Decision points and considerations for identifying rural districts that have closed student achievement gaps

Applied Research Methods

Michael J. Culbertson
Shelley H. Billig
RMC Research Corporation

Rural districts have long faced challenges in closing achievement gaps between subgroups of students. This brief describes key decision points and considerations for decisionmakers interested in identifying rural districts that have closed academic achievement gaps. Examining practices in these districts may suggest activities associated with making achievement gains and narrowing achievement gaps that can be systematically investigated. Examples from recent work with rural stakeholder groups in Colorado and Nebraska are used to highlight key issues in the identification process.

Why this study?

Rural districts face numerous challenges, including federal and state funding inequities, difficulty recruiting and retaining highly effective teachers, and problems coordinating resources stemming from geographic isolation (Alliance for Excellent Education, 2010; Ayers, 2011; Provasnik et al., 2007; U.S. Government Accountability Office, 2004). Many rural districts serve a high percentage of children from economically disadvantaged households, which impacts curriculum content, resource allocation, and student academic outcomes (Bouck, 2004). Rural districts have fewer resources than urban districts to address achievement gaps, including more limited funds, higher per pupil costs, fewer support services, and more instructors teaching outside their specialties. Rural teachers tend to have less competitive salaries and less evaluation support (Harmon, Gordanier, Henry, & George, 2007; Stephens, 1998).

Many challenges and barriers to success in rural education have been well documented in the research literature, but few studies identify how to improve achievement in rural settings, and very few specifically identify practices that close achievement gaps (Alliance for Excellent Education, 2010; Barley & Beesley, 2007). One starting point for examining practices that may be associated with closing achievement gaps

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences
U.S. Department of Education

REL CENTRAL
Regional Educational Laboratory
At Marzano Research

in rural districts is to identify districts that have narrowing achievement gaps and then ask stakeholders in those districts to develop hypotheses about what practices they believe are associated with closing achievement gaps in their district. These hypotheses can then be tested to gain understanding of effective strategies for closing achievement gaps in rural districts.

Identifying rural districts with narrowing achievement gaps requires careful consideration because of the demographic features and unique challenges of rural districts. For example, because rural districts in some states enroll fewer English learner students than urban districts do, examining achievement gaps for those students may not be meaningful for rural districts. Moreover, the smaller size of most rural districts creates statistical challenges not present in studies of achievement gaps with larger samples. As a result, measuring reductions in student achievement gaps in rural districts requires different procedures from those used in urban or suburban districts, in states, or in the nation as a whole.

This report describes key decision points and considerations for states to use in developing a distinctly rural and locally meaningful approach to identifying districts that have closed academic achievement gaps. Although the approach taken will depend on local conditions and priorities, the considerations and examples presented here may guide decisionmakers in developing a similar process for identifying rural districts in their own state that have experienced the greatest reductions in student achievement gaps.

Considerations for measuring reductions in achievement gaps in rural districts

This report outlines six key considerations for identifying rural districts with narrowing achievement gaps. Each consideration is accompanied by examples from work that Regional Educational Laboratory Central's Rural Education Research Alliance has conducted with decisionmakers (primarily state education agency staff and rural superintendents) in Colorado and Nebraska that illustrate how the process may be tailored to local contexts and priorities.

Which subgroup gaps should be examined?

A number of achievement gaps in rural areas may be of policy interest, so deciding which achievement gaps to examine is a key first step. Forty-four percent of rural students are eligible for the federal school lunch program (Southern Education Foundation, 2013), and students from low-income households are of interest to many policymakers (Center on Education Policy, 2011). English learner students tend to be more concentrated in some rural districts than in others, so whether the group is of interest may depend on the proportion of districts in a state that have high concentrations of English learner students.

Specific subgroups of interest can be chosen by examining the demographic distribution of rural students across a state, as well as by conducting a preliminary examination of state- or district-wide student achievement data for potential groups and measures. Decisionmakers in Colorado and Nebraska examined the gap between students from economically advantaged households and economically disadvantaged households. Colorado decisionmakers also identified English learner students as a subgroup of interest, while Nebraska decisionmakers noted that few rural districts in the state had substantial English learner student populations. In addition to these groups, decisionmakers in other states may be interested in examining gaps for particular racial/ethnic groups, such as Native American or Hispanic students, depending on their representation in the state.

What counts as rural?

