JANUARY 12, 2021

A Study on Geographic Education Cost Variations and School District Transportation Costs

TEA CONTRACT #: 4077

LORI L. TAYLOR, TIMOTHY J. GRONBERG,
DENNIS W. JANSEN & CAROLINE S. BARTLETT
TEXAS A&M UNIVERSITY

The authors thank the University of Texas at Dallas Education Research Center for research support on this project. All views expressed are those of the authors alone. The conclusions of this research do not necessarily reflect the opinion or official position of Texas A&M University, the Texas Education Agency, or the State of Texas.

Table of Contents

Study on Geographic Education Cost Variations and School District Transportation Costs

Executive Summary

In accordance with House Bill 3 (section 48.012), 86th Texas Legislature, 2019, the Texas Education Agency entered a Memorandum of Understanding and Agreement with Texas A&M University to conduct a study on geographic variations in known resource costs and costs of education due to factors beyond the control of school districts; and school district transportation costs. This report presents the results of that study.

The report was divided into four chapters. Chapter 1 of this report describes geographic differences in the cost of education that arise from uncontrollable differences in wages and salaries. Chapter 2 describes variations in the cost of education that arise from uncontrollable differences in cost factors other than wage levels. Chapter 3 describes differences in the cost of student transportation. Chapter 4 concludes the report by describing strategies for adjusting the Foundation School Program and Transportation Allotment protocols to address the cost differences identified in the previous chapters.

Executive Summary of Chapter 1: Geographic Variations in Wages and Salaries

Differences in the cost of living and the availability of amenities can lead to geographic differences in the prices that school districts must pay for their most important resource—people. Recent estimates from the National Center for Education Statistics (NCES) and the US Census Bureau indicate that the cost of hiring a college graduate can differ by as much as 50% from one part of Texas to another. Similar estimates from Texas Smart Schools (TSS) indicate that the cost of hiring a worker with a high school diploma but no bachelor's degree can differ by up to 31%.

Inequalities in school district purchasing power can lead to inequities in school funding formula. Therefore, a dozen US states have responded to evidence about geographic differences in labor cost by implementing geographic cost adjustments to their state school funding formulas. Six states have adopted a regional cost adjustment that relies on some sort of Comparable Wage Index (CWI). The basic premise of a CWI is that all types of workers—including teachers and other educators—demand higher wages in areas where the cost of living is high or there is a lack of desirable local amenities (such as good climate, low crime rates, or access to beaches, museums, or fancy restaurants). As a result, it should be possible to measure most of the uncontrollable variation in educator pay by observing systematic, regional variations in the earnings of comparably educated workers who are not educators. Intuitively, if accountants in Austin are paid 5 percent more than the national average accounting wage; Austin engineers are paid 5 percent more than the national average engineering wage; Austin nurses are paid 5 percent more than the national average nursing wage; and so on, then a CWI would predict that the wage level for Austin teachers is also 5 percent more than the national average teacher wage. The new estimates from NCES/Census and TSS are both CWIs.

Another four states rely on a Teacher Cost Index (TCI) for their regional cost adjustments. A TCI is a labor cost index that has been based on an analysis of teacher compensation within the state. Researchers use regression analysis to separate the observed variation in teacher salaries into the part that is explained by school district decisions (such as teacher demographics or teaching assignments) and the part that is systematically related to factors outside of school district control (such as the cost of living, the degree of geographic isolation or student demographics). The researchers then use their regression model to predict the salary that each district would need to pay to hire a teacher with an identical set of characteristics. Finally, they construct a TCI as the local salary prediction divided by some reference salary (such as the state average prediction or state minimum prediction). Because all of the decision factors are held constant in the construction of a TCI, the resulting index is purely a function of labor market characteristics and other uncontrollable cost factors (like student demographics). The Texas Cost of Education Index (CEI), which was an element in the Foundation School Program from 1991 until 2019, was a TCI that had been estimated using teacher salary data from the 1988-1989 school year.

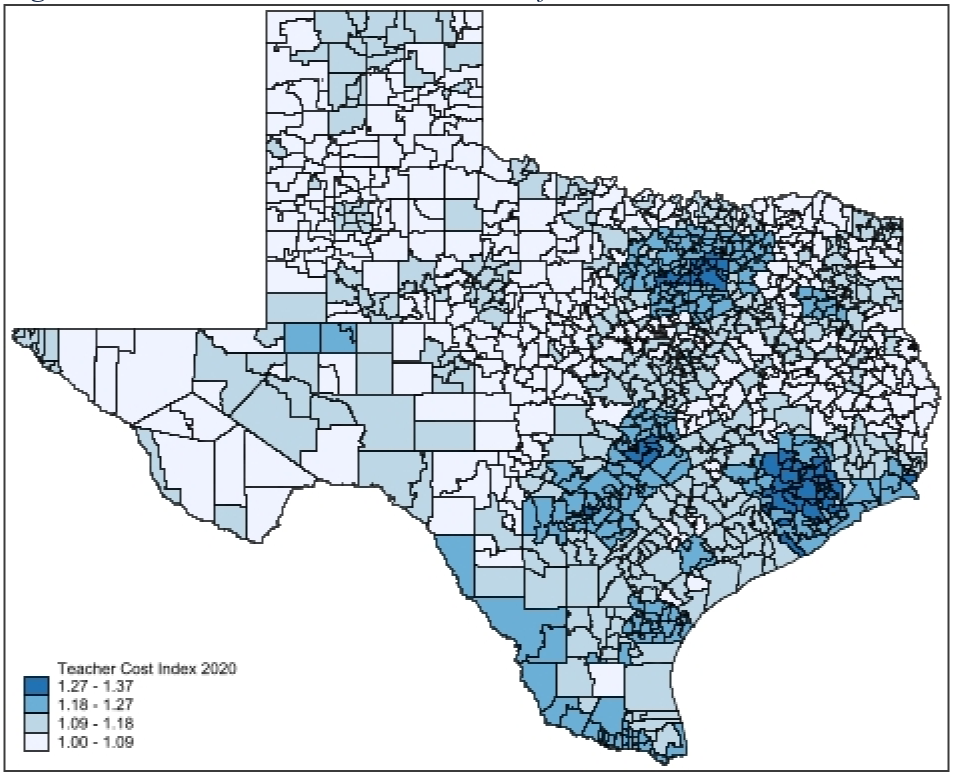
A New Texas Teacher Cost Index

This chapter presents a new TCI for Texas that was estimated using data from the six most recent school years (2014-15 through 2019-20). The Texas TCI embeds the new index from NCES/Census (the American Community Survey-Comparable Wage Index, or ACS-CWIFT) as one of the key cost factors outside of school district control. Other uncontrollable cost factors included in the index are fair market rents and unemployment rates, as well as various measures of geographic isolation, climate, student demographics, district type, and county type.

One of the keys to constructing a successful TCI is the inclusion of sufficient controls for teacher characteristics, because a failure to adequately account for differences in teacher quality can lead to measurement errors that misidentify high spending districts as high cost districts. The salary model used in the construction of the Texas TCI included a particularly rich set of demographic controls, including measures of teacher experience; teacher training; teacher years of service in the district; teaching assignments, and indicators for whether or not the teacher had administrative or support duties in addition to teaching. The ability to control for a wide array or teacher and assignment characteristics helped to ensure that the Texas TCI measures costs outside of school district control

The Texas TCI for 2019-20 ranged from 1.00 to 1.37, meaning the cost of hiring teachers was 37% higher in highest-cost districts than the lowest-cost districts. As Figure E-1 illustrates, the Texas TCI was highest in the Houston metropolitan area, and lowest in a district on the outskirts of the El Paso metropolitan area (i.e., a place where teachers can have easy access to urban amenities while enjoying a relatively low cost of living).

Figure E-1: The New Teacher Cost Index for Texas, 2019-20



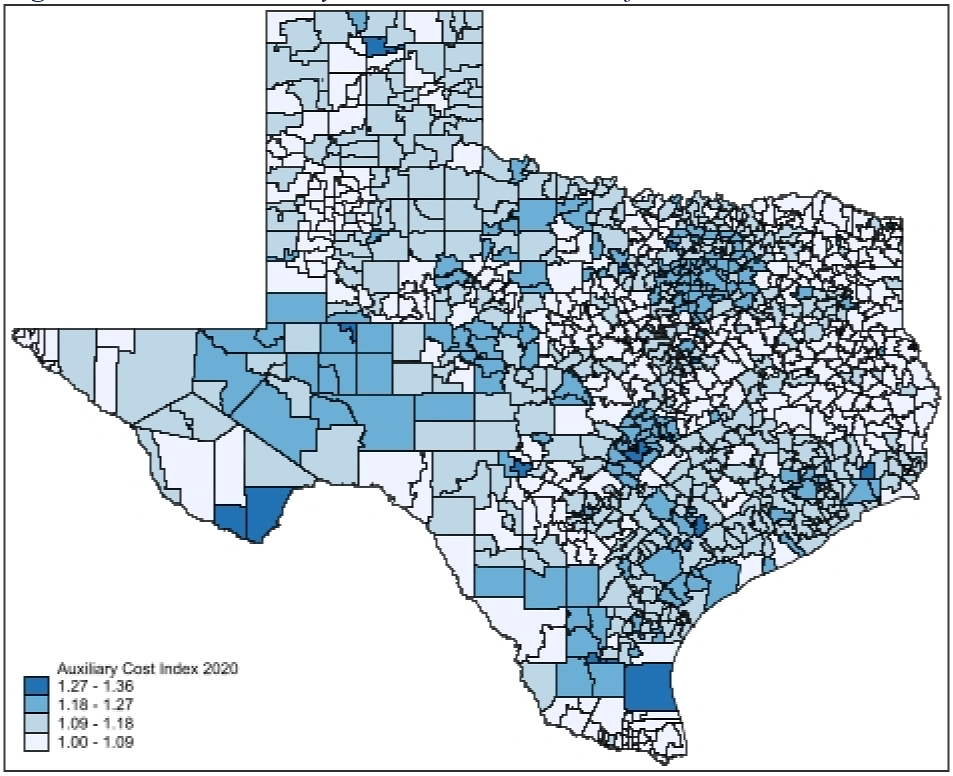
Source: Authors' calculations from PEIMS and other data sources.

A Cost Index for Auxiliary Personnel (APCI)

Geographic differences in the salaries for administrators, counselors and other certified personnel are likely to be highly correlated with those of teachers, and therefore align with the geographic patterns indicated by the Texas TCI. However, the wages of auxiliary personnel—those holding jobs that do not typically require a professional license or other form of certification (such as bus drivers, clerical workers or cafeteria staff) —may follow a different geographic pattern. This chapter also develops an Auxiliary Personnel Cost Index (APCI) that was constructed using the same methodology as the Texas TCI, but was based on data about the earnings of auxiliary workers in Texas school districts. Due to data availability issues, the APCI was estimated from only three years of data (2017-18 through 2019-20). Again the cost index was based on a wage model that controlled for worker demographics and job assignments, so that the cost index was a function of labor market conditions, working conditions outside of school district control and county type.

The APCI for 2019-20 ranged from 1.00 to 1.36, meaning the cost of hiring auxiliary workers was 36% higher in highest-cost districts than the lowest-cost districts. As Figure E-2 illustrates, the APCI was highest among K8 districts in the Alice and Austin metropolitan areas, and lowest in Hale, Lamb, Sabine and Starr Counties.

Figure E-2: The Auxiliary Personnel Cost Index for Texas, 2019-20



Source: Authors' calculations from PEIMS and other data sources.

Policy Implications from Chapter 1

In the past, Texas has incorporated geographic cost adjustments into the school finance formula. This analysis suggests that adjustments for differences in the price of labor are still needed in Texas. Such adjustments level the playing field so that all school districts can recruit and retain the same sort of high quality personnel despite local conditions that make some districts more attractive to teachers than others. Just as inflation adjustments allow the state to equalize school district purchasing power over time, regional cost adjustments allow the state to equalize purchasing power over locations. As such, regional cost adjustments can greatly enhance the equity of a school funding formula.

Executive Summary of Chapter 2: Geographic Variation in Costs of Education other than Wages

The school funding and finance literature has identified three main drivers of uncontrollable variation in educational cost: input prices, student needs, and economies of scale. All three of these drivers can vary geographically. Therefore, any analysis of geographic variation in the cost of education must be able to handle multiple dimensions of cost.

Cost-function analysis is the strategy best suited to an examination of geographic differences in the cost of education, and is the method used here. In the educational context, a cost function describes the relationship between school spending and school outputs, given the prices of educational inputs (such as teachers or school supplies), student characteristics, and other determinants of the educational environment (such as school district size or population density).

The Educational Cost Function

When properly specified and estimated using stochastic frontier analysis (SFA), the education cost function is a theoretically and statistically reliable method for estimating the relationship between the cost of education and various cost drivers, both those that are under the control of school districts and those that are considered uncontrollable by school districts.

The outcomes of the educational process are cost drivers that are under the control of school districts. Here, the campus-level educational outcomes have both a quantity and a quality dimension. Quantity is measured using the number of students in fall enrollment at the campus. Quality is measured using a campus-average, normalized gain score in mathematics and reading. The gain scores (which measure changes in the performance of an individual student from one year to the next) were based on student performance on the State of Texas Assessments of Academic Readiness (STAAR®) Grades 3–8 and end-of-course (EOC) exams.

The uncontrollable cost drivers examined in this campus-level analysis were input price measures and environmental cost factors. Key input prices included the two measures of labor price developed in Chapter 1 (the Texas TCI and the APCI). Key environmental cost factors included school district size, student demographics, and county population density.

The basic approach was to use SFA to estimate a campus-level cost function using data from the five most-recent school years with actual financial data (2014-15 through 2018-19). SFA explicitly allows for the possibility that spending could be systematically higher than cost. If schools are behaving efficiently, then SFA generates the same cost function estimates as other estimation techniques. Therefore, SFA can be thought of as a more general approach.

Findings

The analysis supported a number of key findings:

1. There were significant economies of scale at the campus level. For example, the cost function indicated that all other things being equal, a 200-student campus cost 4% more to operate (per pupil) than a 400-student campus, which in turn costs 2.5% more to operate (per pupil) than an 800 student campus. Per-pupil costs were minimized at a campus size of 1,500 students. However, the economies of scale at the campus level were largely exhausted once campus enrollment reached 1,000. The difference in the predicted, per-pupil cost of education between a campus of 1,000 students and a campus of 1,500 students was only 0.3%.
2. There were also significant economies of scale at the district level. Assuming that each campus in the district had the average campus-level enrollment for that district, and holding all other factors at their statewide means, the cost function indicated that a district with 300 students cost 15% more to operate (per pupil) than a district with 1,000 students. Similarly, a district with 1,000 students was predicted to cost 10% more to operate (per pupil) than a district with 5,000 students. As district size increased, costs per

pupil tended to fall, but most of the cost savings from increases in district size were largely exhausted at relatively low levels of district enrollment.

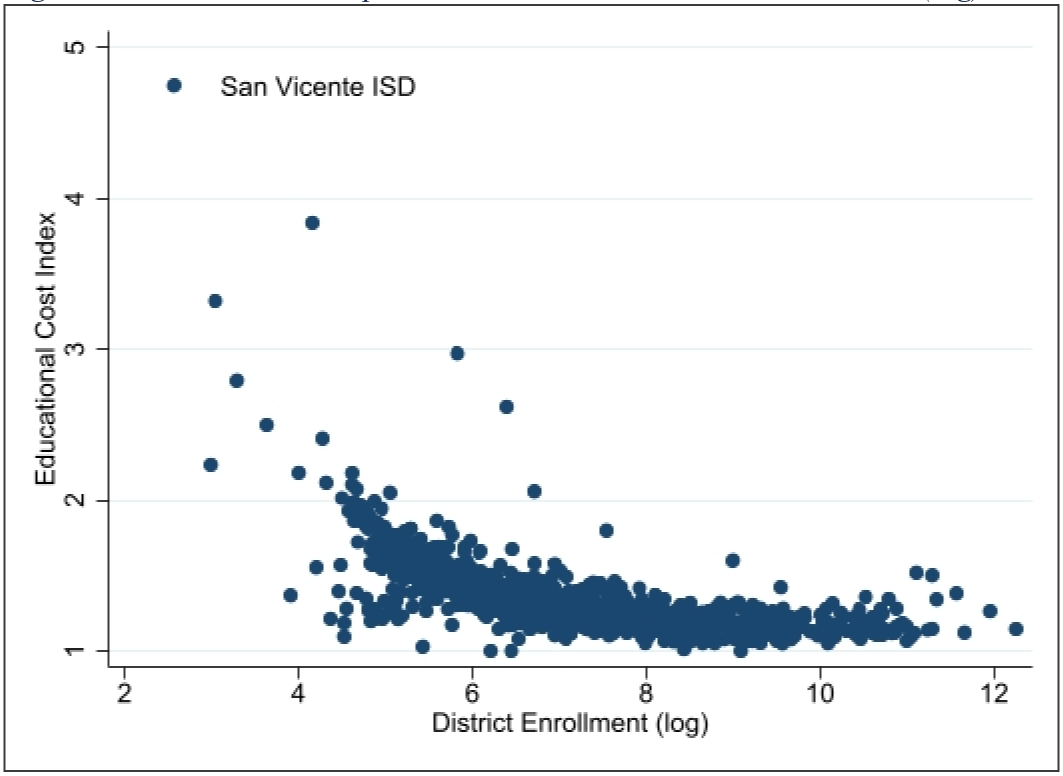
3. The cost of education was highly sensitive to differences in teacher wage levels. On average, a 10% increase in the Texas TCI was associated with a 6.6% increase in predicted cost per pupil, all other things being equal. In contrast, a 10% increase in the APCI (holding teacher wages constant) was only associated with a 0.5% increase in predicted cost.
4. The cost of serving an additional economically disadvantaged student was sharply higher for campuses that already had a high percentage of economically disadvantaged students. Among campuses with very low percentages of economically disadvantaged students, the marginal cost of serving an additional student who was economically disadvantaged was negative. This pattern of increasing intensity leading to sharply increasing cost was also observed for students who had ever been identified as an English Language Learner (ELL).
5. Expenditures exceeded what would be expected if campuses were operating efficiently. The average cost efficiency score was 0.93, indicating that campuses were producing 93% of their potential output. Given that inefficiency in this context means unexplained expenditures, not necessarily waste, and that many campuses may have been producing outcomes that were not reflected in test scores, the average efficiency level was high. On the other hand, efficiency was measured relative to the best practice in Texas, and that may still fall short of the ideal. Furthermore, the minimum efficiency scores were below 50%, suggesting that some campuses spend much more than could be explained by measured outcomes, input prices or student need. As a general rule, campus efficiency was higher in locations where educational choice was also higher.

The Educational Cost Index

The findings above all describe how the cost of education changes as one uncontrollable cost factor changes, holding all other cost factors constant. However, the cost function can also be used to estimate how much more or less it costs to produce educational outcomes in specific districts. Essentially, one uses the cost function to predict how much each district must spend, each year, in order to produce the normal (i.e., state average) level of output, assuming the district was making cost-minimizing choices about campus size. The Educational Cost Index (ECI) is the ratio of the predicted cost for the district, divided by the state minimum predicted cost.

Figure E-3 illustrates the relationship between district size and the ECI. As the figure illustrates, the ECI ranges from 1.00 to 4.74. In other words, the cost model predicted that the per-pupil cost of producing an average level of academic performance in the highest- cost district—San Vincente ISD with its total enrollment of 13 students—was more than 4.7 times the per-pupil cost of producing the same level of performance in the districts with the lowest cost of education. The median of the ECI was 1.29, so given their district-specific uncontrollable factors, half of the districts had to spend more than 29% above the minimum just to provide the state average level of educational output.

Figure E-3: The Relationship between the ECI and District Enrollment (log), 2018-19



Source: Authors' calculations from Appendix F.

Table E-1 provides another perspective on the ECI. As the table illustrates, the average ECI was higher in rural counties than in metropolitan or micropolitan areas. The average rural district had an ECI of 1.46 while the average metropolitan district had an ECI of 1.28. It appears that the generally lower wages in rural areas were more than offset by the district size and sparsity adjustments built into the ECI. The average high-poverty district had an ECI that was more than 16 percentage points higher than the average low-poverty district. Districts in sparsely populated counties had significantly higher ECIs than districts in more populous counties. All of these patterns were consistent with reasonable expectations about the costs of education in Texas.

Table E-1: The Educational Cost Index, by Location and School District Type, 2019-20

School District Type	Number of Districts	Mean	Minimum	Maximum
Metropolitan	493	1.283	1.000	2.410
Micropolitan	200	1.350	1.000	3.837
Rural	329	1.462	1.151	4.736
Very Sparsely Populated County	165	1.608	1.099	4.736
Sparsely Populated County	110	1.405	1.061	1.913
Other County	747	1.290	1.000	3.315
Small district	640	1.444	1.000	4.736
Midsized district	201	1.219	1.010	1.783
Large district	181	1.185	1.000	1.586
Highest Poverty Quintile	205	1.439	1.061	2.950
Lowest Poverty Quintile	204	1.277	1.000	4.736

Source: Authors' calculations from Appendix F.

Policy Implications from Chapter 2

The overarching takeaway from this analysis of the educational cost function is that the cost of education is far from uniform. Texas has a big and highly diverse educational landscape. Those differences in educational context drive differences in educational cost. Wages differ by up to 37% from one district to another and those differences drive significant differences in cost. The largest district in the state is more than 15 thousand times as large as the smallest district in the state. Small districts with correspondingly small campuses face significantly higher costs than other districts. The percentage of economically disadvantaged students ranges from zero to 100%. Such dramatic differences in the educational environment lead to dramatic differences in the cost of education.

Educational costs are higher in some parts of the state because the prices those districts must pay for educational resources—like teachers—are particularly high. But the cost function analysis suggests that other external cost drivers—namely student need, sparsity and a lack of economies of scale—require some districts to use real resources more intensively than others. Thus, the analysis suggests a need for adjustments to the funding formula in all these dimensions.

Executive Summary of Chapter 3: Geographic Variations in Transportation Cost

Transportation costs may vary between districts based on a number of factors outside of district control, including district size and location. Districts that are sparsely populated, for example, may have to transport fewer students across longer distances, thus generating large per-pupil transportation costs. On the other hand, densely populated, urban districts may face lower per-pupil transportation costs if they transport many students across a short distance.

Cost function analysis is a popular tool for analyses of transportation cost. This cost function analysis provides estimates of geographic variation in transportation costs, based on the five most-recent school years with actual financial data (2014-15 through 2018-19). As was the case for the educational cost function described in Chapter 2 of this report, the transportation cost function was estimated using SFA because, unlike other statistical techniques, SFA explicitly allows for the possibility that spending could be systematically higher than cost.

Key components of any cost function analysis are expenditures, outcome measures, input prices and environmental factors.

Expenditures

State administrative records provide two alternative sources of annual transportation expenditures—the Transportation Operations Report and the Public Information Management System (PEIMS). Each data system has strengths and weaknesses, and the data are not directly comparable. For consistency with the educational cost function analysis in Chapter 2, this analysis relied on the PEIMS data on current operating expenditures for student transportation (Function 34). Because the expenditure data reported under Function 34 exclude transportation expenses associated with extracurricular and co-curricular activities, the cost model also excluded such transportation activities. In other words, this was a cost function analysis of route transportation services.

Outputs

The definition of outputs is critical to any cost function analysis. In the case of school bus transportation, there are two common measures of output. One measure is bus miles; the second is the number of student passenger trips. Following the literature, this analysis included both bus miles and riders per mile (as a measure of passenger trips) as output measures. This specification also corresponded to the two dimensions—total miles and riders per mile (a.k.a. linear rider density)—that were used in the transportation allotment formula in place during the 2014-15 through 2018-19 time period. As with the expenditures data, the bus miles were restricted to Route miles and the student passenger trips were restricted to Route trips.

In addition, the model included two measures of the bus fleets—the percentage of buses that were less than five years old and the total number of buses. Although the number of buses is a direct measure of bus capital input, it also serves as a meaningful proxy for the number of bus routes, which can be viewed as a quality dimension of transportation output.

