Abstract Title Page
Not included in page count.

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

Authors and Affiliations: Michael Hansen, American Institutes for Research / CALDER
mhansen@air.org

Kilchan Choi, American Institutes for Research / CALDER
kchoi@air.org

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual 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. The Federal School Improvement Grants (SIG) and Race to the Top (RTT) initiatives, 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. 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.

School-based accountability measures, which have dominated the relatively small literature on empirically identifying low-performing and turnaround schools, are problematic on three points:

1. Percentage measures (e.g., percent proficiency, graduation rates, etc.) ignore important variations in tested outcomes that occur on either side of the cutpoint.
2. School-based measures ignore the implied measurement error when calculating across schools of differing sizes.
3. School-based measures ignore compositional changes in the student body from cohort to cohort (Kane & Staiger, 2002).

One may reasonably expect these problems to be exacerbated 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. Therefore, 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.

A second critical issue in identifying CLP and TA schools is that accountability measures are generally status measures, which confound pre-existing differences among students with differences in school quality. Growth measures 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 are less correlated over time (Kane & Staiger, 2002; Linn & Haug, 2002). The implication for TALPS is that a focus on status measures would underidentify TA schools or identify improvements only several years after the turnaround actually occurred. A focus on both status and growth would be more likely to capture movement toward improved outcomes. Kane and Staiger (2002) contend that using student-level growth measures can overcome the criticisms of school-based accountability measures raised above: growth measures are generally 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.

Purpose / Objective / Research Question / Focus of Study:
Description of the focus of the research.

The primary research question for this substudy is as follows: What proportion of CLP schools display specific performance trajectories over time? In particular, what proportion of CLP schools display: quick, dramatic improvement (TA schools); weak net improvement (MI schools); or negatively sloped performance trajectories (NI schools)?

Setting:
Description of the research location.
(May not be applicable for research methods or panel submissions)

Administrative data from three states on student test scores were used (Florida, North Carolina, and Texas), from which two separate data samples were created in each state for elementary and middle schools. The data spanned six school years (2002-02 through 2007-08).

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.
(May not be applicable for research methods submissions)

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. Table 1 presents descriptive statistics of the universe of data and the schools used in the Florida elementary school sample, one of six data samples used in the study.
<Insert Table 1 Here>

Intervention / Program / Practice:
Description of the intervention, program, or practice, including details of administration and duration.
(May not be applicable for research methods submissions)

Not applicable.

Research Design:
Description of the research design.

Through a series of simulations and iterative data analysis using the data samples from three states, we develop a set of rules to empirically identify CLP schools, and then classify schools as TA/MI/NI based on subsequent performance. 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 a representative 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 cut point.

Schools were classified as chronically low performing 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 chronically low-performing school.

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 policy-relevant 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) were statistically significant but did not meet the policy-relevant thresholds, the school was categorized as a MI school. Schools that exhibited no statistically significant improvement in either parameter estimate were labeled NI schools.

Statistical, Measurement, or Econometric Model:

Description of the proposed new methods or novel applications of existing methods.

(May not be applicable to submissions in sections other than research methods)

Not applicable.

Usefulness / Applicability of Method:
Demonstration of the usefulness of the proposed methods using hypothetical or real data.
(May not be applicable submissions in sections other than research methods)

The methods developed in this study are broadly applicable, and are helpful for states in prioritizing intervention efforts to turnaround low-performing schools, and in recognizing signs of improvement among CLP schools.

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.
(May not be applicable for research methods submissions)

No supplemental data collection.

Findings / Results:
Description of the main findings with specific details.
(May not be applicable for research methods submissions)

Overall findings for the paper are summarized in Table 2. Supplementary evidence comparing performance trajectories of CLP schools by TA, MI, and NI status are presented in Figures 1, 2, and 3; these and other investigations suggest the identification method developed here is valid and viable.
<Insert Table 2, Figures 1, 2, and 3 Here>

Conclusions:
Description of conclusions, recommendations, and limitations based on findings.

This paper summarizes the lessons learned from this exercise of empirically identifying CLP schools and binning them into performance categories based on their trajectories. This paper provides guidance for others charged with a similar task. Specifically, we learned the critical importance of using student-level data (rather than school-level aggregate measures), using growth-based measures in conjunction with status-based performance metrics, the stability of these performance metrics over time, and how to empirically recognize turnaround in schools as it occurs.

Also, we learned that 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. Similar rates were observed in North Carolina—13% and 16%, respectively; and even higher in Texas—29% and 31%, respectively.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

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

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

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

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

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

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

Appendix B. Tables and Figures

Not included in page count.

Table 1. Descriptive Statistics of FL Elementary Schools

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.			

Table 2. Summary Results

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.					

Figure 1.

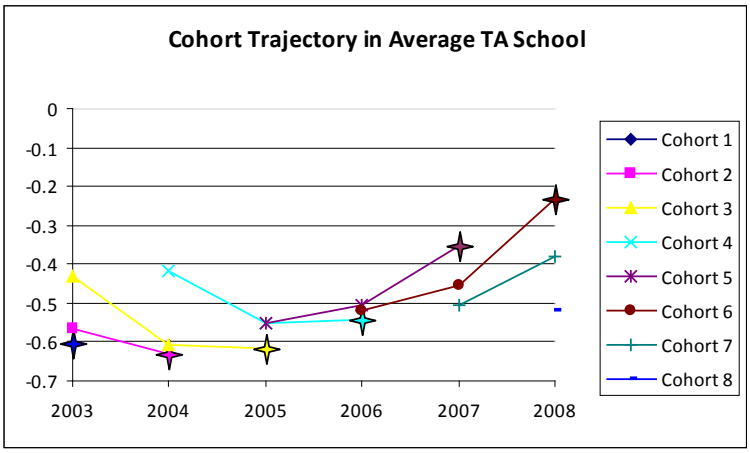

Figure 2.

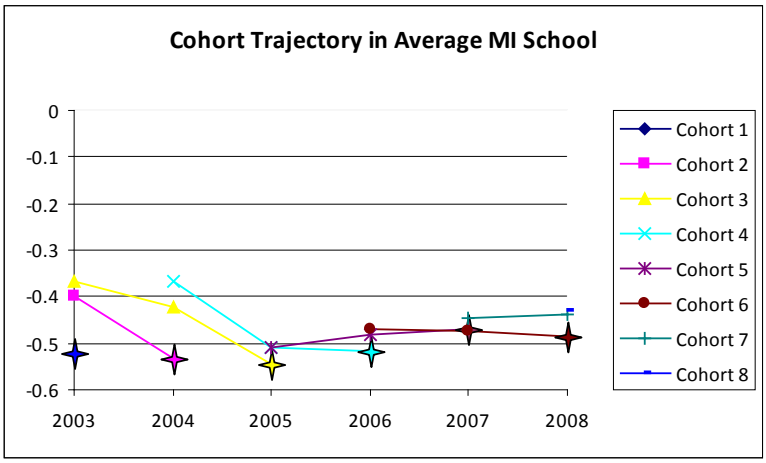

Figure 3.