# Texas Smart Schools Methodology 2019

## Executive Summary

## Contents



# Texas Smart Schools Methodology 2019

## Executive Summary

### Apples to Apples

The key to understanding and improving schools and school districts is being able to make fair and helpful comparisons among them. It is not especially useful to tell schools with a high proportion of English Language Learners (ELL) to emulate the best practices of schools that do not have any ELL students. It also does not help much to tell a district with fewer than 1,000 students to adopt staffing practices that work well in a district with 75,000 students. For schools to improve, they need to learn from the best practices of similarly situated schools, not be told to adopt some one-size-fits-all notion of the academic ideal.

Two key dimensions for comparison include academic progress and real expenditures. The Texas Smart Schools (TSS) methodology aims to identify schools and districts where:

- Students perform better than would be expected given their demographics and previous performance;
- Educational expenditures are lower than would be expected given their cost environment.

Schools and districts that stand out in both dimensions—high performing and low spending—are highlighted as best practitioners.

Because raw data seldom provide sufficient insight for effective decision-making and differences in educational context have to be taken into consideration to transform data into information, Texas Smart Schools uses recognized statistical methods to create better, more apples-to-apples, comparison measures for spending and academic growth. Those methods are described below.

### Academic Progress Measures

Capturing how much schools and districts have contributed to the learning growth of their students is as complex as the number and variety of students served by those schools. Some students test very well with little effort, while other students struggle. If a very high-achieving student and a very low-achieving student each bring home the identical average test result, it would be a source of concern for one of them and a reason to celebrate for the other. There are many non-school factors—like family income, language proficiency, and prior achievement—that also contribute to student performance. It can be challenging to separate the school effects from non-school effects.

For the Texas Smart Schools academic progress calculations, we follow the scholarly literature by using a value-added model to generate our academic progress measures. A value-added model measures the extent to which student performance in a school (or district) differs systematically from what would have been expected had the students attended school somewhere else. Schools where students perform better than expected, given their prior performance and demographic characteristics, have high academic progress. Schools where students perform worse than expected have low academic progress. For example, a school where none of the students are passing standardized tests would have high academic progress if the students are improving more rapidly than similar students elsewhere in the state. Similarly, a school where all of the students are passing standardized tests would have low academic progress if the students are failing to improve as much as their peers in other schools.

To generate the academic progress measures, we use HLM (a form of regression analysis) to predict the math and reading performance of each individual student. Those predictions are based on the personal characteristics and prior performance in reading and math for each individual student. The demographic characteristics include sex, race, socio-economic status, limited English proficiency (LEP) status, special education status, and grade level. Furthermore, the prediction model allows for interactive effects, so that being both economically disadvantaged and LEP can have a compound effect on student performance, above and beyond the impact of either one alone. Students who are excluded from the "accountability subset" when the Texas Education Agency (TEA) calculates its performance measures are also excluded here, as are non-required tests.

The predictions capture the share of student performance that can be explained by non-school factors. The average gap between actual test scores and predicted test scores represents the best available measure of the school's or district's current contribution to academic performance (which we report as Z-scores). The academic progress score in math for a school is the average gap between actual math scores and predicted math scores for the students in that school; the academic progress score in reading is the average gap in reading scores. Our composite index of academic progress is an average of the academic progress scores in math and reading.

An academic progress score greater than zero indicates that the students in a school or district are performing better than students with similar characteristics statewide (on average), while an academic progress score below zero indicates that the students are performing worse than students with similar characteristics statewide. To ease interpretation, Texas Smart Schools converts the academic progress scores into academic progress percentiles. The academic progress percentiles range from 0 to 99. Schools in the 99$^{th}$ percentile had academic progress scores that were better than 99 percent of Texas schools.

## Fiscal Measures

Schools that operate in high cost-of-living communities must spend more dollars to provide the same level of real resources as other schools. Similarly, schools that serve more challenging student bodies must deploy more real resources to accomplish the same results as other schools. Economies of scale make the per-pupil cost of education lower in large school districts than in small ones. All of these factors—labor cost, student need, and size—combine to form an educational environment that shapes the decisions school districts make.

