ACT Research Report Series 2016-11

Comparisons of Student Achievement Levels By District Performance and Poverty

Chrys Dougherty, PhD
Teresa Shaw

ACT®

Chrys Dougherty is a principal research scientist in Statistical and Applied Research at ACT.

Teresa Shaw is a statistical project consultant with Pastor, Behling & Wheeler, an environmental consulting firm. From 2007 to 2015, she worked as a senior software engineer and statistical analyst at ACT and researched performance growth and value-added modeling for schools and districts.

Acknowledgments

The authors would like to thank Jeff Allen, Kurt Burkum, Ty Cruce, Justine Radunzel, Edgar Sanchez, and Richard Sawyer for their helpful comments on drafts of the report. In addition, we would like to thank the staff of the Arkansas Department of Education for making available the data that made this report possible.

Contents

Abstract 1

Introduction 1

Data 3

1. Creation of Student Cohorts and Calculation of Student-Level Statistics 4
2. District-Level Statistics and Inclusion of Districts in the Analysis. 6
3. Combining Student- and District-Level Data 8

Methods 8

1. Statistical Models Used to Create District Performance Statistics 8
2. Classification of Districts into Relative Performance Categories 10
3. District-Level Student Achievement Statistics 11
4. Aggregation of Performance and Achievement Statistics by District Performance and Poverty. 11

Limitations 11

Results 12

Conclusion 22

References 24

Appendix A 27

Appendix B 30

Appendix C 35

Appendix D 51

Abstract

This report looks at student achievement levels in Arkansas school districts disaggregated by district poverty and by the district's performance relative to other districts. We estimated district performance statistics by subject and grade level (4, 8, and 11–12) for longitudinal student cohorts, using statistical models that adjusted for district demographics and the percentage of students in the analysis. We found that differences in these performance statistics across districts at each poverty level were large enough to be of practical importance to educators and policymakers. Variation in district performance statistics was sometimes greater in lower-poverty districts, but this result was not consistent across subjects and grade levels.

We also calculated unadjusted descriptive student achievement statistics—average scores and percentages of students On Track or Far Off Track—for districts classified as above or below average based on the district performance statistics. Differences in these unadjusted statistics were also large enough to be of practical importance. However, even in above-average districts, the majority of students in moderate- and high-poverty districts did not reach On Track benchmarks for college readiness in mathematics, reading, and science. This is reason to pay increased attention to promising but often underemphasized approaches to improving student outcomes.

Introduction

A substantial body of literature focuses on the role of school districts in supporting improvement in student learning (Hightower, Knapp, Marsh, & McLaughlin, 2002; Marsh *et al*., 2005; Supovitz, 2006; Mac Iver & Farley-Ripple, 2008; ACT, 2012a; Daly & Finnigan, 2016). As such, school districts have been a significant focus of education reform efforts (Whitehurst, Chingos, & Gallaher, 2013; Chingos, Whitehurst, & Gallaher, 2013) and public recognition (Broad Foundation, 2014).

School districts are potentially important because they are often the administrative unit closest to students that can oversee improvement strategies spanning preschool through grade 12 (ACT, 2012a; Dougherty, 2016). For example, working closely with teachers and school leaders, district leaders can work to ensure students receive a content-rich curriculum in each subject aligned across elementary, middle, and high school levels; establish assessment and data systems to monitor student progress and follow students as they change schools; promote educators' use of those systems (Dougherty, 2015a, 2015b); ensure that time is set aside in every school for teachers to collaborate; develop a coaching system for teachers; and lead efforts to involve parents and community leaders.

In theory, we would expect differences in educator practices across districts to account for a substantial share of cross-district differences in student outcome indicators such as test scores, high school graduation, and college enrollment. In practice, the relationship between practices and student outcomes is difficult to investigate. First, good information on what practices are being implemented how well in which districts is generally missing.[1] Second, practices are not generally introduced separately in randomized trials, so that observed

[1] A separate ACT report (Dougherty, 2016) discusses how school system leaders might keep track of the implementation of specific practices in their districts.

relationships between changes in practices and changes in performance are correlational, with causal relationships more difficult to determine. Third, practices may be implemented inconsistently within districts, so that the practice is found only in some schools and classrooms and the overall impact of the practice on district-wide student outcomes is diluted.

As a fallback, it is worth asking whether differences in student learning across districts, as assessed by test scores, are large enough to matter when districts with similar student populations are compared. Does there appear to be a large enough "district effect" to make it worthwhile to investigate the cause? One approach is to look at the share of total variance in student test scores accounted for by differences between districts. This approach makes school districts appear to be unimportant. For example, Whitehurst, Chingos, & Gallaher (2013) found that differences across districts accounted for only about 1–2% of the total variance in student scores, depending on the subject and the state whose data were analyzed.

However, it is possible for score differences between the highest- and lowest-performing districts with similar student populations to be large enough to matter, even if districts overall account for only a small share of the total variance in individual scores. Looking into the issue using student data from Florida and North Carolina, Whitehurst, Chingos, & Gallaher (2013) found that differences across districts were indeed large enough to matter: a one standard-deviation difference in district performance statistics amounted to a difference in test scores similar to the gains associated with about 7–12 weeks of instruction. Measured in standard deviation units, their estimates were roughly similar to those found by Dougherty & Shaw (2016) using Arkansas data.[2]

In our previous report (Dougherty & Shaw, 2016), we estimated district performance statistics using four sets of statistical models that controlled for different combinations (depending on the model) of student and district demographics and prior achievement scores. For each grade level, subject area, and statistical model, we classified districts into above-average, average, or below-average performance categories based on the performance statistics estimated by the model for that grade and subject. We found that differences in these performance statistics between above- and below-average districts were generally large enough to be of practical importance. However, the size of these differences and the specific districts identified as above- and below-average varied substantially by grade, subject, and model.

District-wide efforts to improve teaching and learning at all grade levels from preschool through twelfth grade are likely to be particularly important for economically disadvantaged students, who are more likely to start out far behind academically and have more trouble catching up when they are behind (Stanovich, 1986; ACT, 2012b; Dougherty, 2014). As such, this report focuses on comparisons of districts by the poverty level of the students being served. This report adds to our previous report in three ways. First, we classified districts into three poverty levels—high-poverty, medium-poverty, and lower-poverty—and compared the model-generated performance statistics of above- and below-average districts at each poverty level. Second, we compared districts not only on these performance statistics, but also on unadjusted descriptive student achievement statistics such as average student scores and the

[2] See Dougherty & Shaw (2016), Appendix B, Tables B1–B10, bottom row of table labeled "SD of Random Effect (std)."

percentages of students who were On Track or Far Off Track.[3] Third, we examined whether at least half of the students were On Track in districts in different poverty and performance categories. If student performance in many of these district categories falls short of this relatively modest goal, that should emphasize the importance of stronger support for educators and community leaders to help them identify and implement promising approaches to improve student outcomes.[4]

Specifically, this report addresses the following five questions:

1. Was the difference in district performance statistics between above- and below-average districts at each poverty level large enough to be of practical importance?
2. Did district performance vary more among high-poverty districts? If so, was this difference in variation large enough to be of practical importance?
3. How much did the percent of students who were academically On Track differ between above- and below-average districts at each poverty level? Were these differences large enough to be of practical importance?
4. How much did the percent of students who were academically Far Off Track differ between above- and below-average districts at each poverty level? Were these differences large enough to be of practical importance?
5. In what grades and subjects were specified percentages of students On Track in districts at different poverty and performance levels?

In this report, we use the phrase “district performance statistics” to refer to model-generated estimates of differences in student scores between districts that control for student and district demographics and other measured variables that might affect those scores. These statistics were used to answer Questions 1 and 2. We use “student achievement statistics” to refer to unadjusted descriptive statistics on student performance as measured by test scores. These statistics were used to answer Questions 3 through 5. In general, unadjusted student achievement statistics such as average scores and percentages of students On Track are better at addressing the question, “How well are the students doing?” District performance statistics are better at addressing the question, “How well is the district doing?” In all cases, we classified districts as above-average, average, or below-average based on their performance statistics, not their unadjusted student achievement statistics.[5]

Data

This report used longitudinal cohorts created from student-level enrollment and test data supplied by the Arkansas Department of Education for the 2006–07 through the 2013–14 school years. The test data we used included student scores on the Arkansas Benchmark Exams (ABE) in grade 4, the ACT Explore® tests in grade 8, and the ACT® in grades 11 and 12. All enrollment and test datasets contained state-encrypted student IDs so that records for the same students could be linked anonymously across enrollment and test datasets.

[3] See the Data section for definitions of “On Track,” “Off Track,” and “Far Off Track.”

[4] The conclusion to this report discusses several of these approaches.

[5] Cross-district comparisons may look different depending on whether district performance statistics or unadjusted student achievement statistics are used. For example, a district with above-average performance statistics, but more disadvantaged students, may have lower student achievement levels than a district with average performance statistics but more advantaged students. We discuss an example of this in Appendix C.

Our statistical analysis used data both on students' individual demographic characteristics and the demographics of their districts. This required three steps to create the necessary datasets: (1) create longitudinal cohorts of students to be included in the analysis, coding individual students into demographic and academic achievement level categories; (2) calculate district-level statistics and apply rules for including districts in the analysis; and (3) merge student- and district-level datasets together based on the district in which each student was enrolled.

1. Creation of Student Cohorts and Calculation of Student-Level Statistics

This section describes how longitudinal cohorts of students were created for the analysis, how their demographic characteristics were derived from the Arkansas state enrollment and test data, and how students were classified into achievement levels (On Track, Off Track, and Far Off Track) based on their test results.

Creation of student cohorts. To determine what student cohorts to create, we focused on academic achievement in the waypoint grades of 4, 8, and 12. For the analysis using fourth grade test scores, we began with students enrolled in kindergarten in the initial cohort year (for example, the 2006–07 school year) and followed them forward for four subsequent school years, keeping students who took both fourth-grade ABE tests four years later and who were continuously enrolled in the same district during the entire period. Likewise, we followed fourth-grade students forward for four subsequent years, keeping students who were enrolled in the district the entire time and who took both fourth-grade ABE tests in the initial cohort year and all four ACT Explore tests in eighth grade in the final cohort year. We followed initial-year eighth grade students forward for four subsequent years, keeping students who were twelfth-graders four years later who had been enrolled in the district the entire time, and who took all four ACT Explore tests in the initial cohort year and all four ACT tests in grade 12 in the final cohort year or in grade 11 in the next to final year. We used the most recent score for students who took the ACT more than once in grades 11 and 12.

This process created four longitudinal cohorts at each level, referred to as the 2007–11, 2008–12, 2009–13, and 2010–14 cohorts based on the initial cohort year and the final cohort year four years later.[6] Thus, there were twelve total student cohorts (four each for grades K–4, 4–8, and 8–12). At each grade level (K–4, 4–8, and 8–12), we concatenated the four student-level cohorts into a single dataset in order to avoid double-counting students who were retained in the initial cohort grade and to create an indicator variable for those retained students. In all, the 71,977 students in our four K–4 cohorts, 73,633 students in our four 4–8 cohorts, and 36,377 students in our four 8–12 longitudinal cohorts comprised respectively 49%, 51%, and 25% of the students enrolled in the initial cohort years (Appendix A, Tables A1–A3).

We separated out students who had been enrolled in the same district for multiple years—as opposed to using snapshot data on all students enrolled in the final year—in order to focus on those students whose test results would be more likely to reflect the instructional program in

[6] In this nomenclature, school years are named after their spring semesters, so that students in the 2007–11 cohort were present in the district from the collection of enrollment data in the fall of the 2006–07 school year to the collection of test data in the spring of the 2010–11 school year.

the district where they were tested. We followed students forward from an initial cohort year—rather than starting with tested students in the final year and looking backward—in order to account for student attrition between grades 8 and 12, when many students in some districts may drop out.[7] We required that students in the 4–8 and 8–12 cohorts have test scores from the initial year, in order to make the set of students in the analysis for this report match those in a subsequent report on student growth.

Identification of students' demographic and program participation status. Students' characteristics may vary naturally over time: for example, a student's family may qualify for the free and reduced price lunch program when the student is in fourth grade but not when the same student is in eighth grade. Likewise, a student's special education and English language learner status or a student's self-identified ethnicity may change over time. Because an indicator of low-income, English language learner, or special education status may signal a level of disadvantage even if the status is not consistent every year, we identified students as low-income, English language learners, or special education if they had that status in either the initial or final cohort grade level (e.g., either kindergarten or fourth grade for the students in the grade 4 analysis). Because no such logic applies to inconsistent reporting of student ethnic status, we used the student's reported ethnicity in the earliest cohort year as the determining factor for the student's overall ethnic status.[8] Students whose ethnicity, low-income status, special education status, or English language learner status could not be ascertained using these criteria were dropped from the analysis. Fewer than 1% of records were dropped based on incomplete demographic data (footnotes 2, 4, and 6 in Appendix A, Tables A1–A3).[9]

Classification of students by academic achievement level. In grades 8 and 11–12, we classified students as On Track in a given subject if they met or exceeded the College Readiness Benchmark on ACT Explore (in grade 8) or the ACT (in grades 11–12) in the subject in question.[10] In fourth grade, we used the ABE On-Track targets calculated for a previous study (Dougherty, Hiserote, & Shaw, 2014) which identified the fourth grade ABE score in literacy and mathematics associated with a 50% or better probability of meeting or exceeding the eighth grade Benchmark on ACT Explore in the corresponding subject.[11] In turn, Off-Track students were defined as those missing the On-Track Level by one standard deviation or less in the grade and subject in question, while Far-Off-Track students scored more than a full standard deviation below the On-Track Level. These criteria resulted in the definitions for On Track, Off Track, and Far Off Track achievement levels shown in Table 1.

---

[7] A backward- or forward-looking cohort selection process can create the same student cohort, with the only difference being the denominator to which the size of the cohort is compared.

[8] The only exception was for a student with missing ethnic data for the earliest cohort grade (e.g., kindergarten in the fourth grade analysis) but ethnic data present for the final cohort grade level (e.g., grade 4 in the fourth grade analysis), in which case we used the data from the final grade level.

[9] The percentage of records dropped due to incomplete data was around 0.3% in K–4 and 0.1% in 4–8 and 8–12.

[10] The College Readiness Benchmarks on the ACT, updated in 2013, identify the ACT scores associated with a 50% probability of earning a B or approximately a 75% chance of earning a C in entry-level college courses corresponding to the ACT subject tested (Allen & Sconing, 2005; Allen, 2013). In turn, the ACT Explore Benchmarks identify the scores on that test associated with a 50% probability of reaching the Benchmark in the corresponding subject on the ACT (Allen, 2013).