Input prices

Although one tends to think of fuel costs as the most important price for a cost analysis of transportation services, salary and benefit costs are around 80% of variable operating costs for the districts that run their own student transportation operation (i.e., those that do not contract out their transportation services). Variations in the price that districts must pay to hire their transportation employees are thus expected to be key drivers of variations in transportation costs across districts. Given the pivotal role of labor price differences in understanding operating cost differences, and the need to use a price measure that is outside of school district control, the researchers relied on the Auxiliary Personnel Cost Index (APCI) described in Chapter 1.

Fuel costs are the second most important price to include in a model of transportation costs. Unfortunately, TEA does not collect data on the prices that districts are paying to fuel up their buses and other student transport vehicles. Therefore, the researchers purchased a dataset from Oil Price Information Services (OPIS) of average annual diesel fuel prices by county in Texas for the five years under analysis. These data were collected by OPIS on a daily basis from a sample of reporting suppliers. The OPIS data were retail prices that included federal and state diesel taxes. In Texas, during the sample period, the state tax on diesel was 20 cents/gallon and the federal tax was 24.4 cents/gallon. Since school districts are exempt from these taxes, 44.4 cents were subtracted from the average county diesel fuel prices reported in the OPIS data. Clearly, such tax-adjusted retail prices are imperfect measures of the actual prices paid by districts. Undoubtedly, districts purchase fuel under a variety of contracting arrangements with fuel suppliers, many of which are with wholesale suppliers. However, the underlying market conditions that generate the quite persistent retail diesel price differences across counties in Texas (e.g., the transportation and distribution cost differentials) should lead to matching uncontrollable variations in the wholesale contract prices at which districts actually transact.

Environmental Factors

The prices for labor and fuel can vary geographically, giving rise to uncontrollable differences across districts in the cost of providing student transportation. Environmental factors—such as population density—can also influence the cost of student transportation in ways that are beyond school district control. To capture these sources of uncontrollable variation in cost, the model also included district population density, and a measure of roadway utilization (vehicle miles per lane mile) that was developed for this study by the Texas A&M Transportation Institute. Other environmental factors in the model included the percentage of route riders who were special education students and the percentage of route miles that were designated as special education route miles.

Findings

The cost function analysis provided a quite reasonable picture of the supply of route student transportation services. Costs were increasing in outputs and in input prices. Differences in the density of the distribution of potential riders and in the congestion features of the district road system also impacted the costs of hauling kids from home to school (and back again). The estimates indicated that costs increased as density *decreased*, which matched the finding in Hutchinson and Pratt (1999) from their study of school transportation costs in Tennessee, but was opposite of their finding for school transportation costs in Louisiana (2007).

Districts were spending more than would have been expected if they were operating efficiently. The extent of the estimated inefficiency was, however, quite modest. On average, the cost efficiency score was 0.94, indicating that districts were producing 94% of their potential output. The relatively strong efficiency estimates are not altogether surprising. By the end of the analysis period, the transportation allotment –which was the mechanism through with the Legislature provided school districts with financial aid for transportation—covered less than 25% of district outlays for route services. A potential benefit of the low level of state support for school transportation was the strong incentive for districts to run their transportation operations efficiently. It is also helpful that the technology for producing bus miles is relatively straightforward and well-known, reducing the scope for managerial error.

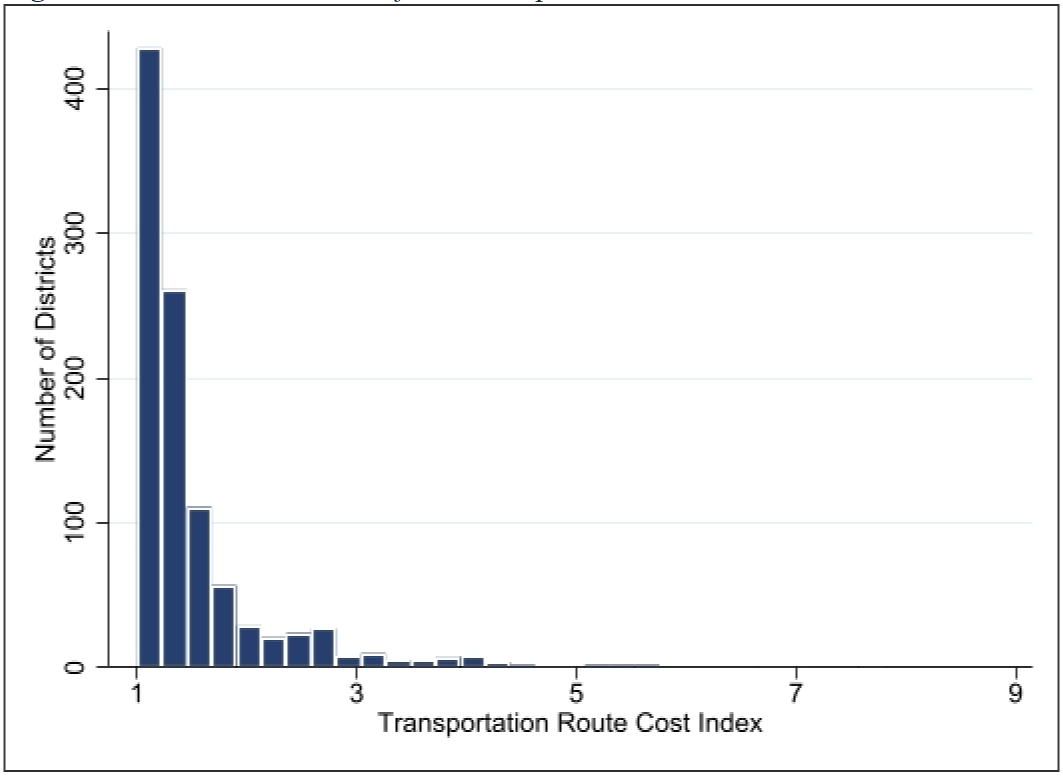
The Transportation Route Cost Index

Once the transportation cost function had been estimated, transportation cost indices could be generated. These cost indices indicated how much more or less it costs to produce bus miles in Houston than in Hutto. Essentially, one uses the cost function to predict how much each district must spend, each year, in order to produce the state average level of transportation output, given the state average quantity and quality of buses. For the other cost factors, which are treated as uncontrollable, the cost model was evaluated at the actual values in each district.

We estimated the cost index for each district by dividing the predicted spending level for each district by the minimum predicted spending level among the sample population of districts. The Transportation Route Cost Index (TRCI) is the ratio of the predicted cost for the district, divided by the state minimum predicted cost. The index values provide a measure of the uncontrollable cost in a district relative to the cost in a district with the most cost-favorable characteristics. For example, an index value of 1.5 indicates that a district is predicted to require 50 percent more dollars per mile than the least cost district to achieve the same level of output. Other normalizations are, of course, possible. For example, the reference cost level could be the predicted cost of producing the standardized outputs for a district with the average values of the uncontrollable cost factors.

Figure E-4 illustrates the frequency distribution of TRCI for the 2018-19 school year, which ranged from 1.00 to 7.80. The median of the TRCI was 1.29, so half of the district values of costs per mile, given their district-specific uncontrollable factors, were between 100% and 129% of the minimum. The TRCI distribution was rather heavily skewed, with a long right tail of districts with TRCI values greater than 2.00. Still, these extreme values of the TRCI distribution were outliers. More than 85% of districts had an index value less than 2.00, and only 5% of the values were more than 3.00.

Figure E-4: The Distribution of the Transportation Route Cost Index, 2018-19



Source: Authors' calculations from Appendix G.

Policy Implications from Chapter 3

This chapter develops and estimates a model of the costs to school districts of transporting students to and from school. The model is grounded in the academic literature on bus transportation, both school bus services and municipal transit services. The cost of producing school bus miles depends on the number of miles the buses are covering, the prices of bus mile inputs (such as labor and fuel), the number and spatial distribution of student riders, and upon the environment in which the bus miles are being produced (such as features of the road infrastructure). The analysis demonstrates that there are important and uncontrollable differences among school districts with respect to the cost of providing route transportation services.

Executive Summary of Chapter 4: Strategies to Address Geographic Cost Differences

The above analyses demonstrate clearly that there are large geographic differences in the cost of providing educational and transportation services in Texas. Those differences arise from a lack of population density and economies of scale in rural Texas, higher labor costs in urban Texas, and district-by-district differences in uncontrollable cost factors like student need.

Uncontrollable Cost Adjustment Using the ECI and TRCI

The cost function analyses also generated cost indices that could be used to adjust the Foundation School Program (FSP) and the transportation allotment for those differences. The ECI could be used as an adjustment for the basic allotment in the FSP. If the legislature chose to go that route, the ECI would replace the small and midsized allotments, the compensatory education allotments, the dyslexia allotment and the special education allotments, as well as the bilingual/ESL allotment. Because those allotments largely define weighted average daily attendance (WADA), the ECI would also largely redefine WADA in Tier II of the FSP.

Similarly, the TRCI could be used to adjust the transportation allotment for geographic differences in the cost of education. During the 2014-15 through 2018-19 time period, the regular program allotment was determined using a linear density-based formula that provided a higher rate per mile for districts with a larger number of riders per mile. The formula for determining the regular transportation allotment was amended under House Bill 3 (HB 3) in June 2019. Under HB 3, the regular program allotment will be based on a flat rate per mile to be set by the Legislature in the General Appropriations Act (GAA). The rate adopted for 2020-21 under the current GAA is $1 per mile.

The TRCI could be used directly to adjust the base allotment rate per mile for the estimated differential costs associated with the different uncontrollable cost environments facing district transportation planners. A district with an estimated TRCI of 1.29 would be assigned a regular program allotment rate of 1.29 times the base allotment rate per mile. Assuming no change in the base allotment, then at the median the current HB3 rate of $1 per mile would be increased to $1.29/ mile. Of course, the legislature could also use its discretion to make a revenue-neutral, downward adjustment to the base allotment per mile. A base allotment of $0.82 per mile multiplied by the TRCI would have the same predicted impact on the state's total transportation allotment as the flat $1 per mile under HB3. Under a TRCI-driven model, relatively more transportation funding would flow to the districts where uncontrollable transportation costs are higher.

One key to successful long-term implementation of either the ECI or the TRCI would be the development of a strategy for regularly updating the indices. Although many of the factors that drive differences in the costs of education and transportation are unlikely to change over time, other factors—such as wage levels outside of education, fuel costs and student enrollments—are sensitive to changing economic and socioeconomic conditions. To ensure that the cost indices are functioning as intended, the ECI and TRCI should be updated regularly, either by using the estimated cost models to generate new cost predictions corresponding to new values for the various cost factors, or by re-estimating the cost models themselves.

Uncontrollable Cost Adjustment Using Individual Cost Factor Adjustments

While it would be straightforward and analytically sound to use the ECI and TRCI as black-box cost adjustments, the legislature may instead choose to use the information provided herein to refine key components of the two funding models. For example:

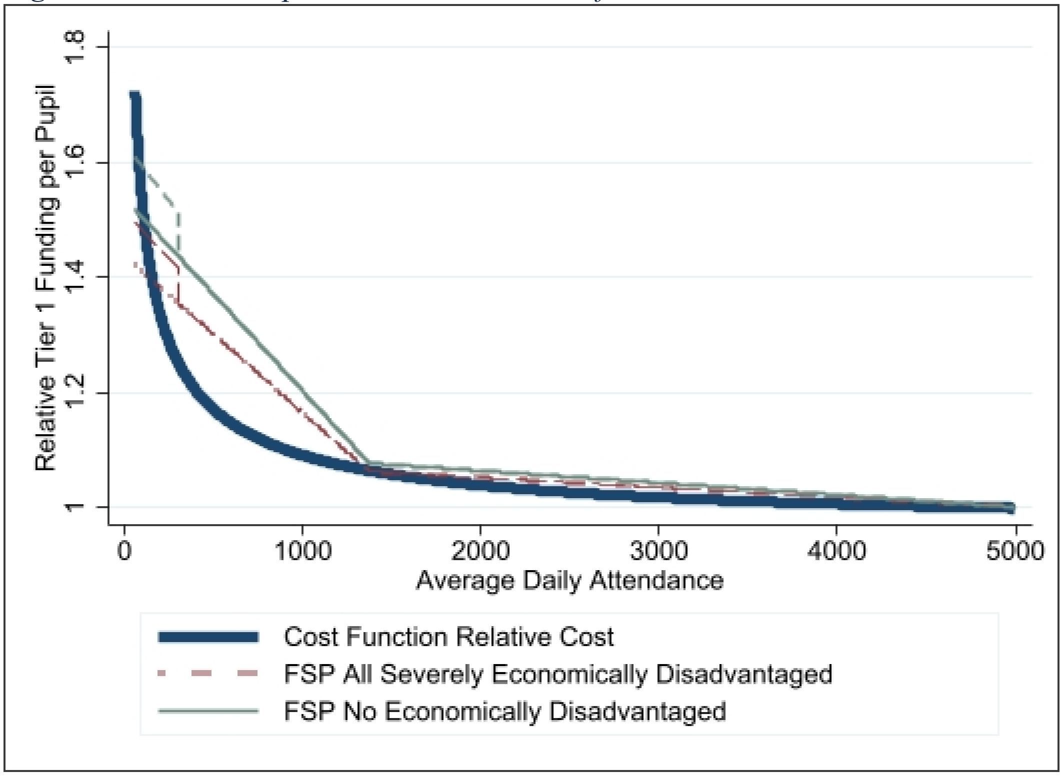
The Compensatory Program Allotments: HB3 instructed the Commissioner of Education to develop new measures of student socioeconomic status. The new measures were to be based on the demographics of the Census block where each educationally disadvantaged student resides. The funding formula weight was increased by 2.5 percentage points for economically disadvantaged students who live in Census blocks that were the least disadvantaged; and increased by 7.5 percentage points for economically disadvantaged students who live in severely disadvantaged Census blocks.

The cost function estimates suggest that the concentration of poverty in the school –not just the concentration of poverty in the student's neighborhood—can have a significant impact on the cost of education. The higher the percentage of economically disadvantaged students, the higher the increase in cost associated with an additional disadvantaged student. The legislature may wish to consider adding an intensity adjustment, perhaps modeled after the concentration grants or the targeted assistance grants that are part of the Title 1 program.

The Small and Midsized Allotments: HB3 replaced the small and midsized adjustments in the funding formula with small and midsized allotments. This change treated the scale adjustments in a manner analogous to the allotment for compensatory education. "Instead of flowing funds to small and mid-size districts as an adjustment that occurs before other funding adjustments, the funding now flows as an allotment under Tier I at the same time as other funding adjustments, such as the compensatory education allotment and the bilingual allotment." (TEA 2019). As a result, the small and midsized adjustments no longer have a multiplicative effect on the other allotments, such as the compensatory or bilingual program allotments. This change reduced the funding differential for small and midsized districts.

The cost function estimates suggest that the small and midsized allotments still overstate the relationship between school district size and the cost of education for all but the smallest districts. Figure E-5 compares the small and midsized allotments expected under HB3 (as a percentage of the funding for an otherwise identical district that was not eligible for the size adjustments) to those implied by the cost function analysis. (The dashed line indicates the supplemental allotment provided to districts with fewer than 300 students when the district is the only one in the county.) There are two alternatives for the FSP—one in which all the students are economically disadvantaged and live in a severely disadvantaged Census block, and one in which none of the students are economically disadvantaged. As the figure illustrates, the cost function estimates indicate that a district with 300 students costs 25% more to operate than a school district with 5,000 students, whereas the funding formula provides an additional 35% to 44%, depending on the percentage of economically disadvantaged students. The gap between the FSP and the cost function-based estimates is even wider for districts with between 300 and 1,000 students.

Figure E-5: Two Perspectives on Economies of Scale



Source: Authors' calculations from the FSP and Appendix F.

The Cost of Education Index: HB3 removed the Cost of Education Index (CEI) that had been part of the FSP since 1991. While the CEI was clearly outdated, Chapter 1 of this report provides evidence that significant regional differences in labor cost persist, and offers a ready-made replacement for legislative consideration, namely the Texas TCI.

Because the non-labor components of a school district's budget are unlikely to have the same geographic pattern as labor costs, the legislature may wish to embed the Texas TCI or the ACS-CWIFT in a regional cost index. As discussed in Taylor (2015) a regional cost index can be constructed as a weighted average of the various price indices (here, the Texas TCI/ACS-CWIFT and APCI) where the weights are the shares of each input in the total budget of a typical district. Prior to HB3, the CEI was operationalized into Tier I of the funding formula in a way that was equivalent to a regional cost index with a labor weight of 0.71 (Taylor 2015b).

Should the legislature choose to adopt the Texas TCI or APCI, it would be prudent to also adopt a process by which the indices would be updated, so that the indices would continue to perform as intended when economic conditions changed.

Transportation Cost Adjustments: As an alternative to the TRCI, one could use the transportation cost function to derive cost-function based weights that can be applied to each of the uncontrollable cost factors and then added up to generate an adjusted allotment rate for regular program miles. This is similar in spirit to the use of weights to adjust the base allotment per pupil in determining the Tier I education allotment. For the three principal uncontrollable factors in the model—fuel price, labor price, and population density—we generated the estimated change in cost from a small change in the designated cost factor (i.e., the marginal effect) holding all other factors constant at their statewide means.

We then used the estimated marginal effects to generate a set of cost allotment adjustment factors for each of the three key cost factors. Using 2018-19 data, we first divided each cost factor into four or five groups. The groups corresponded to quartiles for the input prices. Because of the large differences within the top density quartile, we further subdivided that quartile into the values above the 90th percentile and values below the 90th percentile. For each group, we then calculated a predicted percentage increase in cost per mile due to the higher fuel price, higher wage level, or higher population density, respectively. We end up with four fuel price adjustment factors, four wage level adjustment factors, and five density adjustment factors. The first quartile is the category with the lowest fuel price, the lowest wage level or the lowest population density.

Table E-2: Transportation Regular Program Allotment Rate Adjustment Factors

Quartiles	Fuel index Adjustment Factor	Wage level Adjustment Factor	Population Density Adjustment Factor
First Quartile	0.03	0.02	0.35
Second Quartile	0.04	0.03	0.33
Third Quartile	0.05	0.05	0.31
Fourth Quartile up to 90th percentile	0.07	0.07	0.13
Fourth Quartile above 90th percentile	0.07	0.07	0.00

Source: Authors' calculations.

A district could be in the first quartile for the fuel index, the second for the wage and the third for population density (or vice versa). The total input cost factor adjustment to the basic transportation allotment would simply be the sum of the three adjustment factors for each district.

List of Tables

List of Figures

List of Abbreviations

ACS – American Community Survey, a survey of US households conducted annually by the US Census Bureau

ACS-CWIFT – ACS-Comparable Wage Index for Teachers

ADA – Average Daily Attendance

APCI – Auxiliary Personnel Cost Index

ADA – Average Daily Attendance

CDP – Census-Designated Place

CWI – Comparable Wage Index

AEA – Alternative Education Accountability

APCI – Auxiliary Personnel Cost Index

AR – Autoregressive Model

BLS – Bureau of Labor Statistics

CBSA – Core-Based Statistical Area

CEI – Cost of Education Index

CLM – Centerline miles

CPI – Consumer Price Index

CWI – Comparable Wage Index

ECI—Educational Cost Index

EOC – End-of-Course

ELL – English Language Learner

EverELL – Ever identified as ELL by the Texas public school system

FSP – Foundation School Program

FTE – Full Time Equivalent

HB3 – House Bill 3, 86th Texas Legislature, 2019

HS-CWI – High School Comparable Wage Index

ISD – Independent School District

LEA – Local education agency

LM – Lane miles

NCE – Normal curve equivalent

NCES – National Center for Education Statistics

NOAA – National Oceanic and Atmospheric Administration

OPIS – Oil Price Information Service

PEIMS – Public Education Information Management System, a TEA data collection

SFA – Stochastic Frontier Analysis

STAAR® – State of Texas Assessments of Academic Readiness

TAPR – Texas Academic Performance Report

TCI – Teacher Cost Index

TEA – Texas Education Agency

TOR – Transportation Operations Report

TPS – Traditional public school

TRCI – Transportation Route Cost Index

TRS – Texas Retirement System

TSS – Texas Smart Schools

TTI – Texas Transportation Institute

VMT – Vehicle miles traveled

WADA – Weighted Average Daily Attendance

Glossary of Terms

Auxiliary Personnel Cost Index (APCI): An APCI is a labor cost index that has been based on an analysis of auxiliary personnel compensation within the state.

Control Function: A statistical technique used to control for bias generated by a potential correlation between an independent variable and the error term in a regression analysis. A control function is an alternative strategy for specifying an instrumental variables model.

Comparable Wage Index (CWI): A CWI is a labor cost index that has been based on an analysis of nonteacher compensation. The basic premise of a Comparable Wage Index (CWI) is that one should be able to measure regional variations in the cost of hiring educators by observing variations in the earnings of comparable workers who are not educators.

Core-Based Statistical Area (CBSA): A term used by the US Office of Management and Budget and US Census Bureau to refer collectively to all metropolitan and micropolitan areas. A metropolitan area is a county or cluster of counties with a central, urbanized area of at least 50,000 people. A micropolitan area is a county or cluster of counties with a central city of at least 10,000 people. Two counties are considered part of the same CBSA whenever commuting patterns indicate that the counties are part of the same integrated labor market area. In Texas, College Station-Bryan is a metropolitan area, and Nacogdoches is a micropolitan area.

Cost Function: A mathematical description of the relationship between the inputs, outputs and costs of operating a fully efficient firm. In the educational context, a cost function describes the relationship between (efficient) school spending and student performance, given the price of educational inputs (such as teachers or school supplies), student characteristics, and other determinants of the educational environment such as school district size.

Cost Function Analysis: The estimation of a cost function using statistics or some other data-driven technique.

Economies of scale: Economies of scale exist when it is possible to reduce per-pupil costs by increasing the size of the school or district.

Educational Cost Index (ECI): To generate an ECI, one uses a cost function to predict the cost of producing a designated level of output in all jurisdictions. Here, the ECI is the ratio of the cost function's predicted cost for the school district, divided by the state minimum predicted cost.

Efficient: A school or district is efficient (i.e., behaving efficiently) when it is not possible to increase measured educational or transportation outputs without increasing expenditures on purchased inputs.

Environmental Factors: Characteristics of the school district or location that influence the cost of education or the cost of transportation, but are neither purchased inputs nor outputs. Common environmental factors include district size, student demographics and population density.

Hedonic Wage Analysis: A regression-based analysis of the relationship between observed pay and variables representing worker characteristics, job characteristics and location characteristics.

Herfindahl Index: A measure of the amount of competition in a market. In the education context, it is defined as the sum of the squared local education agency (LEA) enrollment shares, where an LEA's enrollment share is its own enrollment divided by the total enrollment in the CBSA. The Herfindahl index increases as the level of enrollment concentration increases (i.e., as the level of competition decreases). A Herfindahl index of 1.00 indicates a metropolitan area with a single LEA; a Herfindahl index of 0.10 indicates a metropolitan area with 10 LEAs of equal size.

Inefficient: A school or district is inefficient when it is possible to increase measured educational or transportation outputs without increasing expenditures on purchased inputs.

Inputs: The equipment, personnel or raw materials used to produce outputs/outcomes.

Labor Cost Index: A labor cost index describes geographic variations in the prevailing wage for a designated type of worker. It is a measure of the price employers pay for labor, normalized relative to sum reference wage level, such as the state average wage or the state minimum wage for the worker type.

Outputs/Outcomes: The goods or services produced. In the education context, the primary outcomes are some measure of student performance and the number of students served (the product of which yields total output). In the bus transportation context, the primary outcomes are bus miles and passenger trips.

