0.12	0.99	0.16	1.00			
Cutpoint	Growth	-0.10	0.45	-0.08	0.56	-0.06	0.54	0.10	0.89	0.14	1.00		
Math 2006	Status	0.91	0.27	0.92	0.13	0.93	0.13	0.98	0.15	0.99	0.14	1.00	
Cutpoint	Growth	-0.14	0.40	-0.12	0.53	-0.11	0.56	0.05	0.79	0.09	0.94	0.11	1.00

Correlation matrix reporting correlations between pre-period parameter estimates ranking schools under each specification. The sample is comprised of 1,599 unique elementary schools.

Appendix Table 8: Correlation Matrix for FL Middle School Sample

Correlation Matrix on Pre-period Parameter Estimates (FL Middle Schools)

		Reading 2004 Cutpoint		Reading 2005 Cutpoint		Reading 2006 Cutpoint		Math 2004 Cutpoint		Math 2005 Cutpoint		Math 2006 Cutpoint	
		Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth
Reading 2004	Status	1.00									$.75<\rho<1$		
Cutpoint	Growth	-0.08	1.00								$.30<\rho<.75$		
Reading 2005	Status	1.00	-0.07	1.00							$0<\rho<.30$		
Cutpoint	Growth	-0.11	0.90	-0.10	1.00						$\rho<0$		
Reading 2006	Status	1.00	-0.05	1.00	-0.07	1.00							
Cutpoint	Growth	-0.12	0.84	-0.12	0.96	-0.09	1.00						
Math 2004	Status	0.97	-0.04	0.97	-0.06	0.97	-0.07	1.00					
Cutpoint	Growth	0.11	0.53	0.12	0.51	0.13	0.47	0.19	1.00				
Math 2005	Status	0.97	-0.02	0.97	-0.04	0.97	-0.06	1.00	0.21	1.00			
Cutpoint	Growth	0.02	0.53	0.03	0.58	0.04	0.55	0.10	0.87	0.12	1.00		
Math 2006	Status	0.97	0.00	0.97	-0.03	0.97	-0.04	0.99	0.23	1.00	0.14	1.00	
Cutpoint	Growth	-0.02	0.48	-0.01	0.55	0.00	0.55	0.06	0.74	0.08	0.95	0.09	1.00

Correlation matrix reporting correlations between pre-period parameter estimates ranking schools under each specification. The sample is comprised of 535 unique middle schools.

Appendix Table 9: Correlation Matrix for NC Elementary School Sample

Correlation Matrix on Pre-period Parameter Estimates (NC Elementary Schools)

		Reading 2004 Cutpoint		Reading 2005 Cutpoint		Reading 2006 Cutpoint		Math 2004 Cutpoint		Math 2005 Cutpoint		Math 2006 Cutpoint	
		Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth
Reading 2004	Status	1.00									.75<ρ<1		
Cutpoint	Growth	0.20	1.00								.30<ρ<.75		
Reading 2005	Status	1.00	0.22	1.00							0<ρ<.30		
Cutpoint	Growth	0.21	0.87	0.23	1.00						ρ<0		
Reading 2006	Status	0.99	0.23	1.00	0.24	1.00							
Cutpoint	Growth	0.22	0.79	0.24	0.95	0.25	1.00						
Math 2004	Status	0.94	0.22	0.93	0.23	0.93	0.24	1.00					
Cutpoint	Growth	0.10	0.55	0.11	0.53	0.11	0.49	0.21	1.00				
Math 2005	Status	0.93	0.23	0.94	0.24	0.93	0.25	0.99	0.23	1.00			
Cutpoint	Growth	0.06	0.51	0.07	0.60	0.08	0.59	0.19	0.87	0.22	1.00		
Math 2006	Status	0.93	0.24	0.93	0.25	0.94	0.26	0.98	0.23	0.99	0.23	1.00	
Cutpoint	Growth	0.03	0.47	0.04	0.57	0.05	0.61	0.17	0.75	0.20	0.93	0.22	1.00

Correlation matrix reporting correlations between pre-period parameter estimates ranking schools under each specification. The sample is comprised of 1,095 unique elementary schools.