[11] The analysis linked student-level fourth grade ABE scores in 2007 and 2008 with the same student's ACT Explore scores in the 2010–11 and 2011–12 school years. The eighth grade ACT Explore reading test was treated as the closest same-subject match to the fourth grade ABE literacy test, which covers both reading and writing.

Table 1. Scale Score Ranges for On Track, Off Track, and Far Off Track[12]

	On Track	Off Track	Far Off Track
Grade 4 Arkansas Benchmark Exam			
Literacy	772 and above	586–771	585 or below
Mathematics	675 and above	575–674	574 or below
Grade 8 ACT Explore			
English	13–25	9–12	8 or below
Mathematics	17–25	14–16	13 or below
Reading	16–25	12–15	11 or below
Science	18–25	15–17	14 or below
Grade 11/12 ACT			
English	18–36	12–17	11 or below
Mathematics	22–36	17–21	16 or below
Reading	22–36	16–21	15 or below
Science	23–36	18–22	17 or below

2. District-Level Statistics and Inclusion of Districts in the Analysis

Our calculation of district-level statistics began with 238 K–12 school districts that were in existence continuously from the 2006–07 through the 2013–14 school years.[13] Since our focus was on traditional K–12 school districts, charter schools that were not part of such a district were omitted from the analysis.[14] For these 238 districts, we calculated statistics on district-wide demographics and the district's number and percentage of students included in the analysis. Next, we classified districts as rural or non-rural, identified districts that were eligible for the analysis, and divided the eligible districts into the three poverty categories used in this report.

District-wide demographics. We used each district's fall student-level enrollment data for kindergarten through twelfth grade for each year from 2006–07 through 2009–10 to derive annual statistics on the overall district-wide percentage of low-income, African American, Hispanic, Asian, White, Native American, English language learner, and special education students. These statistics were used as district-level predictors in our statistical models.[15]

[12] Standard deviations on the ACT Explore tests were 4.2 points in English, 3.5 in mathematics, 3.9 in reading, and 3.3 in science (ACT 2013, Table 4.11). Standard deviations on the ACT were 6.5 in English, 5.3 in mathematics, 6.3 in reading, and 5.3 in science (ACT 2014, Table 5.4). Standard deviations on the grade 4 ABE, calculated for all students tested in the 2006–07 and 2007–08 school years, were 186.37 scale score points in literacy and 100.93 in mathematics.

[13] If District A consolidated into District B at any time between the 2006–07 and 2013–14 school years, then A's students were combined with B's for the years prior to the consolidation and everyone was treated as part of District B. Thus, basing the analysis on the 238 districts that existed after consolidation did not, in itself, reduce the number of students in the analysis.

[14] Omitting students in charter schools that were not part of a K–12 district reduced the number of students in the analysis—after the other rules for inclusion were applied—by 336 students in grades K–4, 199 students in 4–8, and 86 students in 812.

[15] The district-wide demographic statistics calculated this way differ from ones that would be calculated by aggregating our cohort data, which do not cover all grades. As was the case in our student cohorts, we dropped students with missing demographic data when calculating the district-wide statistics.

District percentage of students in the analysis. For each cohort, we calculated the size of the cohort in each district as a percentage of the total number of students enrolled in the district in the initial cohort grade and year. If a district's percentage of students in the analysis is low compared with other districts—reflecting a higher rate of student mobility or a lower percentage of students taking the test—this might either raise or lower the relative performance of the district. For example, if students whose families face the most challenges leave the sample in disproportionate numbers, that could bias the results in favor of districts with high attrition.[16] High attrition could also result from the presence of a nearby military base or from the district's being less effective at retaining and educating students. In the last of these cases, controlling for attrition rates in the statistical model picks up some of the district performance we are trying to measure. Further research may explore the variables that are associated with student attrition to identify when attrition should be treated as a district performance indicator (e.g., as in the case of high school dropout rates) and when it is simply an aspect of the environment in which the district operates.

Rural district status. Using school-level information from the 2013 Common Core of Data, we defined as "rural" any district in which all schools have a two-digit NCES locale code beginning with 3 (small town) or 4 (rural).

Selection of districts for inclusion in the analysis. We applied two additional criteria to identify which of the 238 continuously existing regular K–12 districts should be included in the analysis:

Accuracy of low-income statistics. The use of students' low-income status as an important control in the statistical models made the accuracy of this classification an important consideration. To assess the accuracy of each district's low-income statistics in a given year, we regressed the district's overall percentage of low-income students in that year on Census estimates of poverty rates of individuals age 5–17 in the district to get a statewide relationship between the two variables, which in turn yielded a Census-predicted district low-income percentage for each year.[17] To have its students included in the analysis, a district's percentage of low-income students in kindergarten through grade 12 had to fall within 20 percentage points of its Census-predicted value in each school year from 2006–07 through 2009–10, the starting years for the cohorts in this report. 202 out of 238 Arkansas K–12 districts met this requirement.

Number of students in the analysis. To be included in the analysis for a given grade level (4, 8, or 11–12), districts were required to have at least 20 students in the four combined longitudinal cohorts for that grade level. In grades 4 and 8, all 202 districts that met the income data requirement also met this criterion, despite the fact that Arkansas has many small districts (Appendix B, Figure B1 and Table B1). Because of low ACT Explore and ACT participation

[16] The poverty measure based on students' free and reduced price lunch status is an imperfect measure of those challenges, so using this measure as a predictor in the statistical models does not completely adjust for this possible bias.

[17] Census-defined poverty uses a lower income threshold than the state definition of low-income, which is based on federal eligibility requirements for the free and reduced price school lunch program. Thus we needed to derive a predicted low-income percentage from the Census data rather than just using the Census percentage. We hypothesized that a district with accurate low-income data would have a relationship between the two poverty measures that is not too different from the state average relationship between the two measures.

rates, 33 of the 202 districts meeting the low-income data criterion had fewer than 20 students eligible for the grades 11–12 analysis, leaving 169 eligible districts at that grade level.[18]

Disaggregation of districts into poverty categories. We calculated district-wide percentages of low-income students across the initial cohort years, 2006–07 through 2009–10, and used these cross-year percentages to classify school districts in the study into three poverty categories:

- Lower poverty: >20–50% low-income students;
- Medium poverty: >50%–70% low-income students; and
- High poverty: >70% low-income students.

We selected these categories because the bottom and top categories each accounted for just under one-quarter of the 202 Arkansas districts included in the grades 4 and 8 analysis, while just over half of the eligible districts were contained in the middle category (Appendix B, Figure B2). Arkansas had no districts in the lowest poverty category, with 20% or fewer low-income students.

3. Combining Student- and District-Level Data

At each grade level (K–4, 4–8, and 8–12), we merged the concatenated file containing student-level data on the four cohorts with the district-level data created in the previous step, based on the district in which each student was enrolled. This process created a single dataset at each level with matched student- and district-level data.

Methods

Once the datasets for the study were built, our analysis had four steps: (1) use statistical models to estimate district performance statistics for each subject and grade level; (2) use these performance statistics to classify districts into above-average, average, or below-average performance categories by subject and grade level; (3) calculate additional district-level student achievement statistics; and (4) aggregate the district performance statistics and student achievement statistics by district performance and poverty category to address the research questions in the study. We describe these steps here.

1. Statistical Models Used to Create District Performance Statistics

We used similar sets of student- and district-level predictors to predict student-level scores on each of the two fourth-grade Arkansas Benchmark Exams (ABE), four eighth-grade ACT Explore tests, and four ACT tests for students in grades 11–12 (Table 2).[19] The models (one per subject and grade level) contained student-level predictors on students' low-income, ethnic, English language learner (ELL), and special education status, and district-level averages of these predictors. The district-level averages might be related to a school district's academic

[18] The differences between the third and fourth data columns in Appendix A, Tables A1–A3 show the effect that removing ineligible districts (and charter schools) had on the number of students in the analysis. The percentages in the last column are based on the number of students in eligible districts.

[19] Except for the addition of a rural district status indicator, the predictors in these models were the same as those used in Model 2 in Dougherty & Shaw (2016).

Table 2. Predictors in the Models for Grades 4, 8, and 11–12

Type of Data	Predictor
Student-Level	Intercept
	Low-income status
	African American status
	Hispanic status
	Asian status
	Native American status
	ELL status
	Special education status
	Flag for retained student*
	Flag for cohort ending in 2012
	Flag for cohort ending in 2013
	Flag for cohort ending in 2014
District-Level	% low-income students
	% African American students
	% Hispanic students
	% Asian students
	% Native American students
	% ELL students
	% special education students
	Number of students in model
	% of students in model
	Flag for rural district

* Retained students had enrollment records in consecutive initial cohort years. If the student met the other criteria for inclusion in the analysis based on the second initial cohort year, the student was included and assigned a flag as a retained student.

culture, funding, and priorities; these influences might in turn affect students' test scores. The models also contained predictors on the district's number and percentage of students included in the analysis, in order to explore the effects of district size and cohort attrition, respectively. We also included a dummy variable for whether the district was located in a rural area, on the theory that that might affect teacher recruitment and thus, indirectly, student performance. In addition, dummy variables for three of the four student cohorts were included to allow for shifts in average test scores across years.

We refer to the statistical models used in this report as "status models" because they used information on student demographics but not on their prior achievement, as value-added models do. The statistics generated by the models may be thought of as answering the question: How did students in this district perform relative to what would have been predicted for students with the same demographics in districts with the same demographics? In comparison, the statistics created by value-added models, which take students' prior achievement into account, may be thought of as answering the question: How did students in this district perform relative to what would have been predicted for students with the same demographics and prior scores in districts with the same demographics and average prior scores? District value-added performance statistics are more appropriate for classifying

districts as above- or below-average when comparing districts' student growth statistics, whereas statistics from status models are more appropriate for classifying districts when comparing student achievement statistics that are not conditioned on students' prior test scores.

The district performance statistics produced by the status models in this report differed from those produced by state accountability reports on student achievement levels in two major ways. First, we used student scores, not proficiency status, as the dependent variable. Second, accountability reports often do not adjust for students' and districts' demographic characteristics in rating district performance, as policymakers do not want to convey the message that educators are free to aim for lower test results for disadvantaged students. The focus in this report is not on accountability targets, but on indicators that do a better job of reflecting relative district effectiveness. For that purpose, it is important to take the degree of difficulty in educating students into account (as reflected by the regression coefficients associated with student characteristics), just as degree of difficulty is taken into account in scoring Olympic gymnastics or diving events.

Because the models used in the study contained both student- and district-level predictors, we estimated them as hierarchical linear models.[20] We used the district-level random effects estimated by the model as the district performance statistics for the grade and subject in question. In fourth grade, this process generated two sets of district performance statistics, one for each ABE subject. In eighth and twelfth grades, this process generated four sets of district performance statistics at each grade level, one for each ACT Explore or ACT subject.

2. Classification of Districts into Relative Performance Categories

For a given subject and grade level, we classified as "above-average" those districts whose performance statistics fell in the top quintile for the grade and subject in question and also were statistically different from average at the .05 confidence level. Similarly, "below-average" districts were those in the bottom quintile whose performance statistics were different from average at the .05 confidence level. Districts not meeting these requirements—i.e., in the middle three quintiles, or in the top or bottom quintile but not statistically different from average at the .05 level—were classified as average.[21]

Although each district was classified in the same poverty category throughout, the same district might fall into different performance categories in different subjects and grade levels. For example, a district could be above average in grade 4 mathematics and below average in grades 11–12 reading.[22] Thus, a performance category such as "above-average districts" comprised different districts depending on the grade and subject.

[20] Appendix C shows the fixed effects of the predictors in Table 2 estimated in each of these statistical models. SAS Proc Mixed was used for all of the statistical models in this report. Information on the SAS code used for the models is available on request.

[21] Top- or bottom-quintile districts were more likely to be classified as average if they had smaller numbers of students in the analysis. Districts were classified into quintiles based on comparing their random effects with those of all the districts in the analysis, not just those in their poverty category. The use of student and district poverty levels as predictors in estimating these random effects was intended to level the playing field in this comparison, so high- and lower-poverty districts would have roughly the same probability (size considerations aside) of being classified as above or below average.

[22] We explored the frequency with which this sort of inconsistency occurs in Dougherty & Shaw (2016).

3. District-Level Student Achievement Statistics

In the next step, we created datasets containing these district-level student achievement statistics in grades 4, 8, and 11–12:

1) the district's percentage of students in the analysis who were classified into On Track, Off Track, and Far Off Track achievement levels; and

2) the average student score of the district's students in the analysis. In order to produce a measure that could be compared across tests with different score scales (such as the ABE in grade 4 and ACT Explore in grade 8), we converted student scores into the number of standard deviations above or below the On-Track level for the grade and subject in question, and averaged those standardized scores across the students in a district.[23]

4. Aggregation of Performance and Achievement Statistics by District Performance and Poverty

For each grade and subject, we calculated weighted averages of the district performance statistics and student achievement statistics across districts in each poverty and performance category. We used the average district performance statistics to address the first question in the study, and the average student achievement statistics to address the third, fourth, and fifth questions. To address the second question, we calculated the unweighted average variance of the district performance statistics within each poverty category. We then did pairwise statistical tests (F-tests) of whether these variances differed at a .05 significance level between pairs of poverty categories in a given subject (for example, whether the variances were different in high- and medium-poverty districts in eighth grade mathematics).

Limitations

Though this report looked at district performance statistics, we were not able to differentiate "district effects" from "school effects." Thus, we did not make a distinction between "performance of the district in grade X" and "performance of the district's school(s) in grade X." The majority of Arkansas school districts are small and rural, and many districts have only one school at a given level. For example, in 2014, 156 (92%) of the 169 districts in the grades 11–12 analysis had only one high school serving grades 11 and 12. Likewise, 180 (89%) of the 202 districts in the grade 8 analysis had only one school serving eighth grade, and 154 (76%) of the 202 districts in the grade 4 analysis had only one school serving fourth grade. Thus, for the great majority of Arkansas districts, the performance statistic in grades 11–12 could also be thought of as an indicator of the performance of the district's single high school, building cumulatively on the performance of its feeder elementary and middle school(s). The comparable statistic for eighth grade could be used as an indicator of

[23] For example, the On-Track cutscore in grade 8 ACT Explore mathematics was 17 (Table 1), and the standard deviation of student scores in ACT Explore mathematics was 3.5 (footnote 12). So if the student's ACT Explore mathematics score was 13, the student's standardized score was (13−17)/3.5 = -1.14. (Students with standardized scores below -1 were classified as Far Off Track.) These standardized student scores were averaged across the students in a given district, and ultimately across the students in districts in a given poverty and performance category (e.g., above-average high-poverty districts).

the performance of the district's single middle or junior high school and its feeder elementary school(s).[24] The value of treating the district as the unit of analysis is to draw attention to the fact that students' earlier schools are likely to have contributed to their achievement levels in grades 8 and 11–12, and to focus attention on the district's potential to improve its schools systematically across the elementary, middle, and high school levels.