Stochastic Frontier Analysis (SFA): SFA is a statistical technique used to describe the best—as opposed to average—practice in the data. In this project, the cost function is estimated using SFA. Other statistical approaches to cost function estimation assume that, on average, school spending equals the cost of education. SFA explicitly allows for the possibility that spending could be systematically higher than cost. If school districts are behaving efficiently, SFA yields the same cost function estimates as other techniques.

Teacher Cost Index (TCI): A TCI is a labor cost index that has been based on an analysis of teacher compensation within the state.

Transportation Route Cost Index (TRCI): To generate a TRCI, one uses a cost function to predict the cost of producing a designated level of transportation output in all jurisdictions. Here, the TRCI is the ratio of the cost function's predicted cost for the school district, divided by the state minimum predicted cost.

Introduction

Expenditures vary from one school district to another for two main reasons: uncontrollable differences in the cost of education, and controllable differences in the choices school districts make.

Some school districts spend more than other districts for reasons that are clearly outside of their control. For example, school districts that face high market prices for teachers due to the cost of living will tend to spend more than other districts just to be able to hire the same type of personnel. School districts that serve a student population that is particularly challenging, may need to offer higher salaries to recruit and retain effective teachers, and may need to employ more teachers per pupil than other districts. School districts with widely dispersed student populations may need to operate smaller, less cost-effective schools and may need to spend more per pupil on student transportation.

On the other hand, some school districts spend more than other districts for reasons within their control. For example, a district may choose to provide educational services or enrichment activities that are not provided by other districts, or may make other operational decisions that tend to increase costs. Although such decisions may be motivated by the desire to achieve better student outcomes, in some cases the result may be a less efficient use of available resources.

Separating uncontrollable causes of observed differences in spending from controllable causes is the fundamental challenge facing researchers and policymakers who are interested in comparing or equalizing the purchasing power of school districts. If the challenge is not met, high spending districts may be misinterpreted as high-cost districts, policymakers may misallocate scarce educational resources, and researchers may be misled about the relationship between school resources and educational outcomes.

In accordance with House Bill 3 (section 48.012), 86th Texas Legislature, 2019, the Texas Education Agency entered a Memorandum of Understanding and Agreement with Texas A&M University to conduct a study on geographic variations in known resource costs and costs of education due to factors beyond the control of school districts; and school district transportation costs. This report presents the results of that study. Chapter 1 of this report describes geographic differences in the cost of education that arise from uncontrollable differences in wages and salaries. Chapter 2 describes variations in the cost of education that arise from uncontrollable differences in cost factors other than wage levels. Chapter 3 describes differences in the cost of student transportation. Chapter 4 concludes the report by describing strategies for adjusting the Foundation School Program to address the cost differences identified in the previous chapters.

Findings indicate that there are significant variations in labor costs across the state that fall outside of school district control. There are also significant and uncontrollable differences in transportation cost. These findings indicate that districts in high-cost environments require additional funding to purchase the same resources available to other districts at a lower cost. As the state continues to grow and diversify, cost differences across school districts will widen. Failure to account for these cost differences in the school finance formula may lead to inequities in district purchasing power and, ultimately, student outcomes. Fortunately, adjustments for uncontrollable cost differences can be included in the Foundation School Program. This report presents the results of a study on geographic variations in the cost of education and recommends addition of a geographic cost adjustment to the current school finance formula in order to more effectively meet real resource goals outlined by the Foundation School Program.

Chapter 1: Geographic Variations in Wages and Salaries

Differences in the cost of living and the availability of amenities can lead to geographic differences in the prices that school districts must pay for their most important resource—labor. The geographic labor cost indices developed in this chapter describe regional differences in the wages that must be paid to recruit and retain the same sort of high quality personnel in every school district. As such, they serve the same purpose as inflation indices: they reflect the real purchasing power of school districts when prices are different.

The ACS-CWIFT

The National Center for Education Statistics (NCES) recently collaborated with the US Census Bureau to publish a new index designed to measure regional variation in the cost of education—the ACS-Comparable Wage Index for Teachers (ACS-CWIFT).

The ACS-CWIFT is a labor cost index that was based on millions of responses to the American Community Survey (ACS). The ACS is a survey of US households conducted annually by the US Census Bureau. The ACS is the primary source of demographic information about the US population. It provides information about the earnings, age, occupation, industry, and other demographic characteristics for millions of US workers.

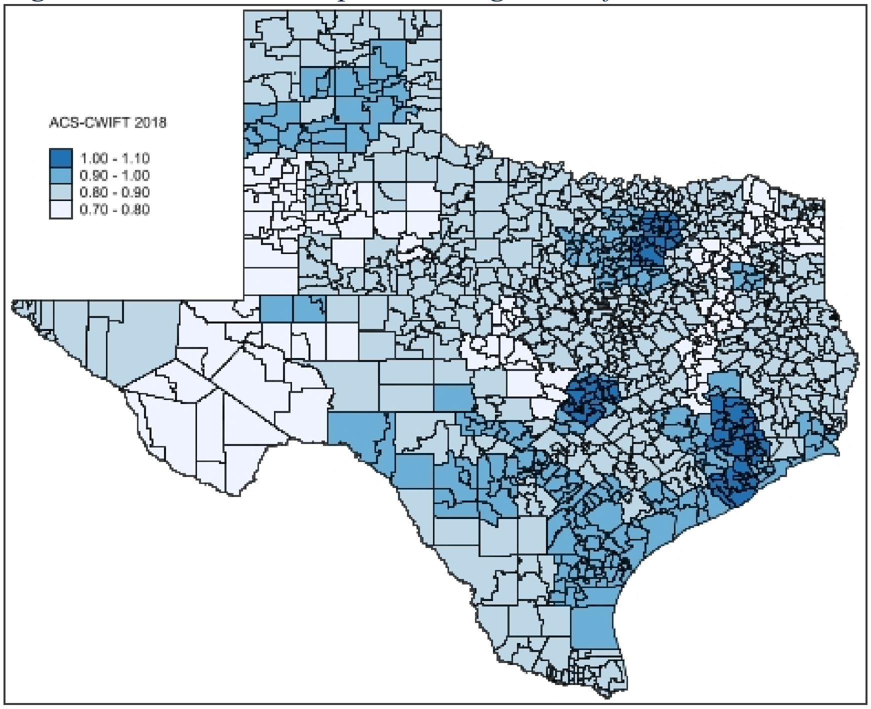
The ACS-CWIFT measures geographic variation in the prevailing wage for college-educated workers who are not educators. This focus on non-educators ensures that the ACS-CWIFT is measuring variations in labor cost that are beyond school district control. The basic premise of the ACS-CWIFT is that all types of workers—including teachers and other educators—demand higher wages in areas where the cost of living is high or there is a lack of desirable local amenities (such as good climate, low crime rates, or access to beaches, museums, or fancy restaurants). As a result, it should be possible to measure most of the uncontrollable variation in educator pay by observing systematic, regional variations in the earnings of comparably educated workers who are not educators. Intuitively, if accountants in Austin are paid 5 percent more than the national average accounting wage; Austin engineers are paid 5 percent more than the national average engineering wage; Austin nurses are paid 5 percent more than the national average nursing wage; and so on, then a comparable wage index (CWI) like the ACS-CWIFT would predict that the wage level for Austin teachers is also 5 percent more than the national average teacher wage.

The ACS-CWIFT was estimated at the county level using data from three consecutive years of the ACS. The most recent index was estimated from over one-million survey responses and covers the three-year period from 2016 through 2018.

As a general rule, the ACS-CWIFT for a Texas school district is the ACS CWIFT for the corresponding county. However, some Texas districts span multiple counties. In those cases, the ACS-CWIFT for the district is a weighted average of the ACS-CWIFTs for each county in the district, and the weights reflect the shares of school-aged children in the district who live in each county.

Figure 1-1 maps the geographic distribution of the ACS-CWIFT in Texas. Values range from 0.712 in the lowest cost districts to 1.067 in the highest cost districts, implying that the cost of hiring college educated workers can differ by as much as 50 percent (1.067/0.712) from one part of Texas to another. As the map illustrates, the ACS-CWIFT indicates that the cost of hiring college-educated workers is highest in the Houston metropolitan area, and lowest in rural west Texas.

Figure 1-1: The ACS Comparable Wage Index for Teachers, 2018



Source: National Center for Education Statistics.

The TSS High School Comparable Wage Index

Because school districts employ non-professional staff as well as professional staff, and the wages of workers without a college degree may have a different geographic pattern than do the wages of college graduates, the Texas Smart Schools research team used the publically available ACS data to estimate a CWI for high school graduates who do not have a bachelor's degree. This High School CWI (HS-CWI) serves as the indicator for regional differences in the prices paid for non-professional staff.

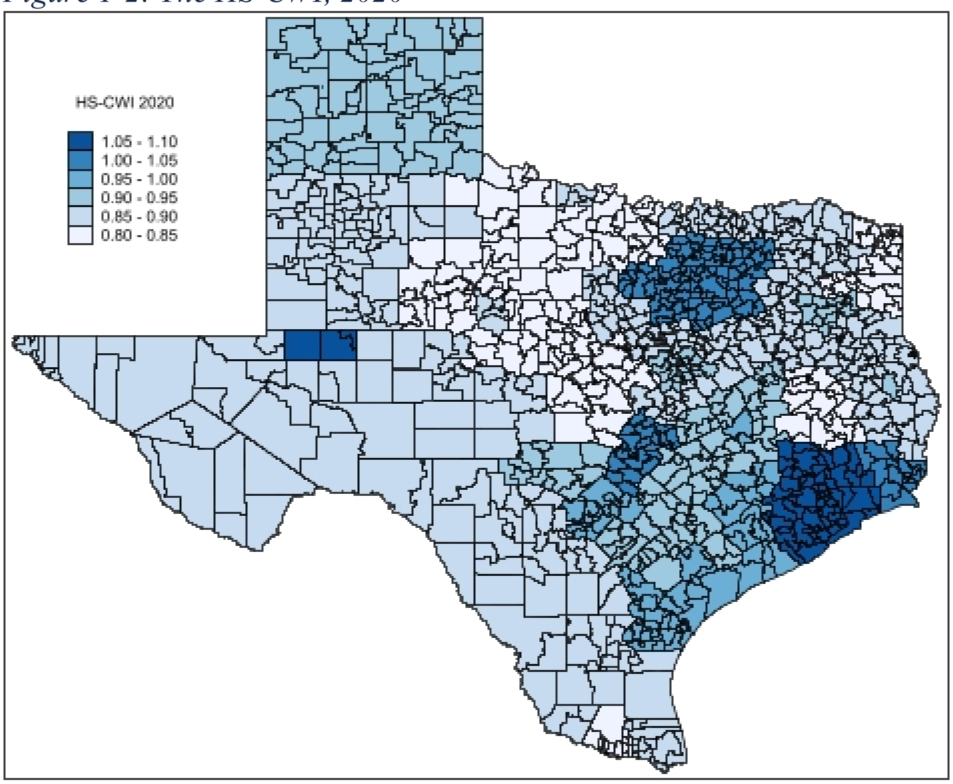
The HS-CWI was modeled after the ACS-CWIFT. As such, the HS-CWI was also based on millions of responses from three consecutive years of the American Community Survey. However, unlike the ACS-CWIFT (which derived from survey responses from college-educated workers), the HS-CWI was estimated from the survey responses of workers who had at least a high school diploma but did not have a bachelor's degree. Again, survey respondents who were employed in K-12 education were excluded, to ensure that the HS-CWI was measuring labor cost variations beyond school district control. (For more on the HS-CWI, see Texas Smart Schools, 2019).

Because the HS-CWI was estimated from publically available data (and therefore some geographic information was suppressed to ensure privacy), it lacks some of the geographic detail available with the ACS-CWIFT. The HS-CWI was based on "place-of-work areas" as defined by the Census Bureau. Census place-of-work areas are geographic regions designed to contain at least 100,000 persons while not crossing state boundaries. In sparsely populated parts of a state, one place-of-work could comprise multiple counties; in densely populated parts of a state, there could be multiple places-of-work within a single county. The geographic units used in the HS-CWI analyses were either single places of work, or a cluster of the places-of-work that comprise a metropolitan area.

The predicted wage level in each labor market area captured systematic variations in labor earnings while controlling for worker demographics, industrial and occupational mix, and amount of time worked. Dividing each local wage prediction by the corresponding national average yielded the HS-CWI. The HS-CWI for a Texas school district is the HS-CWI for the corresponding place of work (either county or metropolitan area).

Figure 1-2 maps the geographic distribution of the HS-CWI in Texas. Values range from 0.84 in the lowest cost districts to 1.10 in the highest cost districts, implying that the cost of hiring workers with a high school diploma can differ by as much as 31 percent (1.10/0.84) from one part of Texas to another. The tighter range on the HS-CWI (when compared with the ACS-CWI) likely arises from the more limited level of geographic detail in the HS-CWI. As the map illustrates, the HS-CWI indicates that the cost of hiring college-educated workers is highest in Midland, Odessa, and the major Texas metropolitan areas, and lowest in rural areas outside of the panhandle of Texas.

Figure 1-2: The HS-CWI, 2020



Source: Texas Smart Schools.

Geographic Cost Adjustments in Other States

Texas is not the only state to grapple with the challenges posed by geographic variation in labor cost. In an effort to level the playing field, so that all school districts can afford to hire high-quality teachers, a dozen states have incorporated regional cost adjustments into their school finance formulae (see Table 1-1).

Table 1-1: Geographic Cost Adjustment Strategies, by State in 2018–2019

State	Name of Index	Labor Cost Strategy
Alaska	District Cost Factor	Teacher Cost Index
Colorado	Cost of Living Factor	Cost of Living Index
Florida	District Cost Differential	Comparable Wage Index
Maine	Regional Labor Market Area Adjustment	Teacher Cost Index
Maryland	Geographic Cost of Education Index	Teacher Cost Index
Massachusetts	Wage Adjustment Factor	Comparable Wage Index
Missouri	Dollar Value Modifier	Comparable Wage Index
New Jersey	Geographic Cost Adjustment	Comparable Wage Index
New York	Regional Cost Index	Comparable Wage Index
Virginia	Cost of Competing Adjustment	Comparable Wage Index
Washington	Regionalization Factor	Cost of Living Index
Wyoming	Regional Cost Adjustment	Cost of Living Index and Teacher Cost Index

Note: Sources listed in Appendix A.

The most common strategy for regional cost adjustment has been the CWI. Six states—Florida, Massachusetts, Missouri, New Jersey, New York and Virginia—use some sort of CWI in their state funding formula. Most of the six estimate their CWI using data collected by their state agencies as part of the Bureau of Labor Statistics' Occupational Employment Survey.

Three states—Colorado, Washington and Wyoming—use a cost of living index for regional cost adjustments. A cost-of living index measures the cost of purchasing a designated basket of consumer goods—such as food, housing or clothing—in each locality. Researchers construct an educational cost-of-living index using the same basic methodology as is used in the construction of common inflation measures like the consumer price index (CPI).

Finally, four states use a teacher cost index (TCI) to measure regional differences in labor cost. A TCI is based on an analysis of teacher compensation within the state. Researchers use regression analysis to separate the observed variation in teacher salaries into the part that is explained by school district decisions (such as teacher demographics or teaching assignments) and the part that is systematically related to factors outside of school district control (such as the cost of living, the degree of geographic isolation or student demographics). The researchers then use their regression model to predict the salary that each district would need to pay to hire a teacher with an identical set of characteristics. Finally, they construct a TCI as the local salary prediction divided by some reference salary (such as the state average prediction or state minimum prediction).

A New Texas Teacher Cost Index

There are a number of benefits to using a TCI to measure geographic differences in labor cost. While other models rely on indirect measures of labor cost, TCIs arise from regression analysis of existing teacher salaries. As such, TCIs are directly relevant measures of educational costs. In addition, TCIs can be measured at the school or school district level, allowing them to identify systematic cost differences across districts in the same labor market.

On the other hand, TCIs can also have several disadvantages. Teacher characteristics are generally treated as controllable in such analyses, and a failure to adequately account for differences in teacher quality can lead to measurement errors that misidentify high spending districts as high cost districts (Goldhaber 1999; Rothstein and Smith 1997). TCIs have also been criticized as biased by noncompetitive teacher labor markets (Goldhaber, Destler, and Player 2010; Hanushek 1999), susceptible to school district manipulation (McMahon 1996) and vulnerable to errors of estimation (Taylor and Keller 2003).

The Texas Cost of Education Index (CEI), which was an element in the Foundation School Program from 1991 until 2019, was a TCI that had been estimated using teacher salary data from the 1988–1989 school year. Although researchers updated the analysis many times over the intervening 25 years (e.g., Alexander et al. 2000, Taylor 2004, Taylor 2015a), the CEI remained unchanged until it was removed from the Foundation School Program by HB3.

This report presents a new TCI for Texas. The salary model used in this analysis updated and extended the model used in Taylor, Gronberg, and Jansen (2017) and described the observed pattern of teacher salaries in Texas as a function of labor market characteristics, job characteristics, observable teacher characteristics, and unobservable teacher characteristics.[1]

Using the model, one can predict how much each district must pay, each year, in order to hire a teacher with standard characteristics (i.e., a master's degree and 10 years of experience, or a bachelor's degree with zero years of experience). The TCI for each district (each year) is the predicted salary in that district for a teacher with a standard set of characteristics who was assigned to a standard campus, divided by a minimum predicted salary (for that year).[2] Because teacher and

[1] Formally, the model can be expressed as:

$$\ln(W_{idjt}) = D_{dt}\beta + T_{it}\delta + M_{jt} + \eta_i + \varepsilon_{idjt}$$

where the subscripts i, d, j and t stand for individuals, districts, labor markets and time, respectively, W_{idjt} is the teacher's full-time-equivalent monthly salary, D_{dt} is a vector of job characteristics that could give rise to compensating differentials, T_{it} is a vector of individual teacher characteristics that vary over time, M_{jt} is a vector of labor market characteristics, and the η_i are individual teacher random effects. The model was estimated assuming the individual random effects follow an autoregressive (AR(1)) time series process.

[2] The reference prediction used in the construction of the TCI is the prediction at the one-quarter percentile (so that only one quarter of one percent of the districts have a predicted wage below the reference wage). The TCI was set to

campus characteristics are standardized (i.e., held constant) in the construction of the TCI, the resulting index is purely a function of labor market characteristics and other uncontrollable cost factors. As such, it represents a measure of geographic variations in labor costs that arise from factors beyond school district control.

Estimating a New Texas TCI

Data used in this analysis came from the Texas Education Agency, the National Center for Education Statistics (NCES), the US Bureau of Labor Statistics, the National Weather Service, and the US Census Bureau. The analysis covered the six school years from 2014–2015 through 2019–2020, and included all teachers with complete data who worked at least half time for a traditional public school district in a traditional classroom setting.[3]

The data on teacher salaries, teaching assignments and individual teacher characteristics came from TEA's Public Education Information Management System (PEIMS). Following Taylor, Gronberg, and Jansen (2017), this analysis used the log of total, full time equivalent (FTE) annual salary as the measure of teacher compensation.[4] A focus on salary rather than salary and benefits is appropriate in this instance because the most important benefits—pensions and health insurance—do not vary for the vast majority of Texas school districts. All Texas school districts participated in the Texas Retirement System (TRS) and more than 92 percent of districts provided in health insurance through TRS (TRS 2018). Furthermore, Alexander et al. (2000) found that extending their salary analysis to include health insurance benefits had very little impact on the geographic pattern of salary predictions in Texas.

Table 1-2 describes the factors (i.e., variables) included in the salary model to explain observed variations in FTE salaries. The controllable factors capture variations in salary that arise from differences in the teachers themselves and differences in the jobs their districts assign them to do; the uncontrollable factors capture differences in the places where they work and the students they serve. The New Texas TCI is a function of the uncontrollable cost factors in Table 1-2.

1.00 for the handful of districts with predicted wages below the reference wage. This approach ensures that the reference wage is not an extreme outlier.

[3] Thus, data about teachers from open-enrollment charter campuses, virtual campuses, and alternative education campuses have been excluded.

[4] By definition, the FTE annual salary was the observed total salary divided by the percent FTE. FTE annual teacher salaries less than 90% of the state's statutory minimum were deemed implausible and treated as missing data, as were FTE annual teacher salaries in excess of $200,000.

Table 1-2: Controllable and Uncontrollable Cost Factors from the Teacher Salary Model

Controllable Cost Factors	Uncontrollable Cost Factors
Teacher Experience	Working Conditions
Teacher Educational Attainment	– Student Need – Social Security status
New Hire Indicator	Labor Market Conditions
Teaching Assignment – Subject Matter Assignment – Grade-level assignment – Campus Type	– ACS-CWIFT – Fair Market Rents – Unemployment Rate – Geographic Isolation – Climate
Other Duties – Department Head – Administrator – Support Staff	County Type Indicators

Note: Variable definitions provided in Appendix B.

Controllable Cost Factors

School districts clearly have a choice when it comes to the people that they hire, so teacher characteristics are controllable factors (at least in the long run). One of the keys to a successful TCI is the inclusion of sufficient controls for teacher characteristics. The salary model used in the construction of the Texas TCI included a particularly rich set of demographic controls, including measures of teacher experience (log of years of experience, log of years of experience squared, log of experience, cubed and an indicator for first year teaching); teacher training (indicators for whether the teacher held a master's degree, doctorate degree, or did not hold at least of bachelor's degree) and teacher years of service in the district (an indicator for whether the teacher is a new hire).[5]

School districts also controlled the jobs to which teachers were assigned. Therefore, the salary model included indicators for subject-matter assignment (elementary subjects, language arts, mathematics, science, social studies, health and physical education, foreign languages, fine arts, computers, technical/vocational, special education, and tested subjects or grades) and grade-level assignment (elementary or secondary, non-graded, pre-kindergarten, or kindergarten). Teachers could have multiple teaching assignments or serve multiple grade levels. For example, if an individual taught both science and math, or kindergarten and pre-kindergarten students, both of their assignments were recorded in the analysis.

School districts controlled the schools to which teachers were assigned, and some types of schools were conceivably more attractive to teachers than other assignments. Therefore, the model also

[5] The inclusion of multiple controls for teacher experience and educational attainment improves the flexibility of the salary model, thereby improving the extent to which the model controls for district choices about teacher characteristics. This approach is standard in the literature. See, for example, Taylor (2020).

included indicators for school type (elementary, middle, multi-grade, large high school, or other high school).

Finally, school districts controlled whether or not teachers were assigned other duties in addition to teaching. Therefore, the model included indicators for whether or not the teacher served as a department head, a school administrator, or a member of the support staff.

Clearly, there are other factors within the control of school districts that influence the attractiveness of a teaching position. Workplace culture can have considerable impact on teacher job satisfaction, as can the quality of principal supervision or the availability of high quality mentoring (e.g., Nguyen et al. 2020; Bogler, 2001; Borman and Dowling, 2008; Ingersoll, 2001; Tillman and Tillman, 2008). Unfortunately, reliable data on such factors were not available for this analysis. Given the richness of the salary model and the goodness-of-fit between the model predictions and the observed teacher salaries (see below) it is unlikely that omitting these difficult-to-measure controls has led to significant bias in the new Texas TCI.

Uncontrollable Cost Factors

A substantial literature suggests that student demographics are factors outside of school district control that can have a significant influence on the attractiveness of a teaching position (e.g., Loeb, Darling-Hammond, and Luczak, 2005; Borman and Dowling, 2008; or Erichsen and Reynolds, 2019) and therefore on the salaries that districts must pay to attract and retain high quality personnel. The underlying premise is that teaching is more difficult where student needs are greater, so salaries must be higher to compensate for the increased difficulty.

Previous work using data on nonrural teachers in Texas (Taylor, Gronberg and Jansen 2017) found that teacher salaries were higher in campuses with a higher percentage of English Language Learners (ELLs) or special education students, but—contrary to expectations—lower in campuses with a higher percentage of economically disadvantaged students. The multidimensional correlation between local labor market conditions and student poverty likely explains this counterintuitive result.