Appendix Table 10: Correlation Matrix for NC Middle School Sample

Correlation Matrix on Pre-period Parameter Estimates (NC Middle Schools)

		Reading 2004 Cutpoint		Reading 2005 Cutpoint		Reading 2006 Cutpoint		Math 2004 Cutpoint		Math 2005 Cutpoint		Math 2006 Cutpoint	
		Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth
Reading 2004	Status	1.00									.75<ρ<1		
Cutpoint	Growth	0.11	1.00								.30<ρ<.75		
Reading 2005	Status	1.00	0.12	1.00							0<ρ<.30		
Cutpoint	Growth	0.18	0.83	0.19	1.00						ρ<0		
Reading 2006	Status	0.99	0.13	1.00	0.20	1.00							
Cutpoint	Growth	0.17	0.71	0.17	0.95	0.18	1.00						
Math 2004	Status	0.91	0.10	0.91	0.15	0.90	0.13	1.00					
Cutpoint	Growth	0.00	0.34	0.01	0.23	0.01	0.18	0.14	1.00				
Math 2005	Status	0.92	0.11	0.92	0.17	0.91	0.15	0.99	0.15	1.00			
Cutpoint	Growth	0.08	0.33	0.09	0.37	0.09	0.36	0.20	0.87	0.23	1.00		
Math 2006	Status	0.92	0.12	0.92	0.18	0.92	0.16	0.98	0.13	0.99	0.23	1.00	
Cutpoint	Growth	0.12	0.32	0.13	0.41	0.13	0.44	0.22	0.70	0.25	0.92	0.27	1.00

Correlation matrix reporting correlations between pre-period parameter estimates ranking schools under each specification. The sample is comprised of 504 unique middle schools.

Appendix Table 11: Correlation Matrix for TX Elementary School Sample

Correlation Matrix on Pre-period Parameter Estimates (TX Elementary Schools)													
		Reading 2004 Cutpoint		Reading 2005 Cutpoint		Reading 2006 Cutpoint		Math 2004 Cutpoint		Math 2005 Cutpoint		Math 2006 Cutpoint	
		Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth
Reading 2004	Status	1.00									.75<ρ<1		
Cutpoint	Growth	0.20	1.00								.30<ρ<.75		
Reading 2005	Status	0.99	0.19	1.00							0<ρ<.30		
Cutpoint	Growth	0.42	0.81	0.43	1.00						ρ<0		
Reading 2006	Status	0.98	0.18	0.99	0.43	1.00							
Cutpoint	Growth	0.43	0.71	0.44	0.95	0.45	1.00						
Math 2004	Status	0.91	0.21	0.89	0.41	0.88	0.41	1.00					
Cutpoint	Growth	-0.04	0.57	-0.04	0.48	-0.04	0.43	0.09	1.00				
Math 2005	Status	0.91	0.21	0.91	0.43	0.91	0.43	0.98	0.10	1.00			
Cutpoint	Growth	0.07	0.49	0.07	0.58	0.08	0.56	0.17	0.86	0.21	1.00		
Math 2006	Status	0.91	0.21	0.91	0.44	0.92	0.45	0.96	0.10	0.99	0.22	1.00	
Cutpoint	Growth	0.04	0.43	0.05	0.55	0.06	0.59	0.14	0.74	0.17	0.93	0.20	1.00

Correlation matrix reporting correlations between pre-period parameter estimates ranking schools under each specification. The sample is comprised of 2,662 unique elementary schools.