Definitions of rural can range from precise statistical statements to personal subjective sensibilities (Rural Information Center, 2014). The definition of rural used for identifying reductions in achievement gaps

should be meaningful for local stakeholders to increase buy-in and enhance the relevance of the results. In some cases, using external definitions of rural may produce results that do not align with decisionmakers' understanding of rural education. For example, Nebraska stakeholders initially chose to follow the definition of rural used by the U.S. Census Bureau (Keaton, 2014; U.S. Office of Management and Budget, 2000) but were surprised that a number of districts that they considered rural were not represented in preliminary results because the districts were classified as towns under the Census Bureau definition. The Nebraska stakeholders decided to expand the list of districts for their study to include those that met the Census Bureau definition of rural or town, because this more closely reflected their sense of rural Nebraska. Because definitions of rural vary widely, reports on closing achievement gaps in rural districts should clearly state which definition is used and why.

Which measures should be used?

Although achievement gaps are often measured by student performance on standardized assessments of math and reading, decisionmakers may wish to expand the list of measures to provide a broader picture of student academic success, potentially including multiple, diverse measures of achievement, academic growth, and postsecondary readiness. However, any measure needs to meet several criteria to meaningfully add to the analysis of achievement gaps, including reflection of academic success, sensitivity to change, and wide availability. Issues related to size and access in rural schools can affect each of these three aspects of the suitability of measures.

Measures of achievement need to be consistently meaningful across the state or region being analyzed. If a given measure is not available in all areas or if it is not meaningful for all districts, analysis based on that measure will not yield a valid comparison of rural districts. For example, decisionmakers in Colorado and Nebraska initially wanted to include college and career readiness measures related to dual credit, reasoning that concurrent enrollment in college coursework must indicate stronger academic preparation. However, after additional reflection and discussion, they decided that dual credit did not uniformly reflect academic success in all districts because of considerable variability in the accessibility and grading practices of community colleges across the state.

Measures of achievement gaps must be sensitive to policy and instructional change. To close a gap, the achievement of the lower performing group must increase faster than the achievement of the higher performing group. But if a given measure is not sensitive to policy and instructional change, the measured achievement of the two groups, along with the corresponding gap, could remain static. For example, because the ACT measures college readiness very broadly, it is not particularly sensitive to many instructional changes. Although stakeholders in Colorado and Nebraska included ACT scores in their analyses, ACT scores did not exhibit enough change over time to allow the identification of reductions in student achievement gaps.

Alternate measures based on the same data source may have different sensitivity to change. For example, scaled scores from standardized assessments are more sensitive to modest improvements in academic achievement than proportions of students who achieve proficiency cutscores. However, using scaled scores requires that scores across grades be on a similar scale relative to proficiency in each grade. In Nebraska, state assessment scores are scaled so that the proficiency cutscores are the same in each grade. In other situations it may be necessary to standardize scores within each grade before aggregating them, which may not be possible with publicly available data. For example, limitations of publicly available data in Colorado necessitated the use of the proportion of students proficient or advanced in each content area.

Data expressed as a proportion may be insensitive to changes in achievement if the proportion is close to 0 percent or 100 percent. For example, stakeholders in Colorado were interested in including graduation

and dropout rates in the gap analysis; however, most rural districts in Colorado have achieved great success at increasing their graduation rates and lowering their dropout rates. When few students do not complete high school, there is little room to measure improvement.

Data availability concerns may severely restrict the number of alternative or innovative measures used to identify achievement gaps. Including measures that are available in only a subset of rural districts may be viewed as unfair, because districts that happen to have more measures have additional opportunities to demonstrate reductions in achievement gaps (and greater risk of not demonstrating reductions in achievement gaps). For example, decisionmakers in Colorado and Nebraska were interested in including data from the Northwest Evaluation Association Measures of Academic Progress because many districts in the two states use this product for interim or benchmark assessment and many educators view the results as particularly meaningful for their instructional practice. However, the measure was eventually dropped, in part because more than a third of the rural districts in each state did not administer the test every year. One reason standardized assessment data are so frequently used in analyzing achievement gaps is that the data are readily available for nearly all schools and districts annually. While decisionmakers in Colorado and Nebraska were eager to expand the list of measures, in the end their achievement gap analyses largely reflected traditional achievement gap measures based on state standardized test scores because of the many technical challenges associated with other measures.

What additional criteria should be considered?

