

Chapter 3

Study Findings—Can Existing State Education Databases Be Used to Investigate Education Resources and Student Performance?

The purpose of this study was to explore the capacity of existing state data to answer policy questions about the relationship between instructional resources and student performance. In this chapter we discuss how researchers can use state education data to answer policy questions about three resource areas: (a) instructional expenditures, (b) staff characteristics and teacher quality, and (c) instructional staffing patterns. We also describe student performance data as a desired outcome measure when examining the effectiveness of resource allocation. We describe student, school, and district characteristics data that researchers must consider due to the influence of socioeconomic and other factors on the availability of resources and the level of student outcomes. This analysis represents a regional overview of the four states. State-by-state descriptions of state education data in Arkansas, Louisiana, New Mexico, and Texas are provided in appendixes B–E. We expect that this information will guide researchers and policymakers to make effective use of existing data to answer policy questions. SEDL also expects to use this information to plan and conduct new research to support the improved allocation of resources in diverse state and local environments.

SEDL researchers have particular interest in investigating policy issues in the southwestern region. Our focus for this study was to examine existing state data related to three broad areas of education resources: (a) the allocation of instructional dollars, (b) the importance of staff characteristics, and (c) the distribution of instructional staff. The allocation of instructional dollars addresses policy questions regarding the adequacy and equity of state funding formulas, compensation for teachers and administrators, and other fiscal considerations.

Policy questions about staff characteristics address how the quality of teachers and other staff relates to student success. Resource allocation research on the distribution of instructional staff investigates whether staff with different roles, qualifications, and demographics relate to student learning in different regions, districts, and schools. This chapter discusses the data available in state education databases that are relevant to each of these three resource areas. We provide an overview of overall data quality and discuss the questions that can be answered with existing state data as well as the questions that could be answered if data were improved. We also present the student outcome measures and demographic information available in the four study states.

Instructional Expenditures

Instructional expenditures are collected in state education databases and are generally defined as funds spent to support teaching and learning that occurs in the classroom, such as the cost of teachers, classroom materials and supplies, and contracted instructional services. Financial information can be used to examine spending patterns in instructional and other large expenditure categories such as general administration or support services (Pan et al., 2003).

We investigated whether education data in the four study states could reveal more detailed information regarding how funds are spent within the category of instruction than federal data allow; we found state financial data to be of great potential benefit. As listed in Table 1, in two states, Arkansas and Louisiana, expenditures are organized into function categories that align to federal (Census F-33 form) categories. New Mexico and Texas divide instructional expenditures into two separate function categories. In all four states, financial databases break down instructional expenditures into object-level categories for each school district (see Table 1). Although the objects used in each state vary, these can be loosely categorized into salaries, benefits, contracted services, supplies and materials, and other

expenditures. Using these data, researchers can investigate whether different allocation patterns of object expenditures are related to district performance. Some states have more refined fiscal data. In Texas, instructional object expenditures can be investigated at the school level, and in Texas, Louisiana, and Arkansas, instructional expenditures are further divided into programmatic categories (regular education, special education, etc.). Also, as described in greater detail in appendixes B–E, states have additional subfunction and/or subobject categories that allow even more refined measurement of instructional spending.

Table 1

Description of the Instructional Expenditure Data Collected by the Four Study States by Function, Object, and Unit of Analysis

	Arkansas ^a	Louisiana ^a	New Mexico	Texas
Instruction-related function categories	<ul style="list-style-type: none"> • Instruction • Student support • Instructional staff services 	<ul style="list-style-type: none"> • Instruction • Student support • Instructional staff services 	<ul style="list-style-type: none"> • Direct instruction • Instructional support 	<ul style="list-style-type: none"> • Instruction and instruction-related services • Instructional and school leadership • Support services—student
Object categories	<ul style="list-style-type: none"> • Salaries • Benefits • Professional purchased services • Supplies and materials • Other objects 	<ul style="list-style-type: none"> • Salaries • Benefits • Professional purchased services • Supplies and materials • Other objects 	<ul style="list-style-type: none"> • Personnel services • Employee benefits • Purchased services • Supplies and materials • Travel and training • Capital outlay 	<ul style="list-style-type: none"> • Payroll costs • Professional and contracted services • Supplies and materials • Other operating costs • Debt service • Capital outlay—land, buildings, and equipment
Unit of analysis	Program (for instruction only) District	Program District	District	Program School District

^aFunction and object categories align with federal functions and objects (Census form F-33).

Researchers also considered whether state education data could support an investigation of spending on individual staff salaries. In addition to the fiscal object category of salaries that is collected at the district and/or school level for all states, state education agencies collect information about the salary paid to each individual staff member (see Table 2). Also contained in staff databases are other characteristics of individual staff, such as the position a staff person holds (teacher, principal, etc.), education level, and demographic information. These data provide a rich source of information for researchers interested in analyzing how instructional staff salaries correspond to staff qualifications or other characteristics.

The benefit of having individual-level data on staff salaries is the flexibility they afford researchers to conduct analyses on subgroups of staff, such as classroom teachers, principals, or aides. They also allow researchers to aggregate salary information to multiple levels (district, school, and grade in some states). Data on the amount spent for individual staff benefits are also available in Arkansas state databases. In New Mexico and in Arkansas prior to 2003, an individual's contracted salary is recorded, but this amount cannot be linked with actual expenditure data recorded in the fiscal database. In Arkansas (since 2003), Louisiana, and Texas, individual staff salary information is recorded with financial function and object category labels so that this information aligns with aggregated salary information in the fiscal database.