Second, we did not attempt to compare the wide range of statistical models that could be used to generate district performance statistics. Our goal was to examine results from relatively straightforward status models that control for generally available student- and district-level demographic statistics. We did not refine the models to eliminate variables that did not add much explanatory power to the models.

Third, we studied measured district performance differences with the understanding that these differences may reflect the effects both of educator practices and of unmeasured student, parent, and community influences that were not picked up as controls in the statistical analysis (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010). For example, some districts may operate in more favorable community environments than other districts with similar student demographics. Thus, the measured performance differences such as those discussed in this report should be treated as the starting point for further inquiry into why these differences exist and what can be done to improve student outcomes in all school systems.

Fourth, our focus on goals of at least 40%, 50%, or 60% of students getting on Track (Question 5) was for illustrative purposes. Educators and policymakers may choose to set different goals—ideally, goals that are both challenging and attainable.

Fifth, because the data in this report are for Arkansas students and districts, further research in other states is needed to determine how the results generalize across states. One goal of this study is to encourage this research.

Finally, because of concerns about the statistical reliability of results from small student groups, we did not report results for groups of less than 20 students.

Results

Question 1: Was the difference in district performance statistics between above- and below-average districts at each poverty level large enough to be of practical importance?

Tables 3–5 compare the performance statistics of above- and below-average districts by district poverty level. For example, in fourth grade literacy, students in the eight above-average lower-poverty districts performed on average 0.19 standard deviations above predicted based on the variables in the model (Table 3).[25] Likewise students in the ten below-average lower-poverty districts scored an average of 0.27 standard deviations below predicted. Thus, the performance gap between the two groups of districts (based on the unrounded statistics) was 0.47 standard deviations.[26]

[24] In a state with a number of larger districts, one could partition the variance in performance across schools in those districts into the variance across districts and the variance across schools within districts.

[25] The size of a single standard deviation in score points is shown by subject and test in footnote 12.

[26] As noted in the table, the apparent discrepancy between these two statistics and their difference is due to rounding.

To assess the importance of these performance differences, they may be compared with the number of standard deviations per year that students typically grow between grades 4 and 8 or between grades 8 and 12. For example, the fourth and eighth grade performance differences shown in Tables 3 and 4 may be compared with average growth per year between grades 4–8 on the ABE exam ranging from 0.24 to 0.30 standard deviations in literacy and from 0.27 to 0.35 standard deviations in mathematics (depending on student cohort), calculated for the three study cohorts in Dougherty & Shaw (2016). Alternatively, the eighth or eleventh and twelfth grade performance differences shown in Tables 4 and 5 may be compared with average growth per year between the ACT Explore and ACT exams of 0.26 standard deviations in English, 0.28 in mathematics, 0.31 in reading, and 0.24 in science (ACT, 2012c), using the average of the ACT Explore and ACT standard deviations to convert typical growth in score points to standard deviations.[27] Thus, differences in performance statistics between above- and below-average districts in the same poverty category, measured in standard deviation units, frequently exceeded a year's typical student growth.

Table 3. Performance of Above- and Below-Average Districts in Grade 4 by District Poverty

		Lower Poverty		Medium Poverty		High Poverty	
Subject	**District Performance Category**	**# of Districts**	**Average of District Performance Statistics**	**# of Districts**	**Average of District Performance Statistics**	**# of Districts**	**Average of District Performance Statistics**
Literacy	Above average	8	0.19	20	0.18	8	0.18
	Below average	10	-0.27	19	-0.17	6	-0.20
	Difference		0.47		0.36		0.38
Mathematics	Above average	9	0.21	19	0.21	9	0.25
	Below average	6	-0.21	19	-0.21	10	-0.20
	Difference		0.42		0.43		0.44

Apparent discrepancies in the differences between numbers shown in this and subsequent tables in this report are due to rounding; these statistics are based on differences between the unrounded numbers, as opposed to the rounded numbers shown in the table.

[27] For example, average growth per year in mathematics = (4.7/3.77)/[(3.5 + 5.3)/2], where 4.7 is average growth across the average period of 45.2 months (= 3.77 years) between the ACT Explore and ACT tests, and 3.5 and 5.3 are the standard deviations of student scores on the two tests.

Table 4. Performance of Above- and Below-Average Districts in Grade 8 by District Poverty

		Lower Poverty		Medium Poverty		High Poverty	
Subject	District Performance Category	# of Districts	Average of District Performance Statistics	# of Districts	Average of District Performance Statistics	# of Districts	Average of District Performance Statistics
English	Above average	9	0.16	14	0.13	2	0.12
	Below average	8	-0.14	17	-0.14	6	-0.14
	Difference		0.30		0.27		0.26
Mathematics	Above average	16	0.20	16	0.21	4	0.21
	Below average	11	-0.18	14	-0.17	6	-0.15
	Difference		0.38		0.37		0.36
Reading	Above average	8	0.14	18	0.13	1	0.25
	Below average	8	-0.14	13	-0.15	6	-0.19
	Difference		0.29		0.28		0.44
Science	Above average	12	0.17	14	0.15	5	0.19
	Below average	8	-0.20	17	-0.17	7	-0.20
	Difference		0.37		0.33		0.39

Question 2: Did district performance vary more among high-poverty districts? If so, was this difference in variation large enough to be of practical importance?

This question was prompted by the authors' recollection of scatter plots in which the variation of *school* performance was higher for high-poverty schools. However, those plots were generally based on student proficiency rates, not average scores, so that the location of the proficiency cutscore relative to typical student performance might have an impact on the relative variation across groups of schools in the percentage of proficient students.[28] In general, theory does not provide us with a clear expectation of whether performance differences should be larger among high- or lower-poverty districts. We might hypothesize that in lower poverty districts, a larger share of student learning takes place outside of school—which is why average student achievement is generally higher in those districts. In that case, in-school factors might account for a smaller share of learning differences across districts. On the other hand, we don't know how much out-of-school learning varies across districts, or whether this variation is greater in lower-poverty districts. Nor do we know whether educator practices vary more across high-poverty than across lower-poverty districts, or how these differences in practices are likely to affect variations in district performance.

[28] For example, if the proficiency cutscore is located near the middle of the score distribution among schools in Group A and near the tail of the distribution among schools in Group B, we would expect schools in Group A to show more variation in the percentage of proficient students.

Table 5. Performance of Above- and Below-Average Districts in Grades 11–12 by District Poverty

		Lower Poverty		Medium Poverty		High Poverty	
Subject	District Performance Category	# of Districts	Average of District Performance Statistics	# of Districts	Average of District Performance Statistics	# of Districts	Average of District Performance Statistics
English	Above average	8	0.20	8	0.16	2	0.14
	Below average	6	-0.21	5	-0.17	2	-0.20
	Difference		0.41		0.33		0.34
Mathematics	Above average	10	0.23	12	0.17	6	0.17
	Below average	8	-0.20	14	-0.16	4	-0.19
	Difference		0.43		0.32		0.36
Reading	Above average	10	0.15	8	0.12	0	n/a
	Below average	6	-0.20	6	-0.12	2	-0.18
	Difference		0.34		0.24		n/a
Science	Above average	7	0.18	6	0.15	2	0.12
	Below average	5	-0.20	9	-0.14	2	-0.19
	Difference		0.38		0.29		0.31

Tables 3–5 do not appear to show a consistent pattern in the size of performance differences across districts in the three poverty level categories. For example, looking from left to right across Table 3, performance difference in fourth grade literacy between above- and below-average districts was larger for lower-poverty districts (0.47 vs. 0.36 and 0.38 for lower-, medium-, and high-poverty districts, respectively); whereas in fourth grade mathematics, the size of these differences was similar across poverty categories (0.42, 0.43, and 0.44). In eighth grade, performance differences between above- and below-average districts were larger for lower-poverty districts in English (0.30 vs. 0.27 and 0.26), but larger for high-poverty districts in reading (Table 4). In eleventh and twelfth grades, lower-poverty districts appeared to have larger performance differences (Table 5). However, several of these comparisons in grades 8 and 11–12 were based on small numbers of districts. For example, the performance gap for high-poverty districts in eighth grade reading was based on a single above-average district. In other cases only two districts were involved on one or both sides of the comparison.

To explore this issue in a way that examines variation across all districts, including those in the middle three quintiles or in the top and bottom quintiles but not statistically different from average, we calculated the unweighted standard deviation and variance (the square of the standard deviation) of the random effect statistics for districts in each poverty level by grade and subject, and conducted a set of statistical significance tests comparing these variances between district poverty levels. Table 6 shows the standard deviations of the random effect statistics and the results of the statistical tests.

Table 6. Comparison of Within-Category Performance Variation between Poverty Levels

		Performance Variation within Poverty Levels			Statistical Significance of Comparison Between Levels		
Grade	Subject	Lower-Poverty Districts	Medium-Poverty Districts	High-Poverty Districts	Lower vs. High	Lower vs. Medium	Medium vs. High
4	Literacy	0.152	0.141	0.146			
	Mathematics	0.155	0.150	0.180			
8	English	0.111	0.093	0.090			
	Mathematics	0.161	0.117	0.105	***	***	
	Reading	0.106	0.094	0.099			
	Science	0.130	0.102	0.117		**	
11–12	English	0.127	0.101	0.094	*	*	
	Mathematics	0.152	0.116	0.121		**	
	Reading	0.119	0.081	0.076	***	***	
	Science	0.112	0.088	0.084	*	*	

***Significant at the .01 level. **Significant at the .05 level. * Significant at the .10 level.

From Table 6, we can see that the variation in performance across districts was largest for *lower*-poverty districts in all cases except for fourth grade mathematics, where the variation was largest for high-poverty districts. However, these comparisons were not statistically significant in fourth grade literacy and mathematics and eighth grade English and reading, indicating that we have only limited evidence that this pattern of differences would continue to recur in repeated sampling (for example, in subsequent time periods). There were no statistically significant differences between the variation in performance in medium- and high-poverty districts (last column of Table 6). In eighth grade science and eleventh and twelfth grade mathematics, the difference between lower- and medium-poverty districts was significant at the 0.05 level, but the (smaller) difference between lower- and high-poverty districts was not. The patterns of performance variation did not always match by subject in different grades: for example, in eighth grade, performance differences are significantly larger in lower-poverty districts in mathematics but not in reading, whereas in eleventh and twelfth grades, performance differences are larger in lower-poverty districts in reading but not statistically different from those in high-poverty districts in mathematics.

Looking at practical importance, the largest statistically significant difference in performance variation between lower- and high-poverty districts was about 0.056 of a standard deviation in eighth grade mathematics, comparing the first and third data columns of Table 6. If this is interpreted as the differences in the variation of the district influence on the typical student's ACT Explore mathematics score, this difference amounts to about 0.2 of an ACT Explore score point, which may be considered of relatively low practical importance.

Question 3: How much did the percent of students who were academically On Track differ between above- and below-average districts at each poverty level? Were these differences large enough to be of practical importance?

Looking across all grade levels, differences in the percentage of On-Track students between above- and below-average districts in the same poverty category ranged from nine percentage points for medium-poverty districts in eighth grade reading to 36 percentage points for high-poverty districts, also in eighth-grade reading (Tables 7–9). The largest difference with more than one district on each side of the comparison was 26 percentage points for lower-poverty districts in mathematics in eleventh and twelfth grade (Table 9), and the median difference among all 29 comparisons of above- and below-average districts in Tables 7–9 was 20 percentage points.[29]

Looking separately by grade level, in fourth grade the differences in On-Track rates ranged from ten percentage points among medium-poverty districts in literacy to 25 percentage points among high-poverty districts in mathematics (Table 7). In eighth grade, differences that involved more than one district on each side of the comparison ranged from nine percentage points among medium-poverty districts in reading to 22 percentage points among high-poverty districts in English (Table 8). In eleventh and twelfth grades, differences ranged from 12 percentage points among high-poverty districts in science to 26 percentage points among lower-poverty districts in mathematics (Table 9).[30]

To consider how important these differences are, we can simulate how many additional students in the analysis in the below-average districts would have been On Track had those districts had the same On-Track rates as the above-average districts. Results of these simulations are shown in the rightmost column of Tables 7–9. For example, increasing the percentage of students On Track in fourth-grade literacy in the six below-average high-poverty districts from 22% to 43% to match the percentage in the eight above-average high-poverty districts would result in 286 additional students On Track (Table 7).[31] In fourth grade, the number of additional students who would be On Track in these simulations ranges from 276 students in lower-poverty districts to 1,183 students in medium-poverty districts, both in mathematics (Table 7). In eighth grade, the numbers range from 274 additional students in lower-poverty districts to 1,206 students in medium-poverty districts, both in English (Table 8). In eleventh and twelfth grades, differences range from 27 additional students in science in high-poverty districts to 871 students in mathematics in medium-poverty districts (Table 9).

[29] No comparison was available for high-poverty districts in reading in grades 11–12 (Table 9).

[30] Tables C1–C5 in Appendix C provide information on the number of students in each district poverty and performance category, average scores for those students, and the percentages of those students who were On Track, Off Track, and Far Off Track. These tables can be used to make comparisons within and across district performance and poverty categories. For example, average scores might be compared between above-average high-poverty districts and average lower-poverty districts.

[31] Apparent discrepancies between the percentages in Tables 7–12 and their differences are due to rounding; the differences are calculated using the unrounded percentages. Also, changes in the numbers of simulated On-Track students in Tables 7–9 or Far-Off-Track students in Tables 10–12 cannot be added up across subjects without double-counting students who would change their status in more than one subject. They can, however, be added up across district poverty categories in the same subject.

Table 7. Percentage of On-Track Students in Above- and Below-Average Districts by District Poverty: Grade

Subject	District Poverty Category	Number of Above-Average Districts	% On Track	Number of Below-Average Districts	% On Track	Difference in On-Track Rates	Students in Analysis in Below-Average Districts	Simulated Additional On-Track Students in Below-Average Districts*
Literacy	Lower	8	61%	10	37%	24%	2,285	540
	Medium	20	47%	19	37%	10%	5,259	537
	High	8	43%	6	22%	21%	1,356	286
Mathematics	Lower	9	51%	6	29%	22%	1,259	276
	Medium	19	45%	19	32%	13%	8,843	1,183
	High	9	41%	10	16%	25%	2,190	537

* Equals the number of students in the analysis in below-average districts multiplied by the difference in On-Track rates between above- and below-average districts.