Even if a particular achievement gap is meaningful for many rural districts in a region, the gap may not be meaningful for every rural district in the state. If a district has few students in a subgroup of interest, measures of achievement gaps for that subgroup may not be meaningful. When the proportion of students in the selected group varies across districts in a state, a threshold percentage of students may be desirable. For example, decisionmakers in Colorado and Nebraska considered economic achievement gaps only in rural districts where at least 35 percent of students were eligible for the federal school lunch program. Moreover, districts should not be identified when a higher performing subgroup's achievement decreases, resulting in a smaller difference between the two subgroups. Decisionmakers in Colorado and Nebraska determined that reductions in achievement gaps should not be at the expense of overall achievement levels and excluded districts with decreasing achievement. Similarly, the decisionmakers disqualified districts with low graduation rates or high dropout rates to ensure that apparent reductions in gaps were not due to increasing dropout among the lowest performing students.

How should reductions in achievement gaps be measured?

Changes in achievement gaps can be most simply measured by computing the difference in the size of the gap at the beginning and end of a given time period. However, the small student populations in most rural districts create substantial variability in district-wide scores over time. When changes are compared from one year to the next, achievement gaps in a small district may appear to close dramatically one year, only to widen just as dramatically the next. Choosing slightly different years to measure the change in achievement gaps could radically change the list of districts identified as having closed achievement gaps.

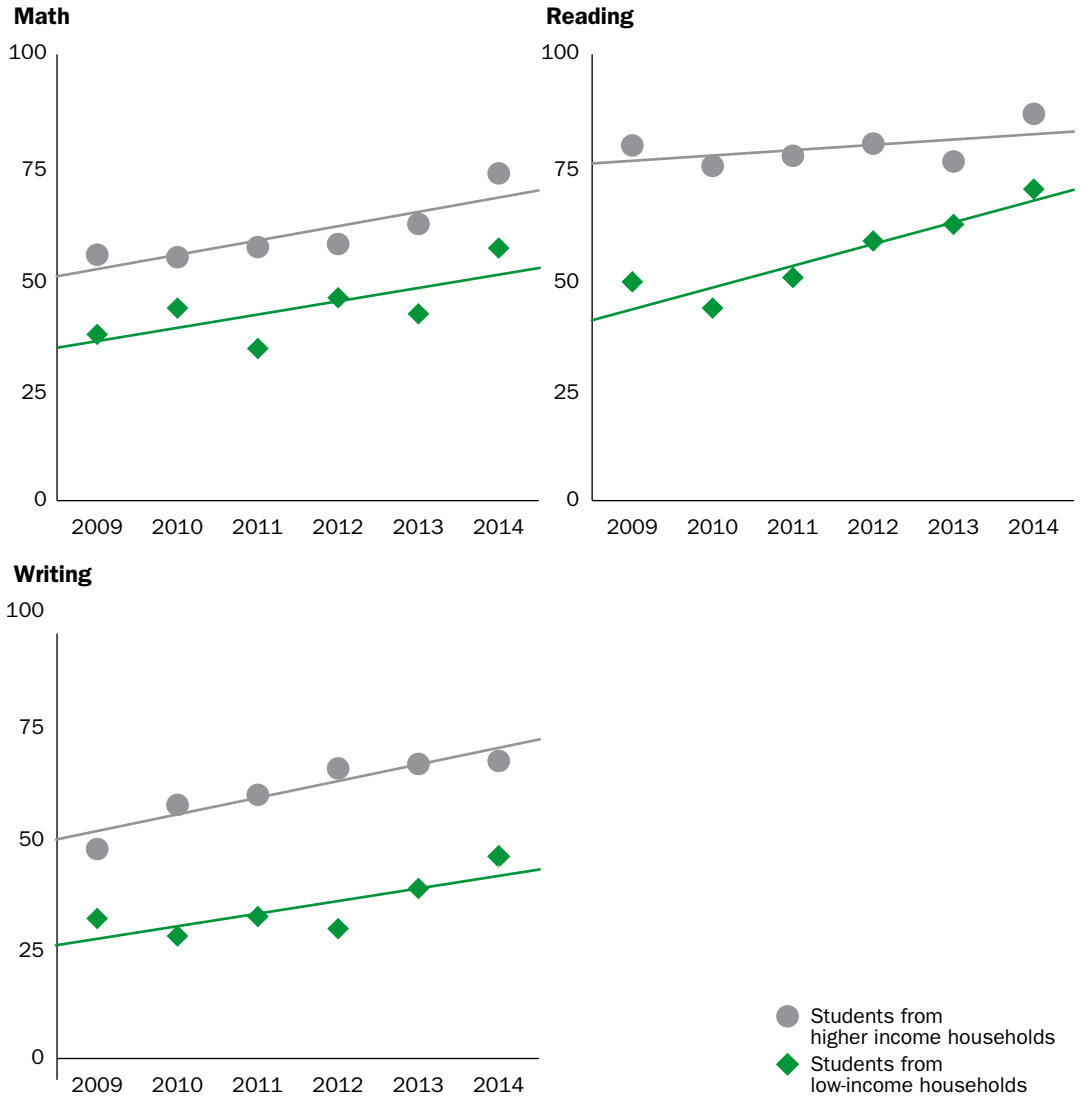
One strategy for overcoming the variability inherent in small districts' scores is to average achievement information across several years. While there are several ways such an average could be computed, the best-fitting trend line averages a district's scores for each group of students over an extended period in a visually intuitive way, to clearly communicate the achievement gap analysis to a wide audience. Graphs of the data and corresponding trend lines allow stakeholders without statistical training to intuitively understand how the analysis was conducted without needing to follow detailed statistical calculations, because