In order to construct a reliable TCI, it was crucial that the measures of student demographics be outside of school district control. In Texas (as in other states), a student's ELL status is a function of his or her academic performance, and districts clearly influence academics. Students who pass the English reading/ English Language Arts STAAR® test are, by definition, no longer ELL students. Any student who succeeds academically (at least in this dimension) is subsequently removed from the category of ELL students, making the percentage of ELL students subject to school district influence. Therefore, rather than rely on the percentage of students who were ELL, the researchers constructed a measure of student need that was clearly outside of school district control—the percentage of students in the district who had ever been considered English Language learners (ELL). Using data from the Education Research Center at the University of Texas at Dallas, the research team traced each student's academic history to identify those students who had been ELL at some point during their experience in Texas schools (EverELL).[6] On average during the analysis period, 30% of Texas students were EverELL.

Preliminary analysis of the salary model suggested that teacher salaries were higher where the percentage ELL was higher for both metropolitan and nonmetropolitan school districts in Texas. However, the preliminary model also indicated that the relationship between teacher salaries and the percentage economically disadvantaged or the percentage special education was counterintuitive. In other words, estimated labor costs were lower where student need was higher. Because the goal of the analysis was an intuitive, reliable TCI, and the resulting TCI was generally insensitive to the inclusion or exclusion of the measures of student socio-economic status and special education status, the final specification did not include the percentage special education or the percentage economically disadvantaged. [7]

Another uncontrollable factor likely to influence labor cost was whether or not the district's teachers participated in the Social Security System. Most district do not participate, relying on the TRS to protect teachers in their retirements. However, a small number of Texas districts participate in Social Security, not just for their auxiliary workers but also for their teachers. Teachers in those districts pay Social Security taxes on their earnings, and will be eligible for partial Social Security benefits upon retirement. (Federal rules limiting Social Security benefits for individuals with a government pension prevent those teachers from receiving full benefits from the Social Security system.) Unlike other districts, the districts that participate in the Social Security system for teachers are also required to pay the employer's share of social security taxes, making their labor costs higher than other districts for reasons beyond their control.[8] Therefore, the salary model included an indicator for social security status.

The model also included a number of variables designed to capture local variation in labor market conditions. The ACS CWIFT reflected the prevailing wage for college graduates.[9] The US Bureau of Labor Statistic's measure of the county unemployment rate captured additional information about job prospects outside of teaching.

Previous work by Taylor, Gronberg, and Jansen (2017) suggests that teacher wages are not as high as nonteacher wages in metropolitan locations with relatively high housing costs. Therefore, the analysis included the US Department of Housing and Urban Development's estimate of Fair Market Rents for a two-bedroom apartment in the county (and the interaction between the fair market rents and the ACS-CWIFT). The HUD data indicated that rents in the Austin metropolitan

[6] This analysis uses the term EverELL to refer to students who had ever been designated as Limited English Proficient (LEP) in the PEIMS data collection. To avoid the statistical noise associated with anomalous blips in student demographics, a three-year moving average of the percentage EverELL was used in the estimation and construction of the new TCI.

[7] The correlation between the TCI derived from a model including the percentages of economically disadvantaged and special education students, and the TCI derived from a model excluding those variables was 0.998.

[8] For purposes of estimation, the salaries in districts where teachers participate in the social security system were adjusted upward to reflect the employer's share of Social Security taxes (6.2 percent).

[9] The ACS-CWIFT is only available for four of the six years of the analysis period. The 2018 values (which correspond to the 2018–19 school year) were used to fill in for the 2019–20 school year. The 2015 values were used to fill-in for the 2014–15 school year.

area were the highest in the state in 2020, and more than 90 percent higher than rents in the rural counties with the lowest rents.

Climate influences the cost of living because of its influence on energy costs (particularly air conditioning costs). It could also be a component of the general attractiveness of a locale. Therefore, following Alexander et al. (2000), this analysis included a measure of the 30-year average total number of cooling-degree days at the three weather stations that are closest to each teacher's primary campus.[10]

Geographic isolation can influence the salaries teachers are willing to accept in various locations. Therefore, following Alexander et al. (2000) we incorporated two measures of geographic isolation (measured at the school level). The first was the distance to the nearest approved educator preparation program; the second was the distance to the center of the nearest metropolitan area (in miles). In both cases, distances were measured as the crow flies using data on latitudes and longitudes.[11] In addition, the model included a series of categorical variables based on county-type (rural, micropolitan, outlying metropolitan, and central metropolitan) and county population density (sparse and very sparse).

Previous work has suggested that the relationship between teacher characteristics and teacher salaries is different in metropolitan parts of Texas than it is elsewhere in the state (Alexander et al. 2000). Therefore, the salary model allowed for different estimated coefficients in metropolitan and nonmetropolitan areas (as defined by the US Census Bureau).

Estimation Results

Appendix C presents the coefficient estimates and robust standard errors for two alternative specifications of the salary model. The first model is a teacher fixed effects model. The fixed effects methodology adjusts for any variation in salaries that might arise from persistent, but unmeasured, teacher characteristics such as intelligence or verbal ability. As such, it does the best possible job of controlling for differences in salary that could be attributed to school district choices about their employees. Unfortunately, in so doing, it may over-control for the variation in cost that is driven by stable characteristics of school districts. Stable district characteristics—such as geographic remoteness or a persistently high cost of living—will only register for teachers who change districts. If teachers who change districts are not representative of the teaching population as a whole, the fixed-effects model can be misleading. During the period of analysis, inexperienced teachers who did not have an advanced degree were more likely than other teachers to move between districts, and more than 80 percent of teachers did not move at all, suggesting that mobile teachers may be systematically different from teachers who do not move.

The second model is an autoregressive (AR) random effects model. Like the fixed effects model, the AR random effects model incorporates all of the information in the data and (partially) adjusts

[10] The number of cooling degrees for any given day is the number of degrees that a day's average temperature is above 65 degrees Fahrenheit. Climate was measured as the average number of cooling degree days per year during the 30-year period from 1981–2010.

[11]. Where available, latitude and longitude information for campuses came from the National Center for Education Statistics' Common Core Database. The remaining campuses were assigned latitudes and longitudes according to their street address or (if necessary) the zip codes at their street address.

for persistent but unmeasured differences in teacher quality. Unlike the fixed effects model, the AR random effects model captures the influence of cost factors that are relatively stable over time using data from all teachers, not just the teachers who move between districts. Here, the random effects model has been estimated allowing the residuals to follow the autoregressive pattern found in the data. (An autoregressive pattern to teacher salaries means that if a teacher earns more than the model predicts in one year, she will probably earn more than the model predicts the next year too.) The autoregressive error structure further augments the model's ability to control for unobservable teacher characteristics.[12]

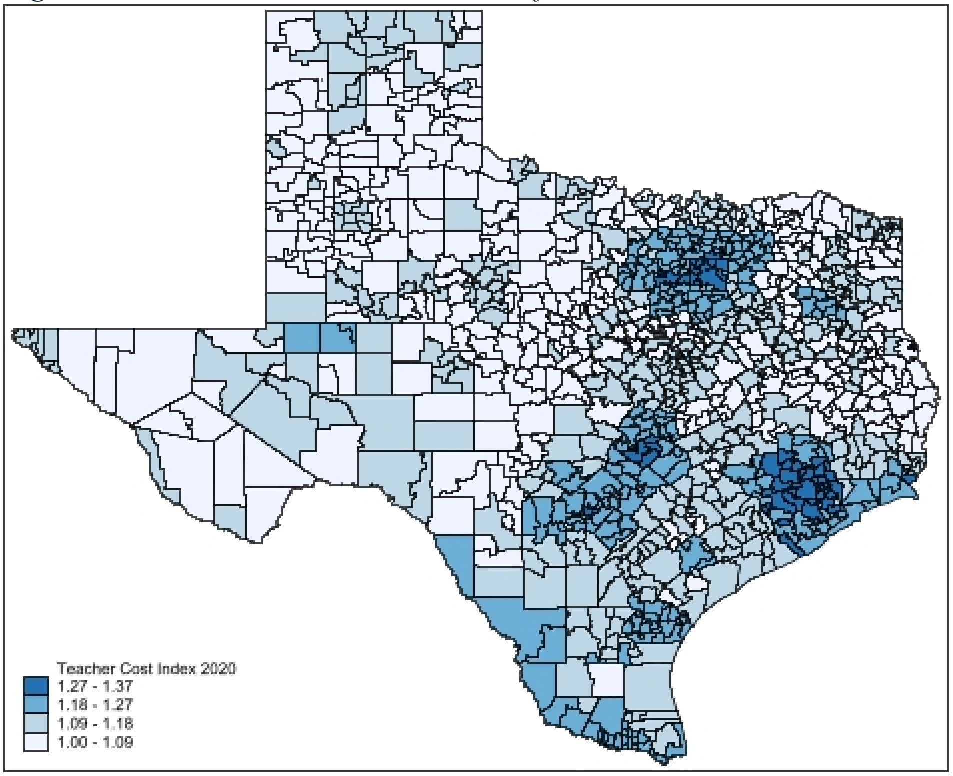
Both models did a very good job of capturing variations in teacher salaries. Including the individual teacher effects, either model explained more than 94 percent of the variation in teacher salaries in Texas. Furthermore, the two models yielded very similar predictions about the salary needed to hire in each school district, and therefore very similar TCIs. (The TCIs generated by the two models had a Pearson correlation of 0.99.) Given the importance of largely time-invariant cost factors (like climate and geographic isolation) in the construction of a TCI and the likelihood that teachers who move between districts are systematically different from those who do not, the AR random effects model represents the preferred specification, and was used to generate the new Texas TCI.

Geography of the New Texas TCI

The Texas TCI for 2019–20 ranged from 1.00 to 1.37, meaning the cost of hiring teachers was 37% higher in highest-cost districts than the lowest-cost districts. As Figure 1-3 illustrates, the Texas TCI was highest in the Houston metropolitan area, and lowest in a district on the outskirts of the El Paso metropolitan area (i.e., a place where teachers have easy access to urban amenities while enjoying a relatively low cost of living).

[12] The Wooldridge test for autocorrelation indicated significant autocorrelation, having generated an F-statistic of 9,536.59 and a probability of a greater F-statistic less than 0.0001.

Figure 1-3: The New Teacher Cost Index for Texas, 2019–20



Source: Authors' calculations.

Figure 1-4 illustrates the responsiveness of the TCI to the various uncontrollable cost factors. For indicator variables, the figure shows the change in the TCI from switching the indicator on or off. For continuous variables, the figure shows the impact of a one standard deviation increase in the cost factor, holding all other cost factors constant at their statewide means. The baseline is the TCI for a rural county where all continuous variables are at the state mean and all indicator variables are switched off.

As Figure 1-4 illustrates, working conditions outside of school district control have a large influence on labor cost. All other things being equal, the TCI is 9 percentage points higher in central metropolitan counties than in rural counties, and 7 percentage points higher in other metropolitan counties than in rural counties. However, labor costs are lower in very sparsely populated metropolitan counties (such as Crosby or Clay) than they are in micropolitan counties.

Teacher wages were systematically higher in locations where nonteacher wages were also higher. A one standard deviation increase in the ACS-CWIFT was associated with a 5 percentage point increase in the TCI in metropolitan areas. The effect was much smaller in nonmetropolitan areas. On the other hand, the positive impact of higher fair market rents on labor costs was stronger in nonmetropolitan counties than in metropolitan counties.

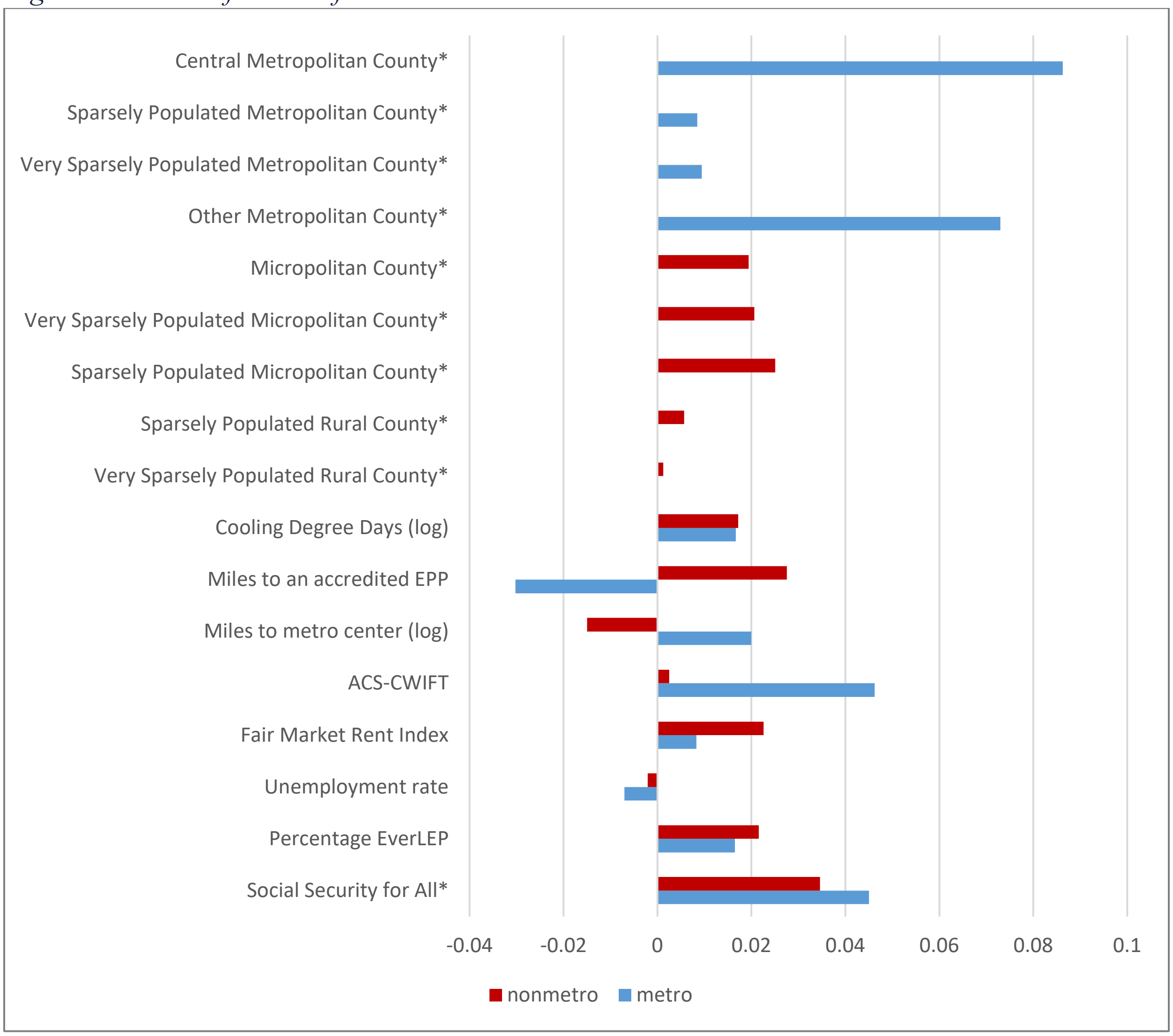
Districts that participate in the social security system face higher labor costs because they must pay the employers share of the social security taxes (6.2 percent). However, the TCI is only 3.5 percent higher in nonmetropolitan districts where teachers participate in social security than in otherwise equal nonmetropolitan districts where teachers do not participate, suggesting that teachers perceive a modest, partially offsetting benefit from participating in the social security system.

Locations where the air conditioning runs all the time have systematically higher labor costs. For both metropolitan and nonmetropolitan locations, a one standard deviation increase in the number of cooling degree days increases the TCI by between 1 and 2 percent.

Districts in nonmetropolitan counties that are far from a metropolitan area have lower wages, but nonmetropolitan districts that are far from an accredited educator preparation program must pay a substantial premium.

Districts with higher student needs also have systematically higher labor costs. A one standard deviation increase in the percentage of students who have ever been designated as ELL is associated with a 2 percentage point increase in the TCI.

Figure 1-4: The Influence of Uncontrollable Cost Factors on the New Texas TCI



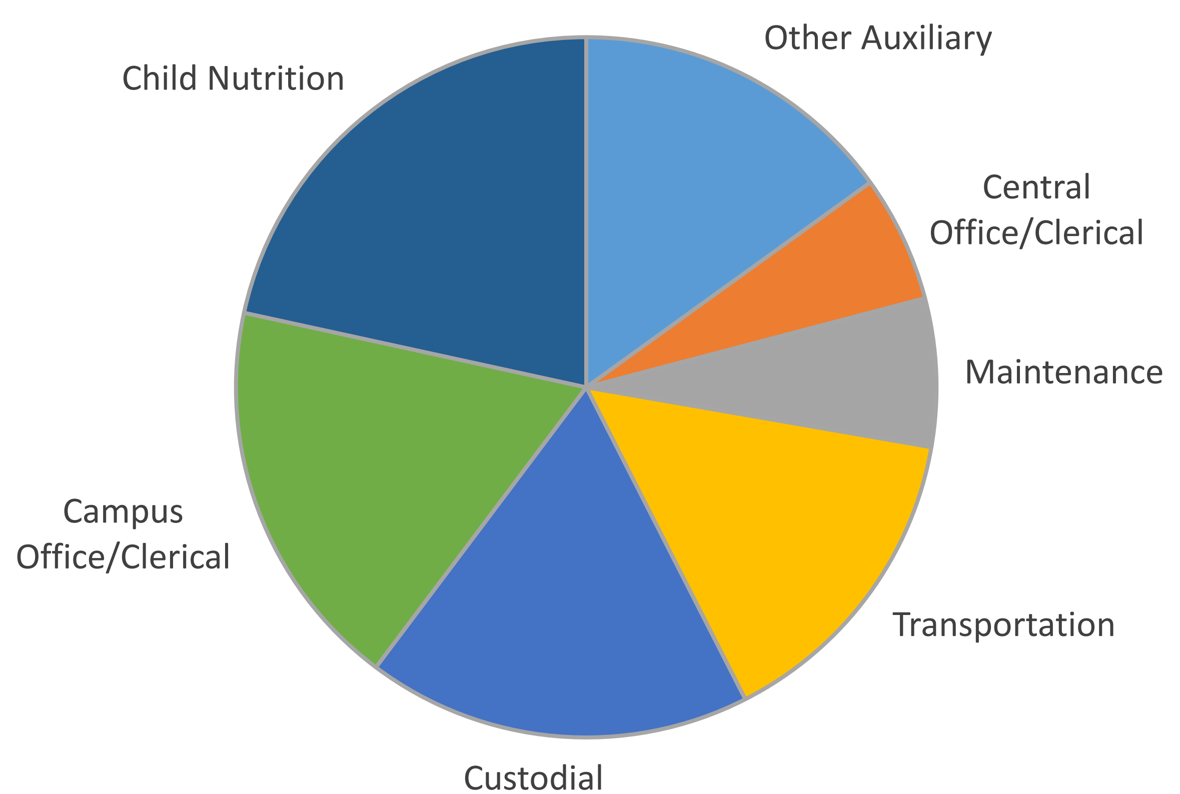
Note: Asterisks indicate indicator variables. Source: Authors' calculations.

A Cost Index for Auxiliary Personnel (APCI)

As the previous section demonstrates, there is substantial geographic variation in wages that school district must pay in order to be able to hire highly qualified teachers. There undoubtedly are also geographic differences in the wages that school districts must pay to hire other school district personnel. The wages for administrators, counselors and other certified personnel are likely to be highly correlated with those of teachers. However, the wages of auxiliary personnel—those holding jobs that do not typically require a professional license or other form of certification such as bus drivers, clerical workers or cafeteria staff—may follow a different geographic pattern. In this section, we examine the geographic pattern in wages for auxiliary personnel.

Figure 1-5 illustrates the wide variety of school district employees who are auxiliary workers. As the figure illustrates, the largest fractions of auxiliary personnel work in child nutrition, clerical, custodial, or transportation service positions. The "other auxiliary" category includes an array of job types such as business office clerks, computer technicians, plumbers, electricians, HVAC personnel, warehouse workers, and safety/security workers.

Figure 1-5: Composition of Employment among Auxiliary School District Personnel 2019–20



Source: PEIMS.

Obviously, there are important differences between computer technicians and transportation workers. The hedonic wage model used in this analysis accommodates those differences using fixed effects for job classifications. The job fixed effects transform the data under analysis into differences from the occupational mean. Thus, the wages of an individual clerk are compared to the state average wage for clerical workers, the wages of an individual maintenance worker are compared to the state average wage for maintenance workers, and so on. Because the geographic patterns captured by the Auxiliary Personnel Cost Index (APCI) reflect differences in these occupationally adjusted wages, it does not matter that wages for computer technician are higher than the wages for custodians, or that one district might have more computer technicians than another. The APCI describes geographic differences that arise when all types of workers in a district earn more (or less) than the state average for their occupations.

Estimating the New APCI

Data used in this analysis came from the Texas Education Agency, the National Center for Education Statistics (NCES), the US Bureau of Labor Statistics, the National Weather Service, and the US Census Bureau. The analysis covered the three school years from 2017–2018 through 2019–2020, and included all auxiliary workers with complete data who worked at least half time for a traditional public school district [13] The auxiliary wage analysis only covers the last three years because the data needed to incorporate fixed effects for occupation were only available for the most-recent three years.

The data on worker characteristics came from TEA's Public Education Information Management System (PEIMS). This analysis used the log of the full time equivalent (FTE) daily wage as the measure of worker compensation.[14] Data on individual worker benefits were not available so (as with the above analysis of teacher salaries) the auxiliary wage analysis did not include benefits.

Table 3 describes the factors (i.e., variables) included in the wage model to explain observed variations in FTE daily wages. The controllable factors capture variations in wages that arise from differences in the workers themselves and differences in the jobs they held; the uncontrollable factors capture differences in the places where they worked.

[13] Data from open-enrollment charter schools have been excluded. Auxiliary employees categorized as "Other Non-Exempt Auxiliary" were also excluded because the category includes non-exempt auxiliary volunteers and there were concerns about the quality of the wage data for those volunteers.

[14] The full-time-equivalent daily wage was calculated as the observed total salary divided by the effective number of days worked, where the effective number of days worked was the number of days employed times the percent day worked. Thus, a person who worked 14 days for 50% of the day had the same number of effective work days as a person who worked seven days for 100% of the day. Daily wages below $58 (the federal minimum for an eight hour day) were deemed implausible and treated as missing data.

Table 1-3: Controllable and Uncontrollable Cost Factors from the Analysis of Auxiliary Personnel Wages

Controllable Cost Factors	Uncontrollable Cost Factors
Worker Potential Experience	Working Conditions
Worker Educational Attainment	– Social Security status – K8 District Indicator – Large District Indicator – Large Footprint Indicator
New Hire Indicator	Labor Market Conditions
Time Worked per year – Days per year – Percent time per day	– HS-CWI – Fair Market Rents – Unemployment Rate – Geographic Isolation – Climate
Job category fixed effects	County Type Indicators

Controllable Cost Factors

School districts clearly have a choice when it comes to the people that they hire, so worker characteristics are controllable factors (at least in the long run). Unfortunately, the PEIMS data contain only limited demographic data on auxiliary personnel. In particular, the PEIMS data do not include a direct measure of experience for auxiliary personnel. Because experience is such an important determinant of earnings, a proxy must be used. The best available proxy for worker experience is the worker's potential years of experience. Therefore, following the approach taken by NCES in the construction of the ACS-CWIFT, the model included the age of the worker (and its square). To allow for the possibility that age was a better measure of experience for men than for women (because women frequently interrupt their careers to have children, which leaves them with fewer years of work experience than a man of the same age) the model also included the sex of the worker and the interaction between worker sex and the age variables.

In addition to the control for potential experience, the wage model also controls for worker educational attainment (an indicator whether the individual held a college degree) and tenure in the district (an indicator for whether the worker had worked in the district the previous year).