Table 2

Staff Salary Data Available in the Four Study States, Including Salary Measure, Unit of Analysis, Partial Salary Determination, and Benefits

Salary Data	Arkansas	Louisiana	New Mexico	Texas
Salary measure(s)	• Total salary	• Base salary • Additional compensation (3 types)	• Base pay	• Base pay • Supplemental pay
Unit of analysis	Individual	Individual	Individual	Individual
Can partial salaries be determined for part-time staff?	Yes (since 2003)	Yes	Yes	Yes
Do salary data align with actual expenditures?	Yes (since 2003)	Yes	No	Yes
Are expenditures for benefits available at the individual level?	Yes	No	No	No

Data Quality. Data on instructional expenditures vary in quality from state to state. The information about each state's data systems provides detailed state-by-state descriptions of the quality of these data and are located in appendixes B–E as follows:

- Appendix B: Arkansas State Education Data
- Appendix C: Louisiana State Education Data
- Appendix D: New Mexico State Education Data
- Appendix E: Texas State Education Data

Overall, data that measure instructional spending and individual salaries have been consistent over time within each state. With the exception of changes to the Arkansas salary information, few substantive changes have been made in data variables during the study period (1999–2003). This consistency contributes to greater data reliability. State data managers identified few chronic problems in their financial data with regard to accuracy or completeness,

and districts have improved their data reporting systems over the years.¹ Other fiscal categories can be aligned only at the district level in all states except Texas, where fiscal data are reported at the school and program levels, and Arkansas and Louisiana, where district-level instructional expenditures are broken down into program categories. Salary data can be averaged across schools and districts for alignment with other staff data, student performance data, and other indicators.

Spending on staff benefits is recorded at the individual level only in Arkansas, and benefit costs must be prorated across individual staff to estimate total compensation costs in the other three states. Benefits could be estimated in the remaining three states by prorating the total expenditures that a district spends on benefits for instructional staff. This district-level average would need to be applied universally across all staff included in the function category. In most states, this minimally includes teachers, substitutes, and teacher aides, further decreasing the possibility of extracting a usable measure of benefits for any individual staff. Also, information about bonuses or other incentives for teachers in high-demand subject areas or geographic regions are not separated from base salaries except in Louisiana.

Questions That Can Be Answered. As shown in Table 3, policy research questions about instructional spending can be answered with existing state education data. State fiscal databases contain expenditure function categories that are familiar to state and local policy audiences, and researchers can examine how instructional dollars are allocated in districts of varying levels of student performance. Student performance scores are necessary for this analysis, and these data are available in each of the four study states, as discussed in greater detail beginning on page 38. Each state compiles test scores on criterion-referenced and/or norm-referenced exams for

¹ The identification of other important sources of error in these data such as inconsistency in coding data at the district level were beyond the scope of this study.

students in specific grades and subject areas. These test results can be requested from the state for use in resource allocation research. By merging student performance data and fiscal data, researchers can investigate which object-level expenditure categories are more likely to affect student performance. For these analyses, researchers can examine district-level spending in all four states and school-level spending in Texas. Researchers also can compare function-level expenditures to determine if districts of varying levels of performance allocate instructional and administrative dollars differently.

Table 3

Policy Questions About Instructional Resources That Can be Answered With Existing Data in the Four Study States, Including Relevant Variables and Level of Alignment

A. Do higher performing districts allocate more instructional dollars to salaries and benefits?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> Expenditure objects within instruction function Student performance scores 	<ul style="list-style-type: none"> Expenditure objects within instruction function Student performance scores 	<ul style="list-style-type: none"> Expenditure objects within direct instruction function Student performance scores 	<ul style="list-style-type: none"> Expenditure objects within instruction and instruction-related services function Student performance scores
Level at which data can be aligned	District	District	District	School District
Notes	Salaries and benefits constitute a large proportion of object expenditures in instruction, decreasing potential for identifying variation in spending between object categories.			
B. How do districts of varying levels of performance allocate administrative vs. instructional dollars?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> Expenditure functions: instruction, student support, instructional staff services, general administration, school administration Student performance scores 	<ul style="list-style-type: none"> Expenditure functions: instruction, student support, instructional staff services, general administration, school administration Student performance scores 	<ul style="list-style-type: none"> Expenditure functions: direct instruction, instructional support, administration Student performance scores 	<ul style="list-style-type: none"> Expenditure functions: instruction and instruction-related services, instructional and school leadership, student support services, administrative support services Student performance scores
Level at which data can be aligned	District	District	District	School District

Individual-level salary information provides an even greater potential for research and policy and can be applied to important questions. Table 4 displays the relevant variables available in the four states to address such questions. Existing data can investigate the relationship between teacher pay and such desired outcomes as increased student performance or the retention of qualified teachers. These data also can examine whether salaries are distributed equitably between schools and districts. State education data in all four study states allow researchers to investigate the relationship between teacher salaries and student performance at the district and school levels. Grade-level analysis could be conducted with Louisiana, New Mexico, and Texas salary and student performance data. Individual teacher- and student-level analysis could be conducted with New Mexico data.

Salary information also can be linked to teacher experience and mobility in order to examine the effect of pay on retention. Teacher mobility is not a standard variable in existing databases but can be calculated using information on the school and district assignment for teachers over multiple years. Those teachers who change schools or districts over time, or who leave the system altogether, can be recorded as mobile teachers; those who do not change can be recorded as retained teachers. Teacher salary, teacher experience, and school and district identifier data are collected on an individual level, so researchers can examine the relationship between salary and teacher retention at the individual, school, or district level. Teacher qualification data are collected in all four study states as well. These data can help researchers consider whether the relationship between salary and retention changes for teachers of varying education levels, years of experience, or certification status.

The distribution of salaries for teachers and other staff across schools and districts also can be examined using existing data. Distribution of other instructional expenditures (benefits,

supplies and materials, contracted services) can be examined at the district level for all states and at the school level in Texas. Further, these data can be grouped using characteristics such as high- or low-performing school districts to arrive at comparative information on instructional spending.