trend lines that are drawing closer together reflect achievement gaps that, on average, are closing over time. For example, while the achievement of students in one Colorado district increased in math, the gap in math performance did not close as much as the gap in reading performance (figure 1). The slope of the trend line for each group reflects the rate of change in that group's performance over time, and a district's achievement gap reduction score for a given measure is the difference in the slopes for each group, which reflects the rate of change in the achievement gap over time. Districts with a larger difference in trend line slopes have experienced greater change in the size of the achievement gap, on average, over a given time period.

Averaging over a longer period of time reduces year-to-year variability but can also obscure changes in the performance trajectory of a district. For example, if a district has decreasing achievement for five years followed by increasing achievement for five years, the long-term trend line will appear relatively flat, suggesting incorrectly that performance in the district was stable over the 10-year period. Averaging district

Figure 1. Trend lines for any given district average information across years to reduce variability in academic achievement measures from year to year

Students scoring proficient or advanced on Colorado achievement tests in a sample district (percent)



Source: Student achievement data from the Colorado Department of Education.

performance over time may also obscure the short-term effects of recent policy changes. However, given the small size of most rural districts, it may not be possible to distinguish the effects of a recent policy or practice change from random year-to-year variability, unless the policy's or practice's effects are examined over a longer period of time. The time period for the analysis should be long enough to provide stability to the estimate of the reduction in achievement gaps but not so long as to obscure changes in the trajectory of district performance. Colorado decisionmakers decided that six years of assessment data provided an adequate balance between sensitivity to policy changes and reducing year-to-year variability. Nebraska decisionmakers chose a different span: five years for reading, four years for math, and three years for science, based on the new state requirements for participation in common state assessments.

How should measures be aggregated to create a single ranking of districts?

Once the extent of the reduction in the achievement gap has been calculated for each measure, states may wish to aggregate across all measures to identify districts with the greatest overall reduction in the achievement gap for a given group. Two simple methods of aggregating gap reduction scores are counting and averaging.

Counting the number of measures for which a district has experienced a reduction in the achievement gap is a particularly useful way to aggregate the information across measures because it does not require the measures to be on the same scale and is easily understood by a broad audience. However, counting does not take into consideration the size of the reduction in the achievement gap—a district with a modest reduction in the achievement gap on two measures would appear the same as a district with a large reduction in the achievement gap on the same two measures. Counting also places districts without data on a given measure at a disadvantage relative to districts with complete data.

Averaging the gap reduction scores across measures can account for the size of the reduction in achievement gaps but requires the measures to be on the same scale so that one measure does not disproportionately affect the overall score. Standardizing the original measures or the gap reduction scores can account for differences of scale but increases the statistical complexity of the analysis. This may decrease transparency for stakeholders. Decisionmakers in Colorado and Nebraska chose a combination of counting and averaging: they ranked districts by the number of measures for which the achievement gap decreased and then broke ties based on the average size of the reductions in the achievement gap. Stakeholders may also be interested in weighting some measures more heavily than others to reflect the relative policy importance of different measures. For example, decisionmakers in Colorado chose a weighted average of growth and proficiency measures in which growth measures counted twice as much as proficiency measures to reflect the priorities of the state's district performance framework.