Any evaluation of school district efficiency must take differences in this educational environment into account. Texas Smart Schools accounts for the educational cost environment by evaluating the fiscal performance of each school or district in comparison to that of its fiscal peers. Fiscal peers are schools or districts that operate in a similar labor market, are of similar size, and serve similar students. Our method for identifying Fiscal Peers is described in more detail in the section below labelled *Identifying Fiscal Peers*.

The Fiscal Index for a school district is constructed in three steps.

1. The first step uses data on actual expenditures (which districts self-report to TEA) to calculate the level of core operating expenditures per pupil. The key financial indicator for the TSS methodology is core operating expenditures. Core operating expenditures are current operating expenditures (as defined by TEA) on:

    - Instruction (function 11)
    - Instructional resources (function 12)
    - Curriculum and staff development (function 13)
    - Instructional leadership (function 21)
    - School leadership (function 23)
    - Guidance counseling (function 31)
    - Social work (function 32)
    - Health services (function 33)
    - Extracurricular activities (function 36)
    - General administration (function 41)
    - Facilities maintenance and operations (function 51)
    - Security and monitoring (function 52)
    - Data processing services (function 53) and
    - Fund raising (function 81—charter schools only).

    Unlike TEA's definition of current operating expenditures, our definition of core operating expenditures excludes student transportation (function 34), food service (function 35), the incremental costs associated with the chapter 41 purchase or sale of Weighted Average Daily Attendance (WADA) related to school district wealth sharing

(function 92), and payments to juvenile justice alternative education programs (function 95). We don't consider these categories of spending to be part of core operating expenditures because they represent additional functions of local school districts not directly related to student achievement. Notably, neither definition of operating expenditures includes spending on construction or debt service.

2. The second step is to adjust core operating expenditures. There are two adjustments. First, the payroll part of core operating expenditures is adjusted for regional differences in labor costs. Adjustments for labor cost differences are like adjustments for inflation—they reflect the real purchasing power of school districts when prices are different. Second, core operating expenditures are adjusted to account for the impact of shared service agreements. Shared service agreements are a partnership of sorts among school districts. An agreement might cover a jointly operated special education program or a consulting teacher who works for more than one district. Typically, one district in a shared service agreement acts as a fiscal agent for the group, collecting funds from the other member districts, and making purchases or paying salaries with those shared funds on behalf of the other member districts. Without proper adjustments, the spending by fiscal agents looks artificially high while the spending by member districts looks artificially low. These adjustments are described in more detail in the section below labelled *Adjusting Core Operating Expenditures.*

3. Finally, a three-year average of the adjusted core spending of a school district is compared with a three-year average of the adjusted core spending of its fiscal peers. Districts that spend more than 80% of the districts in their peer group are identified as very high spending districts. Districts that spend more than 60% of the districts in their peer group are identified as high spending districts, and so on. Districts in the lowest-spending 20% are identified as very low spending districts.

The Fiscal Index for a campus is constructed the same way as the Fiscal Index for a district, except that the campus-level index is based on a narrower definition of core operating expenditures—campus-related core operating expenditures—which is defined as operating expenditures for instruction, instructional resources, instructional leadership, school leadership, and student support services (the total of all spending in functions 11-33). Unlike district core operating expenditures, campus-related core operating expenditures *exclude* extracurricular activities, general administration, facility maintenance and operations, security and monitoring services, and data processing services.

### *Identifying Fiscal Peers*

The key to identifying fiscal peers is developing reliable data on the fiscal environment in which each school district operates. Guided by conversations with Texas stakeholders and the scholarly literature on educational productivity, the TSS research team matched schools and

districts on the basis of two labor cost indicators, two size measures, and five measures of student needs.

Labor Costs. The education sector is labor-intensive, requiring professional staff such as teachers and administrators as well as nonprofessional staff such as clerks, educational aides, and maintenance workers. To capture regional differences in the prices paid for professional staff, the TSS research team used the American Community Survey Comparable Wage Index for Teachers (ACS-CWIFT) which has been newly created by the National Center for Education Statistics (NCES). A Comparable Wage Index (CWI) measures regional variations in the prevailing wage for college graduates. The basic premise of the ACS-CWIFT is that all types of workers demand higher wages in areas with a higher cost of living or a lack of amenities. Thus, if Dallas accountants are paid 15 percent more than the state average accounting wage, Dallas engineers are paid 15 percent more than the state average engineering wage, Dallas nurses are paid 15 percent more than the state average nursing wage, and so on, then a ACS-CWIFT predicts that Dallas teachers would need to be paid 15 percent more than the state average teacher salary, and that Dallas principals would need to be paid 15 percent more than the state average principal salary.