Table 4

Policy Questions About Salaries That Can be Answered With Existing Data in the Four Study States, Including Relevant Variables and Level of Alignment

C. How is teacher pay linked to student performance?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Total pay for teachers • Benefits for teachers^a • Student performance scores 	<ul style="list-style-type: none"> • Base pay or total pay for teachers • Student performance scores 	<ul style="list-style-type: none"> • Base pay for teachers • Student performance scores 	<ul style="list-style-type: none"> • Base pay for teachers • Student performance scores
Level at which data can be aligned	School District	Grade School District	Individual Grade School District	Grade School District
Notes	^a 2002–2003 data only			
D. What is the effect of salary on retaining qualified teachers?				
Relevant variables	<ul style="list-style-type: none"> • Total pay for teachers • Years of experience • Mobility of teachers (calculated) • Certification • Education level 	<ul style="list-style-type: none"> • Base pay or total pay for teachers • Years of experience • Mobility of teachers (calculated) • Certification • Education level 	<ul style="list-style-type: none"> • Base pay for teachers • Years of experience • Mobility of teachers (calculated) • Certification • Education level 	<ul style="list-style-type: none"> • Base pay for teachers • Years of experience • Mobility of teachers (calculated) • Certification • Education level
Level at which data can be aligned	Individual School District	Individual School District	Individual School District	Individual School District
Notes	Certification information is limited due to lack of issue dates.			
E. Are teacher salaries distributed equitably between schools and districts?				
Relevant variables	Total pay for teachers	Base pay for teachers	Base pay for teachers	Base pay for teachers
Level at which data can be aligned	School District	School District	School District	School District
Notes	Data from non-education databases could also be used to support this analysis, such as cost-of-living indicators.			

Questions That Could be Answered With Additional Data. Additional data that would enable policy researchers to address a wider range of questions regarding instructional spending include better measures for total compensation and professional development. As mentioned above, salary information is available for individual staff persons; however, data on dollars spent on benefits are available at the individual staff level in only one state. If all states were to collect individual-level data on the cost of benefits, bonuses, and incentives, researchers could calculate a more accurate estimate of total staff compensation. These more refined data also could be used to examine the influence of benefits and incentives on teacher recruitment or the successful use of incentives for hiring and retaining teachers in shortage areas (bilingual education, special education, etc.).

Professional development is another area of instructional spending and a policy issue about which decision makers need information. No fiscal spending category or individual staff data elements currently exist to help policy researchers understand how investments in professional development for teaching staff improve student outcomes. Researchers need actual or dollar-equivalent measures for teacher time, stipends, travel expenses, and costs for teacher substitutes related to professional development that can be prorated or directly assigned to individual teachers. The No Child Left Behind legislation emphasizes the importance of tracking the number of hours a teacher spends in professional development. Information on the content of professional development for teachers is also needed. Researchers could use these data to address questions about the effectiveness of professional development, its relative costs, and the distribution of professional development resources across schools and districts.

Staff Characteristics Data and Teacher Quality

Data on staff characteristics are important for policy audiences who need to understand the contribution of instructional and administrative staff in the learning process. A major emphasis within staff-related policy concerns is the qualifications of teachers. A number of critical policy questions need answers so that states can make research-based decisions to improve teacher quality. Questions that can be investigated using existing state data include those related to teacher experience, education, and certification.²

As displayed in Table 5, teacher characteristics data available across the four study states are relatively similar. For each of the states, researchers can measure basic demographics of teachers, educational attainment, years of experience, and scores on state teacher tests. Each of the four states also collects, minimally, data on teacher certification type, grade level, and subject area. All states except New Mexico also collect the issue date and expiration date of the certification, and Arkansas and Louisiana track national board certification. Arkansas, Louisiana, and Texas record teachers' route to certification (traditional or alternative). All states except Arkansas provide data to measure the proportion of a full-time equivalency (FTE) position a teacher holds. All four study states collect data on individual teachers, and these data can be linked to school and district identifiers for analysis at these aggregated levels. In New Mexico, teachers are linked to students through class codes and identification numbers.³

² These areas represent some important and easily quantifiable indicators of teacher quality. However, it is important to note that other measures of quality such as application of pedagogical techniques, teacher motivation, and classroom management skills are not collected via state databases.

³ Louisiana began linking individual students and teachers beginning in 2004-2005; however, these data are beyond the time period of this study and were not examined for this report.

Table 5

Teacher Characteristics Data Available in State Databases in the Four Study States, Including Characteristics Variables, Certification Data, and Level of Alignment

	Arkansas	Louisiana	New Mexico	Texas
Teacher characteristics	<ul style="list-style-type: none"> • Gender • Race/ethnicity • Birth date • Years of experience • Lowest and highest degree • Highest degree institution • Score on teacher tests 	<ul style="list-style-type: none"> • Gender • Race/ethnicity • Full-time equivalency of position • Years of experience • Highest degree institution • Score on teacher tests 	<ul style="list-style-type: none"> • Gender • Race/ethnicity • Birth date • Full-time equivalency of position • Years of experience • Highest degree institution • Bachelor’s degree and highest institution • Score on teacher tests 	<ul style="list-style-type: none"> • Gender • Race/ethnicity • Birth date • Full-time equivalency of position • Years of experience • Highest degree institution • Bachelor’s degree institution • Score on teacher tests
Teacher certification data	<ul style="list-style-type: none"> • Type • Grade level • Subject area • Effective date • National board • Route to certification 	<ul style="list-style-type: none"> • Type • Grade level • Subject area • Effective date • National board • Route to certification 	<ul style="list-style-type: none"> • Type • Grade level • Subject area 	<ul style="list-style-type: none"> • Type • Grade level • Subject area • Effective date • Route to certification
Level at which data are available or can be aligned	Individual School District	Individual School District	Individual School District	Individual School District

Data Quality. Data measuring staff characteristics are, for the most part, of sufficient quality to be used for policy research purposes. Data on teachers are more extensive than other staff data and include individual-level information on experience, education levels, certification status, and teacher test scores. Teacher data are relatively accessible by special request to state education agencies because individual staff data are protected with a lower level of confidentiality than individual student data.⁴ A number of challenges related to teacher data,

⁴ For those states that use Social Security numbers as unique identifiers, these numbers must be stripped or scrambled by the education agency before release to the outside users.

however, limit the usability of these data to answer policy questions. According to interviews with data managers across the four states, teacher experience data are of limited reliability due to reporting error at the level of local collection. Also, teacher certification information is problematic for three reasons:

1. Teacher certification data are collected and managed by a separate department from that which collects and manages other teacher data, increasing the probability of misalignment when merging these two sets of data. This problem is most apparent in states where different agencies use different unique identifiers or must scramble unique identifiers for outside data users.