Decisionmakers may also be interested in examining rankings of districts for each measure separately, instead of aggregating across measures to identify districts that have closed achievement gaps across the board. For example, after examining rankings that aggregated across content areas, decisionmakers in Nebraska conducted a follow-up analysis to identify districts with narrowing achievement gaps in each content area separately, reasoning that the strategies used to close achievement gaps in reading may be different from those used for math.

Next steps

Identifying districts that have closed the achievement gap is only the first step in helping rural districts because the analysis does not explain why the reduction occurred. Changes in achievement gaps may be due to district policies and practices or due to factors not under the district's control, such as demographic

shifts. The next step is to systematically examine the policies and practices currently implemented in the districts that experienced the greatest reductions in achievement gaps to identify promising policies and practices for further study. For example, rural superintendents in Colorado and Nebraska who have closed the gap have presented information in webinars and conference sessions and have started to identify common organizational and instructional practices for further investigation through case study analysis or other systematic means. The activity itself has resulted in rich dialogue, and members of the Rural Education Research Alliance hope to develop a research agenda based on the hypotheses being generated.

References

Alliance for Excellent Education. (2010). *Current challenges and opportunities in preparing rural high school students for success in college and careers: What federal policymakers need to know.* Washington, DC: Author. http://eric.ed.gov/?id=ED509184

Ayers, J. (2011). *Make rural schools a priority: Considerations for reauthorizing the Elementary and Secondary Education Act.* Washington, DC: Center for American Progress. http://eric.ed.gov/?id=ED535987

Barley, Z. A., & Beesley, A. D. (2007). Rural school success: What can we learn? *Journal of Research in Rural Education, 22*(1), 1–16. http://eric.ed.gov/?id=EJ751592

Bouck, E. (2004). How size and setting impact education in rural schools. *The Rural Educator, 25*(3), 38–42. http://eric.ed.gov/?id=EJ783816

Center on Education Policy. (2011). *Is achievement improving and are gaps narrowing for Title I students? State test score trends through 2008–09, Part 4.* Washington DC: Author. http://eric.ed.gov/?id=ED522866

Harmon, H. L., Gordanier, J., Henry, L., & George, A. (2007). Changing teacher practices in rural schools. *The Rural Educator, 28*(2), 8–12.

Keaton, P. (2014). *Documentation to the NCES Common Core of Data Local Education Agency Universe Survey: School year 2011–12 provisional version 1a* (NCES No. 2014–035). National Center for Education Statistics Working Paper. Washington, DC: U.S. Department of Education.

Provasnik, S., KewalRamani, A., Coleman, M. M., Gilbertson, L., Herring, W., & Xie, Q. (2007). *Status of education in rural America* (NCES No. 2007–040). National Center for Education Statistics Working Paper. Washington, DC: U.S. Department of Education. http://eric.ed.gov/?id=ED497509

Rural Information Center. (2014). *What is rural?* Washington, DC: U.S. Department of Agriculture. Retrieved May 19, 2015, from http://ric.nal.usda.gov/what-is-rural.

Southern Education Foundation. (2013). *A new majority: Low income students in the south and nation.* Atlanta, GA: Author. http://eric.ed.gov/?id=ED555829

Stephens, E. R. (1998). *Expanding the vision: New roles for educational service agencies in rural school district improvement.* Charleston, WV: AEL, Inc. http://eric.ed.gov/?idED418822

U.S. Government Accountability Office. (2004). *No Child Left Behind Act: Additional assistance and research on effective strategies would help small rural districts* (GAO-04–909). Washington, DC: Author. http://eric.ed.gov/?id=ED483246

U.S. Office of Management and Budget. (2000). Standards for defining metropolitan and micropolitan statistical areas; Notice. *Federal Register* 65(249). Washington, DC: Author.

REL 2016–130

The National Center for Education Evaluation and Regional Assistance (NCEE) conducts unbiased large-scale evaluations of education programs and practices supported by federal funds; provides research-based technical assistance to educators and policymakers; and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

April 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0007 by Regional Educational Laboratory Central administered by Marzano Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Culbertson, M. J., & Billig, S. H. (2016). *Decision points and considerations for identifying rural districts that have closed student achievement gaps* (REL 2016–130). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Central. Retrieved from http://ies.ed.gov/ncee/edlabs.

This report is available on the Regional Educational Laboratory website at http://ies.ed.gov/ncee/edlabs.

The Regional Educational Laboratory Program produces 7 types of reports