Because the wages of workers without a college degree may have a different geographic pattern than do the wages of college graduates, the TSS research team used a CWI that measures regional variations in the prevailing wage for high school graduates who do not have a bachelor's degree as the indicator for regional differences in the prices paid for non-professional staff.

Size Measures. Differences in school district size are a primary determinant of variations in the cost of education. Districts with small enrollments are much more expensive to operate (on a per pupil basis) than are larger school districts, for a host of reasons. Small enrollment districts have higher administrative costs per pupil and may have classrooms that are too small to be cost effective, simply because there aren't enough students in a grade level to fill all the seats. On the other hand, districts with large geographic areas may be more expensive to administer because the students, teachers, and schools are highly dispersed. The school finance formula of the state of Texas recognizes the inherent cost differences of small enrollment districts by providing additional revenue to small and midsized school districts. Additional funding adjustments are also provided to small districts that serve a geographic area of more than 300 square miles. To reflect these potential cost drivers, our analysis includes two measures of school district size—the number of students in fall enrollment, and the number of square miles in the district.

Student Need. To capture variations in costs that derive from variations in student needs, districts were matched based on five measures of student demographics, the percentages of students in each district who were:

- High needs special education students (available only at the district level)
- Other special education students
- Limited English proficient (LEP) students
- Economically disadvantaged students
- High mobility students (those who missed six or more weeks at a particular school)

Schools are expected to need more resources (for example, specialized teachers and supplies, or smaller required class sizes) as the share of students in each category increases.

Matches are based on a three-year average of school and district characteristics. Using a three-year average reduces the influence of one-time events on the selection of fiscal peers.

Matching Strategies. TSS uses a well-regarded research strategy to identify the fiscal peers for each school district—propensity score matching. Propensity score matching is a statistical strategy used to construct a control group for experiments that do not use random assignment. For example, if you want to know the effect of a jobs training program, you need to compare the program participants to a group of nonparticipants who are as similar as possible to the participant group, so that you can be reasonably confident that differences in employment outcomes are the result of the training, and not a result of some other difference between the two groups. Propensity score matching identifies the best available potential controls for any given member of the treatment group. The TSS research team used propensity-score matching to identify the up to 40 school districts that are most similar to each Texas school district with respect to the common determinants of school district cost—labor costs, school district size, and student demographics. The team used a similar methodology and campus-level data to identify the fiscal peers for individual campuses.

However, some Texas school districts are unusual enough in at least one cost dimension to limit their number of potential peers. For example, seven Texas districts are designated by the TEA as residential treatment facilities. Arguably, these seven districts should be matched only with one another. Similarly, while most school districts serve a full range of grade levels, some have no high school and others have no elementary schools. It seems most appropriate to match these restricted grade-level districts only to districts offering similar grade ranges. Still another group, districts in the Alternative Education Accountability (AEA) system serving at-risk youth, seems to match poorly with other K-12 districts. Finally, a small number of districts in Texas are very large — more than 1,000 times larger than some other districts. It seems inappropriate to match a very large district with a very small one, no matter how similar they are in other respects.

To accommodate these unusual cases, the districts were stratified before applying the propensity score matching technique. Each district was assigned to one of seven strata based on various student population characteristics, and propensity score matching was used as

needed to identify fiscal peers within each stratum. If the stratum contained no more than 40 districts, then all districts in the stratum were designated as potential fiscal peers, and propensity score matching was not used. The district strata used for this analysis were:

- Residential Treatment Facilities
- Very Small K-12 districts (those with 125 students or fewer)
- Very Large K-12 districts (those with more than 50,000 students)
- AEA Districts
- All Other K-12 Districts
- No Elementary Grades
- No High School Grades.