School districts also controlled the jobs to which personnel were assigned. Therefore, the salary model included measures of the time worked (the number of days employed per year and percent time per day) as well as indicators for the type of auxiliary position (Business/Finance, Campus Office/Clerical, Central Office/Clerical, Child Nutrition, Human Resources, Information Technology, Campus Technology Specialist, Custodial, Maintenance, Plumber, Painter, HVAC, Electrician, Warehouse, Safety/Security, or Transportation).

As was the case for teachers, there are clearly other factors within school district control that affect the attractiveness of the position. Workplace culture and the quality of supervision are likely to be at least as important for workers in auxiliary positions as they are for teachers. Unfortunately, reliable data on such factors were not available for this analysis. To the extent that these omitted controllable factors were positively correlated with the uncontrollable factors in the model (so that, for example, bosses are better behaved in locations where their employees have lots of job alternatives) then the ACPI probably somewhat understates the geographic variation in wages for auxiliary workers.

Uncontrollable Cost Factors

The uncontrollable factors used in this analysis described working conditions outside of school district control and local labor market conditions.

The first aspect of working conditions was an indicator for whether or not the district participated in the Social Security system for some or all of their employees. Unlike other districts, the districts that participate in the Social Security system are required to pay the employer's share of social security taxes, making their labor costs higher than other districts for reasons beyond their control.[15] A district can participate for some of their employees even if they do not participate for teachers.

The second aspect of working conditions was an indicator for whether or not the district served high school grades. Including this cost factor in the model allowed for the possibility that school districts that only serve students in grades K¬8 could be more (or less) attractive to auxiliary workers than other districts.

The final aspect of working conditions was a pair of indicators for school district size. The first, an indicator for whether or not the district had more than 5,000 students in fall enrollment, was included to capture the relative attractions of a district where auxiliary personnel were likely to be able to specialize rather than wear many hats. The second, an indicator for whether or not the district's geographic footprint covered at least 400 square miles, was included to capture the relative complexity of working for a district where campuses are likely to be highly dispersed.

The model included two variables designed to capture local variation in labor market conditions: the county unemployment rate and the number of potential employers in the vicinity. Theory suggests that school district jobs would be more difficult to fill (and therefore that upward pressure on wages would be stronger) in locations with low unemployment rates and many alternative employers.

The HS-CWI was included in the model to reflect the prevailing wage for high school graduates. The US Department of Housing and Urban Development's estimate of Fair Market Rents for a two-bedroom apartment in the county and the interaction between the fair market rents and the HS-CWI were also included to further capture differences in amenities and the cost of living.

Climate influences the cost of living because of its influence on energy costs (particularly air conditioning costs). It could also be a component of the general attractiveness of a locale. Therefore, this analysis included a measure of the 30-year average total number of cooling-degree

[15] For purposes of estimation, the salaries in districts where auxiliary personnel participate in the social security system were adjusted upward to reflect the employer's share of Social Security taxes (6.2 percent).

days at the three weather stations that are closest to each teacher's primary campus and the average number of heating degree days at those same sites.[16]

As was the case with teachers, geographic isolation could influence the wages that auxiliary employees would be willing to accept in various locations. Here, the analysis incorporated the average distance to the center of the nearest metropolitan area (in miles), and a series of categorical variables based on county-type (rural, micropolitan, outlying metropolitan, and central metropolitan) and county population density (sparse and very sparse).[17]

Appendix D presents the coefficient estimates and robust standard errors for the auxiliary wage model. The model was estimated as an autoregressive (AR) random effects model and, including the individual random effects, explained more than 97 percent of the variation in auxiliary wages.

Geography of the APCI

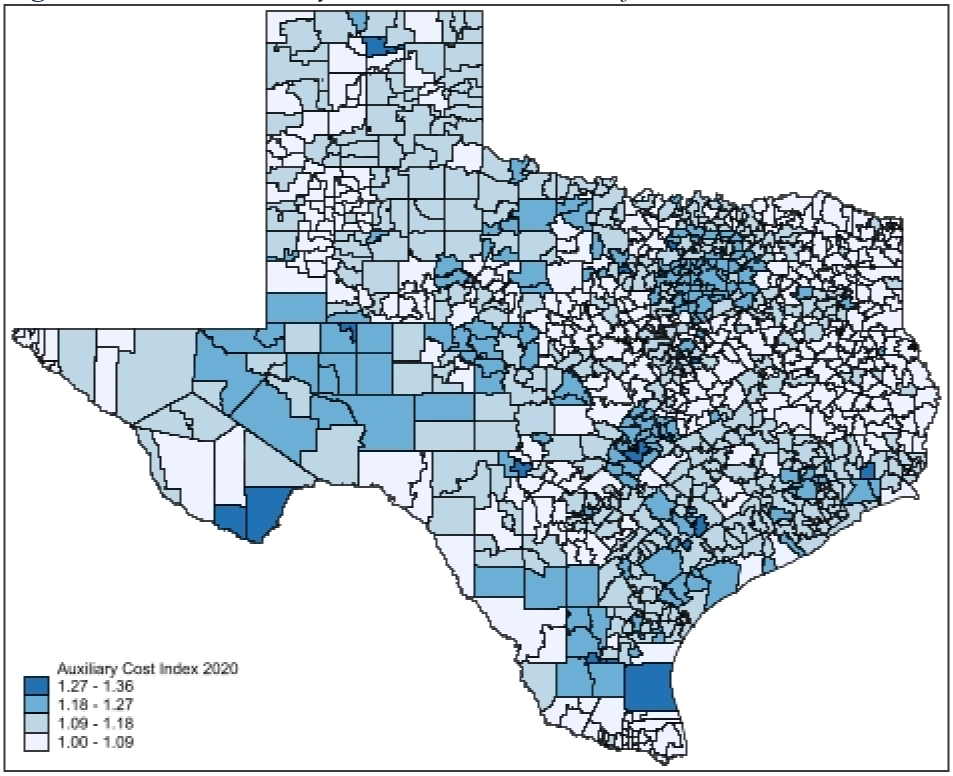
The APCI is the ratio of the predicted auxiliary wage for each district, divided by a state minimum predicted auxiliary wage (for that year).[18] The APCI for 2019–20 ranged from 1.00 to 1.36, meaning the cost of hiring auxiliary workers was 36% higher in highest-cost districts than the lowest-cost districts. As Figure 1-6 illustrates, the APCI was highest among K¬8 districts in Brewster, Duval, and Williamson Counties, and lowest in Hale, Lamb, Sabine, and Starr Counties.

[16] The number of cooling degrees for any given day is the number of degrees that a day's average temperature is above 65 degrees Fahrenheit. Climate was measured as the average number of cooling degree days per year during the 30-year period from 1981–2010.

[17]. Where available, latitude and longitude information for campuses came from the National Center for Education Statistics' Common Core Database. The remaining campuses were assigned latitudes and longitudes according to their street address or the zip codes at their street address.

[18] As with the TCI, the reference prediction used in the construction of the APCI is the prediction at the one-quarter percentile (so that only one quarter of one percent of the districts have a predicted wage below the reference wage). The TCI was set to 1.00 for the handful of districts with predicted wages below the reference wage. This approach ensures that the reference wage is not an extreme outlier.

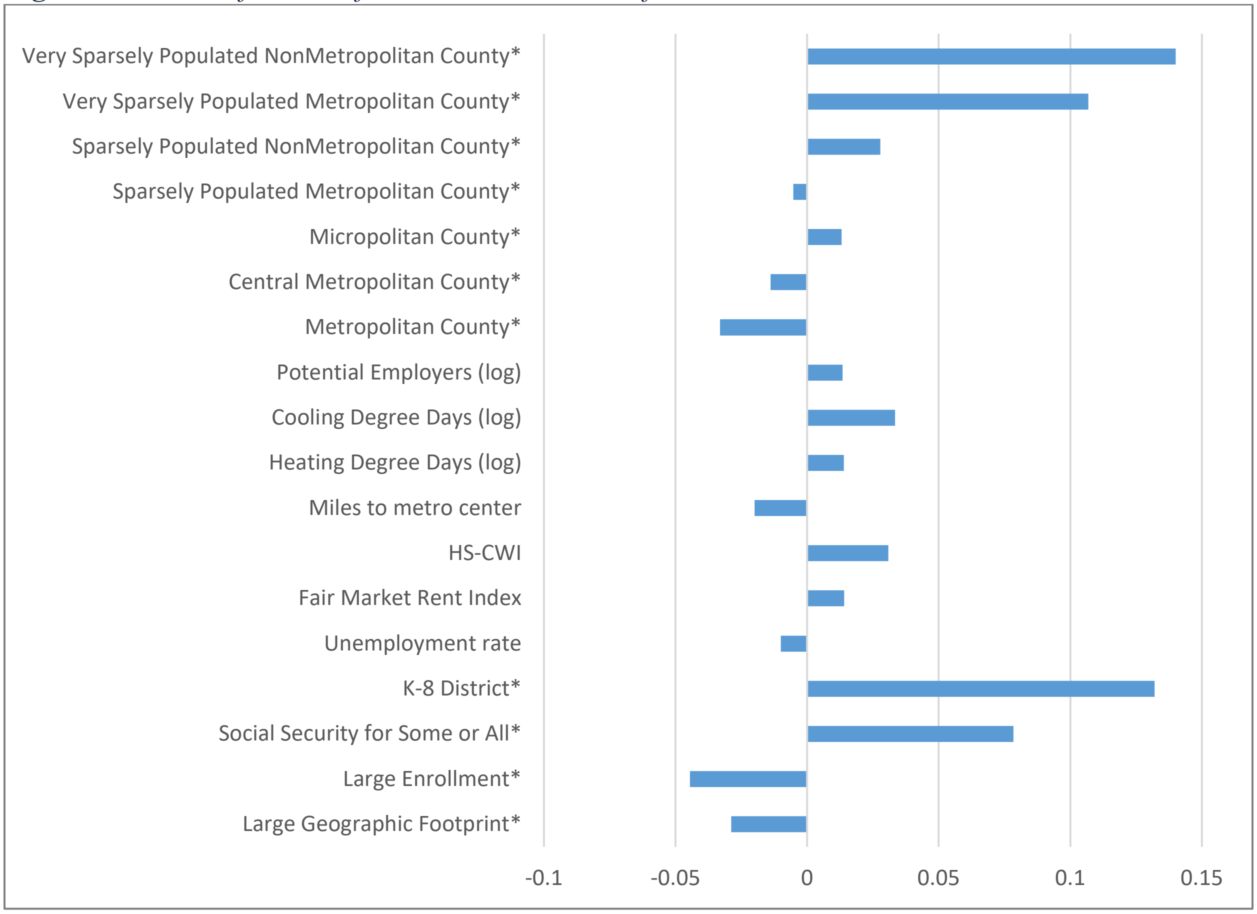
Figure 1-6: The Auxiliary Personnel Cost Index for Texas, 2019–20



Source: Authors' calculations.

Figure 1-7 illustrates the responsiveness of the APCI to the various uncontrollable cost factors. For indicator variables, the figure shows the change in the APCI from switching the indicator on or off. For continuous variables, the figure shows the impact of a one standard deviation increase in the cost factor, holding all other cost factors constant at their statewide means. The baseline is the APCI for a rural county where all continuous variables are at the state mean and all indicator variables are switched off.

Figure 1-7: The Influence of Uncontrollable Cost factors on the APCI



Note: Asterisks indicate indicator variables. Source: authors' calculations.

As Figure 1-7 illustrates, working conditions outside of school district control have a large influence on the cost of hiring auxiliary workers. All other things being equal, the APCI is 3 percentage points lower in metropolitan counties than in rural counties, and 11 percentage points higher in very sparsely populated nonmetropolitan counties than in rural counties. On the other hand, most sparsely populated nonmetropolitan counties were far from the center of a metropolitan area, and the APCI fell as the miles to a metro center increased.

The APCI was systematically higher in locations where general wage levels were also higher. A one standard deviation increase in the HS-CWI was associated with a 3 percentage point increase in the APCI. Higher fair market rents also increased the value of the APCI.

The APCI was lower in districts that were large in either respect, suggesting that such locations are relatively attractive to auxiliary personnel. On the other hand, the APCI was sharply higher in K¬8 districts than in districts that served a full range of grades, suggesting that auxiliary positions are less attractive in those districts and a K¬8 school district must pay a premium to attract high quality auxiliary personnel.

On average, districts that participate in the social security system face a labor cost differential that exceeds the employer's share of the social security tax. This pattern suggests that, on average, auxiliary workers prefer to work at a school district that does not participate in social security.

Implications for School Districts

The TCI and APCI clearly demonstrate that the cost of hiring varies significantly from one Texas district to another. School districts in some parts of the state must pay up to 36 percent more than other districts, just to be able to hire comparable personnel.

Hiring is not the only dimension affected by such differences, however. Regional differences in wage levels also affect district costs associated with teacher turnover and retention.

Although some turnover is undoubtedly beneficial (Billingsley, 1993; Ingersoll and Smith, 2003; Roseman, 1981; Smith and Ingersoll, 2004), most researchers have concluded that, on net, turnover imposes costs on schools. Those costs include the administrative costs of employee separation, as well as the costs associated with employee search, recruitment, and induction. Pinkovitz, Moskal, and Green (1997) emphasized the intangible costs of turnover such as "the uncompensated increased workloads other workers assume due to vacancies, the stress and tension turnover causes, declining employee morale, and decreased productivity due to loss of work group synergy" (p. 71). In addition, new hires frequently takes months to get up to speed in their new environment, leading to lower classroom effectiveness among new hires. Synar (2010) calculated that lost productivity represented nearly 41% of the total costs of teacher turnover.

Rough rules of thumb based on analyses by the US Department of Labor suggest that the cost of teacher turnover equals 30 percent of the leaver's salary (e.g., Nweke et al., 2006; Alliance for Excellent Education, 2004). As a result, turnover represents a significant source of cost for Texas school districts.

Districts where wages are high relative to the TCI are locations where teacher pay is higher than needed to attract and retain the typical Texas teacher. As such, those districts are expected to have either below-average teacher turnover or above-average teacher quality (or both). Similarly, districts where wages are low relative to the TCI are expected to have either higher levels of turnover or lower levels of teacher quality.

An analysis of the relationship between the TCI and teacher quality was beyond the scope of this report. However, a simple analysis of the relationship between teacher turnover and the TCI strongly suggested that districts where salaries were below those implied by the TCI had elevated turnover rates during the period from 2014–15 through 2018–19. (Data on turnover for the last year of the salary analysis, 2019–20, were not available, so the turnover analysis included one fewer year than the teacher salary analysis.)

Table 1-4 compares the turnover rates for districts where salaries were high relative to the TCI with the turnover rates for districts where salaries were low relative to the TCI.[19] As the table illustrates, turnover rates were substantially higher for districts where salaries were low relative to the TCI. The differential was particularly large for beginning teachers, where the districts with the lowest salaries, relative to the TCI, had an average annual turnover rate that was nearly 50% higher (i.e., 7.4 percentage points higher) than the districts with the highest salaries, relative to the TCI.

[19] Relative salaries were determined by comparing the district salary level to the salary implied by the TCI. The district salary level was determined by estimating the hedonic wage index using ordinary least squares after replacing all of the controllable cost factors with district-by-year fixed effects.

Table 1-4: Teacher Turnover for Districts with High and Low Salaries, Relative to the TCI, 2014–15 through 2018–19

Salary Quintiles	Overall Turnover Rate	Beginning Teacher Turnover Rate	Experienced Teacher Turnover Rate
Observed Turnover Rates			
Lowest Relative Salary Quintile	15.8%	23.5%	13.6%
Low Relative Salary Quintile	13.9%	20.7%	12.0%
Average Relative Salary Quintile	12.9%	18.4%	11.3%
High Relative Salary Quintile	12.8%	17.2%	11.3%
Highest Relative Salary Quintile	12.4%	16.1%	11.0%
Demographically Adjusted Turnover Rates			
Lowest Relative Salary Quintile	13.4%	19.0%	12.1%
Low Relative Salary Quintile	13.2%	19.0%	11.7%
Average Relative Salary Quintile	12.7%	18.4%	11.2%
High Relative Salary Quintile	12.3%	17.4%	10.9%
Highest Relative Salary Quintile	11.8%	16.5%	10.5%

Note: To avoid confusing normal retirement patterns with other sources of turnover, this analysis has been restricted to teachers who are no more than 65 years old. Beginning teachers have fewer than 3 years of professional experience; experienced teachers have at least 3 years of experience. Demographically adjusted turnover rates were determined using a Probit analysis, wherein turnover was the dependent variable, and the independent variables were the district salary level and teacher characteristics (i.e., the uncontrollable characteristics in Table 1-2 plus sex, race, age, and age squared).

One might wonder if the pattern above simply reflected higher retention rates where teacher salaries were generally high, not just high relative to the TCI. The second set of turnover estimates in Table 1-4 indicate the turnover rates adjusted for teacher demographics (the uncontrollable characteristics from Table 1-2 plus sex, race, age, and age squared) and the district average salary levels. Although the differentials narrow, the evidence still suggests that districts where salaries were low relative to the TCI had significantly higher turnover than average, and districts where salaries were high relative to the TCI had significantly lower turnover than average.[20] In other words, the evidence suggests that a failure to pay salaries consistent with the TCI has had real consequences for the ability to retain high quality teachers.

Conclusions

In the past, Texas has incorporated geographic cost adjustments into the school finance formula. For the 27 years from 1991–92 through the passage of House Bill 3, the Cost of Education Index (CEI)—which was a teacher cost index similar in spirit to the new TCI—was used to enhance the purchasing power of school districts in high labor cost areas, and thereby enhance the equity of Texas' Foundation School program.

This analysis suggests that adjustments for differences in the price of labor are needed in Texas. Such adjustments level the playing field so that all school districts can recruit and retain the same

[20] The differences among the quintiles are statistically significant at the 1-percent level in all three cases.

sort of high quality personnel despite local conditions that make some districts more attractive to teachers than others. All other things being equal, regions with a high cost of living are less attractive to teachers than regions with a low cost of living, so districts in high cost of living areas must pay higher wages if they want to attract and retain highly qualified teachers. Similarly, regions that have a lot of natural beauty or other local amenities are more attractive to teachers than other regions, so districts without such amenities may need to offer a salary premium to attract teachers. Just as inflation adjustments allow the state to equalize school district purchasing power over time, regional cost adjustments allow the state to equalize purchasing power across locations.

Addition of a geographic cost adjustment in the state funding formula can advance the equity and adequacy goals of the Foundation School Program. Incorporating a geographic cost adjustment into the current funding formula is feasible and can be accomplished through the use of either the CWIFT or the TCI. Through cost adjustments, the Texas legislature can direct additional funding to districts in high-cost environments to ensure all districts can afford the same caliber of teachers regardless of uncontrollable costs.

Chapter 2: Geographic Variation in Costs of Education other than Wages

Schools can provide education services without providing transportation services, and transportation services can be provided apart from education services. The Texas school funding formula addresses transportation service funding separately from funding of education services. Much of the literature on estimating the cost of education services abstracts from transportation service costs either explicitly or implicitly. Our work studies these two components of school services in separate chapters. This chapter focuses on operating costs of providing non-transportation education services. Chapter 3 will address the cost of providing transportation services to and from the location of education services.

The school funding and finance literature has identified three main drivers of uncontrollable variation in educational cost: input prices, student needs, and economies of scale. All three of these drivers can vary geographically. As the previous chapter indicates, differences in the cost of living and the availability of amenities can lead to geographic differences in the wage level. The percentage of students needing compensatory education services can also differ from one district to the next, as can the percentage of students who are English language learners or in need of special education services. Per pupil, smaller districts are more expensive to operate than larger districts for a variety of reasons, and many small districts are located in rural areas. In addition, rural districts can face costs arising from the lack of population density. Sparsely populated districts tend to operate smaller schools than other districts of comparable size, leading to a lack of economies of scale at the school level to higher operating costs in such areas.

The literature has also identified two broad approaches to measuring the impact of geographic differences on the cost of education: bottom-up strategies and top-down strategies. Bottom-up strategies start with the construction of prototype schools and then ask how the characteristics of those schools might vary with school size or student need. As a general rule, bottom-up strategies rely on professional judgement (either that of the researcher or that of local practitioners) to make ad hoc geographic or demographic adjustments to the prototypes. The final step in a bottom-up analysis is to calculate the cost of replicating their prototype schools given the observed, geographic variation in input prices.

Top-down strategies start with the observed inputs and outputs of schools in various locations, and then examine the extent to which differences in circumstance explain differences in expenditures or outcomes. Top-down strategies use statistical analysis to identify the appropriate adjustments for geographic differences in prices, demographics, or economies of scale. The final step in a top-down analysis is to use the statistical model to predict the cost of producing a desired outcome given the observed, uncontrollable characteristics of school districts.

Cost-function analysis is the strategy best suited to an examination of geographic differences in the cost of education, and is the method used here. A cost function is a top-down strategy that estimates the relationship between educational inputs that are purchased (such as teachers, administrators, software, and air conditioning) with an array of environmental factors that are not purchased (such as student abilities) to produce educational outcomes (such as test scores). (For more on the cost function methodology, see Appendix E.)

Previous Analyses of Educational Cost Functions

Cost functions have been widely studied in evaluating the cost of education and in evaluating school funding formulas including possible economies of scale and the cost implications of other differences across educational units.

Many researchers have used cost function analysis to examine educational economies of scale, frequently in the context of potential school district consolidation.[21] Andrews, Duncombe, and Yinger (2002) surveyed 10 cost studies that were published between 1985 and 1999, and concluded that per-pupil cost was very high for school districts with fewer than 500 students, lowest for school districts in the 2,000 to 5,000 student range, and somewhat higher for school districts with more than 5,000 students. More recent cost-function analyses have reached similar conclusions about the high cost of operating small districts (e.g., Imazeki and Reschovsky, 2006, and Eom et al., 2014). Researchers using Texas data have found evidence that many of the apparent economies of scale at the district level actually arise from substantial economies of scale at the campus level (e.g., Taylor et al., 2017, and Gronberg, Jansen, and Taylor, 2017).

In addition to economies of scale, researchers have also used cost function analyses to explore the additional costs associated with variations in student need. As discussed in Golebiewski (2011), Rumberger and Gandara (2008), and Baker, Taylor, and Vedlitz (2008), cost function estimates of the cost associated with serving economically disadvantaged students varied widely. Some of the studies they surveyed found that no additional funding would be needed (Downes and Pogue, 1994) while other studies suggested that economically disadvantaged students require more than twice the funding of students who are not disadvantaged (Duncombe and Yinger, 2005a).

Many cost-function researchers have estimated the additional funding needed to achieve the same level of performance with English language learners (ELL) as with students who are already proficient in English. Recent reviews of the literature include Jimenez-Castellanos and Topper (2012), Golebiewski (2011), and Rumberger and Gandara (2008). They all found that the estimated range of costs is even wider for ELL students than for economically disadvantage students. For example, Duncombe and Yinger (2005b) estimated that the cost of serving an ELL student in Kansas was a statistically significant, but tiny, 0.14 percent higher than the cost of serving a student who was not ELL. At the other end of the spectrum, Duncombe and Yinger (1997) estimated that the cost of serving an ELL student in New York was four times the cost of serving a student who was not ELL. Taylor et al. (2014) and Taylor, Gronberg, and Jansen (2017) found that in Texas the cost of serving a student who had ever been identified as ELL was between 9 percent and 13.5 percent higher than the cost of serving a student who had never been identified as ELL.

A large literature has developed regarding the cost of serving special education students. Recent reviews of the literature include Golebiewski (2011) who notes that there is little consensus as to how to measure the extent of student disabilities, and even less consensus regarding the associated costs. A number of researchers have found that costs were systematically higher for students with more profound disabilities. For example, Gronberg et al. (2004) and Imazeki and Reschovsky (2004) found that the cost of serving students with speech and learning disabilities were much

[21] For example, see Dodson and Garrett (2004); Duncombe, Miner and Ruggiero (1995); Zimmer, DeBoer and Hirth (2009); Gronberg et al. (2015); Taylor et al. (2014); or Karakaplan and Kutlu (2019).

lower than the costs of serving other special education students, although they were still significantly higher than the costs of serving students in regular education programs.