Table 8. Percentage of On-Track Students in Above- and Below-Average Districts by District Poverty: Grade 8

Subject	District Poverty Category	Number of Above-Average Districts	% On Track	Number of Below-Average District	% On Track	Difference in On-Track Rates	Students in Analysis in Below-Average Districts	Simulated Additional On-Track Students in Below-Average Districts
English	Lower	9	76%	8	64%	12%	2,257	274
	Medium	14	70%	17	57%	12%	9,691	1,206
	High	2	61%	6	40%	22%	1,984	428
Mathematics	Lower	16	51%	11	31%	20%	4,468	891
	Medium	16	44%	14	24%	20%	3,517	688
	High	4	34%	6	16%	18%	1,802	324
Reading	Lower	8	44%	8	31%	13%	2,460	324
	Medium	18	37%	13	28%	9%	8,154	735
	High	1	46%	6	9%	36%	2,045	746
Science	Lower	12	47%	8	27%	20%	1,578	313
	Medium	14	36%	17	26%	10%	9,136	920
	High	5	30%	7	10%	20%	2,146	429

Table 9. Percentage of On-Track Students in Above- and Below-Average Districts by District Poverty: Grades 11–12

Subject	District Poverty Category	Number of Above-Average Districts	% On Track	Number of Below-Average District	% On Track	Difference in On-Track Rates	Students in Analysis in Below-Average Districts	Simulated Additional On-Track Students in Below-Average Districts
English	Lower	8	82%	6	61%	21%	1,203	254
	Medium	8	70%	5	53%	17%	545	92
	High	2	41%	2	22%	18%	231	43
Mathematics	Lower	10	57%	8	30%	26%	1,416	374
	Medium	12	45%	14	25%	21%	4,222	871
	High	6	26%	4	9%	17%	434	76
Reading	Lower	10	57%	6	34%	23%	1,203	274
	Medium	8	45%	6	31%	13%	1,030	139
	High	0	#N/A	2	17%	#N/A	200	#N/A
Science	Lower	7	47%	5	24%	23%	984	228
	Medium	6	38%	9	19%	19%	1,175	222
	High	2	18%	2	5%	12%	219	27

Question 4: How much did the percent of students who were academically Far Off Track differ between above- and below-average districts at each poverty level? Were these differences large enough to be of practical importance?

Students who are Far Off Track in grades 4 or 8 have trouble getting On Track by grades 8 or 11–12, so it is important for districts to reduce the percentage of students who are Far Off Track (ACT, 2012c; Dougherty & Fleming, 2012; Dougherty, 2014; Dougherty, Hiserote, & Shaw, 2014). Looking across all grade levels, differences in the percentage of Far-Off-Track students between above- and below-average districts in the same poverty category ranged from one percentage point for lower- and medium-poverty districts in eighth-grade English to 47 percentage points for high-poverty districts in eighth-grade reading (Tables 10–12). The largest differences in Far-Off-Track rates with more than one district on each side of the comparison were 26 percentage points for high-poverty districts in eighth grade science and eleventh and twelfth grade mathematics (Tables 11 and 12), and the median difference among all 29 comparisons of above- and below-average districts in Tables 10–12 was 13 percentage points. At a given poverty level, fewer students were Far Off Track in English than in the other subjects.

Looking separately by grade level, in fourth grade the differences in Far-Off-Track rates ranged from six percentage points among medium-poverty districts in literacy to 25 percentage points among high-poverty districts in mathematics (Table 10). In eighth grade, differences that involved more than one district on each side of the comparison ranged from one percentage point for lower- and medium-poverty districts in English to 26 percentage points for high-poverty districts in science (Table 11). In eleventh and twelfth grades, differences ranged from seven percentage points among lower- and medium-poverty districts in English to 26 percentage points among high-poverty districts in mathematics (Table 12).

As was the case with On-Track students, we simulated how many fewer students in below-average districts would have been Far Off Track had those districts had the same Far-Off-Track rates as the above-average districts. Results of these simulations are shown in the rightmost column of Tables 10–12. For example, reducing the percentage of Far-Off-Track

students in fourth-grade literacy in the six below-average high-poverty districts from 38% to 15% to match the percentage in the eight above-average high-poverty districts would result in 306 fewer Far-Off-Track students (Table 10). In fourth grade, the reduction in the number of Far-Off-Track students in these simulations ranges from 147 students in lower-poverty districts to 827 students in medium-poverty districts, both in mathematics (Table 10). In eighth grade, the numbers range from 24 fewer Far-Off-Track students in lower-poverty districts in English to 958 fewer students in high-poverty districts in reading, or 790 fewer students in medium-poverty districts in science in comparisons with more than one district on each side (Table 11). In eleventh and twelfth grades, differences range from 25 fewer Far-Off-Track students in high-poverty districts in English to 934 fewer students in medium-poverty districts in mathematics (Table 12).

Table 10. Percentage of Far-Off-Track Students in Above- and Below-Average Districts by District Poverty: Grade 4

Subject	District Poverty Category	Number of Above-Average Districts	% Far Off Track	Number of Below-Average District	% Far Off Track	Difference in Far-Off-Track Rates	Students in Analysis in Below-Average Districts	Simulated Change in Far-Off-Track Students in Below-Average Districts
Literacy	Lower	8	7%	10	20%	-13%	2,285	-305
	Medium	20	17%	19	22%	-6%	5,259	-295
	High	8	15%	6	38%	-23%	1,356	-306
Mathematics	Lower	9	10%	6	22%	-12%	1,259	-147
	Medium	19	14%	19	23%	-9%	8,843	-827
	High	9	17%	10	42%	-25%	2,190	-548

Table 11. Percentage of Far-Off-Track Students in Above- and Below-Average Districts by District Poverty: Grade 8

Subject	District Poverty Category	Number of Above-Average Districts	% Far Off Track	Number of Below-Average District	% Far Off Track	Difference in Far-Off-Track Rates	Students in Analysis in Below-Average Districts	Simulated Change in Far-Off-Track Students in Below-Average Districts
English	Lower	9	1%	8	2%	-1%	2,257	-24
	Medium	14	2%	17	3%	-1%	9,691	-104
	High	2	2%	6	8%	-6%	1,984	-116
Mathematics	Lower	16	12%	11	23%	-11%	4,468	-481
	Medium	16	15%	14	31%	-16%	3,517	-566
	High	4	21%	6	41%	-20%	1,802	-364
Reading	Lower	8	22%	8	34%	-12%	2,460	-303
	Medium	18	28%	13	37%	-9%	8,154	-750
	High	1	19%	6	66%	-47%	2,045	-958
Science	Lower	12	14%	8	29%	-16%	1,578	-245
	Medium	14	19%	17	28%	-9%	9,136	-790
	High	5	26%	7	51%	-26%	2,146	-551

Table 12. Percentage of Far-Off-Track Students in Above- and Below-Average Districts by District Poverty: Grades 11–12

Subject	District Poverty Category	Number of Above-Average Districts	% Far Off Track	Number of Below-Average District	% Far Off Track	Difference in Far-Off-Track rates	Students in Analysis in Below-Average Districts	Simulated Change in Far-Off-Track Students in Below-Average Districts
English	Lower	8	2%	6	9%	-7%	1,203	-82
	Medium	8	6%	5	13%	-7%	545	-38
	High	2	18%	2	29%	-11%	231	-25
Mathematics	Lower	10	12%	8	30%	-18%	1,416	-254
	Medium	12	20%	14	42%	-22%	4,222	-934
	High	6	38%	4	64%	-26%	434	-113
Reading	Lower	10	10%	6	23%	-13%	1,203	-152
	Medium	8	18%	6	26%	-8%	1,030	-84
	High	0	#N/A	2	45%	#N/A	200	#N/A
Science	Lower	7	13%	5	32%	-19%	984	-190
	Medium	6	23%	9	36%	-14%	1,175	-159
	High	2	42%	2	59%	-17%	219	-36

Question 5: In what grades and subjects were specified percentages of students On Track in districts at different poverty and performance levels?

In this section, we show the percentages of students who were On Track in districts at different poverty and performance levels (Table 13), color-coding the table to show cases in which these percentages met or exceeded specified targets. Somewhat arbitrarily, we chose three target percentages: that at least 60%, 50%, or 40% of students be On Track in the grade in question. By varying these target percentages and observing the resulting pattern of who has met the targets, educators and policymakers can get an idea of what kinds of targets might be realistic for the near future. For example, a goal that at least 75% of students be On Track was not met in any group of districts in any subject except for English, and thus might be considered an aspirational goal in the other subjects.

Table 13. Percentages of Students On Track in Districts with Different Poverty and Performance Levels[32]

		Lower-Poverty			Medium-Poverty			High-Poverty		
Grade	**Subject**	**Above**	**Average**	**Below**	**Above**	**Average**	**Below**	**Above**	**Average**	**Below**
4	Literacy	61	55	37	47	44	37	43	30	22
	Mathematics	51	45	29	45	34	32	41	24	16
8	English	76	73	64	70	61	57	61	49	40
	Mathematics	51	43	31	44	32	24	34	23	16
	Reading	44	41	31	37	30	28	46	21	9
	Science	47	37	27	36	29	26	30	20	10
11–12	English	82	76	61	70	63	53	41	50	22
	Mathematics	57	50	30	45	36	25	26	20	9
	Reading	57	50	34	45	39	31	N/A	24	17
	Science	47	42	24	38	30	19	18	17	5

N/A = Not Available because there were no statistically significantly above-average districts in this subject and poverty category. Apparent discrepancies in the table are due to rounding: for example, below-average high-poverty districts had slightly fewer than 40% of students On Track in eighth-grade English, and just under 50% of students in average high-poverty districts were On Track in grades 11–12 English.

Criterion ≥ 60% ≥ 50% ≥ 40%

Table 13 can be used to illustrate subjects and groups of districts where student performance is of concern because students have not met relatively modest goals. For example, fewer than half of students were On Track in the great majority of cases in mathematics, reading, and science in grades 8 and 11–12 and in literacy and mathematics in grade 4, and in many of these cases, the percentage of On-Track students was lower than 40%. On-Track percentages were especially low in below-average high-poverty districts. Even in above-average high- and moderate-poverty districts, the majority of students did not reach On-Track benchmarks for college readiness in mathematics, reading, and science.

Conclusion

When performance was adjusted for differences in student demographics, performance differences between above- and below-average Arkansas districts were large enough to be of practical importance, often as much as the difference made by a year's growth for a typical student. These differences occurred within each district poverty category, so that for students in high-poverty school districts, it mattered what high-poverty district they were in, and the same for students in lower-poverty districts. There was a slight tendency for performance differences to be larger in lower-poverty districts, but this tendency was not consistent across grades and subjects.

Likewise, differences in unadjusted student achievement statistics were large enough to be of practical significance—based on the larger number of students who would have been On Track (and the smaller number Far Off Track) had student achievement levels in below-average districts matched those in above-average districts in the same poverty category.

[32] These percentages are also shown for above- and below-average districts in Tables 7–9, and for all groups of districts in Appendix C, Tables C1–C5.

However, even in above-average districts, the majority of students in medium- and high-poverty areas did not meet On Track benchmarks for college readiness in mathematics, reading, or science. In thinking about how to improve student outcomes, educators and policymakers should consider four basic approaches:

1. *Start early.* College and career readiness does not begin in high school, or even in middle school (ACT, 2008). Gaps in student learning begin in early childhood and are well established by kindergarten (Hart & Risley, 1995; West, Denton, & Germino-Hausken, 2000). But "starting early" is not confined to improving early childhood and preschool programs. Improvements must be made in the early elementary grades as well. These improvements can include strengthening the early reading and mathematics program, promoting better student behaviors, and emphasizing a content-rich curriculum in the early grades that includes science, history/social studies, and the fine arts (Dougherty, 2013).
2. *Monitor and improve implementation of practices in key areas*. These areas should be chosen based on their ability to improve a district's capacity to address a wide range of problems related to student learning. Based on research by ACT and others, these practice areas might include (Dougherty, 2016):

   a) Develop or adopt, refine, and use a written district curriculum that describes what students should learn in each grade/course and subject.

   b) Teach a content-rich curriculum in the early grades.

   c) Use data from multiple sources to guide improvements in teaching and learning.

   d) Encourage teachers to collaborate routinely around curriculum, instruction, and assessment.

   e) Develop a coaching system for teachers.

   f) Communicate with parents about their children's academic progress and what their children are expected to learn.

   When district leaders target improvement in a given practice area, they should systematically gather information on what practices are actually being implemented and how that implementation correlates with gains in student learning (Dougherty, 2016). They should treat teachers and school leaders as partners in figuring out how to improve practices (Knight, 2007).
3. *Form networks among practitioners and researchers to share learning about improvement.* These networks can connect educators in different districts working on the same problem, in addition to connecting educators in different schools in the same district. The creation of cross-district knowledge-sharing networks can be particularly important in a state such as Arkansas with many small geographically dispersed districts. The creation of such a network can be facilitated by researchers and practitioners in a state education agency, university, regional education laboratory, or nonprofit organization (Bryk, Gomez, Grunow, & LeMahieu, 2015).
4. *Work with policymakers and community leaders to strengthen out-of-school supports for students and their families.* This approach can be particularly valuable in high-poverty communities, where students face out-of-school challenges that distract them from

learning (Willingham, 2012). Strengthening support for students and their families can require better coordination among social service agencies and between social service agencies and schools (Broader, Bolder Approach to Education, 2016).

By using these four approaches and keeping track of associations between the implementation of specific practices and improvements in student outcomes, educators and policymakers can increase their effectiveness in improving student learning.

References

ACT. (2008). *The forgotten middle*. Iowa City, IA: Author. http://www.act.org/content/dam/act/unsecured/documents/ForgottenMiddle.pdf.

ACT. (2012a). *Rising to the challenge of college and career readiness: A framework for effective practices*. Iowa City, IA: Author. http://www.act.org/content/dam/act/unsecured/documents/RisingToChallenge.pdf.

ACT. (2012b). *Catching up to college and career readiness*. Iowa City, IA: Author. http://www.act.org/content/dam/act/unsecured/documents/Catching-Up-To-College-and-Career-Readiness.pdf.

ACT. (2012c). *Principles for measuring growth towards college and career readiness*. Iowa City, IA: Author. http://www.act.org/content/dam/act/unsecured/documents/GrowthModelingReport.pdf.