2. Teacher certification data are collected in databases that are cumulatively updated without year-to-year archive information on teachers.

3. Certification requirements and teacher tests change frequently, making longitudinal analysis of certified teachers extremely difficult.

Policy Questions That Can Be Answered. As presented in Table 6, existing data on teachers can answer policy questions about teacher education, years of experience, certification, and teacher test scores. Teacher information is used in standard reporting by state education agencies, and teacher years of experience and education data are used to determine teacher salary levels. All four study states collect the total number of years a teacher has been in the profession, which can be used to examine the relationship between teacher experience and student achievement. The quality of these data varies from state to state and is dependent on unverified self-reports from either individual teachers or school districts. As summarized in Table 6, these data can be used to explore the relationship between teacher experience and student achievement. In order to address the relative unreliability of the self-reported experience information,

researchers may choose to crosscheck these data across a span of years. Since teachers and students are not aligned in state databases, except in New Mexico, researchers need to aggregate data to school or district levels in the remaining three states.

Researchers also can use state data to address the relationship between teacher education and student achievement (see Table 6). Teacher education variables in the four study states include highest degree and degree institution. Degree data usually include bachelor's, master's, and doctoral levels. Degree institution data are limited because in all four states only the names of in-state institutions attended are collected, and states use a universal label for out-of-state degrees.

Teacher certification is another important area of policy research that can be supported with state data, as shown in Table 6. Certification data in the four states are compiled in a separate database from other staff data and are managed by a department or agency whose prime responsibility is overseeing the licensure or certification of teachers. Measures available for teacher certification in each of the four study states include the certification type, subject area, and grade level. Arkansas, Louisiana, and Texas also track teachers' route to certification, and Texas provides an estimate of the number of teachers in each school who are teaching in their field of certification. This estimate of in-field teachers, however, is imprecise because elementary teachers do not teach specific subject areas and secondary teachers potentially teach multiple subjects. Also, the in-field determination does not consider teachers' degree major, which is not currently collected by any of the four study states. Arkansas and Louisiana also collect data on whether a teacher is national board certified. It is important to note that teacher certification data in New Mexico are limited with regard to their usability for policy research. As described above, teacher certification databases are organized in a cumulative fashion, so the

issue date and expiration date of certifications are essential for conducting analysis on these data. In New Mexico, the certification database lacks dates when certifications were granted or expired, making it impossible to identify whether or not a teacher was certified during any given study period.

Scores on teacher tests are also available in state databases for teachers in all study states. These data represent an array of test scores because teachers have had the option to take one of a number of tests or required tests have changed from year to year. Researchers can conduct analysis of subsets of teachers who took the same exam or explore the possibility of standardizing scores across different exams. Teachers in some states have no test score information if they were certified before testing requirements were instituted. Also, since teachers are assessed on a pass/fail basis and test takers may not be motivated to score as high as possible, test scores may reflect an inaccurate level of teachers' knowledge.

Table 6

Policy Questions About Teachers That Can be Answered With Existing Data in the Four Study States, Including Relevant Variables and Level of Alignment

F. What is the relationship between teacher experience and student achievement?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Years of experience in the profession • Student performance scores 	<ul style="list-style-type: none"> • Years of experience in the profession • Student performance scores 	<ul style="list-style-type: none"> • Years of experience in the profession • Student performance scores 	<ul style="list-style-type: none"> • Years of experience in the profession • Student performance scores
Level at which data can be aligned	School District	School District	Individual School District	School District
Notes	Experience data are not fully reliable due to self-reporting error. Researchers may choose to crosscheck these data across a span of years.			
G. What is the relationship between teacher education and student achievement?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Highest degree • Degree institution • Student performance scores 	<ul style="list-style-type: none"> • Highest degree • Degree institution • Student performance scores 	<ul style="list-style-type: none"> • Highest degree • Degree institution • Student performance scores 	<ul style="list-style-type: none"> • Highest degree • Degree institution • Student performance scores
Level at which data can be aligned	School District	School District	Individual School District	School District
Notes	<ul style="list-style-type: none"> • Degree major not available • Out-of-state degree institutions are not specified 			

H. What is the relationship between teacher certification and student achievement?

	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • National board certification • Route to certification • Student performance scores 	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • National board certification • Route to certification • Student performance scores 	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • Student performance scores 	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • Route to certification • In-field teacher • Student performance scores
Level at which data can be aligned	School District	School District	School District	School District
Notes			Certification status data are problematic due to lack of issue dates.	In-field teacher data are estimated and may not be a reliable measure.

I. What is the relationship between teacher test scores and student achievement?

	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Teacher test score • Student performance score 	<ul style="list-style-type: none"> • Teacher test score • Student performance score 	<ul style="list-style-type: none"> • Teacher test score • Student performance score 	<ul style="list-style-type: none"> • Teacher test score • Student performance score
Level at which data can be aligned	School District	School District	School District	School District
Notes	Teacher test score data are an inconsistent measure due to variability in tests over time and the pass/fail nature of the tests. This question can be answered using data on subsets of teachers, which may bias results.			

Data on teacher experience, education, and certification in the four study states can be merged with other variables to answer additional policy questions (see Table 7). Researchers can, for example, merge teacher characteristics and teacher salary information at the individual staff level in all four states and ask whether higher teacher salaries buy teachers with more experience, higher education levels, and advanced certification status. Researchers also can use teacher data to examine the distribution of teachers who were educated at different teacher education institutions or to determine whether rural areas have a higher rate of uncertified

teachers. Arkansas, Louisiana, and Texas data can inform studies regarding the relationship between teacher retention and route to certification.