The Educational Landscape in Texas

Size is the most important dimension over which Texas school districts vary. Table 2-1 provides information on the remarkable variation in district enrollments, district operating expenditures per pupil, and transportation expenditures per pupil.

Table 2-1: Total Operating Expenditures per Pupil and Transportation Expenditures per Pupil for Traditional Public School Districts, All Funds, 2018–19, by Enrollment Size Category

Fall Enrollment	Number of Districts	Total Operating Expenditures per Pupil	Transportation Expenditures per Pupil	Transportation as a Share of Total Operating
10,000 and Above	111	$9,800	$307	3.15%
5,000 to 9,999	70	$9,804	$333	3.40%
1,000 to 4,999	332	$10,365	$327	3.17%
500 to 999	194	$11,532	$320	2.78%
Less than 500	315	$13,965	$426	3.01%

Source: Authors' calculations from Texas Education Agency (2020).

There are important features of school districts in Texas that are related to enrollment. Table 2-1 shows that as size increases, both operating expenditures per pupil and transportation expenditures per pupil tend to decrease. This is especially apparent for small districts. There are a large number of districts in Texas with fewer than 500 students. To put this in perspective, many larger districts in Texas have more than 500 students in a single campus, yet 315 districts in Texas—just under one-third of our sample—have fewer than 500 students enrolled. These districts have high operating expenditures per pupil, and high transportation costs per pupil. If we move to districts with enrollment between 500 and 1000, operating expenditures per pupil drop by over $2,400 per pupil, or by over 17% of the value for districts with fewer than 500 students enrolled. Transportation costs fall by $106 per pupil, or by 25%.

The largest districts in Texas, those having enrollments between 5,000 and 10,000 students, and those very large districts with enrollments over 10,000, have the lowest operating costs per pupil and the lowest transportation costs per pupil. The very large districts have average operating costs per pupil nearly identical to the average for districts with enrollment between 5,000 and 10,000, but even at that, their operating costs are just over $4,100 per pupil lower, 30% lower, than the operating costs of the smallest districts. Moreover, the very large districts have average transportation costs per pupil that are $119, or 28%, lower than the average transportation costs per pupil of the smallest districts.

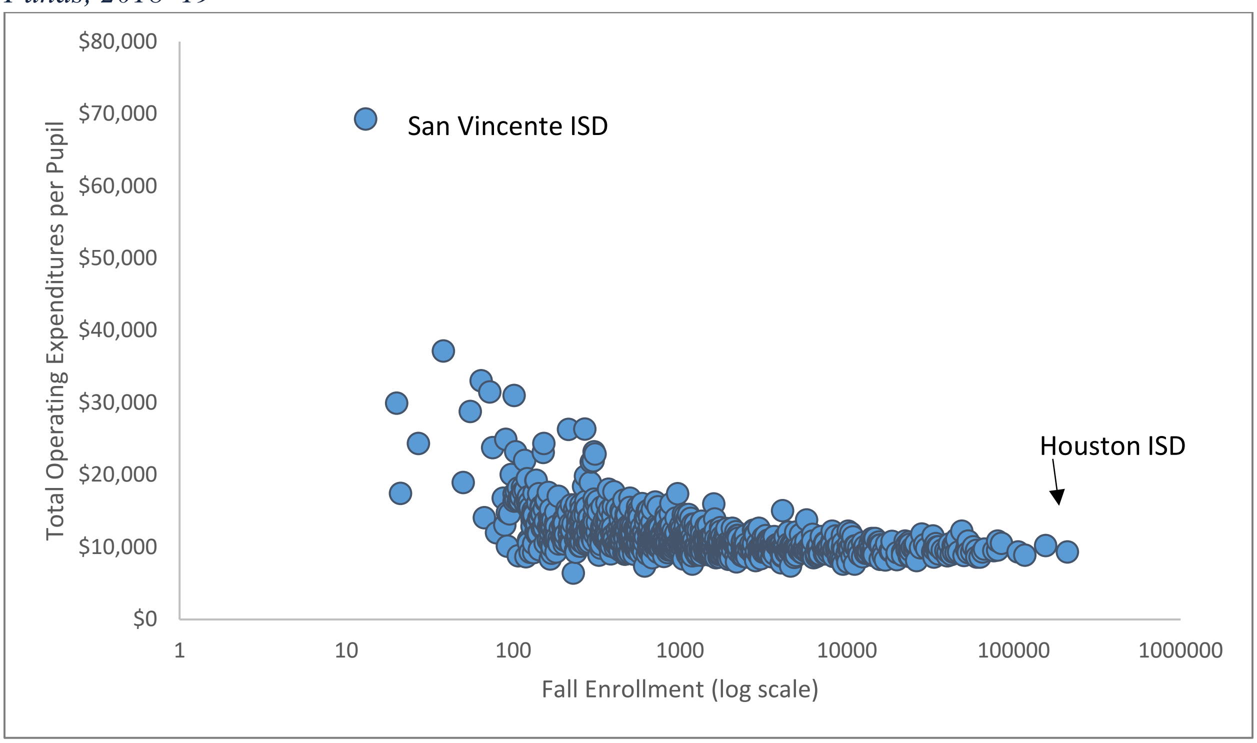
The decline in operating costs per pupil continues as enrollment increases up to that 5,000–10,000 student range, and then holds constant as we move to those districts with enrollment above 10,000 students. This flattening in per pupil spending at the largest districts suggests that there may be limits to the economies of scale as district enrollments increase.

Transportation costs per pupil are quite high for the smallest districts. The decline as we move to districts with enrollments of 500–1,000 is steep, but then costs per pupil change little up to districts of size 5,000–10,000. They decline again for the largest districts with over 10,000 students, indicating some possibility of continuing economies of scale, or a change in the fraction of students receiving transportation services.

It is important to note that the smallest districts and the largest districts vary in many dimensions other than just enrollment. The smallest districts also tend to be in sparsely populated and remote areas where school choice is limited. These geographic differences cannot be addressed by simple consolidation, so proposals to consolidate very small districts in order to lower cost per pupil may be ineffective and could conceivably result in higher, not lower, costs (Taylor et al., 2014 & 2017).

This wide dispersion in district size, especially enrollment, always raises issues about possible economies of scale, and how operating expenditures per pupil varies with district size. There are well-recognized economies of scale in education, and the per-pupil cost of operating a very small district is much more than the cost of operating a larger district. Information on this issue can be gleaned, on an initial pass, by looking at Figure 2-1, which plots total operating expenditures per pupil against district size. Smaller school districts in Texas clearly tend to spend more per pupil on operations than larger ones. This is also shown in Table 2-1, as total operating expenditures per pupil are lowest (on average) for districts with at least 5,000 students, and highest (on average) for districts with fewer than 500 students. In 2018–19, the smallest district in the state, San Vicente Independent School District, spent more than four times as much per pupil as did the largest district in the state, Houston Independent School District (ISD). (See Figure 2-1).

Figure 2-1: Total Operating Expenditures per Pupil for Traditional Public School Districts, All Funds, 2018–19



Source: TAPR and PEIMS.

A Cost Function for Educational Services.

Cost function analysis has been widely used in a wide variety of contexts for over half a century, and in the education context for at least the last three decades. When properly specified and estimated using stochastic frontier analysis (SFA), the education cost function is a theoretically and statistically reliable method for estimating the relationship between the cost of education and various cost drivers, both those that are under the control of school districts and those that are considered uncontrollable by school districts. The cost drivers include measures of enrollment and measures of density such as population per square mile. Other cost drivers that vary by geography include labor costs, materials costs (as proxied by distance from the nearest metro area), and insurance costs (as proxied by distance from the coast). See Gronberg et al. (2015) and Taylor et al. (2014) for a discussion of the use of cost functions within a SFA.

The key components of the cost function analysis are summarized in Table 2-2 and described in the sections below. For a technical description of the cost function analysis, see Appendix F.

Table 2-2: Key Components of the Educational Cost Function

Component	Measured by
Units of Analysis	All Standard Campuses in Traditional Public School Districts
	Five Most Recent School Years (2014–15 through 2018–19)
Expenditures	Operating Expenditures Excluding Food and Transportation
Outcomes	Average Conditional NCE Scores on State Assessments
	Campus Number of Students Enrolled
Input Prices	Teacher Cost Index
	Auxiliary Personnel Cost Index?
	Distance to the Center of the Nearest Metropolitan Area
Environmental Factors	% Economically Disadvantaged
	% Ever Limited English Proficient (Ever-ELL)
	% Special Education
	% High-Needs Special Education
	Campus Type (high school, middle school, multi-grade school)
	K–8 District Indicator
	Metropolitan and Micropolitan Area Indicators
	County Population Density (sparse and very sparse indicators)
	High Windstorm Risk County Indicator
Controls for Inefficiency	Stochastic Frontier Analysis
	Degree of Educational Competition

Units of analysis

This study looks at individual campuses within a district as the main unit of analysis. The data covers a time period ranging from 2014–15 through 2018–19.

To develop the best possible estimates of the size-cost relationship, the cost-function analysis includes all standard accountability campuses in traditional public school districts.[22,23] Standard accountability campuses are subject to all the rules and regulations pertaining to the Texas Accountability Rating System and therefore share a similar set of goals, objectives, and educational processes (TEA, 2014). Alternative Education Accountability (AEA) campuses (e.g., juvenile justice campuses, disciplinary education campuses, residential campuses, and all other alternative education campuses) have been excluded because they are subject to different accountability requirements and may have different cost structures than other campuses. Because they operate under a different set of rules and regulations than traditional public school districts and consolidation does not imply deregulation, open-enrollment charter schools have also been excluded from the data set.

Expenditures

The educational cost function seeks to explain variations in educational expenditures using data on educational outcomes, input prices, and environmental factors. Here, educational expenditures are measured as operating expenditures per pupil, excluding food and student transportation expenditures. It is customary to exclude food and transportation expenditures from the measure of expenditures used in cost function analyses because those categories of expenditures are unlikely to be explained by the same factors that explain student performance, and therefore add unnecessary noise to the analysis.[24]

The actual expenditures data come from the Public Education Information Management System (PEIMS) and have been adjusted to account for school districts that serve as a fiscal agent for another school district or group of districts.[25] All expenditures have also been adjusted to account for the fact that districts differ in the percentage of their total spending they attribute to specific campuses. Some districts provide maintenance services centrally, for example, whereas other districts assign maintenance personnel to specific buildings. To ensure that all of the educational resources in a district are accounted for, school district expenditures that were not associated with a specific campus have been allocated to the district's campuses on a per pupil basis.[26] Thus, for

---

[22] Although many Texas school districts cross county lines, TEA officially associates each school district with a single county. Those official designations have been used to identify Core Based Statistical Area (CBSA) locations for campuses in traditional public school districts, using the July 2015 CBSA definitions developed by the US Office of Management and Budget and published by the US Census Bureau. A metropolitan area is a county or cluster of counties with a central, urbanized area of at least 50,000 people. A micropolitan area is a county or cluster of counties with a central city of at least 10,000 people. Two counties are considered part of the same CBSA whenever commuting patterns indicate that the counties are part of the same integrated labor market area. In Texas, College Station-Bryan is a metropolitan area, and Nacogdoches is a micropolitan area.

[23] Virtual campuses and campuses that lack reliable data on student performance (such as elementary education campuses that serve no students in tested grades, or very small campuses) have also been excluded.

[24] For examples, see Gronberg, Jansen, and Taylor (2011a, 2011b), Gronberg, Jansen, Taylor, and Booker (2004, 2005); or Imazeki and Reschovsky (2006).

[25] Fiscal agents collect funds from member districts in a shared service agreement, and make purchases or pay salaries with those shared funds on behalf of the member districts. As a result, spending of fiscal agents is artificially inflated while the spending by member districts is artificially suppressed. See Appendix F.

[26] Taylor et al. (2014) and Gronberg et al. (2012) also followed this approach.

example, if Little Elementary serves 20% of the students in its district, it is presumed to be responsible for 20% of the unallocated spending.

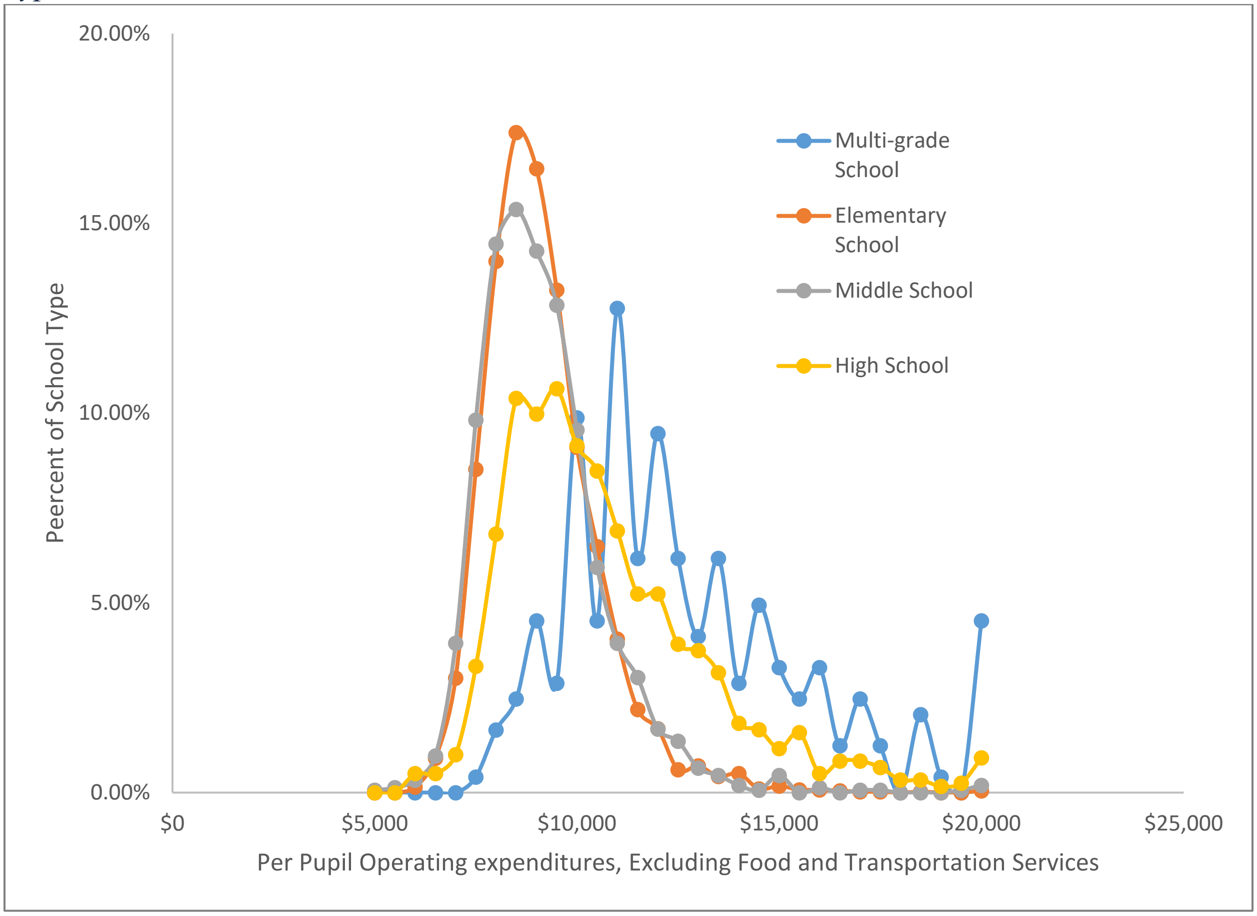
Figure 2-2 illustrates the distribution of operating expenditures per pupil for the standard accountability campuses used in this analysis.[27] Note that because these operating expenditures exclude food and transportation services, and have been adjusted both for shared service agreements and differences in the percentage of spending attributed to campuses, they may not align with what TEA publishes on TXschools.gov.

As the figure illustrates, operating expenditures in 2018–19 ranged from $5,000 to more than $20,000, per pupil. Expenditures per pupil were significantly higher for multi-grade campuses (those that could not be classified as elementary, middle, or high schools) than for any other type of campus, largely because this category includes a number of small, single campus districts such as Harrold ISD in the Vernon, Texas, micropolitan area.[28] On average, spending was significantly higher in high schools (where the mean in 2018–19 was $10,716) than in elementary schools (where the mean was $9,132) or middle schools (where the mean was $9,160). The difference in average spending between elementary and middle schools was not statistically significant.

[27] Per-pupil operating expenditures less than $3,500 or more than $33,000 were deemed implausible and treated as missing in this analysis.

[28] Throughout this report, the term "significantly" indicates something that is statistically significant at the 5% level, meaning that there is less than a 5% chance that the difference is due to chance alone.

Figure 2-2: Operating Expenditures per Pupil for Standard Accountability Campuses by School Type, 2018–19

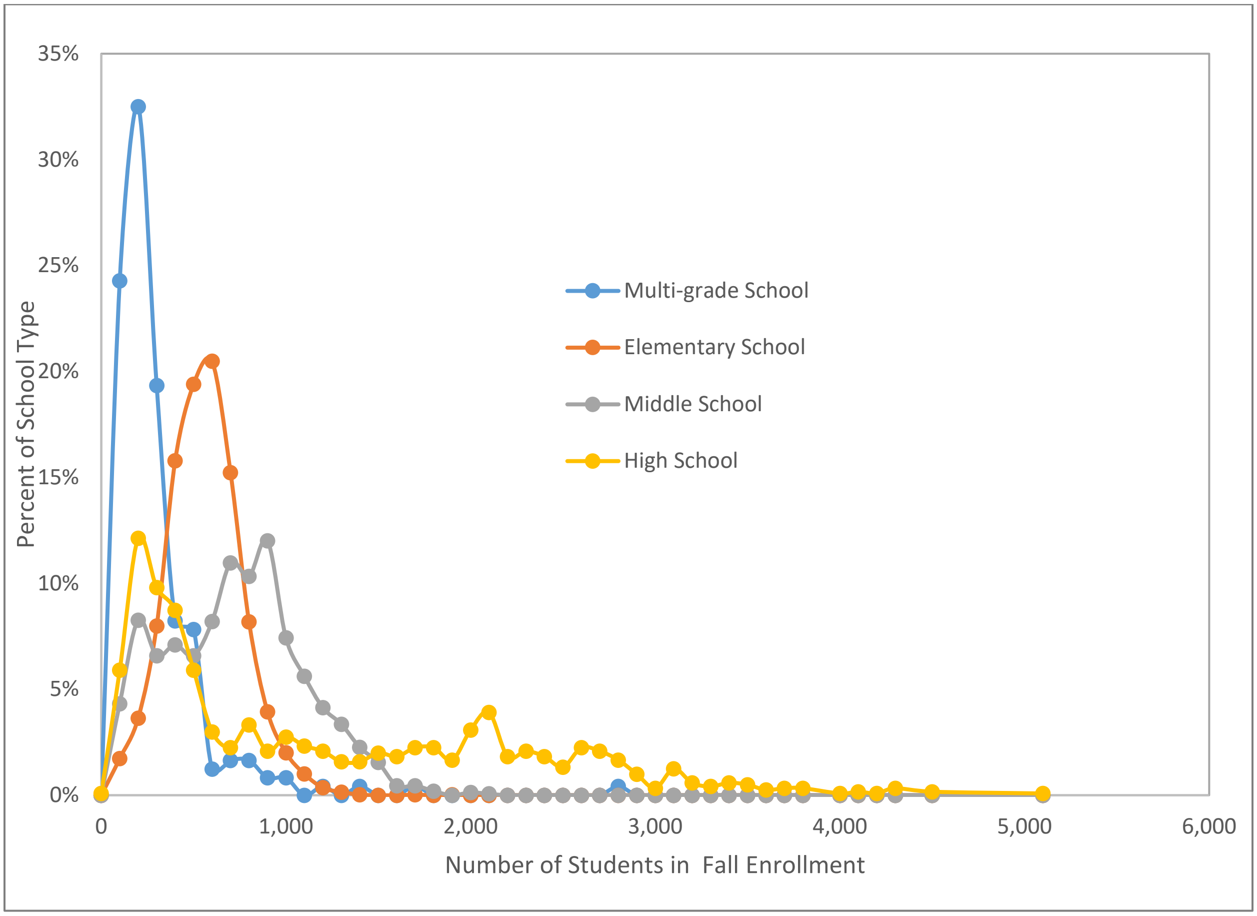


Source: PEIMS.

Outcomes

Educational outcomes have both a quantity and a quality dimension. Quantity is measured using the number of students in fall enrollment at the campus. In 2018–19, campus enrollment in the estimation sample ranged from 41 to 5,098 students; the average campus had 696 students (Figure 2-3). On average, elementary schools were significantly smaller than middle schools which in turn were significantly smaller than high schools. Typically, multi-grade schools were the smallest type of all, but there were a few exceptions to this rule. For example, Benbrook Middle/High School in Fort Worth ISD (which serves Grades 6–12) was a multi-grade campus with an enrollment above 1,700 in 2018–19.

Figure 2-3: Campus Enrollment for Standard Accountability Campuses, by School Type, 2015–16



Source: TAPR.

The quality measure used in this analysis captures differences in average student performance in reading and mathematics. This measure is based on student performance on the required State of Texas Assessments of Academic Readiness (STAAR®) Grades 3–8 and end-of-course (EOC) exams.[29] Although schools clearly produce outcomes that may not be reflected in mathematics and reading test scores, these are performance measures for which districts are held accountable by the state, and the most common measures of school district outcome in the literature.[30] Therefore, they are reasonable outcome measures for cost analysis.

STAAR® Grades 3–8 and EOC scores can be difficult to compare across grades, years or testing regimes. Therefore, the various test scores have been transformed into conditional normal curve equivalent (NCE) scores.[31] A conditional NCE score describes a student's performance relative to what would have been expected given his or her prior test score (i.e., conditional on the prior test

[29] Only state-mandated assessments in reading and mathematics are included.

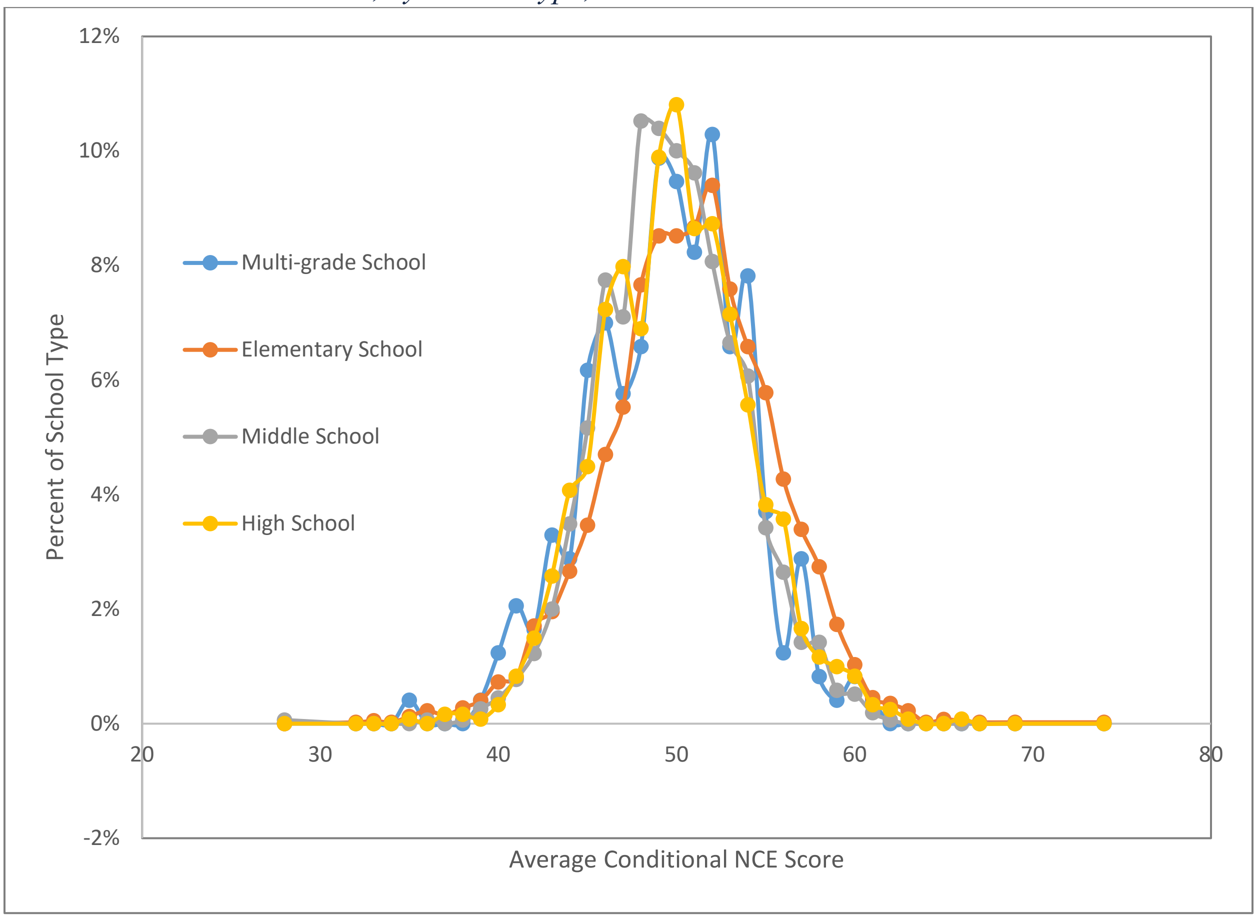
[30] For example, see Gronberg et al. (2011a, 2011b); Grosskopf et al. (2013); Grosskopf, Hayes, and Taylor (2014); or Imazeki and Reschovsky (2006).