ACT. (2013). *ACT Explore technical manual*. Iowa City, IA: Author. http://www.act.org/content/dam/act/unsecured/documents/Explore-TechManual.pdf.

ACT. (2014). *The ACT technical manual*. Iowa City, IA: Author. http://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf.

Allen, J. (2013). *Updating the ACT college readiness benchmarks*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-6.pdf.

Allen, J., Bassiri, D., & Noble, J. (2009). *Statistical properties of accountability measures based on ACT's Educational Planning and Assessment System*. Iowa City, IA: ACT. http://forms.act.org/research/researchers/reports/pdf/ACT_RR2009-1.pdf.

Allen, J. & Sconing, J. (2005). *Using ACT assessment scores to set benchmarks for college readiness*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2005-3.pdf.

Bassiri, D. (2015). *Statistical properties of school value-added scores based on assessments of college readiness*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-5.pdf.

Broad Foundation. (2014). *The Broad prize for urban education*. Web page. www.broadprize.org.

Broader, Bolder Approach to Education. (2016). *Case studies: Bright Futures* (Pea Ridge, AR). Washington, D.C.: Author. http://www.boldapproach.org/case-studies/.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G., (2015). *Learning to improve: How America's schools can get better at getting better*. Cambridge, MA: Harvard Education Press.

Chingos, M. M., Whitehurst, G. J., & Gallaher, M. R. (2013). *School districts and student achievement*. Washington, D.C.: The Brookings Institution Brown Center on Education Policy.

Daly, A. J. & Finnigan, K. S., eds. (2016). *Thinking and acting systemically: Improving school districts under pressure*. Washington, D.C.: American Educational Research Association.

Dougherty, C. (2013). *College and career readiness: The importance of early learning*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ImportanceofEarlyLearning.pdf.

Dougherty, C. (2014). *Catching up to college and career readiness: The challenge is greater for at-risk students*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/CatchingUp-Part3.pdf.

Dougherty, C. (2015a). *How school district leaders can support the use of data to improve teaching and learning*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/Use-of-Data.pdf.

Dougherty, C. (2015b). *Use of data to support teaching and learning: A case study of two school districts*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-1.pdf.

Dougherty, C. (2016). *Keeping track of improvement in educational practices*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/5516-insights-keeping-track-of-improvement-in-educational-practices.pdf.

Dougherty, C. & Fleming, S. (2012). *Getting students on track to college and career readiness: How many catch up from far behind?* Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2012-9.pdf.

Dougherty, C., Hiserote, L., & Shaw, T. (2014). *Catching up to college and career readiness in Arkansas*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2014-3.pdf.

Dougherty, C. & Shaw, T. (2016). *Size and consistency of school district performance measures across models, subjects, and grade levels*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/5293-Research-Report-2016-2-Size-and-Consistancy-of-Performance-Measures.pdf.

Epstein, J. & Sheldon, S. (2006). Moving forward: Ideas for research on school, family, and community partnerships. In C. F. Conrad & R. Serlin (Eds.) *SAGE Handbook for research in education: Engaging ideas and enriching inquiry*. Thousand Oaks, CA: Sage Publications, Chapter 7, pp. 117–138.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in everyday experience of young American children*. Baltimore: Paul H. Brookes.

Hightower, A., Knapp, M., Marsh, J., & McLaughlin, M., (Eds.) (2002). *School districts and instructional renewal*. New York: Teachers College Press.

Hirsch, E. D. (2003). Reading comprehension requires knowledge—of words and the world. *The American Educator*. Washington, D.C.: American Federation of Teachers. http://www.aft.org/sites/default/files/periodicals/Hirsch.pdf.

Hirsch, E. D. (2006). *The Knowledge Deficit: Closing the Shocking Education Gap for American Children*. New York: Houghton Mifflin.

Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction*. Thousand Oaks, CA: Corwin Press.

Mac Iver, M. A. & Farley-Ripple, E. (2008). *Bringing the district back in: The role of the central office in instruction and achievement*. Alexandria, VA: Educational Research Service.

Marsh, J. A., Kerr, K. A., Ikemoto, G. S., Darilek, H., Suttorp, M., Zimmer, R. W., & Barney, H. (2005). *The role of districts in fostering instructional improvement: Lessons from three urban districts partnered with the Institute for Learning*. Santa Monica: Rand Corporation.

Sawyer, R. (2013). *Interpreting changes over time in high school average ACT® college readiness assessment composite scores and ACT College Readiness Benchmark attainment rates*. Iowa City, IA: ACT. http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2013-9.pdf.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 31*(4), 360–407. http://www.psychologytoday.com/files/u81/Stanovich__1986_.pdf.

Supovitz, J. A. (2006). *The case for district-based reform: Leading, building, and sustaining school improvement*. Cambridge, MA: Harvard University Press.

West, J., Denton, K., & Germino-Hausken, E. (2000). *America's kindergartners*. Washington, D.C.: National Center for Education Statistics. http://nces.ed.gov/pubs2000/2000070.pdf.

Whitehurst, G. J., Chingos, M. M., & Gallaher, M. R. (2013). *Do school districts matter*? Washington, D.C.: The Brookings Institution Brown Center on Education Policy.

Willingham, D. T. (2012). Why does family wealth affect learning? *American Educator,* Washington, D.C.: American Federation of Teachers, Spring. http://www.aft.org/sites/default/files/periodicals/Willingham.pdf.

Appendix A

Descriptive Statistics on Students in the Analysis

Tables A1, A2, and A3 show the number and percentage of students from each cohort who were included in the statistical analysis. Table A4 explicitly compares student attrition in the grades 4 and 8 analyses, using the information from Tables A1 and A2. Table A5 does a similar summary comparison of attrition between the grades 8 and 11–12 analyses, using the information from Tables A2 and A3.

Comparing grades 4 and 8 (Table A4), enrollment attrition, defined as students not enrolled in the expected grade four years later, was higher between kindergarten and grade 4 than between grades 4 and 8. Higher enrollment attrition in the grade 4 analysis was likely due to a larger number of retained students in kindergarten and first grade who were not picked up in a later cohort. However, this effect was offset by a higher percentage of students in the eighth grade analysis not taking the ACT Explore test than students in the fourth grade analysis not taking the ABE, so the overall percentages of students in the analysis were similar in grades 4 and 8.

Comparing grades 8 and 11–12 (Table A5), higher enrollment attrition in high school was likely a result of students dropping out. In addition, fewer students in the grades 11–12 analysis took ACT Explore in grade 8 than the percentage of students in the grade 8 analysis who took the ABE in grade 4, resulting in attrition of 28% from the grades 11–12 analysis versus only 4% for the grade 8 analysis (Table A5). The percentage of students taking ACT Explore rose substantially in 2010–11 when the State of Arkansas began paying the districts' costs of giving students the ACT Explore test in eighth grade. This policy increased the number and percentage of students included in the grade 8 analysis (Table A2), but not in the grades 11–12 analysis (Table A3), as students in the 11–12 analysis were eighth graders prior to the 2010–11 school year.

Table A6 illustrates how student attrition affected the percentages of students in various at-risk groups in the study cohorts. As would be expected, the students in the longitudinal cohorts—who were continuously enrolled in the same district, progressed by four grades in four years, and took all tests—were less at-risk than the general enrolled population in the initial cohort years. The percentages of low-income, African American, and special education students were lower in the study cohorts in all three levels (4, 8, and 11–12) than in the population from which the cohorts were drawn. Likewise, the percentages of Hispanic students and English language learners were lower in the high school study cohorts than in the population from which these cohorts were drawn. In addition, the impact of attrition on the percentage of low-income, English language learner, and special education students in the analysis was greater in grades 11–12 than the other two levels (Table A6).

Table A1. Percentage of Arkansas Kindergarten Students in Grade 4 Analysis

Student Cohort[1]	Total Kindergarten Enrollment	Students Tested in 4th Grade[2]	Students Eligible for Statistical Analysis[3]	Eligible Students in Eligible Districts	Percent of Students in Statistical Analysis
2007–2011	33,072	25,783	19,810	17,570	53%
2008–2012	35,950	27,499	21,030	18,051	50%
2009–2013	37,354	28,183	21,206	18,156	49%
2010–2014	39,672	27,255	21,142	18,200	46%
Total	146,048	108,720	83,188	71,977	49%

[1] For example, the 2007–2011 cohort consists of students who were enrolled in kindergarten in the 2006–07 school year and who took the Arkansas Benchmark Exams in fourth grade in the 2010–11 school year.

[2] The attrition of 37,328 students between the first two data columns of this chart includes 437 kindergarten students with incomplete demographic data, 33,603 who were not enrolled in fourth grade four years later, and 3,288 students who were enrolled in fourth grade but did not take both state tests.

[3] The attrition of 25,532 students between the second and third columns of this chart consists of students who were enrolled in kindergarten and tested in both subjects four years later in grade 4, but who were not enrolled throughout grades K–4 and tested in grade 4 in the same district.

Table A2. Percentage of Arkansas 4th Grade Students in Grade 8 Analysis

Student Cohort	Total 4th Grade Enrollment	Students Tested in 4th and 8th Grade[4]	Students Eligible for Statistical Analysis[5]	Eligible Students in Eligible Districts	Percent of Students in Statistical Analysis
2007–2011	34,570	25,512	20,282	17,600	51%
2008–2012	35,418	26,499	21,133	18,340	52%
2009–2013	37,954	28,018	21,883	18,870	50%
2010–2014	37,732	26,465	21,332	18,823	50%
Total	145,674	106,494	84,630	73,633	51%

[4] The attrition of 39,180 students between the first two data columns of this chart includes 115 students with incomplete demographic data; 23,706 students who were not enrolled in 8th grade four years later; 4,590 students who were enrolled four years later but had not taken both fourth grade state tests; and 10,769 students enrolled and tested in grade 4 and enrolled in grade 8, but who did not take the ACT Explore test in eighth grade.

[5] The attrition of 21,864 students between the second and third columns of this chart consists of students enrolled and tested in both grades 4 and 8 but who were not enrolled throughout grades 4–8 and tested in grades 4 and 8 in the same district.

Table A3. Percentage of Arkansas 8th Grade Students in Grades 11–12 Analysis

Student Cohort	Total 8th Grade Enrollment	Students Tested in 8th and 11th or 12th Grade[6]	Students Eligible for Statistical Analysis[7]	Eligible Students in Eligible Districts	Percent of Students in Statistical Analysis
2007–2011	34,810	10,020	8,469	7,082	20%
2008–2012	35,421	10,768	9,180	8,172	23%
2009–2013	36,769	14,078	11,461	9,843	27%
2010–2014	36,882	15,328	13,248	11,280	31%
Total	143,882	50,194	42,358	36,377	25%

[6] The attrition of 93,688 students between the first two data columns of this chart includes 155 students with incomplete demographic data, 37,833 students who were enrolled in eighth grade but not in twelfth grade four years later; 39,162 students enrolled in grade 12 four years later but who had not taken ACT Explore in 8th grade; and 16,538 students enrolled and taking ACT Explore in grade 8 and following a normal grade progression between grades 8 and 12 but not taking the ACT.

[7] The attrition of 7,836 students between the second and third data columns of this chart consists of students who met the requirements for inclusion in the second column but who were not enrolled throughout grades 8–12 and tested in grades 8 and 11 or 12 in the same district.

Table A4. Comparing Attrition in the Grade 4 and Grade 8 Analysis

	Grade 4		Grade 8	
Student Population	**Number of Students**	**% of Students in Initial Grade**	**Number of Students**	**% of Students in Initial Grade**
All enrolled students in the initial grade	146,048	100%	145,674	100%
. . . with complete demographic information	145,611	99.7%	145,559	99.9%
. . . and enrolled in final grade four years later	112,008	77%	121,853	84%
. . . and taking all tests in initial grade	N/A	N/A	117,263	80%
. . . and taking all tests in final grade	108,720	74%	106,494	73%
. . . and continuously enrolled in the district	83,188	57%	84,630	58%
. . . and in an eligible district	71,977	49%	73,633	51%

Table A5. Comparing Attrition in the Grade 8 and Grade 11–12 Analysis

	Grade 8		Grades 11–12	
Student Population	**Number of Students**	**% of Students in Initial Grade**	**Number of Students**	**% of Students in Initial Grade**
All enrolled students in the initial grade	145,674	100%	143,882	100%
. . . with complete demographic information	145,559	99.9%	143,727	99.9%
. . . and enrolled in final grade four years later	121,853	84%	105,894	74%
. . . and taking all tests in initial grade	117,263	80%	66,732	46%
. . . and taking all tests in final grade	106,494	73%	50,194	35%
. . . and continuously enrolled in the district	84,630	58%	42,358	29%
. . . and in an eligible district	73,633	51%	36,377	25%

Table A6. Demographics of Arkansas Student Cohorts

	Grades K–4		Grades 4–8		Grades 8–12	
Demographic Category	**All Students in Initial Grade[8]**	**Students in Cohort**	**All Students in Initial Grade**	**Students in Cohort**	**All Students in Initial Grade**	**Students in Cohort**
% low-income	70%	63%	67%	60%	61%	46%
% African American	21%	18%	22%	18%	22%	19%
% Hispanic	11%	12%	9%	10%	8%	6%
% Asian	2%	2%	2%	2%	1%	1%
% Native American	1%	1%	1%	1%	1%	1%
% ELL	9%	11%	7%	7%	5%	2%
% special education	16%	13%	14%	11%	12%	5%

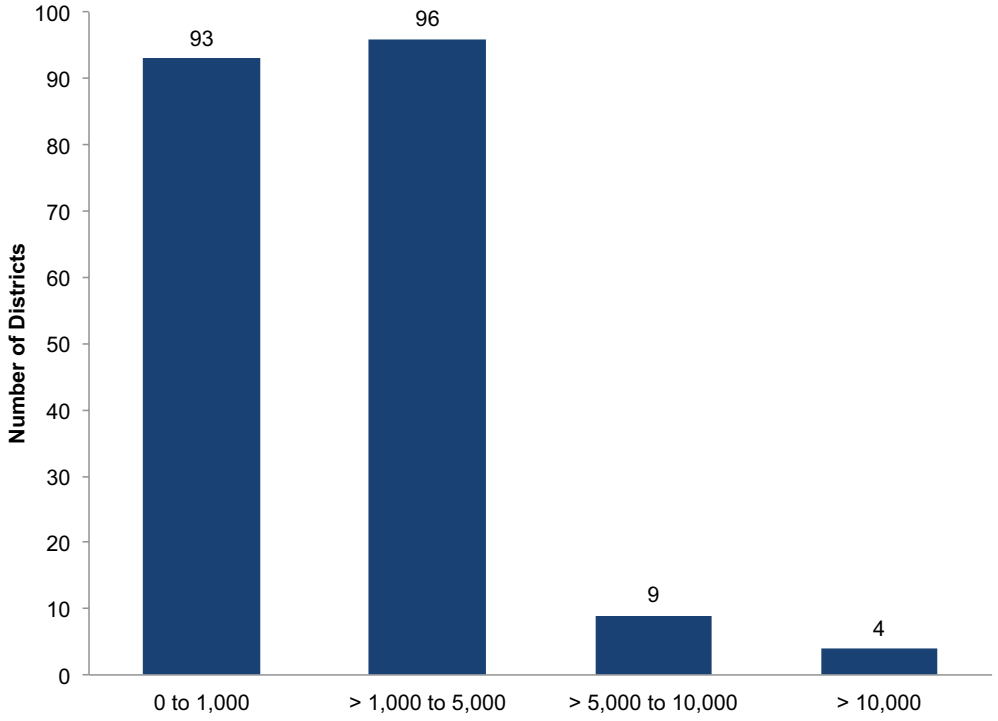
[8] The denominators for the percentages of "All students in initial grade" are students with complete demographic data.