Table 7

Additional Policy Questions About Teachers That Can be Answered With Existing Data in the Four Study States, Including Relevant Variables and Level of Alignment

J. Do higher teacher salaries buy teachers with more experience, higher education levels, and advanced certification status?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Total pay for teachers • Years of experience in the profession • Highest degree • Certification type, subject area, and grade level • National board certification • Student performance scores 	<ul style="list-style-type: none"> • Base pay for teachers • Years of experience in the profession • Highest degree • Certification type, subject area, and grade level • National board certification • Student performance scores 	<ul style="list-style-type: none"> • Base pay for teachers • Years of experience in the profession • Highest degree • Certification type, subject area, and grade level • Student performance scores 	<ul style="list-style-type: none"> • Base pay for teachers • Years of experience in the profession • Highest degree • Certification type, subject area, and grade level • Student performance scores
Level at which data can be aligned	School District	School District	Individual School District	School District
Notes	Certification status data are problematic due to lack of issue dates. Experience data are not fully reliable due to self-reporting error. Researchers may choose to crosscheck these data across a span of years.			

K. What is the pattern of distribution of teachers who were educated at different teacher education institutions?

	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Degree institution • School code • District code 	<ul style="list-style-type: none"> • Degree institution • School code • District code 	<ul style="list-style-type: none"> • Degree institution • School code • District code 	<ul style="list-style-type: none"> • Degree institution • School code • District code
Level at which data can be aligned	School District	School District	School District	School District
Notes	Data do not designate specific out-of-state degree institutions but use a generic code for all such institutions.			

L. Do rural areas have a higher rate of uncertified teachers?

	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • National board certification • Locale code 	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • National board certification • Locale code 	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • Locale code 	<ul style="list-style-type: none"> • Certification type, subject area, and grade level • Locale code
Level at which data can be aligned	School District	School District	School District	School District
Notes	<p>Certification status data are problematic due to lack of issue dates.</p> <p>Locale codes for schools and districts are located in federal Common Core of Data.</p>			

M. What is the relationship between teacher retention and route to certification?

	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • School code (multiple years) • District code (multiple years) • Route to certification 	<ul style="list-style-type: none"> • School code (multiple years) • District code (multiple years) • Route to certification 	<ul style="list-style-type: none"> • N/A 	<ul style="list-style-type: none"> • School code (multiple years) • District code (multiple years) • Route to certification
Level at which data can be aligned	School District	School District	N/A	School District
Notes	<p>Teacher mobility/retention can be inferred by comparing teachers' school and district assignment codes over multiple years.</p> <p>During the study period New Mexico did not have an alternative route to certification.</p>			

Questions That Could be Answered With Additional Data. Improved and expanded collection of teacher characteristics would improve policy research on teacher issues. For example, although teacher experience is measured in all four study states, these states could improve the reliability of the measures for research use by helping districts better understand reporting definitions of this variable—especially for teachers who transfer between districts. Also, none of the four study states collect information to estimate accurately whether a teacher is teaching in-field based on his or her degree major. This measure is a critical component of “highly qualified teachers” as defined by the No Child Left Behind Act. If states would collect data on teachers’ degree major, researchers could align these data with the subject area the teacher is certified to teach and the subject the teacher actually teaches to estimate the number, distribution, and effectiveness of in-field teachers.

Teacher retention and mobility are important measures that researchers currently can calculate using teacher school and district assignment codes over time. If states included a teacher mobility measure in state education databases, this would provide a consistent measure for researchers and would reduce error and variability in the way outside data users calculate mobility. Teacher test score data are limited for research purposes and must be used with caution for reasons beyond data collection and management. These data reflect multiple teacher tests that have changed over time, making comparisons between teachers difficult except for subsets of teachers and for specific years. The pass/fail scoring structure also reduces the accuracy of teacher test scores as a measure of teachers’ knowledge. This measure could be improved by creating a universal standard for teacher tests and by adding incentives for teachers to score as high as possible on the tests.

Instructional Staffing Patterns

As discussed above, policy questions about teacher resources can be investigated using existing state education databases. As presented in Table 8, staff counts and ratios are collected by the four study states and can be used to inform policy. The states also collect detailed information on all classified and certified staff, including salaries, position, years of experience, gender, and race/ethnicity. Other characteristics that are collected on all staff in one or more of the four study states include educational attainment and degree institution. Staff databases in Louisiana, New Mexico, and Texas also include information about the full-time equivalency (FTE) of staff positions, allowing researchers to determine staff allocations with more precision than staff head counts permit. In Arkansas, Louisiana, and Texas, for individuals with multiple job positions, researchers can determine what proportion of an individual salary compensates each position. Individual staff records also contain school assignment, making it possible for researchers to build staff profiles for a specific school or subgroup of schools. School-level staff ratios can be calculated or combined with student enrollment data to create student and staff ratios (pupil:teacher, pupil:administrator, teacher:administrator, teacher:aide, etc.).

Class size is an important data variable for conducting policy research. Policy audiences need to understand the relative cost of reducing class size and its relationship to student performance. Class size information can be estimated in a variety of ways with varying accuracy. A calculation of the ratio between the number of students in a school and the number of teachers in the school is the least accurate measure of class size due to (a) the probability that not all teachers in a school are regular classroom teachers, (b) the fact that in most secondary schools teachers are assigned to multiple classes, and (c) the possibility that team teaching or other nontraditional classroom structures exist. As shown in Table 8, in Arkansas and Texas existing

data allow researchers to calculate pupil:teacher ratios (Texas also calculates a pupil:class ratio in secondary schools). However, in Arkansas this ratio must be calculated with a head count of teachers, rather than the more refined full-time equivalency (FTE) count of teachers. A more accurate measure of class size is available in Louisiana, where local school districts track and report actual class sizes to the state. In New Mexico, the databases are structured so that researchers, theoretically, can link teachers to the students they teach in specific classes to calculate a relatively accurate class size measure.⁵

⁵ As of this writing SEDL researchers could not test whether class size could actually be computed due to unavailability of data from the New Mexico Public Education Department.