[31] For more on the construction of conditional NCE scores, see Appendix F.

score). A conditional NCE score of 50 indicates that the student performed at the 50th percentile (i.e., exactly as expected given his or her prior test performance) and a conditional NCE score of 90 indicates that the student performed as well or better than 90% of his or her academic peers. The average conditional NCE score in mathematics and reading for each campus is the quality measure used in this analysis.

Figure 2-4 illustrates the distribution of average conditional NCE scores in 2018–19. As the figure illustrates, the distribution of average conditional NCE scores is bell-shaped, with most standard accountability campuses in CBSAs having average conditional NCE scores between 40 and 60.[32]

Figure 2-4: Campus Average Conditional NCE Scores for Standard Accountability Campuses in Core Based Statistical Areas, by School Type, 2018–19



Source: Authors' calculations; PEIMS.

Input Prices

One key to estimating an educational cost function is identifying a measure of the price schools must pay for their most important inputs—educators. This analysis uses the new Texas TCI as the

[32] In the interests of statistical reliability, campuses with fewer than 25 students for whom a conditional NCE could be calculated were excluded from the analysis.

price index for profession staff (i.e., teachers, administrators and professional staff) and the APCI as the price index for other staff (i.e., auxiliary personnel and instructional aides).

Ideally, the analysis would also include direct measures of local prices for instructional equipment and classroom materials. Unfortunately, such data are not available. However, prices for pencils, paper, computers, and other instructional materials are largely set in a competitive market (and therefore unlikely to vary across schools), and prices for nonprofessional labor or building rents are largely a function of school location (and therefore likely to be highest in the central cities and lowest in the suburbs or the micropolitan areas). Therefore, as in in Gronberg et al. (2015) and Taylor et al. (2014) the cost analysis includes the distance to the center of the nearest metropolitan area as a proxy for differences in the cost of non-labor inputs.[33]

Electricity is another important input to the educational process, and energy prices have clearly been a source of volatility over the five years included in the cost analysis. However, despite significant year-to-year changes in energy prices, there is little evidence that electricity prices vary geographically within Texas. Recent work by Woerman (2018) found that all of Texas basically faces one price for electricity—especially during the months of the traditional academic school year. Therefore, differences in energy prices are unlikely to lead to significant geographic differences in the cost of education—although they do contributed to significant geographic differences in transportation costs (see Chapter 3).

Environmental Factors

There are several environmental factors that influence the cost of education but are not purchased inputs. One such factor is the size of the school district. As Figure 2-1 and Table 2-1 illustrate, district enrollment for the campuses used in this analysis ranges from fewer than 1,000 students to more than 200,000 students. The median school district in the analysis sample has fewer than 1,700 students and three quarters of the districts have fewer than 5,000 students.

Another such factor is the grade range of the school district. Some Texas school districts serve only elementary grades. Although this analysis was conducted at the school level, there might still be systematic differences between K–12 districts and K–8 districts that influence the cost of education. A district without a high school could specialize more than other districts of similar size, for example, which could lead to lower overall costs. Therefore, this analysis included an indicator for whether or not the district served only elementary grades. (The one traditional public school district that does not serve elementary grades, South Texas ISD, has been excluded from the analysis.)

Population density and metropolitan status are factors that constrain district choices about campus size and could influence other aspects of the educational technology. For example, districts in sparsely populated counties cannot take advantage of the school-level economies of scale available to other districts of similar size because their populations are so dispersed. Instead, such districts must operate smaller schools than other districts, which drives up costs. In addition, districts in metropolitan areas may incur costs (such as school security costs) that are not incurred by districts in other parts of the state. Therefore, this analysis includes indicators for whether or not the district

[33] Miles to the center of the nearest metropolitan area was calculated as-the-crow-flies for each campus using latitude and longitude information.

is located in a metropolitan or micropolitan county, and indicators for whether or not the district is located in a sparsely or very sparsely populated county.[34]

Another geographic cost factor is insurance risk. The Texas Department of Insurance designates 14 Texas counties along the gulf coast as potential windstorm catastrophe areas.[35] Districts in those counties (and in the cities of Morgan's Point, La Porte, Shoreacres, Pasadena, and Seabrook) have elevated risk of damage from a hurricane or tropical storm, and therefore face higher costs to purchase insurance or self-insure. Therefore, this analysis included an indicator for whether or not the district was in a designated catastrophe area.

The other factors identified as influencing the educational environment are student need and school type. To capture variations in cost that derive from variations in student need, the analysis includes three measures of student demographics—the percentages of students who were identified as economically disadvantaged, special education, or ever identified as English language learners (EverELL).,[36] In addition, because previous work by Gronberg et al. (2005) suggested it was important, the analysis also included the percentage of special education students in the district with relatively high needs.[37] To capture differences in the cost of education that arise from differences in mandatory class sizes, or the scope of instruction, the analysis also includes indicators for high schools, middle schools, and multi-grade schools.(Elementary schools were the baseline against which the other school types were measured.)

Controlling for Inefficiency

If schools are behaving efficiently, then increases in educational outcomes will require increases in educational expenditures…but there is no guarantee that all school districts are behaving efficiently. This analysis relies on SFA because, unlike other statistical techniques, SFA explicitly allows for the possibility that spending could be systematically higher than cost. If schools are behaving efficiently, then SFA generates the same cost function estimates as other estimation techniques. Therefore, SFA can be thought of as a more general approach.

When the educational cost function is estimated using SFA, school spending is presumed to depend not only on the direct determinants of educational cost (outcomes, input prices and environmental factors) but also designated factors that could lead one school district to behave more efficiently than another. Previous analyses of Texas data have found that school districts in communities

---

[34] A sparsely populated county has a population density of fewer than 20 persons per square mile; a very sparsely populated county has a population density of fewer than 10 persons per square mile.

[35] The First Tier Counties are: Aransas, Brazoria, Calhoun, Cameron, Chambers, Galveston, Jefferson, Kenedy, Kleberg, Matagorda, Nueces, Refugio, San Patricio, and Willacy. The cities of Morgan's Point, La Porte, Shoreacres, Pasadena, and Seabrook are also considered part of the designated catastrophe area.

[36] For statistical reasons, the measure of ELL status used in this analysis includes not only students who are currently ELL, but also any students who have ever been identified as ELL by the Texas school system. The percentage of students who have ever been identified as ELL greatly exceeds the percentage of students currently identified as ELL in some campuses.

[37] Following Gronberg et al. (2005), high needs special education students are special education students who have any classification other than learning disability or speech-language disability. Due to privacy concerns, these data are not available at the school level.

where parents have more choice about their educational providers tend to behave more efficiently (e.g., Taylor et al. 2014 and 2017). Therefore, the model included a common measure of educational competition—the Herfindahl index—as a possible determinant of school district efficiency. [38] To fully reflect the public school choices available to parents, both traditional public school districts and open-enrollment charter schools were included in the calculation of the Herfindahl index.

Cost Function Results

As detailed in Appendix F, the cost function analysis yields a reasonable picture of the educational process in Texas. According to the cost function estimates, all else equal, increases in average student performance require increases in educational expenditures. Campuses with a higher Texas TCI have a higher cost of education, as do campuses with a higher APCI. Students with greater needs are more costly to educate, K–8 districts are less costly to operate than school districts of similar size that serve the full grade range, and districts in very sparsely populated counties are much more costly to operate than districts in other parts of the state.

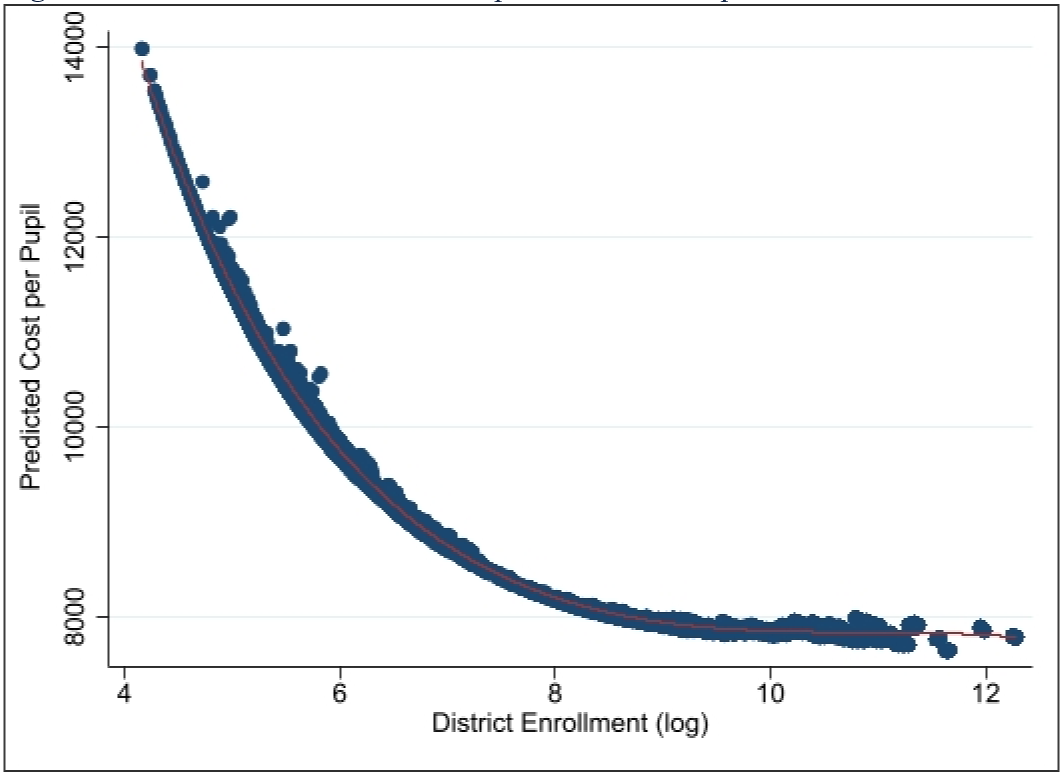
Findings on Economies of Scale

The analysis revealed significant economies of scale for both campuses and districts. As a general rule, increases in campus size led to decreases in the cost of education. For example, the cost function indicated that, all other things being equal, a 200-student campus cost 4% more to operate than a 400-student campus, which in turn costs 2.5% more to operate than an 800 student campus. Costs per pupil were minimized at a campus size of 1,500 students. However, the economies of scale at the campus level were largely exhausted once campus enrollment reached 1,000. The difference in per-pupil cost between a campus of 1,000 students and a campus of 1,500 students was only 0.3%.

The relationship between district enrollment and predicted cost was more complicated, and easiest to understand with the aid of a picture. Figure 2-5 presents predicted per-pupil cost, holding all of the inputs, outputs and environmental factors except campus and district enrollment constant at their sample means. Because it is not possible for a district with 500 students to have an average campus size of 700 (roughly the sample mean), the figure could not credibly be constructed at the sample mean of campus size. It was even less plausible that a district with 500 students could have a campus size of 1,470 (the cost minimizing campus size). Therefore, for the purposes of this illustration, it was assumed that each campus in the district had the average campus-level enrollment for that district. The red line through the middle of the data cloud represents a nonlinear approximation of the combined effect of both campus and district size.

[38] A Herfindahl index is defined as the sum of the squared local education agency (LEA) enrollment shares, where an LEA's enrollment share is its own enrollment divided by the total enrollment in the metropolitan area, micropolitan area or rural county.

Figure 2-5: The Estimated Relationship between Per Pupil Cost and School District Enrollment



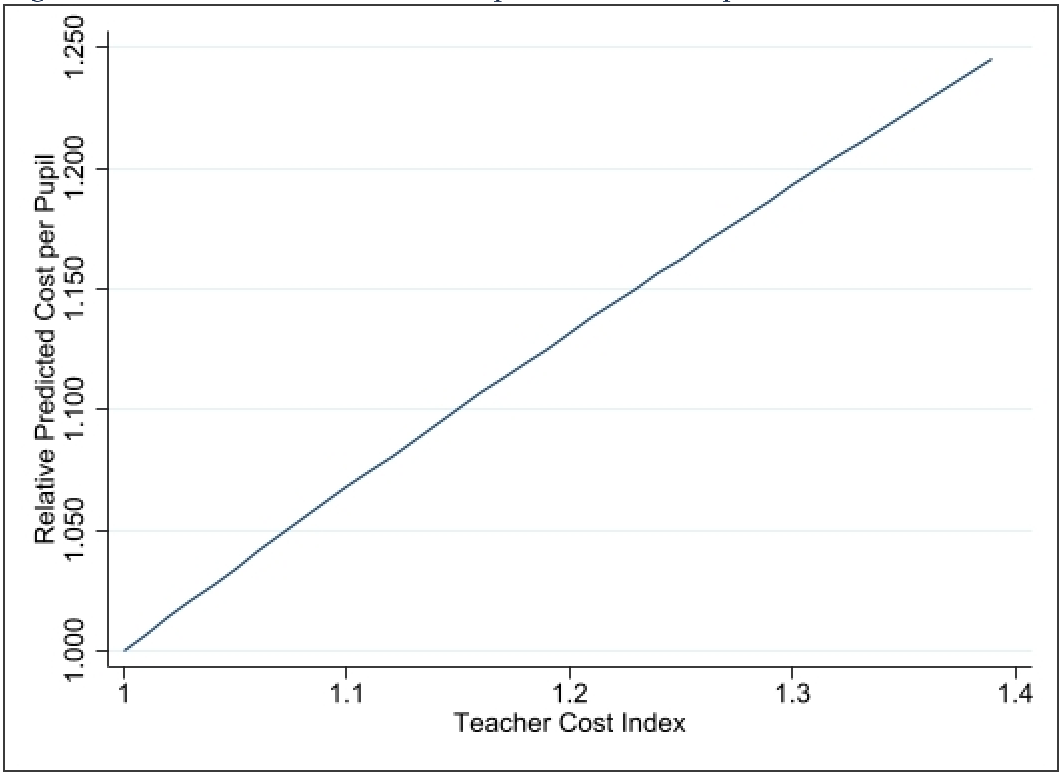
Source: Authors' calculations.

As the figure illustrates, costs are highest for very small districts. A district with 300 students, for example, is predicted to cost 15% more to operate than a district with 1,000 students. Similarly, a district with 1,000 students is predicted to cost 10% more to operate than a district with 5,000 students. As district size increases, costs tend to fall until the log of district enrollment reaches a value of 9.8 (or 18,000 students), at which point it becomes essentially flat. The variation in predicted costs among districts with more than 18,000 students is less than one percent, all other things being equal. Thus, there are clear economies of scale in Texas education, but consistent with the literature discussed above, the cost savings from increases in district size are largely exhausted at relatively modest levels of district enrollment.

Findings on Input Prices

Because education is such a labor-intensive process, geographic differences in the wage level were expected to have a large impact on the cost of education. And, that was indeed the case. Figure 2-6 graphs the impact of the Texas TCI on cost per student, holding all other district characteristics constant. As the figure illustrates, increases in teacher salaries had a positive impact on cost per student throughout the relevant range. On average, a 10% increase in teacher salaries is associated with a 6.6% increase in cost per pupil.

Figure 2-6: The Estimated Relationship between Per-Pupil Cost and the Teacher Cost Index



Source: Authors' calculations.

Changes in the wage level for auxiliary personnel had a much more modest predicted impact on cost per student. On average, a 10% increase in the APCI was associated with a 0.5% increase in the cost of education (holding everything else constant).

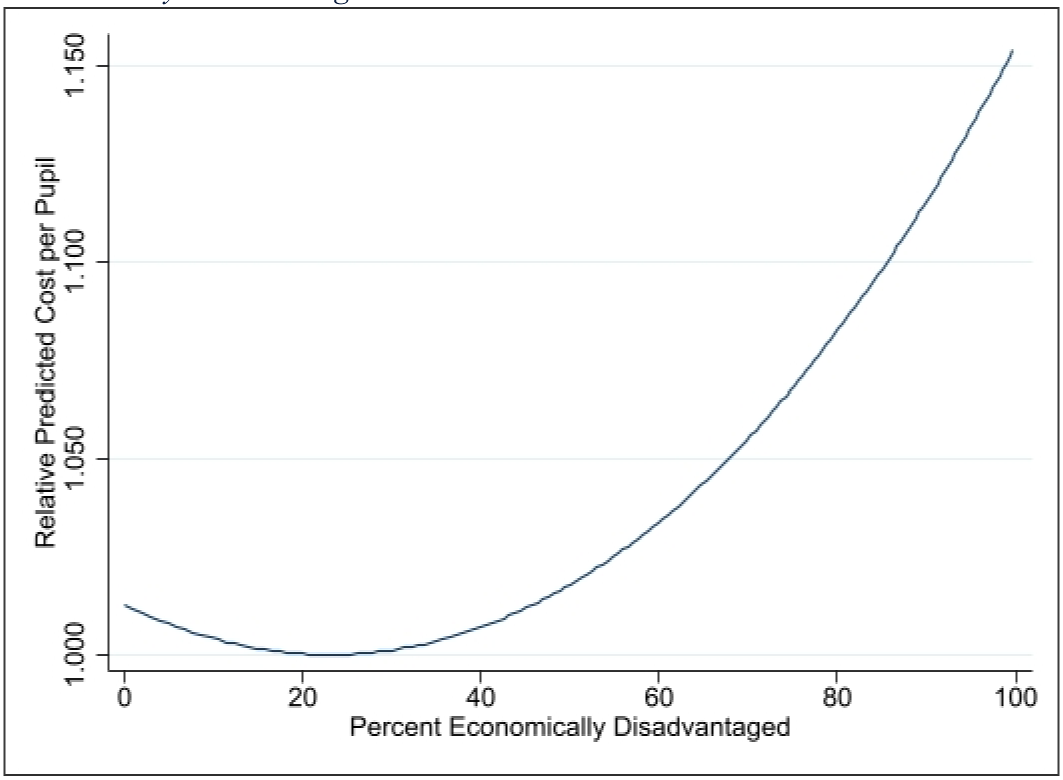
Geographic remoteness (as measured by the distance to the nearest city center) had a statistically significant, but relatively modest impact on the cost of education. With the exception of campuses within 5 miles of a city center, the model predicted that the cost of education rose with distance. On average and holding all other characteristics constant, the cost of education in a district 100 miles from the center of a metropolitan area was 2.3 % higher than the cost of education in a district 20 miles from the center of a metropolitan area.

Findings on Other Environmental Factors

There are several other environmental variables, including the percentage of students classified as Economically Disadvantaged, the percentage of students who have ever been classified as ELL, the percentage of special education students, and the percentage of special education students who had high needs. Increases in each of these four environmental variables all served to increase per student cost.

For example, the analysis indicates that the cost of educating an economically disadvantaged student was, on average, 18% higher than was the cost of educating a student who was not economically disadvantaged. However, the estimated effect was not linear. As Figure 2-7 illustrates, the marginal cost of serving an increased percentage of economically disadvantaged students was sharply higher (i.e., the slope was steeper) for campuses that already had a high percentage of economically disadvantaged students. Among campuses with very low percentages of economically disadvantaged students (which represent less than 10% of the public school campuses under analysis) the marginal cost of serving an additional student who was economically disadvantaged was negative.

Figure 2-7: The Estimated Relationship between Per-Pupil Cost and the Percentage Economically Disadvantaged Students

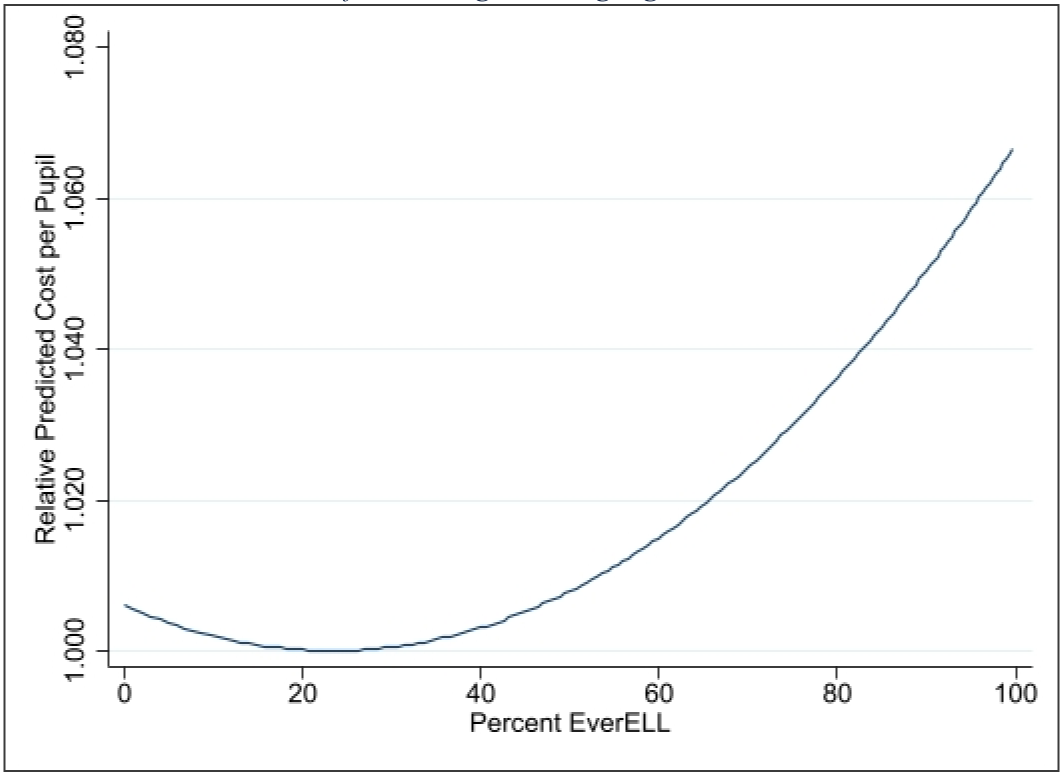


Source: Authors' calculations.

This pattern of increasing intensity leading to sharply increasing cost was also observed for students who had ever been identified as ELL (see Figure 2-8). A district where 100 percent of the students had ever been identified as ELL had predicted costs that were 6.5% above the minimum, all other things being equal.[39]

[39] This marginal effect is not strictly comparable to the Foundation School Program weight for students in bilingual education/English as a second language. The cost function models marginal cost as nonlinear (meaning that the implied funding formula weights are different for different campus configurations), the estimated marginal effect is based on the percentage of students who have ever been designated as ELL, not the percentage of students currently receiving services.