Appendix B

Descriptive Statistics on Districts in the Analysis

Arkansas is a largely rural state whose largest district, the Little Rock School District, had approximately 25,000 K–12 students, averaged across the four initial cohort years.[33] Overall, the majority of Arkansas school districts were small; only 13 (6%) of the 202 eligible districts in the analysis had more than 5,000 students (Figure B1). About half of the districts were in the medium-poverty category with 50–70% low-income students, while about one-quarter of the districts were in each of the lower and high-poverty categories (Figure B2). Higher poverty districts were more likely than their lower-poverty counterparts to be small and located in rural areas or small towns—for example, only one of the 41 high-poverty districts had more than 5,000 students, and 38 of the 41 high-poverty districts were located in rural areas (Table B1). High-poverty districts also had greater concentrations of African American and Hispanic students.

Figure B1. Distribution of eligible districts by total K–12 enrollment (N = 202 districts)



[33] District size and demographic percentages reported in this appendix are based on K–12 statistics averaged across the 2006–07 through the 2009–10 school years.

Figure B2. Distribution of eligible districts by their percentage of low-income students (N = 202 districts)

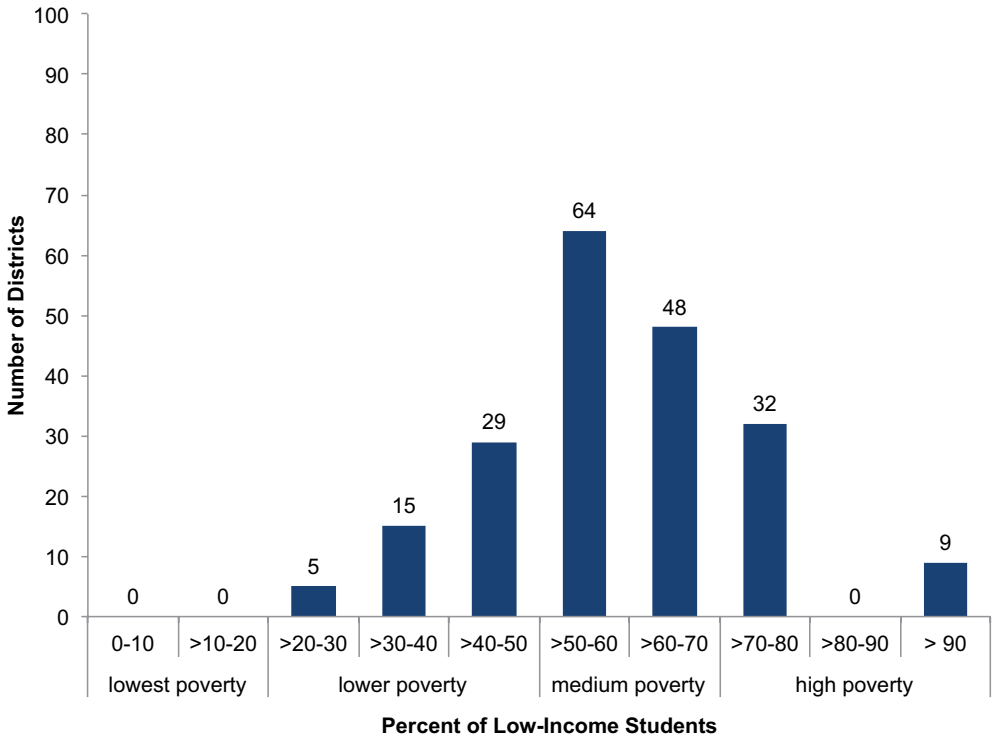


Table B1. District Characteristics by District Poverty Category*

	District Poverty Category (Percentage of Low-Income Students)		
District Characteristics	**Lower (>20–50%)**	**Medium (>50%—70%)**	**High (>70–100%)**
Number of districts	49	112	41
Average size (Number of K–12 students)	2,521	1,996	1,263
% of districts > 5,000 students	10%	6%	2%
% of districts ≤ 1,000 students	37%	45%	61%
% of districts rural	69%	90%	93%
% African American students	5%	11%	40%
% Hispanic students	3%	5%	6%
% special education students	11%	12%	13%

*Statistics are for the 202 districts eligible for the grades 4 and 8 analysis.

Tables B2 and B3 provide district-wide demographic data aggregated across the four initial cohort years (the 2006–07 through the 2009–10 school years). Although three-quarters of districts had 50% or more low-income students, districts with substantial percentages of African American, Hispanic, or English language learner students were in the minority. In addition, Arkansas had relatively few Asian and Native American students. In addition, these tables provide information on the number and percentages of students in the analysis, aggregating across the four student cohorts at each grade level. The median percentages of students in the

analysis shown in Tables B2 and B3 are likely to be higher than the percentages for the same cohorts shown in Tables A1–A3, because the denominators for the percentages in Tables A1–A3 include districts not in the analysis.

Table B2. Descriptive Statistics for Arkansas School Districts in the Grades 4 and 8 Analysis (N = 202)

	District Percentile						
District Statistic	**5th**	**10th**	**25th**	**50th**	**75th**	**90th**	**95th**
% low-income	35%	42%	50%	58%	65%	74%	79%
% African-American	0%	0%	1%	2%	24%	51%	67%
% Hispanic	1%	1%	1%	2%	5%	11%	17%
% Asian	0%	0%	0%	0%	1%	2%	3%
% Native American	0%	0%	0%	0%	1%	1%	3%
% ELL	0%	0%	0%	1%	3%	7%	11%
% special education	9%	9%	10%	12%	14%	15%	17%
# students in analysis—gr4	67	76	102	179	365	688	978
# students in analysis—gr8	73	84	114	196	379	754	984
% students in analysis—gr4	43%	47%	53%	58%	62%	64%	66%
% students in analysis—gr8	43%	50%	56%	61%	66%	68%	70%

The statistics shown are based on district-wide demographic data from the 2006–07 through the 2009–10 school years (the initial cohort years) and the number and percentages of students in the analysis for the four cohorts. In general, reading down the columns, the percentiles refer to different districts: for example, the district with the median percentage of low-income students is not necessarily the district with the median percentage of African American students.

Table B3. Descriptive Statistics for Arkansas School Districts in the Grades 11–12 Analysis (N = 169)

	District Percentile						
District Statistic	**5th**	**10th**	**25th**	**50th**	**75th**	**90th**	**95th**
% low-income	34%	38%	50%	59%	68%	75%	92%
% African-American	0%	0%	1%	2%	27%	56%	75%
% Hispanic	1%	1%	1%	2%	5%	11%	17%
% Asian	0%	0%	0%	0%	1%	2%	3%
% Native American	0%	0%	0%	0%	1%	2%	3%
% ELL	0%	0%	0%	1%	2%	6%	10%
% special education	9%	10%	10%	12%	14%	15%	17%
# students in analysis—gr12	33	40	65	118	215	487	630
% students in analysis—gr12	12%	20%	29%	37%	45%	51%	54%

The statistics shown are based on district-wide demographic data from the 2006–07 through the 2009–10 school years (the initial cohort years) and the number and percentages of students in the analysis for the four cohorts. In general, reading down the columns, the percentiles refer to different districts: for example, the district with the median percentage of low-income students is not necessarily the district with the median percentage of African American students.

Correlation of District-Level Statistics

Tables B4–B6 show correlations of district-level statistics used as predictors in the analysis. To calculate these correlations, we aggregated each statistic over the four initial cohort years, rather than using the yearly values of each statistic.

Only a minority of variable pairs had at least moderately high (.3 or greater in absolute value) correlations. Not surprisingly, districts' percentages of English Language Learners (ELLs) were strongly related to their percentages of Hispanic students (with a correlation of .94 for the 202 districts in the grades 4 and 8 analyses and the subset of 169 districts in the grades 11–12 analysis).[34] In addition, districts' percentages of students in poverty were correlated with their percentages of African American students (correlation of .60 in the 202 districts and .62 in the subset of 169 districts), and districts with more low-income students tended to have lower percentages of cohort students in the analysis, as shown by the negative correlations between those two variables in the bottom row of Tables B4–B6.

Table B4. Pairwise Correlations between District-Level Statistics for Arkansas School Districts: Grade 4 Analysis (grades K–4 student cohorts) (N = 202)

	% Low-Income	% AfrAm	% Hisp	% Asian	% NatAm	% ELL	% Spec Education	# in Analysis
% African-American	**.60**							
	.000							
% Hispanic	.04	-.09						
	.528	*.223*						
% Asian	-.16	-.20	**.37**					
	.023	*.005*	*.000*					
% Native American	-.07	-.24	.16	**.33**				
	.298	*.000*	*.024*	*.000*				
% ELL	.04	-.07	**.94**	**.49**	.17			
	.559	*.298*	*.000*	*.000*	*.018*			
% special education	.29	-.10	-.17	-.13	-.03	-.16		
	.000	*.160*	*.013*	*.056*	*.699*	*.025*		
# students in analysis	-.19	.10	**.42**	**.37**	.01	**.50**	-.21	
	.007	*.141*	*.000*	*.000*	*.934*	*.000*	*.003*	
% students in analysis	**-.49**	**-.35**	.12	.06	.11	.13	-.26	.19
	.000	*.000*	*.100*	*.361*	*.134*	*.056*	*.000*	*.007*

p-values are in italics. Correlations with p-values of .05 or less and with absolute values of .30 or higher are in bold.

[34] This might indicate the possibility of paring down the model by dropping one of those two variables.

Table B5. Pairwise Correlations between District-Level Statistics for Arkansas School Districts: Grade 8 Analysis (grades 4–8 student cohorts) (N = 202)

	% Low-Income	% AfrAm	% Hisp	% Asian	% NatAm	% ELL	% Spec Education	# in Analysis
# students in analysis	-.24	.06	**.42**	**.39**	.02	**.50**	-.22	
	.001	*.361*	*.000*	*.000*	*.793*	*.000*	*.002*	
% students in analysis	**-.52**	**-.49**	.17	.12	.02	.14	-.16	.13
	.000	*.000*	*.018*	*.095*	*.822*	*.046*	*.023*	*.072*

p-values are in italics. The correlations between demographic variables are the same as in Grade 4 (Table D1) because the same district-wide demographic variables were used for all grade levels.

Table B6. Pairwise Correlations between District-Level Statistics for Arkansas School Districts: Grades 11–12 Analysis (grades 8–12 student cohorts) (N = 169)

	% Low-Income	% AfrAm	% Hisp	% Asian	% NatAm	% ELL	% Spec Education	# in Analysis
% African-American	**.62**							
	.000							
% Hispanic	.05	-.08						
	.516	*.315*						
% Asian	-.20	-.19	.21					
	.008	*.012*	*.006*					
% Native American	-.06	-.25	.18	**.39**				
	.412	*.001*	*.019*	*.000*				
% ELL	.05	-.05	**.94**	**.32**	.21			
	.501	*.482*	*.000*	*.000*	*.005*			
% special education	.27	-.10	-.19	-.16	-.02	-.18		
	.000	*.207*	*.012*	*.039*	*.757*	*.016*		
# students in analysis	-.27	.08	.27	.29	.03	**.33**	-.21	
	.000	*.272*	*.000*	*.000*	*.700*	*.000*	*.007*	
% students in analysis	**-.30**	-.15	.13	.12	.00	.07	-.21	.26
	.000	*.046*	*.100*	*.135*	*.975*	*.335*	*.006*	*.001*

p-values are in italics. Correlations with p-values of .05 or less and with absolute values of .30 or higher are in bold.

Appendix C

Supplemental Data Tables

The tables in this appendix provide demographic, student achievement, and district performance information organized by the nine district poverty and performance categories used in this report (three poverty levels × three performance categories). Tables C1–C5 provide student achievement statistics and counts of students and districts; Tables C6–C10 show district-wide K–12 demographics; while Tables C11–C15 show weighted average district performance statistics and the percentage of students in the analysis.

The reader can use these tables to make comparisons not shown in the main body of the paper. In addition, the information in these tables can be used to better understand situations in which comparisons using district performance statistics yielded different results from those using unadjusted student achievement statistics. For example, based on unadjusted student achievement statistics in grades 11–12 English, students in the two above-average high-poverty districts performed worse than those in the 33 average districts in the same poverty category. Forty-one percent of students in the two above-average districts were On Track compared with 50% of students in the average districts, and the average ACT English score in the above-average districts was about a fifth of a standard deviation below that in the average districts (Table C4).

This apparent discrepancy can be better understood by noting the difference in student demographics between the two groups of districts (Table C9). The two above-average districts had higher percentages of economically disadvantaged students (close to 100% vs. 77% for the 33 average districts) and African American students (86% vs. 47%). Adjusting for these variables and the others shown in Table 2 and Appendix D, Table D3, our statistical analysis estimated higher district performance statistics for the two above-average districts (0.14 of a standard deviation averaged across the two districts, versus an average of close to 0 across the 33 average districts) (Table C14).[35]

[35] We can also see from Table C14 that 29% of eighth grade enrolled students from the initial cohort years were eligible for the analysis in grades 11–12 in the above-average districts, compared with 32% in the average districts. The statistical analysis adjusted for this small difference.