Table 8

Information on Classified and Certified Staff Available in State Databases for the Four Study States, Including Counts, Staff Categories, Characteristics, Class Size Estimates, and Ratios

	Arkansas	Louisiana	New Mexico	Texas
Staff counts	<ul style="list-style-type: none"> • Head count 	<ul style="list-style-type: none"> • Head count • Full-time equivalency (FTE) by position 	<ul style="list-style-type: none"> • Head count • Full-time equivalency (FTE) by position 	<ul style="list-style-type: none"> • Head count • Full-time equivalency (FTE) by position
Type of staff included	<ul style="list-style-type: none"> • Classified and certified staff 	<ul style="list-style-type: none"> • All staff 	<ul style="list-style-type: none"> • Classified and certified staff 	<ul style="list-style-type: none"> • Classified and certified staff
Staff characteristics	<ul style="list-style-type: none"> • Position • Years of experience • Education level • Degree institution • Gender • Race/ethnicity • Birth date 	<ul style="list-style-type: none"> • Position • Years of experience • Education level • Gender • Race/ethnicity 	<ul style="list-style-type: none"> • Position • Years of experience • Education level • Degree institution • Gender • Race/ethnicity • Birth date 	<ul style="list-style-type: none"> • Position • Years of experience • Education level • Gender • Race/ethnicity
Class size estimate	Pupil:teacher ratio can be calculated using teacher head count, not full-time equivalency count.	Class size ranges are reported at the school level.	Potential to compute class size using individual-level student, teacher, and class data	Pupil:teacher ratio and pupil:class ratio (secondary) are available in state data.
Staff ratios	Can be calculated (using head count not full-time equivalency count): <ul style="list-style-type: none"> • Pupil:teacher • Pupil:administrator • Teacher:administrator 	Can be calculated: <ul style="list-style-type: none"> • Pupil:teacher • Pupil:administrator • Teacher:administrator 	Can be calculated: <ul style="list-style-type: none"> • Pupil:teacher • Pupil:administrator • Teacher:administrator 	Can be calculated: <ul style="list-style-type: none"> • Pupil:teacher • Pupil:administrator • Teacher:administrator

Data Quality. Staff data in all four study states are collected at the individual level and are of sufficient quality for research use. Individual-level data can be requested from state education agencies, and aggregated reports of these data are available from state agency Web sites. Data collection procedures for staff data have been in place for many years and have not changed substantially during the study period (1999–2003) except in Arkansas. The consistency and relatively clear and detailed reporting instructions for districts contribute to the quality of

staff data. Since data are collected for individual staff and relevant school or district labels are attached to these data, they provide great potential for aligning to student performance data, student characteristics, and school and district characteristics. Researchers also must realize the complexity of staff data. Understanding, for example, how teachers with multiple roles and responsibilities are recorded in each state's databases requires careful study of the data documentation and/or conversations with state data managers.

Questions That Can be Answered. The four study states all collect individual-level staff data that provide flexibility for policy research about staffing patterns. Using staff full-time equivalency (FTE) counts, researchers can estimate staff resources by position at every school. Combined with student performance data, researchers can investigate how staff resources differ across schools with varying student performance levels (see Table 9). Such analyses can be conducted at the school or district level. Researchers also can use demographic data that are collected in each of the four study states (for a description of available demographic data, see the discussion beginning on page 38) and staff characteristics, such as years of experience or race/ethnicity, to better understand the distribution of staff. For example, researchers can examine the distribution of teachers and administrators with varying years of experience in small, medium, and large schools.

Researchers also can use pupil:staff ratios to understand whether allocating more administrators, teachers, or aides is related to student performance (see Table 9). For example, all four study states collect data to study how class size relates to student performance. As discussed above, estimates for class size would vary from pupil:staff ratios to actual reported class sizes, depending on the data available in each state.

Table 9

Policy Questions About Staff Resources That Can be Answered Using State Data in the Four Study States, Including Relevant Variables and Level of Alignment

N. How do staff resources differ across schools with varying student performance levels?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Staff head count • Staff position code • Student performance scores 	<ul style="list-style-type: none"> • Staff full-time equivalency (FTE) count • Staff position code • Student performance scores 	<ul style="list-style-type: none"> • Staff full-time equivalency (FTE) count • Staff position code • Student performance scores 	<ul style="list-style-type: none"> • Staff full-time equivalency (FTE) count • Staff position code • Student performance scores
Level at which data can be aligned	School District	School District	School District	School District
Notes	FTE data are not available.			
O. What is the distribution of teachers and administrators with varying years of experience in small, medium, and large schools?				
	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Teacher head count • Administrator head count • Years of experience • School enrollment 	<ul style="list-style-type: none"> • Teacher full-time equivalency (FTE) count • Administrator full-time equivalency (FTE) count • Years of experience • School enrollment 	<ul style="list-style-type: none"> • Teacher full-time equivalency (FTE) count • Administrator full-time equivalency (FTE) count • Years of experience • School enrollment 	<ul style="list-style-type: none"> • Teacher full-time equivalency (FTE) count • Administrator full-time equivalency (FTE) count • Years of experience • School enrollment
Level at which data can be aligned	School	School	School	School

P. What is the relationship between class size and student performance levels?

	Arkansas	Louisiana	New Mexico	Texas
Relevant variables	<ul style="list-style-type: none"> • Teacher head count • Enrollment • Student performance scores 	<ul style="list-style-type: none"> • Class size range • Student performance scores 	<ul style="list-style-type: none"> • Student, teacher, and class ID • Student performance scores 	<ul style="list-style-type: none"> • Pupil:teacher and pupil:class ratios • Student performance scores
Level at which data can be aligned	School District	School District	Class School District	School District
Notes	<ul style="list-style-type: none"> • Class size is not estimated in state data • Pupil:teacher ratio can be calculated 	<ul style="list-style-type: none"> • Class size ranges: 1–20, 21–26, 27–34, more than 34 	<ul style="list-style-type: none"> • Class database aligns student, teacher, and class ID that can be aggregated to estimate class size 	<ul style="list-style-type: none"> • Pupil:teacher (elementary) and pupil:class (secondary) ratios are in state data

Questions That Could be Answered With Additional Data. The relationship between class size and student performance is of interest at both the state and federal levels of education decision making. If data were collected that link students to teachers and teachers to specific classes in all four study states, this policy issue could be investigated more accurately.