Appendix Table 12: Correlation Matrix for TX Middle School Sample

Correlation Matrix on Pre-period Parameter Estimates (TX Middle Schools)													
		Reading 2004 Cutpoint		Reading 2005 Cutpoint		Reading 2006 Cutpoint		Math 2004 Cutpoint		Math 2005 Cutpoint		Math 2006 Cutpoint	
		Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth	Status	Growth
Reading 2004	Status	1.00									.75<ρ<1		
Cutpoint	Growth	-0.39	1.00								.30<ρ<.75		
Reading 2005	Status	0.99	-0.36	1.00							0<ρ<.30		
Cutpoint	Growth	-0.37	0.91	-0.33	1.00						ρ<0		
Reading 2006	Status	0.99	-0.35	1.00	-0.33	1.00							
Cutpoint	Growth	-0.40	0.85	-0.38	0.96	-0.37	1.00						
Math 2004	Status	0.89	-0.39	0.88	-0.39	0.87	-0.43	1.00					
Cutpoint	Growth	-0.02	0.40	0.00	0.38	0.01	0.34	0.01	1.00				
Math 2005	Status	0.89	-0.37	0.89	-0.37	0.88	-0.41	0.99	0.06	1.00			
Cutpoint	Growth	-0.03	0.38	-0.01	0.43	-0.01	0.40	0.03	0.88	0.10	1.00		
Math 2006	Status	0.89	-0.36	0.89	-0.36	0.89	-0.41	0.98	0.07	0.99	0.11	1.00	
Cutpoint	Growth	-0.06	0.37	-0.05	0.43	-0.04	0.44	0.02	0.75	0.07	0.93	0.09	1.00

Correlation matrix reporting correlations between pre-period parameter estimates ranking schools under each specification. The sample is comprised of 1,023 unique middle schools.

Appendix Table 13: Descriptive Characteristics for FL Elementary School Sample

Descriptive Characteristics of Low-Status vs. Low-Growth Elementary Schools				
	Status		Growth	
	Reading	Math	Reading	Math
Proportion of female students	0.483	0.483	0.487	0.486
	(0.500)	(0.500)	(0.500)	(0.500)
Proportion of white students	0.087	0.153	0.512	0.557
	(0.281)	(0.360)	(0.500)	(0.497)
Proportion of students with limited English proficiency	0.155	0.135	0.059	0.060
	(0.362)	(0.341)	(0.235)	(0.238)
Proportion of students eligible for free or reduced-price lunch program	0.948	0.919	0.617	0.591
	(0.221)	(0.273)	(0.486)	(0.492)

Note: Results represent means (and standard deviations in parentheses) of the variables in low-status or low-growth schools. Data are from elementary schools identified in the bottom 15 percent (status) and 40 percent (growth) in the given model using the 2005 cutpoint. There are 240 schools in each status specification and 640 schools in each growth specification. Percentages are computed over the 6 years of data in the analysis.

Appendix Table 14: Descriptive Characteristics for FL Middle School Sample

Descriptive Characteristics of Low-Status vs. Low-Growth Middle Schools				
	Status		Growth	
	Reading	Math	Reading	Math
Proportion of female students	0.481	0.483	0.490	0.491
	(0.500)	(0.500)	(0.500)	(0.500)
Proportion of white students	0.129	0.149	0.621	0.532
	(0.335)	(0.356)	(0.485)	(0.499)
Proportion of students with limited English proficiency	0.115	0.093	0.030	0.042
	(0.319)	(0.290)	(0.170)	(0.201)
Proportion of students eligible for free or reduced-price lunch program	0.909	0.892	0.558	0.605
	(0.287)	(0.311)	(0.497)	(0.489)

Note: Results represent means (and standard deviations in parentheses) of the variables in low-status or low-growth schools. Data are from middle schools identified in the bottom 15 percent (status) and 40 percent (growth) in the given model using the 2005 cutpoint. There are 81 schools in each stauts specification and 214 schools in each growth specification reported. Percentages are computed over the 6 years of data in the analysis.

Appendix Table 15: Descriptive Characteristics for NC Elementary School Sample

Descriptive Characteristics of Low-Status vs. Low-Growth Elementary Schools in North Carolina

	Status		Growth	
	Reading	Math	Reading	Math
Proportion of female students	0.490	0.490	0.490	0.491
	(0.500)	(0.500)	(0.500)	(0.500)
Proportion of white students	0.149	0.204	0.552	0.572
	(0.356)	(0.403)	(0.497)	(0.495)
Proportions of parents holding BA degree or higher	0.188	0.197	0.334	0.346
	(0.391)	(0.398)	(0.472)	(0.476)
Proportion of students with limited English proficiency	0.111	0.095	0.062	0.068
	(0.314)	(0.293)	(0.241)	(0.251)
Proportion of students ever eligible for free or reduced-price lunch program	0.864	0.839	0.570	0.552
	(0.343)	(0.368)	(0.495)	(0.497)

Note: Results represent means (and standard deviations in parentheses) of the variables in low-status or low-growth schools. Data are from elementary schools identified in the bottom 15 percent (status) and 40 percent (growth) in the given model using the 2005 cutpoint. There are 165 schools in each status specification and 438 schools in each growth specification. Percentages are computed over the 6 years of data in the analysis.