Table C1. Student Achievement Levels and Average Scores by District Poverty and Performance in Grade 4

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	# of Districts	# of Students	% On Track	% Off Track	% Far Off Track	Average Score Relative to On-Track Level*
Literacy	Lower (>20–50%)	Above average	8	4,044	61%	32%	7%	0.13
		Average	31	16,622	55%	34%	11%	0.00
		Below average	10	2,285	37%	43%	20%	-0.34
	Medium (>50–70%)	Above average	20	12,218	47%	36%	17%	-0.16
		Average	73	23,124	44%	39%	17%	-0.22
		Below average	19	5,259	37%	40%	22%	-0.37
	High (>70–100%)	Above average	8	2,147	43%	41%	15%	-0.19
		Average	27	4,922	30%	44%	26%	-0.51
		Below average	6	1,356	22%	40%	38%	-0.76
Mathematics	Lower (>20–50%)	Above average	9	4,517	51%	38%	10%	0.02
		Average	34	17,175	45%	42%	13%	-0.11
		Below average	6	1,259	29%	49%	22%	-0.40
	Medium (>50–70%)	Above average	19	5,619	45%	41%	14%	-0.11
		Average	74	26,139	34%	43%	23%	-0.38
		Below average	19	8,843	32%	45%	23%	-0.41
	High (>70–100%)	Above average	9	2,115	41%	43%	17%	-0.21
		Average	22	4,120	24%	45%	32%	-0.63
		Below average	10	2,190	16%	42%	42%	-0.87

* The average score of cohort students in the districts in question, measured in the number of student-level standard deviations above (+) or below (-) the On-Track score for grade 4.

Table C2. Student Achievement Levels and Average Scores by District Poverty and Performance in Grade 8 English and Mathematics

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	# of Districts	# of Students	% On Track	% Off Track	% Far Off Track	Average Score Relative to On-Track Level*
English	Lower (>20–50%)	Above average	9	8,717	76%	23%	1%	0.65
		Average	32	13,820	73%	26%	1%	0.53
		Below average	8	2,257	64%	34%	2%	0.32
	Medium (>50–70%)	Above average	14	5,628	70%	28%	2%	0.46
		Average	81	25,346	61%	36%	3%	0.27
		Below average	17	9,691	57%	40%	3%	0.19
	High (>70–100%)	Above average	2	477	61%	37%	2%	0.23
		Average	33	5,713	49%	45%	5%	0.00
		Below average	6	1,984	40%	53%	8%	-0.20
Mathematics	Lower (>20–50%)	Above average	16	13,976	51%	37%	12%	-0.06
		Average	22	6,350	43%	42%	15%	-0.25
		Below average	11	4,468	31%	46%	23%	-0.48
	Medium (>50–70%)	Above average	16	4,576	44%	41%	15%	-0.22
		Average	82	32,572	32%	44%	24%	-0.50
		Below average	14	3,517	24%	45%	31%	-0.68
	High (>70–100%)	Above average	4	730	34%	45%	21%	-0.38
		Average	31	5,642	23%	46%	31%	-0.71
		Below average	6	1,802	16%	43%	41%	-0.93

* The average score of cohort students in the districts in question, measured in the number of student-level standard deviations above (+) or below (-) the On-Track score for grade 8.

Table C3. Student Achievement Levels and Average Scores by District Poverty and Performance in Grade 8 Reading and Science

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	# of Districts	# of Students	% On Track	% Off Track	% Far Off Track	Average Score Relative to On-Track Level*
Reading	Lower (>20–50%)	Above average	8	7,649	44%	34%	22%	-0.10
		Average	33	14,685	41%	36%	23%	-0.18
		Below average	8	2,460	31%	35%	34%	-0.43
	Medium (>50–70%)	Above average	18	8,158	37%	36%	28%	-0.29
		Average	81	24,353	30%	34%	36%	-0.47
		Below average	13	8,154	28%	36%	37%	-0.51
	High (>70–100%)	Above average	1	219	46%	35%	19%	-0.08
		Average	34	5,910	21%	34%	45%	-0.69
		Below average	6	2,045	9%	25%	66%	-1.10
Science	Lower (>20–50%)	Above average	12	12,833	47%	40%	14%	-0.13
		Average	29	10,383	37%	44%	19%	-0.35
		Below average	8	1,578	27%	44%	29%	-0.62
	Medium (>50–70%)	Above average	14	4,629	36%	45%	19%	-0.37
		Average	81	26,900	29%	44%	27%	-0.55
		Below average	17	9,136	26%	46%	28%	-0.61
	High (>70–100%)	Above average	5	846	30%	45%	26%	-0.50
		Average	29	5,182	20%	46%	34%	-0.76
		Below average	7	2,146	10%	39%	51%	-1.14

* The average score of cohort students in the districts in question, measured in the number of student-level standard deviations above (+) or below (-) the On-Track score for grade 8.

Table C4. Student Achievement Levels and Average Scores by District Poverty and Performance in Grades 11–12 English and Mathematics

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	# of Districts	# of Students	% On Track	% Off Track	% Far Off Track	Average Score Relative to On-Track Level*
English	Lower (>20–50%)	Above average	8	4,552	82%	16%	2%	0.78
		Average	29	8,131	76%	19%	4%	0.59
		Below average	6	1,203	61%	30%	9%	0.22
	Medium (>50–70%)	Above average	8	2,006	70%	24%	6%	0.50
		Average	76	15,350	63%	28%	8%	0.29
		Below average	5	545	53%	34%	13%	0.04
	High (>70–100%)	Above average	2	558	41%	42%	18%	-0.20
		Average	33	3,801	50%	36%	14%	0.00
		Below average	2	231	22%	49%	29%	-0.51
Mathematics	Lower (>20–50%)	Above average	10	4,605	57%	31%	12%	0.07
		Average	25	7,865	50%	35%	15%	-0.10
		Below average	8	1,416	30%	39%	30%	-0.49
	Medium (>50–70%)	Above average	12	3,183	45%	35%	20%	-0.19
		Average	63	10,496	36%	38%	26%	-0.37
		Below average	14	4,222	25%	33%	42%	-0.63
	High (>70–100%)	Above average	6	832	26%	36%	38%	-0.59
		Average	27	3,324	20%	37%	43%	-0.71
		Below average	4	434	9%	28%	64%	-1.01

* The average score of cohort students in the districts in question, measured in the number of student-level standard deviations above (+) or below (-) the On-Track score for grades 11–12.

Table C5. Student Achievement Levels and Average Scores by District Poverty and Performance in Grades 11–12 Reading and Science

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	# of Districts	# of Students	% On Track	% Off Track	% Far Off Track	Average Score Relative to On-Track Level*
Reading	Lower (>20–50%)	Above average	10	7,059	57%	33%	10%	0.18
		Average	27	5,624	50%	36%	14%	-0.01
		Below average	6	1,203	34%	43%	23%	-0.33
	Medium (>50–70%)	Above average	8	2,850	45%	38%	18%	-0.12
		Average	75	14,021	39%	39%	22%	-0.24
		Below average	6	1,030	31%	43%	26%	-0.40
	High (>70–100%)	Above average	0	0	#N/A	#N/A	#N/A	#N/A
		Average	35	4,390	24%	41%	35%	-0.58
		Below average	2	200	17%	39%	45%	-0.80
Science	Lower (>20–50%)	Above average	7	4,170	47%	40%	13%	-0.11
		Average	31	8,732	42%	41%	16%	-0.24
		Below average	5	984	24%	44%	32%	-0.66
	Medium (>50–70%)	Above average	6	1,843	38%	39%	23%	-0.38
		Average	74	14,883	30%	43%	28%	-0.54
		Below average	9	1,175	19%	45%	36%	-0.78
	High (>70–100%)	Above average	2	702	18%	40%	42%	-0.86
		Average	33	3,669	17%	41%	42%	-0.88
		Below average	2	219	5%	36%	59%	-1.25

* The average score of cohort students in the districts in question, measured in the number of student-level standard deviations above (+) or below (-) the On-Track score for grades 11–12.

Table C6. District-Wide Demographic Statistics by District Poverty and Performance in Grade 4

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Low-Income	African-American	Hispanic	Asian	English Language Learner	Special Ed
Literacy	Lower (>20–50%)	Above average	41%	11%	4%	1%	1%	11%
		Average	36%	5%	5%	2%	3%	11%
		Below average	44%	5%	3%	1%	1%	12%
	Medium (>50–70%)	Above average	61%	35%	12%	1%	9%	12%
		Average	58%	18%	11%	2%	8%	11%
		Below average	57%	8%	6%	1%	4%	13%
	High (>70–100%)	Above average	78%	37%	17%	1%	9%	12%
		Average	79%	55%	4%	0%	2%	12%
		Below average	89%	60%	6%	0%	2%	12%
Mathematics	Lower (>20–50%)	Above average	41%	11%	4%	1%	2%	11%
		Average	37%	5%	5%	2%	3%	11%
		Below average	41%	2%	2%	1%	1%	12%
	Medium (>50–70%)	Above average	59%	18%	6%	1%	3%	12%
		Average	60%	27%	9%	1%	6%	12%
		Below average	57%	5%	21%	4%	19%	10%
	High (>70–100%)	Above average	76%	33%	15%	1%	9%	12%
		Average	83%	53%	5%	1%	3%	12%
		Below average	80%	65%	5%	0%	1%	13%

These statistics consist of district-wide K–12 percentages aggregated across the four initial cohort years (2006–07 through 2009–10). The statistical analysis used yearly percentages for each district to predict each year's test scores, rather than four-year averages.

Table C7. District-Wide Demographic Statistics by District Poverty and Performance in Grade 8 English and Mathematics

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Low-Income	African-American	Hispanic	Asian	English Language Learner	Special Ed
English	Lower (>20–50%)	Above average	37%	10%	5%	2%	3%	12%
		Average	37%	3%	4%	2%	2%	11%
		Below average	43%	4%	3%	1%	1%	11%
	Medium (>50–70%)	Above average	56%	13%	8%	1%	5%	12%
		Average	60%	27%	9%	1%	6%	11%
		Below average	59%	10%	18%	4%	16%	11%
	High (>70–100%)	Above average	77%	42%	5%	0%	3%	12%
		Average	82%	49%	7%	1%	4%	13%
		Below average	75%	62%	9%	0%	4%	11%
Mathematics	Lower (>20–50%)	Above average	36%	7%	6%	2%	3%	11%
		Average	40%	5%	2%	1%	1%	11%
		Below average	39%	2%	3%	1%	1%	12%
	Medium (>50–70%)	Above average	59%	15%	6%	1%	3%	12%
		Average	59%	22%	12%	2%	9%	11%
		Below average	61%	22%	3%	1%	1%	12%
	High (>70–100%)	Above average	75%	34%	6%	1%	4%	12%
		Average	82%	48%	7%	1%	4%	13%
		Below average	76%	68%	8%	0%	3%	11%

These statistics consist of district-wide K–12 percentages aggregated across the four initial cohort years (2006–07 through 2009–10). The statistical analysis used yearly percentages for each district to predict each year's test scores, rather than four-year averages.

Table C8. District-Wide Demographic Statistics by District Poverty and Performance in Grade 8 Reading and Science

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Low-Income	African-American	Hispanic	Asian	English Language Learner	Special Ed
Reading	Lower (>20–50%)	Above average	37%	11%	6%	2%	4%	12%
		Average	37%	3%	4%	2%	2%	11%
		Below average	42%	2%	3%	1%	1%	11%
	Medium (>50–70%)	Above average	57%	20%	7%	1%	4%	12%
		Average	60%	26%	9%	1%	6%	11%
		Below average	58%	6%	22%	5%	19%	11%
	High (>70–100%)	Above average	72%	0%	2%	0%	0%	13%
		Average	80%	46%	8%	1%	4%	12%
		Below average	83%	71%	6%	0%	3%	12%
Science	Lower (>20–50%)	Above average	35%	8%	6%	2%	4%	11%
		Average	40%	3%	2%	1%	1%	11%
		Below average	44%	2%	3%	1%	1%	11%
	Medium (>50–70%)	Above average	59%	16%	7%	1%	4%	13%
		Average	59%	27%	9%	1%	6%	11%
		Below average	58%	5%	19%	4%	17%	11%
	High (>70–100%)	Above average	81%	43%	3%	1%	2%	12%
		Average	80%	45%	8%	1%	4%	12%
		Below average	82%	70%	6%	0%	3%	12%

These statistics consist of district-wide K–12 percentages aggregated across the four initial cohort years (2006–07 through 2009–10). The statistical analysis used yearly percentages for each district to predict each year's test scores, rather than four-year averages.

Table C9. District-Wide Demographic Statistics by District Poverty and Performance in Grades 11–12 English and Mathematics

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Low-Income	African-American	Hispanic	Asian	English Language Learner	Special Ed
English	Lower (>20–50%)	Above average	36%	11%	6%	2%	3%	11%
		Average	37%	3%	4%	2%	2%	11%
		Below average	39%	6%	2%	1%	0%	10%
	Medium (>50–70%)	Above average	60%	30%	5%	1%	2%	12%
		Average	60%	23%	10%	1%	7%	12%
		Below average	60%	17%	3%	2%	2%	12%
	High (>70–100%)	Above average	100%	86%	1%	0%	0%	11%
		Average	77%	47%	9%	1%	5%	12%
		Below average	88%	83%	1%	0%	0%	11%
Mathematics	Lower (>20–50%)	Above average	38%	10%	5%	2%	3%	11%
		Average	36%	3%	5%	2%	2%	11%
		Below average	40%	7%	2%	1%	1%	10%
	Medium (>50–70%)	Above average	59%	19%	10%	1%	6%	11%
		Average	59%	15%	9%	1%	6%	12%
		Below average	63%	48%	7%	2%	6%	11%
	High (>70–100%)	Above average	90%	52%	2%	1%	0%	12%
		Average	79%	53%	9%	0%	5%	12%
		Below average	82%	63%	2%	0%	1%	12%

These statistics consist of district-wide K–12 percentages aggregated across the four initial cohort years (2006–07 through 2009–10). The statistical analysis used yearly percentages for each district to predict each year's test scores, rather than four-year averages

Table C10. District-Wide Demographic Statistics by District Poverty and Performance in Grades 11–12 Reading and Science

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Low-Income	African-American	Hispanic	Asian	English Language Learner	Special Ed
Reading	Lower (>20–50%)	Above average	35%	8%	6%	2%	4%	11%
		Average	39%	4%	3%	1%	2%	12%
		Below average	39%	6%	2%	1%	0%	10%
	Medium (>50–70%)	Above average	59%	36%	5%	1%	2%	11%
		Average	60%	22%	10%	1%	7%	12%
		Below average	59%	7%	7%	2%	4%	14%
	High (>70–100%)	Above average	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
		Average	81%	54%	8%	1%	4%	12%
		Below average	75%	37%	3%	0%	2%	15%
Science	Lower (>20–50%)	Above average	38%	11%	6%	2%	3%	12%
		Average	36%	3%	4%	2%	2%	11%
		Below average	40%	9%	2%	1%	0%	9%
	Medium (>50–70%)	Above average	60%	28%	4%	1%	2%	12%
		Average	60%	24%	10%	1%	7%	12%
		Below average	59%	13%	6%	1%	3%	11%
	High (>70–100%)	Above average	88%	72%	1%	0%	0%	11%
		Average	79%	50%	9%	1%	5%	12%
		Below average	85%	68%	2%	0%	1%	13%

These statistics consist of district-wide K–12 percentages aggregated across the four initial cohort years (2006–07 through 2009–10). The statistical analysis used yearly percentages for each district to predict each year's test scores, rather than four-year averages.