Student Performance Data and Student, School, and District Characteristics

Student performance data and demographic data on students, schools, and districts are critical elements for conducting research on resource allocation and student performance. Student performance data provide important measures of school and district success. Policy analysis that includes student achievement data can uncover resource strategies and inputs that relate to improved student performance. Demographic information is useful for two reasons. First, demographic characteristics such as student poverty and minority status are known to influence student performance measures and must be accounted for when investigating the relationship between resources and student performance. Second, demographic characteristics

can help researchers identify—and policymakers understand—specific resource needs, circumstances, and barriers faced by schools and districts with varying wealth, geographic characteristics, and student populations.

Student performance data described in this report include only scores on student achievement tests compiled and managed by state education agencies in each of the four study states (see Table 10). Other measures of student performance such as attendance and graduation rates are not discussed in this report. Student achievement testing generally reflects instructional goals and priorities established by state policymakers and thus vary from state to state. States use norm-referenced exams, criterion-referenced exams, or both for student accountability purposes. Standardized test scores are available for varying grades and subject areas, depending on state requirements. All states have tests for core subject areas (math, reading or English language arts) and other subjects such as science and social studies. In all four states, tests have changed over the years and grades tested have changed, making longitudinal comparisons difficult over long time spans. Student test scores are recorded for each individual student tested. Additionally, basic demographic information about the student, such as poverty status and race/ethnic group, is attached to performance records. Scores also can be aggregated to grade, school, and district levels.

Table 10

Student Achievement Tests Used in the Four Study States by Test Name and Grades and Subject Areas Tested

	Arkansas	Louisiana	New Mexico	Texas
Tests	<ul style="list-style-type: none"> • State benchmark tests (criterion-referenced) • Stanford Achievement Test, 9th edition (SAT-9) (norm-referenced) 	<ul style="list-style-type: none"> • Louisiana Educational Assessment Program for the 21st Century (LEAP 21) (criterion-referenced) • Iowa Tests of Basic Skills (ITBS)/Iowa Tests of Educational Development (ITED) (norm-referenced) 	<ul style="list-style-type: none"> • Terra Nova (norm-referenced) 	<ul style="list-style-type: none"> • Texas Assessment of Academic Skills (TAAS) (criterion-referenced) • Texas Assessment of Knowledge and Skills (TAKS) (criterion-referenced)
Grades	<ul style="list-style-type: none"> • 4, 6, 8—State benchmark tests • 5, 7, 10—Stanford Achievement Test, 9th edition (SAT-9) 	<ul style="list-style-type: none"> • 4, 8—Louisiana Educational Assessment Program for the 21st Century (LEAP 21) • 3, 5, 6, 7, 9—Iowa Tests of Basic Skills (ITBS)/Iowa Tests of Educational Development (ITED) • 10, 11—Graduation Exit Exam 	<ul style="list-style-type: none"> • 3–9—Terra Nova 	<ul style="list-style-type: none"> • 3–8, 10—Texas Assessment of Academic Skills (TAAS) • 3–11—Texas Assessment of Knowledge and Skills (TAKS)
Subjects	<ul style="list-style-type: none"> • Math, reading, writing—State benchmark tests • Complete battery—Stanford Achievement Test, 9th edition (SAT-9) 	<ul style="list-style-type: none"> • English language arts, mathematics, science, social studies—Louisiana Educational Assessment Program for the 21st Century (LEAP 21) • Reading, language, mathematics, science, social studies—Iowa Tests of Basic Skills (ITBS)/Iowa Tests of Educational Development (ITED) 	<ul style="list-style-type: none"> • Math, reading, language arts, science, social studies 	<ul style="list-style-type: none"> • Math, reading, writing, science, social studies
Notes	<ul style="list-style-type: none"> • For the State Benchmark test, 4th grade was added in 1999, 8th grade was added in 2000, and 6th grade was added in 2001. 	<ul style="list-style-type: none"> • Testing of different grade levels and subject areas have been phased in gradually since 1999–2000. 	<ul style="list-style-type: none"> • The Terra Nova changed formats in 2002–2003. • The state plans to replace the Terra Nova with a criterion-referenced exam (grades 3–9, 11) by 2005–2006. 	<ul style="list-style-type: none"> • The Texas Assessment of Knowledge and Skills replaced the Texas Assessment of Academic Skills in 2002–2003.

Data Quality. Due to increased attention to accountability needs, all four study states have been improving their capacity to measure student performance through standardized tests. While this process has improved the quantity of these data with regard to the number of tests offered and the grades tested, it also has resulted in inconsistency in the test scores available from year to year. In each of the four study states, changes were made during the study period to the tests administered, grades tested, or scoring standards. These changes are important for researchers to consider, especially in applying these data to analysis across time. Data are reported through state agency Web sites in all study states except New Mexico. These publicly available data, however, are limited for use in policy research because they are not disaggregated to reflect individual level scores or, in some states, subgroup populations. These data also are reported using aggregated scoring formats (e.g., percentage of students scoring at or above proficient level) of limited value to researchers. Researchers can request data from state departments of education that provide more refined measures of student achievement than those available on agency Web sites.