Appendix Table 16: Descriptive Characteristics for NC Middle School Sample

Descriptive Characteristics of Low-Status vs. Low-Growth Middle Schools in North Carolina

	Status		Growth	
	Reading	Math	Reading	Math
Proportion of female students	0.490	0.493	0.491	0.491
	(0.500)	(0.500)	(0.500)	(0.500)
Proportion of white students	0.162	0.169	0.489	0.453
	(0.369)	(0.375)	(0.500)	(0.498)
Proportions of parents holding BA degree or higher	0.212	0.220	0.317	0.321
	(0.409)	(0.414)	(0.465)	(0.467)
Proportion of students with limited English proficiency	0.078	0.073	0.040	0.047
	(0.269)	(0.259)	(0.196)	(0.211)
Proportion of students ever eligible for free or reduced-price lunch program	0.814	0.792	0.512	0.526
	(0.389)	(0.406)	(0.500)	(0.499)

Note: Results represent means (and standard deviations in parentheses) of the variables in low-status or low-growth schools. Data are from middle schools identified in the bottom 15 percent (status) and 40 percent (growth) in the given model using the 2005 cutpoint. There are 76 schools in each stauts specification and 202 schools in each growth specification reported. Percentages are computed over the 6 years of data in the analysis.

Appendix Table 17: Descriptive Characteristics for TX Elementary School Sample

Descriptive Characteristics of Low-Status vs. Low-Growth Elementary Schools				
	Status		Growth	
	Reading	Math	Reading	Math
Proportion of female students	0.499	0.499	0.498	0.498
	(0.500)	(0.500)	(0.500)	(0.500)
Proportion of white students	0.037	0.074	0.231	0.321
	(0.190)	(0.261)	(0.421)	(0.467)
Proportion of students with limited English proficiency	0.485	0.414	0.299	0.250
	(0.500)	(0.493)	(0.458)	(0.433)
Proportion of students eligible for free or reduced-price lunch program or otherwise disadvantaged	0.909	0.876	0.729	0.615
	(0.287)	(0.330)	(0.445)	(0.487)
Proportion of students in Title I schools or otherwise receiving Title I assistance	0.987	0.983	0.880	0.745
	(0.113)	(0.130)	(0.325)	(0.436)

Note: Results represent means (and standard deviations in parentheses) of the variables in low-status or low-growth schools. Data are from elementary schools identified in the bottom 15 percent (status) and 40 percent (growth) in the given model using the 2005 cutpoint. There are 400 schools in each status specification and 1,065 schools in each growth specification. Percentages are computed over the 6 years of data in the analysis.

Appendix Table 18: Descriptive Characteristics for TX Middle School Sample

Descriptive Characteristics of Low-Status vs. Low-Growth Middle Schools				
	Status		Growth	
	Reading	Math	Reading	Math
Proportion of female students	0.417	0.416	0.415	0.419
	(0.493)	(0.493)	(0.493)	(0.493)
Proportion of white students	0.047	0.055	0.412	0.311
	(0.211)	(0.227)	(0.492)	(0.463)
Proportion of students with limited English proficiency	0.392	0.379	0.238	0.263
	(0.488)	(0.485)	(0.426)	(0.440)
Proportion of students eligible for free or reduced-price lunch program or otherwise disadvantaged	0.888	0.874	0.490	0.604
	(0.315)	(0.331)	(0.500)	(0.489)
Proportion of students in Title I schools or otherwise receiving Title I assistance	0.779	0.771	0.310	0.477
	(0.415)	(0.420)	(0.462)	(0.499)

Note: Results represent means (and standard deviations in parentheses) of the variables in low-status or low-growth schools. Data are from elementary schools identified in the bottom 15 percent (status) and 40 percent (growth) in the given model using the 2005 cutpoint. There are 154 schools in each status specification and 410 schools in each growth specification. Percentages are computed over the 6 years of data in the analysis.