Table C11. Percentages of Students in the Analysis and Average District Performance Statistics by District Poverty and Performance in Grade 4

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Number of Students in Initial Grade	Number of Students in Analysis	Percent of Students in Analysis	Average District Performance Statistic
Literacy	Lower (>20–50%)	Above average	6,647	4,044	61%	0.19
		Average	27,738	16,622	60%	0.03
		Below average	3,742	2,285	61%	-0.27
	Medium (>50–70%)	Above average	21,343	12,218	57%	0.18
		Average	39,507	23,124	59%	0.01
		Below average	9,001	5,259	58%	-0.17
	High (>70–100%)	Above average	4,229	2,147	51%	0.18
		Average	9,600	4,922	51%	-0.03
		Below average	2,426	1,356	56%	-0.20
Mathematics	Lower (>20–50%)	Above average	7,518	4,517	60%	0.21
		Average	28,566	17,175	60%	0.03
		Below average	2,043	1,259	62%	-0.21
	Medium (>50–70%)	Above average	10,433	5,619	54%	0.21
		Average	45,531	26,139	57%	-0.01
		Below average	13,887	8,843	64%	-0.21
	High (>70–100%)	Above average	4,124	2,115	51%	0.25
		Average	7,953	4,120	52%	0.00
		Below average	4,178	2,190	52%	-0.20

Table C12. Percentages of Students in the Analysis and Average District Performance Statistics by District Poverty and Performance in Grade 8 English and Mathematics

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Number of Students in Initial Grade	Number of Students in Analysis	Percent of Students in Analysis	Average District Performance Statistic
English	Lower (>20–50%)	Above average	13,523	8,717	64%	0.16
		Average	21,456	13,820	64%	0.01
		Below average	3,518	2,257	64%	-0.14
	Medium (>50–70%)	Above average	8,912	5,628	63%	0.13
		Average	44,554	25,346	57%	0.03
		Below average	15,168	9,691	64%	-0.14
	High (>70–100%)	Above average	795	477	60%	0.12
		Average	11,398	5,713	50%	0.02
		Below average	3,739	1,984	53%	-0.14
Mathematics	Lower (>20–50%)	Above average	21,675	13,976	64%	0.20
		Average	9,998	6,350	64%	0.01
		Below average	6,824	4,468	65%	-0.18
	Medium (>50–70%)	Above average	7,363	4,576	62%	0.21
		Average	55,334	32,572	59%	-0.01
		Below average	5,937	3,517	59%	-0.17
	High (>70–100%)	Above average	1,253	730	58%	0.21
		Average	11,185	5,642	50%	0.01
		Below average	3,494	1,802	52%	-0.15

Table C13. Percentages of Students in the Analysis and Average District Performance Statistics by District Poverty and Performance in Grade 8 Reading and Science

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Number of Students in Initial Grade	Number of Students in Analysis	Percent of Students in Analysis	Average District Performance Statistic
Reading	Lower (>20–50%)	Above average	11,922	7,649	64%	0.14
		Average	22,855	14,685	64%	0.02
		Below average	3,720	2,460	66%	-0.14
	Medium (>50–70%)	Above average	13,159	8,158	62%	0.13
		Average	43,128	24,353	56%	0.00
		Below average	12,347	8,154	66%	-0.15
	High (>70–100%)	Above average	302	219	73%	0.25
		Average	11,475	5,910	52%	0.01
		Below average	4,155	2,045	49%	-0.19
Science	Lower (>20–50%)	Above average	19,794	12,833	65%	0.17
		Average	16,069	10,383	65%	0.00
		Below average	2,634	1,578	60%	-0.20
	Medium (>50–70%)	Above average	7,528	4,629	61%	0.15
		Average	47,045	26,900	57%	0.01
		Below average	14,061	9,136	65%	-0.17
	High (>70–100%)	Above average	1,519	846	56%	0.19
		Average	10,061	5,182	52%	0.01
		Below average	4,352	2,146	49%	-0.20

Table C14. Percentages of Students in the Analysis and Average District Performance Statistics by District Poverty and Performance in Grades 11–12 English and Mathematics

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Number of Students in Initial Grade	Number of Students in Analysis	Percent of Students in Analysis	Average District Performance Statistic
English	Lower (>20–50%)	Above average	11,396	4,552	40%	0.20
		Average	21,034	8,131	39%	0.02
		Below average	2,650	1,203	45%	-0.21
	Medium (>50–70%)	Above average	4,789	2,006	42%	0.16
		Average	41,519	15,350	37%	-0.02
		Below average	1,643	545	33%	-0.17
	High (>70–100%)	Above average	1,916	558	29%	0.14
		Average	11,900	3,801	32%	0.00
		Below average	800	231	29%	-0.20
Mathematics	Lower (>20–50%)	Above average	12,430	4,605	37%	0.23
		Average	19,401	7,865	41%	0.04
		Below average	3,249	1,416	44%	-0.20
	Medium (>50–70%)	Above average	7,531	3,183	42%	0.17
		Average	29,368	10,496	36%	0.01
		Below average	11,052	4,222	38%	-0.16
	High (>70–100%)	Above average	2,393	832	35%	0.17
		Average	10,980	3,324	30%	-0.01
		Below average	1,243	434	35%	-0.19

Table C15. Percentages of Students in the Analysis and Average District Performance Statistics by District Poverty and Performance in Grades 11–12 Reading and Science.

Subject	District Poverty Level (% Low-Income Students)	District Performance Level	Number of Students in Initial Grade	Number of Students in Analysis	Percent of Students in Analysis	Average District Performance Statistic
Reading	Lower (>20–50%)	Above average	15,482	7,059	46%	0.15
		Average	16,948	5,624	33%	-0.01
		Below average	2,650	1,203	45%	-0.20
	Medium (>50–70%)	Above average	6,700	2,850	43%	0.12
		Average	39,155	14,021	36%	-0.01
		Below average	2,096	1,030	49%	-0.12
	High (>70–100%)	Above average	#N/A	#N/A	#N/A	#N/A
		Average	14,095	4,390	31%	0.01
		Below average	521	200	38%	-0.18
Science	Lower (>20–50%)	Above average	10,797	4,170	39%	0.18
		Average	22,124	8,732	39%	0.02
		Below average	2,159	984	46%	-0.20
	Medium (>50–70%)	Above average	4,984	1,843	37%	0.15
		Average	39,980	14,883	37%	-0.01
		Below average	2,987	1,175	39%	-0.14
	High (>70–100%)	Above average	1,910	702	37%	0.12
		Average	11,909	3,669	31%	-0.01
		Below average	797	219	27%	-0.19

Appendix D

Fixed-Effect Coefficients from Statistical Models

Tables D1–D3 show the fixed-effect coefficients from the hierarchical models, measured in units of scale score points on the test used as the dependent variable. These are partial effects: For example, in Table D1, the fixed-effect coefficient for "low-income status" of -62.28 in fourth-grade literacy indicates that the predicted score of a low-income student is about 62 points (about one-third of a standard deviation) lower on the grade 4 literacy ABE than the predicted score of a non-low-income student who has the same values of the other variables in the model. No interaction effects were modeled (e.g., we did not model how a student's ethnicity might affect the differences in predicted scores between low- and non-low-income students).

The tables also show the standard deviation of the district performance statistics (random effects) in each model, labeled as "SD of random effects (pts)." This statistic estimates the variation across districts in the true random effect. For example, in Table D1, the "SD of random effects (pts)" is 29.0 for fourth-grade literacy. Given the assumed normal distribution of the random effect in the model, for approximately two-thirds of the districts, the absolute value of the random effect is 29 score points or less; for approximately 95% of the districts, the absolute value of the random effect is 58 score points or less.

We also converted these standard deviations, measured in score points, into standardized form by dividing them by the standard deviation of student scores on the test in question. This facilitates comparisons with the size of the district performance statistics shown in Tables 3–5. For example, for ABE grade 4 literacy, the standard deviation of 29.0 score points translates into a standardized standard deviation of 0.16. Thus, the absolute value of the random effect is 0.16 of a test score standard deviation or less in approximately two-thirds of the districts and 0.32 standard deviations or less in approximately 95% of the districts.

Table D1. Fixed-Effect Coefficients in Regressions Predicting Grade 4 Scores

Variable	Literacy		Mathematics	
Intercept	818.91	***	689.87	***
Low-income status	-62.28	***	-36.97	***
African American status	-60.05	***	-50.57	***
Hispanic status	-0.81		-6.42	***
Asian status	30.40	***	12.59	***
Native American status	-1.76		-4.40	
ELL status	-44.57	***	-18.54	***
Special education status	-170.42	***	-71.67	***
District % low-income	-0.39	**	-0.18	*
District % African American	-0.30	**	0.03	
District % Hispanic	0.21		-0.30	
District % Asian	-0.76		0.24	
District % Native American	-3.04	**	-0.70	
District % ELL	0.69		1.14	**
District % special education	0.97	*	1.12	***
District # students in model	-0.03		0.00	
District % students in model	-0.14		0.01	
Rural district	-5.49		-1.98	
Earlier record deleted	-93.42	***	-46.26	***
Took fourth-grade test in 2012	40.95	***	-0.81	
Took fourth-grade test in 2013	38.31	***	-1.12	
Took fourth-grade test in 2014	35.69	***	-15.73	***
SD of random effect (pts)	29.0		17.1	
SD of random effect (std)	0.16		0.17	

***Significant at the .01 level. **Significant at the .05 level. * Significant at the .10 level.

Table D2. Fixed-Effect Coefficients in Regressions Predicting Grade 8 Scores

Variable	English		Mathematics		Reading		Science	
Intercept	17.09	***	17.89	***	16.76	***	18.36	***
Low-income status	-1.65	***	-1.31	***	-1.40	***	-1.15	***
African American status	-1.86	***	-1.44	***	-1.63	***	-1.26	***
Hispanic status	-0.15	**	0.01		-0.03		0.07	
Asian status	0.81	***	0.89	***	0.90	***	0.81	***
Native American status	-0.33	*	0.07		-0.16		-0.09	
ELL status	-2.02	***	-1.24	***	-1.66	***	-1.24	***
Special education status	-3.49	***	-2.96	***	-2.71	***	-2.28	***
District % low-income	-0.01	**	-0.01	***	-0.01	**	-0.01	***
District % African American	-0.01	**	0.00		-0.01	***	0.00	
District % Hispanic	-0.01		0.02		0.00		0.00	
District % Asian	0.00		-0.01		-0.02		0.01	
District % Native American	-0.03		-0.02		-0.01		0.01	
District % ELL	0.03		0.00		0.01		0.02	
District % special education	-0.01		0.00		-0.02	*	-0.01	
District # students in model	0.00		0.00	*	0.00		0.00	
District % students in model	-0.01	**	-0.01	***	-0.01	**	-0.01	**
Rural district	-0.32	**	-0.07		-0.13		-0.07	
Earlier record deleted	-1.94	***	-1.68	***	-1.67	***	-1.37	***
Took eighth-grade test in 2012	0.22	***	0.21	***	0.31	***	-0.02	
Took eighth-grade test in 2013	0.41	***	0.16	***	0.46	***	0.29	***
Took eighth-grade test in 2014	0.27	***	-0.05		0.26	***	0.19	***
SD of random effect (pts)	0.47		0.49		0.45		0.42	
SD of random effect (std)	0.11		0.14		0.11		0.13	

***Significant at the .01 level. **Significant at the .05 level. * Significant at the .10 level.

Table D3. Fixed-Effect Coefficients in Regressions Predicting Grades 11–12 Scores

Variable	English		Mathematics		Reading		Science	
Intercept	24.64	***	23.05	***	24.27	***	23.47	***
Low-income status	-2.42	***	-1.57	***	-1.71	***	-1.45	***
African American status	-4.18	***	-2.78	***	-3.96	***	-3.17	***
Hispanic status	-1.36	***	-0.73	***	-1.44	***	-1.09	***
Asian status	1.30	***	1.98	***	0.82	***	1.17	***
Native American status	-0.98	***	-0.67	***	-0.60	*	-0.78	***
ELL status	-4.22	***	-2.19	***	-3.80	***	-2.45	***
Special education status	-6.36	***	-3.74	***	-4.99	***	-3.87	***
District % low-income	-0.02	**	-0.02	***	-0.02	***	-0.02	***
District % African American	0.01	**	0.01	**	0.01		0.01	**
District % Hispanic	-0.02		0.00		-0.03		-0.01	
District % Asian	0.06		0.13	**	0.08		0.12	**
District % Native American	-0.03		0.00		0.02		-0.03	
District % ELL	0.07	*	0.03		0.09	**	0.04	
District % special education	-0.02		-0.01		0.01		-0.01	
District # students in model	0.00	*	0.00		0.00		0.00	**
District % students in model	-0.02	***	-0.01	**	-0.02	***	-0.01	***
Rural district	-0.72	***	-0.32		-0.53	**	-0.36	*
Earlier record deleted	-2.61	***	-1.17	*	-2.05	**	-2.38	***
12th grader in 2012	-0.23	**	-0.12	*	-0.09		-0.26	***
12th grader in 2013	-0.22	**	-0.12		-0.13		-0.22	***
12th grader in 2014	-0.07		-0.08		0.19		-0.02	
SD of random effect (pts)	0.83		0.77		0.71		0.61	
SD of random effect (std)	0.13		0.15		0.11		0.11	

***Significant at the .01 level. **Significant at the .05 level. * Significant at the .10 level.

ACT is an independent, nonprofit organization that provides assessment, research, information, and program management services in the broad areas of education and workforce development. Each year, we serve millions of people in high schools, colleges, professional associations, businesses, and government agencies, nationally and internationally. Though designed to meet a wide array of needs, all ACT programs and services have one guiding purpose—helping people achieve education and workplace success.

For more information, visit www.act.org.

* 0 7 0 7 0 5 1 7 0 * Rev 1