Student, School, and District Characteristics. State education databases contain student, school, and district demographic data. Key demographic measures in the four states are listed in Table 11. Student poverty is measured by student participation in the free and reduced-price lunch program. All study states except New Mexico have collected these data in a consistent manner during the study period. Student race/ethnicity, gender, special education status, and participation in limited-English-proficient or gifted and talented programs are collected in education databases in all four states. School characteristics available in state databases include the type of school, the grades taught, student enrollment, school accountability ranking, and aggregated student characteristic data such as percentage of minority or high-poverty students.

District characteristics of interest to policy researchers include district wealth, geographic locale, district enrollment, and district/school accountability rating. District wealth is calculated as part of state funding formulas and is available for all four study states. Geographic locale (urban, rural, suburban, etc.) is available for Texas districts in state databases. A more standardized source of this information, however, is located at the federal level. The National Center for Education Statistics Common Core of Data produces a geographic locale label for each school and district in all states in the nation.

Table 11

Unit of Analysis Possible for Student, School, and District Characteristics in the Four Study States

Measure	Arkansas	Louisiana	New Mexico	Texas
Poverty status (free and reduced-price lunch program participation)	Grade School District	School District	Grade School District	Grade School District
Race/ethnicity	Grade School District	School District	Grade School District	Grade School District
Gender	Grade School District	School District	Grade School District	Grade School District
Special program participation (special ed., bilingual, Title I, etc.)	Grade School District	School District	Grade School District	Grade School District
School type (elementary, high, etc.)	School	School	School	School
Grade range	School	School	School	School
Total enrollment	School District	School District	School District	School District
Per pupil expenditure	District	School (estimated) District	District	School District
District wealth	District	District	District	District
Geographic locale	N/A	N/A	N/A	District
Accountability ranking	N/A	School District	School	School District

Note. Free and reduced-price lunch data in New Mexico were reported inconsistently from 1999–2003.

Community Characteristics. Education databases do not collect information about characteristics of the community within which a school and/or district is located. Census information provides a number of relevant indicators for community well-being, including family income, parent education levels, and receipt of public assistance.

Summary

SEDL researchers found that existing state data can be used for education policy research on instructional resource allocation. Dollars spent to support instruction can be examined using fiscal data broken down by function, object, and program (in three of four study states) categories. These dollars also can be examined through staff databases that contain individual-level staff salaries. The added benefit of using individual staff salary data is the possibility of linking salaries to staff characteristics, such as years of experience, highest degree, and certification information. These staff characteristics, however, must be used with caution due to data quality concerns. Researchers should check staff years of experience data over multiple years to detect and correct errors. Also, since certification databases are managed separately from other staff characteristics, these data must be merged carefully with other staff data to avoid misaligned data. Certification data are also limited because data are continually updated in certification databases, and year-to-year snapshots of certified staff must be derived from certification issue dates (if available). Staff resources can be studied using state databases, and full-time equivalency (FTE) counts (head counts in Arkansas) on a wide range of staff categories are available in staff databases. These counts can be matched with staff demographic or other characteristics or calculated to create ratios such as pupil:teacher or teacher:administrator. Actual class size information, however, is reported only in Louisiana.

Student performance data and information on student, school, and district characteristics are essential for conducting research on resources and student performance. Student performance data in each of the states are unique and have undergone changes in years tested and/or the tests administered, creating difficulty for researchers to attempt regional comparisons or longitudinal analysis. Student performance scores are useful measures because they can be aligned to student characteristics and grade- or school-level resource data. Additionally, these data reflect measures that are familiar to state policy audiences. Student, school, and district characteristics are available in education databases in all four study states. These data are of reliable quality with few exceptions in some states and are of critical value in understanding the relative importance of student, school, and district environments on resources and student performance.

Policy questions of interest to the southwestern region and nationally can be answered using existing state education databases. By merging data on resources, student performance, and student, school, and district characteristics, researchers can answer policy questions such as the following:

- What are effective allocation patterns of instructional dollars in schools and districts with varying characteristics?
- Is teacher pay linked to teacher quality and student performance?
- What are the recommended distribution patterns of staff with differing characteristics?
- Should resources be invested to lower class size?

Although existing state education data offer great potential for informing policy through research, these data also pose opportunities for state policymakers to improve them for future research. Fiscal data could be improved by adding a measure for individual staff benefits in three states (Arkansas already collects these data) so researchers could consider the total compensation

package paid to teachers and other staff. States should collect information on professional development for teachers. Data on professional development spending would greatly increase researchers' ability to understand what investments in this area are effective. In order to better address priorities raised by the federal No Child Left Behind legislation, states could improve the scope and quality of teacher data. State policymakers and data managers should work to improve the reliability of staff experience data and the usability of teacher certification data. In order to better measure teacher quality, states should collect information about teacher degree major and professional development and should standardize data on teacher test scores. To understand the relationship between class size and student performance, researchers need more accurate estimates of actual class size that go beyond pupil:teacher ratios.

Implications for Future Research. Based on this study of existing state education data, we conclude that there is a need for increased attention on and use of these data for policy research purposes. Key features of the fiscal, staff, and student data we identified in this study helped frame SEDL's next resource allocation study. Student performance data are available in the form of criterion-referenced and norm-referenced test scores in all four study states, which provide outcome measures for a resource allocation study. This study also identified teacher salary data as a reliable and flexible subset of state financial data that all states collect at the individual level. Teacher characteristics collected in state databases are critical to understanding salaries and student performance. Several key measures of teacher qualification are available in all four of the study states, including years of experience, highest degree, degree institution, and certification. Finally, data on the socioeconomic environments where teaching and learning occur are collected by state education agencies in the form of student demographics and school and district characteristics. Based on this understanding of data available in existing state databases, we

propose to focus our next study on the question “What is the relationship between teacher salary levels and student achievement in high-need schools?” This study will make use of existing data and provide critical information to state and local policy audiences regarding the allocation of teacher resources.