We took a similar approach to construct the campus-level peer groups. We started with TEA's grade-level classifications (elementary, middle secondary and multi-level) and accountability groupings then refined them further. Propensity score matching then was applied within each stratum containing more than 40 members. The campus strata used for this analysis were:

- Standard Accountability Campuses
    - Early Elementary Schools (those serving student up through the second grade)
    - Intermediate Schools (elementary schools only serving grades 4 and up)
    - Other Elementary Schools
    - Middle Schools
    - Very Large Secondary Schools (those with enrollments at or above 2,000)
    - Other Secondary Schools
    - Multi-Level Schools
- AEA Residential Schools
    - Secondary Schools
    - Other Schools
- AEA Non-Residential Schools
    - Elementary and Middle Schools
    - Secondary Schools
    - Multi-Level Schools
- Juvenile Justice Schools
- Special Education Schools

Although the propensity score matching technique identifies up to 40 school districts or campuses that are a statistically valid comparison group for each district or campus, some of the matches are obviously better than others. And some of the districts or campuses within strata that were not propensity score matched are highly dissimilar. To improve the internal consistency of the comparison groups at both the district and campus levels, the TSS team trimmed out potential fiscal peers that differed from the original by more than six standard

deviations with respect to any of the cost factors. As a result, although the vast majority of districts and campuses have 40 fiscal peers, some have substantially fewer. Districts or campuses with fewer than 4 fiscal peers do not have a Fiscal Index.

### *Adjusting Core Operating Expenditures*

As described above, the core operating expenditures used to construct the TSS Fiscal Index were adjusted for the fact that some school districts act as a fiscal agent for another district or group of districts. Fiscal agents collect funds from the member districts in a shared service agreement, and make purchases or pay salaries with those shared funds on behalf of the other member districts. As a result, the spending of fiscal agents is artificially inflated while the spending by member districts is artificially suppressed.

To correct for this pattern, we relied on TEA data about shared service agreements (SSAs). School districts that serve as fiscal agents are required to indicate the amounts they spent on behalf of the member districts each year. We used this information to allocate the spending by fiscal agents to the member districts on a proportional basis. For example, in 2015-16, Hudson ISD spent $421,891 from shared service funds on instruction, $162,739 on school leadership, $92,939 on facilities maintenance and operations, and $82,692 on miscellaneous other functions. Hudson's SSA report indicated that it spent 19.4% of those funds ($147,795) on its own behalf, 57.1% ($434,185) on behalf of Lufkin ISD, 13.1% ($99,290) on behalf of Diboll ISD, and 10.4% ($78,991) on behalf of Central ISD. Therefore, we allocated 19.4% of Hudson ISD's shared service spending for instruction, 19.4% of its shared service spending for school leadership, 19.4% of its shared service spending on maintenance, and 19.4% of its shared service spending for other functions to Hudson ISD. We similarly allocated 57.1% of Hudson ISD's shared service spending in each category to Lufkin ISD, 13.1% to Diboll ISD, and 10.4% to Central ISD.

The payroll component of core operating expenditures was also adjusted for regional differences in labor cost using the ACS-CWIFT. Adjusting payroll expenditures for differences in the ACS-CWIFT ensured that the Fiscal Index reflected the real resources each district is using to produce academic progress.

## Smart Scores

The final step was to combine the TSS composite academic progress score with the fiscal spending index to create a Smart Score. Smart Scores run from one to five stars. As shown in the table below, a five star score indicates high academic progress and very low spending. One star indicates low academic progress combined with very high spending.

| Composite Academic Progress Percentile | Spending Index | | | | |
|---|---|---|---|---|---|
| | Very High | High | Average | Low | Very Low |
| 80-99 | 3 stars  | 3½ stars  | 4 stars  | 4½ stars  | 5 stars  |
| 60-79 | 2½ stars  | 3 stars  | 3½ stars  | 4 stars  | 4½ stars  |
| 40-59 | 2 stars  | 2½ stars  | 3 stars  | 3½ stars  | 4 stars  |
| 20-39 | 1½ stars  | 2 stars  | 2½ stars  | 3 stars  | 3½ stars  |
| Less than 20 | 1 star  | 1½ stars  | 2 stars  | 2½ stars  | 3 stars  |