Figure 2-8: The Estimated Relationship between Per-Pupil Cost and the Percentage of Students Who Had Ever Been Identified as English Language Learners



Source: Authors' calculations.

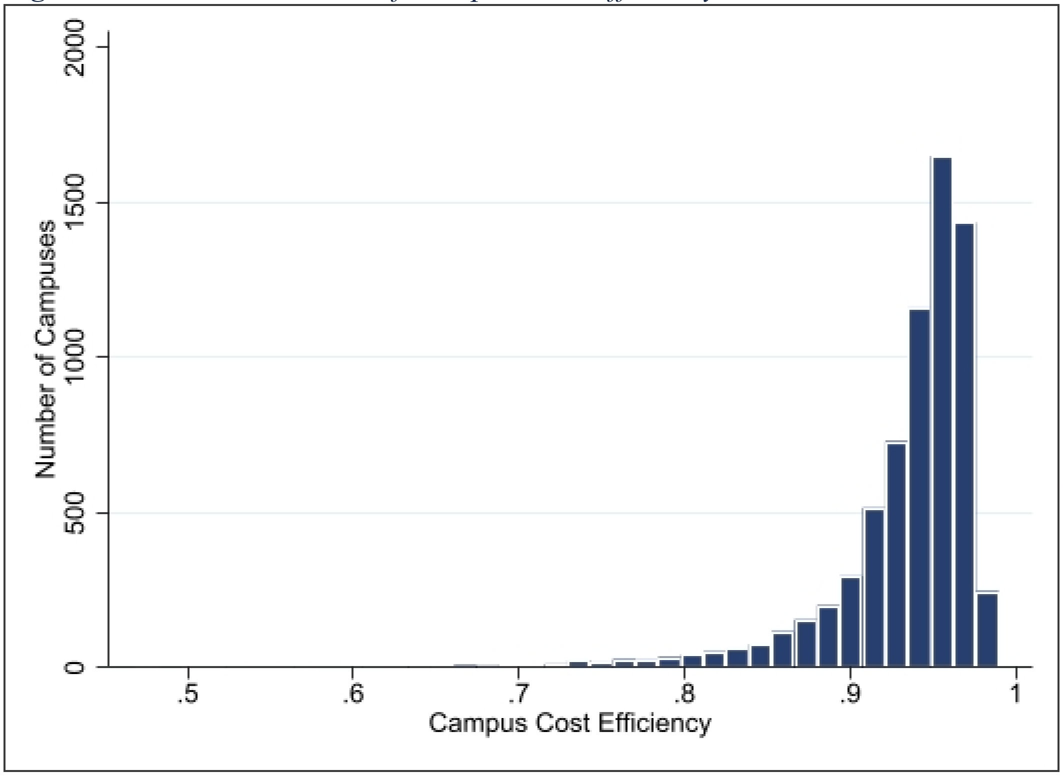
An increase in the percentage of high needs special education students was associated with a percentage increase in per student costs of 1.12 times the increase in the percentage of special education students. In other words, for a campus with average characteristics, the estimated cost of educating a special education student was more than double (112% higher than) the cost of educating a student who was not in the special education program.

Findings on Efficiency

The analysis also found clear evidence that expenditures exceeded what would be expected if campuses were operating efficiently. Figure 2-9 illustrates the distribution of campus cost efficiency for the 2018–19 school year. On average, the cost efficiency score was 0.93, indicating that campuses were producing 93% of their potential output. Given that inefficiency in this context means unexplained expenditures, not necessarily waste, and that many campuses may have been producing outcomes that were not reflected in test scores, the average efficiency level was quite high. On the other hand, efficiency was measured relative to the best practice in Texas, and best practice may still fall short of the ideal. In addition, the minimum efficiency scores were below 50%, suggesting that some campuses spent much more than could be explained by measured outcomes, input prices or student need.

As a general rule, campuses were more efficient in locations where competition for enrollments was more intense. Increases in the Herfindahl index (which measures increases in market concentration with respect to both traditional public schools and open enrollment charter schools) were associated with increase in campus inefficiency.

Figure 2-9: The Distribution of Campus Cost Efficiency, 2018–2019



Source: Authors' calculations.

The Educational Cost Index

Once the educational cost function has been estimated, it can be used to summarize how much more or less it costs to produce educational outcomes from one district to the next. Essentially, one uses the cost function to predict how much each district must spend, each year, in order to produce a standard level of output, assuming it was making cost-minimizing choices about campus size. The Educational Cost Index (ECI) is the ratio of the predicted cost for the district, divided by the state minimum predicted cost. [40]

As is customary in the literature, the level of output was set at the state average (or in other words a Conditional NCE score of 50). Campus size was set at the cost minimizing level for each district. For the other cost factors, which were treated as uncontrollable, the cost model was evaluated at the actual value for these factors in each district. For purposes of this exercise, each district was assigned the average level of efficiency obtained by school districts in Texas.

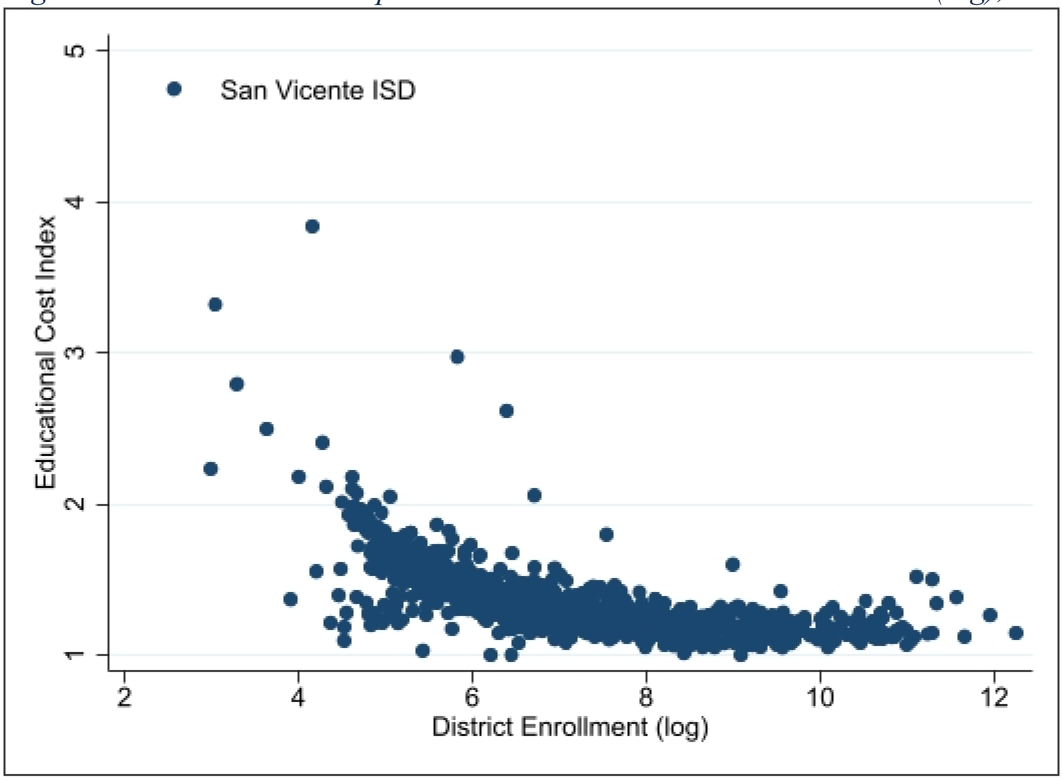
The index values so generated provide a measure of the cost in a district due to its uncontrollable factors relative to the cost in a district with the most cost-favorable characteristics for the uncontrollable factors. For example, an index value of 1.5 indicates that a district is predicted to require 50 percent more per pupil to produce the standardized output levels than the minimum cost

[40] As with the construction of the Texas TCI, the reference prediction used in the construction of the ECI is the prediction at the one-quarter percentile (so that only one quarter of one percent of the districts have a predicted wage below the reference wage). The ECI was set to 1.00 for the handful of districts with predicted costs below the reference level. This approach ensures that the reference level was not an extreme outlier.

district. Other normalizations are, of course, possible. For example, the reference cost level could be the predicted cost of producing the standardized outputs for a district with the average values of the uncontrollable cost factors.

Figure 2-10 illustrates the relationship between district size and the ECI for 2018–19. As the figure illustrates, the ECI ranges from 1.00 to 4.74. In other words, the cost model predicted that the per-pupil cost of producing an average level of academic performance in the highest-cost district—San Vincente ISD with its total enrollment of 13 students—was more than 4.7 times the cost of producing the same level of performance in the district with the lowest cost of education.

Figure 2-10: The Relationship between the ECI and District Enrollment (log), 2018–19



Source: Authors' calculations.

Figure 2-11 illustrates the frequency distribution of the ECI. The median of the ECI was 1.29, so given their district-specific uncontrollable factors, half of the districts had to spend more than 29% above the minimum just to provide the state average level of educational output. The ECI distribution is rather heavily skewed, with a long right tail of districts with ECI values greater than 2.00. Still, these extreme values of the ECI distribution are outliers. More than 95% of districts had an index value less than 1.80.

Figure 2-11: The Distribution of Educational Cost, 2018–19

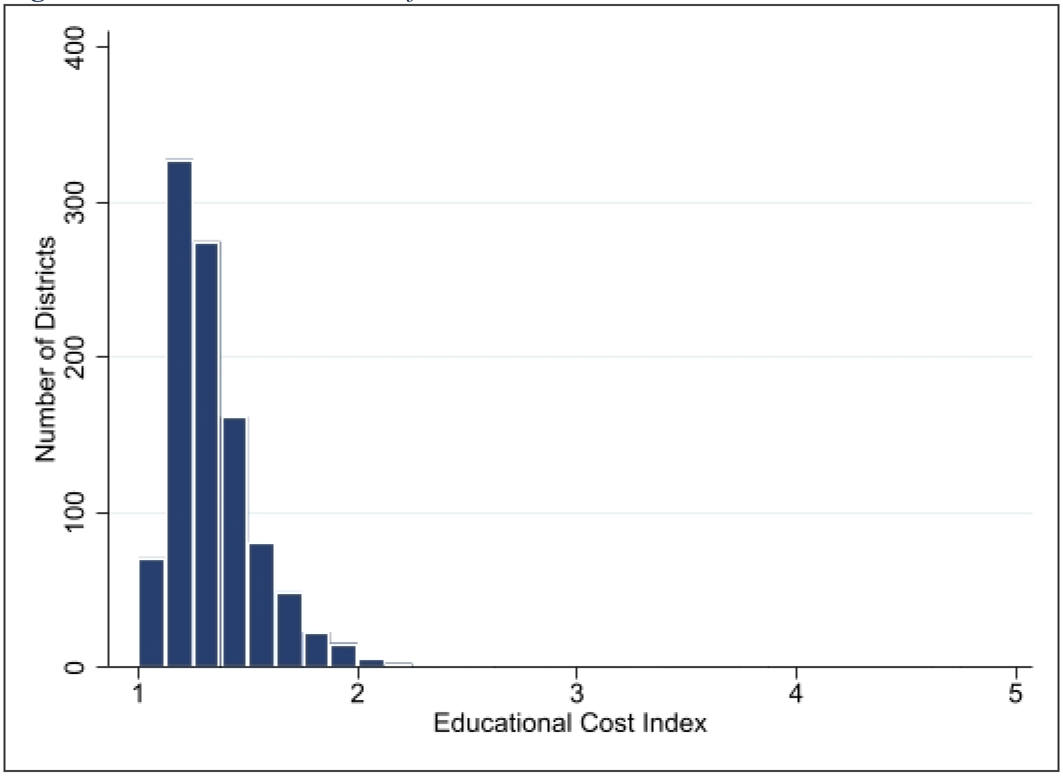


Table 2-3 provides another perspective on the ECI. As the table illustrates, the average ECI was higher in rural counties than in metropolitan or micropolitan areas. The average rural district had an ECI of 1.46 while the average metropolitan district had an ECI of 1.28, suggesting that the generally lower wages in rural areas were more than offset by the district size adjustments built into the ECI. The average high poverty district had an ECI that was more than 16 percentage points higher than the average low poverty district. Districts in sparsely populated counties had higher ECIs than districts in more populous counties.

Table 2-3: The Educational Cost Index, by Location and School District Type, 2019–20

School District Type	Number of Districts	Mean	Minimum	Maximum
Metropolitan	493	1.283	1.000	2.410
Micropolitan	200	1.350	1.000	3.837
Rural	329	1.462	1.151	4.736
Very Sparsely Populated County	165	1.608	1.099	4.736
Sparsely Populated County	110	1.405	1.061	1.913
Other County	747	1.290	1.000	3.315
Small district	640	1.444	1.000	4.736
Midsized district	201	1.219	1.010	1.783
Large district	181	1.185	1.000	1.586
Highest Poverty Quintile	205	1.439	1.061	2.950
Lowest Poverty Quintile	204	1.277	1.000	4.736

Source: Authors' calculations from Appendix F

Conclusions

The overarching takeaway from this analysis of the educational cost function is that the cost of education in Texas is far from uniform. Wages differ by up to 37% from one district to another and those differences drive significant differences in educational cost. On average, the cost of education in high-poverty districts is 44 percent above the state minimum. A lack of economies of scale drives up costs for small districts, and small districts in sparsely populated counties are particularly costly to operate.

Educational costs are higher in some parts of the state because the prices those districts must pay for educational resources—like teachers—are particularly high. But the cost function analysis suggests that other external cost drivers—namely student need, sparsity, and a lack of economies of scale—require some districts to use real resources more intensively than others. Thus, the analysis suggests the need for adjustments to the funding formula in all three dimensions.

Chapter 3: Geographic Variations in Transportation Cost

On the transportation cost front, this study will examine the relationship of transportation costs to staff salaries, to district geographic size and density, to measures of the available road network in the district or county, to the number of campuses and their geographic distribution within the district, and to county disparities in diesel prices. Districts face direct costs to providing student transportation, such as the cost of providing buses, fuel, and drivers to transport students. Transportation costs may vary between districts based on a number of factors outside of district control, including district size and location. Districts that are sparsely populated, for example, may have to transport fewer students across long distances, thus generating large per-pupil transportation costs. On the other hand, densely populated urban districts may face lower per-pupil transportation costs if they transport many students across a short distance. This analysis will provide estimates of geographic variation in transportation costs between school years 2014–2015 and 2018–2019.

The Literature

The academic journal literature on the economics of education has paid surprisingly little attention to the study of school transportation costs. There are only two small school transportation research strands. One set of papers focuses upon the potential of cost advantages to district size and the second set of papers looks at the potential for cost savings from privatization of the school transportation function.

A few researchers have estimated a cost function for transportation as a byproduct of their primary focus on economies of scale in the total operating costs for school districts (Duncombe, Miner, and Ruggiero, 1995; Dodson and Garrett, 2004; and Zimmer, DeBoer, and Hirth, 2009). The distinguishing feature of these papers is that each one disaggregates total operating expenditures into its major subcomponents, including transportation, and then estimates separate cost functions for each subcomponent. The explanatory cost factors in the transportation cost function estimation are identical to those used in the total cost function estimation, meaning that the outputs are measures of student performance and the major input price is a measure of teacher salaries.

The obvious shortcoming in this strand of the literature is that the empirical models are, fundamentally, not developed as transportation cost function models. Only Zimmer, DeBoer, and Hirth (2009) includes a direct transportation output (bus miles) in addition to number of pupils and pupil achievement outputs. None of the papers included appropriate input prices for a transportation cost study. The only labor price is teacher salaries, even though the principal labor price for a transportation cost function study should be bus driver salaries. Fuel prices are not included. As a result, this strand of the literature has limited usefulness for understanding geographic differences in the cost of delivering student transportation services.

The school bus privatization literature, on the other hand, does include papers that estimate credible school transportation cost function models. Lazarus and McCullough (2005) estimated a model of school transportation costs using the number of pupils transported as the measure of transportation output. Their model included prices for bus drivers and for fuel, and controls for the number of miles of road in the district and the number school buses (separating small and large). The percentage of special transportation needs riders was also included in the model to allow for potential differential transportation costs for these rider types. Thompson (2011) expanded the model to include bus miles/student transported as a second output, and make other refinements to the estimation. The main conclusion of both of these papers is that contracting out did not lead to reductions in the cost of pupil transportation services relative to in house provision.

In a pair of papers, Hutchinson and Pratt (1999, 2007) explored the relative cost of contracting out versus in-house production of school bus transportation. The same basic empirical cost function model is used in both of these papers. The cost model assumes two outputs: the average number of students transported daily and the number of one-way bus miles driven, and two input prices: average annual bus driver salary and cost per gallon for fuel. Fixed inputs include the number of Type I and Type II buses and the district population density. Again, the focus of these studies is on the comparative cost of in-house and contracted transportation institutions. Hutchinson and Pratt found that in-house was cheaper in Louisiana but that contracting out was cheaper in Tennessee.

Although the academic literature on school bus transportation functions is sparse, there is a robust academic literature on the costs of municipal bus transit. As discussed in Berechman and Giuliano (1985) some researchers have focused on vehicle-based or technical output measures, such as bus miles or bus-hours; whereas other researchers have focused on passenger-based or demand-based output measures, such as passenger-trips or passenger-hours. However, all of the modern bus transit cost studies have included measures of the price of labor and the price of fuel in their cost function.

In addition to labor and fuel, the third critical input to producing bus services is bus capital. The majority of the transit studies treat the rolling stock of buses as being fixed, and thus the cost function estimates are interpreted as short run bus variable operating cost functions. The number of buses is usually included as an explanatory variable. Some studies include average age of the buses as a measure of capital quality.

An Overview of Student Transportation in Texas

There are three primary service models utilized by school districts in the United States to transport students to and from school. The most common model is district-provided yellow bus service. Under the second most common model, districts contract with private providers for yellow bus service. The third approach is reliance on public transit. District involvement usually comes in the form of district subsidization of fares for student riders. This is a much less common approach that is primarily used in large urban districts that can tap into extensive public transit systems. Across the United States, about two-thirds of all yellow school buses are owned by districts and around one-third of all yellow buses are owned by private contractors.

In Texas, district-provided transportation service is the strongly dominant model. In 2019, over 90% of traditional public school districts reported that they managed transportation in-house. In a pure district-run system, the district makes all the key student transportation decisions. The district handles the route logistics, buys and manages the maintenance of the buses, and is in charge of all human capital management decisions. Under the private contracting out model, the district turns the system management control and responsibility, over both fleet and human resources, to the contractor. Districts can, of course, stipulate service output and input requirements in the contracting process. Under the public transit model, the school district is not involved on the supply side, but may choose to help out students/parents on the demand side.

In addition to the pure models, there are varieties of hybrid-type arrangements as well. Some districts contract out a portion of their student transportation services while providing other transportation services internally. There are also intergovernmental, as opposed to public-private partnership, contracting arrangements. Prominent examples include the (now defunct) Dallas County Schools public transportation agency, which provided student transportation services to a large number of school districts in Dallas County and as well as in neighboring counties, and the Bowie County Schools Transportation Department, an independent governmental unit that provides transportation services to the thirteen school districts in Bowie County.

In 2015, The Mackinac Center for Public Policy surveyed school districts in Texas and four other states to assess the degree of privatization of non-instructional services, including transportation (LaFaive and Hohman 2015). Only 3.7% of traditional public school districts in Texas used privatized transportation services. At 1.7%, Georgia was even less privatized, and Ohio (at 6.5%) was also low on the transportation contracting scale. At the other end of the privatization spectrum, Pennsylvania contracted out at a 66.4% rate. A little over a quarter (26.6%) of Michigan school districts contracted out for school transportation services in 2015, which was slightly more than double the privatization rate in 2011.

Student Transportation Services Supplied

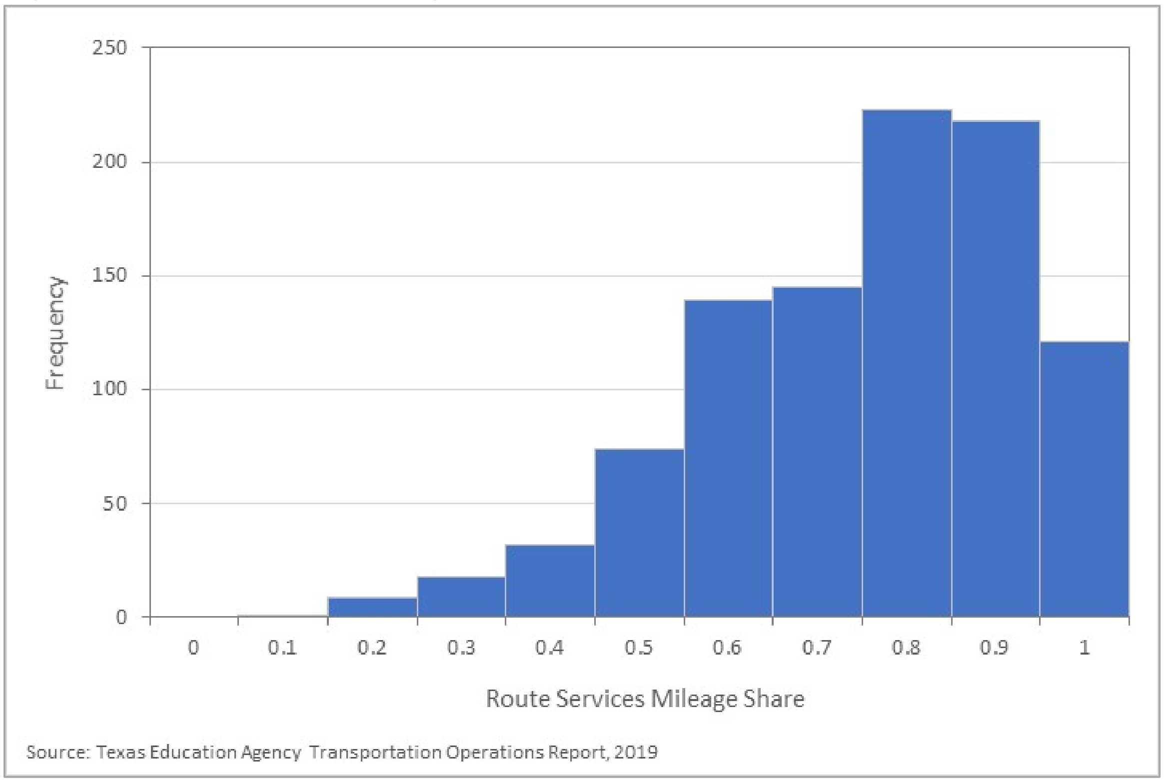
There are five major types of student transportation services provided. Regular route services transport regular program students to and from school as well as to and from alternative academic instruction during the school day. Special route services transport special program students to and from school (regular school year), auxiliary services (during regular school year) and ESY (extended school year) services. The CTE route services transport regular or special-program students to attend a TEA-approved Career and Technology course. Private route services provide to-and-from school transportation, using private or commercial transport, for students facing extreme transportation hardships. Districts also spend considerable resources transporting students for extracurricular and cocurricular activities and events, as well as field trips. It is important to note that only first the four types of district route service miles are eligible, i.e., “count”, in determining State transportation funding support.

For transportation purposes, students are classified as one of two types. Special-program students are those students with a disability who meet the requirements for specialized transportation services. All other students are classified as regular-program students. School districts are only required to provide transportation to special-program students. They may choose to provide transportation services to regular-program students.

A key institutional feature for regular route services is that only a subset of regular-program students are eligible to be counted as riders for State transportation funding purposes. The two most critical (from a number of riders impact standpoint) eligibility criteria are (1) a student rider who lives two or more miles from the student’s campus of regular attendance is eligible to be counted (2) a student rider who lives in a hazardous traffic or a high risk of violence area that is within two mile of the student’s campus of regular attendance is eligible to be counted. The two-mile radius threshold creates an incentive for school districts to restrict their to-and-from school bus stops to only picking up student riders who live outside the two-mile boundary.

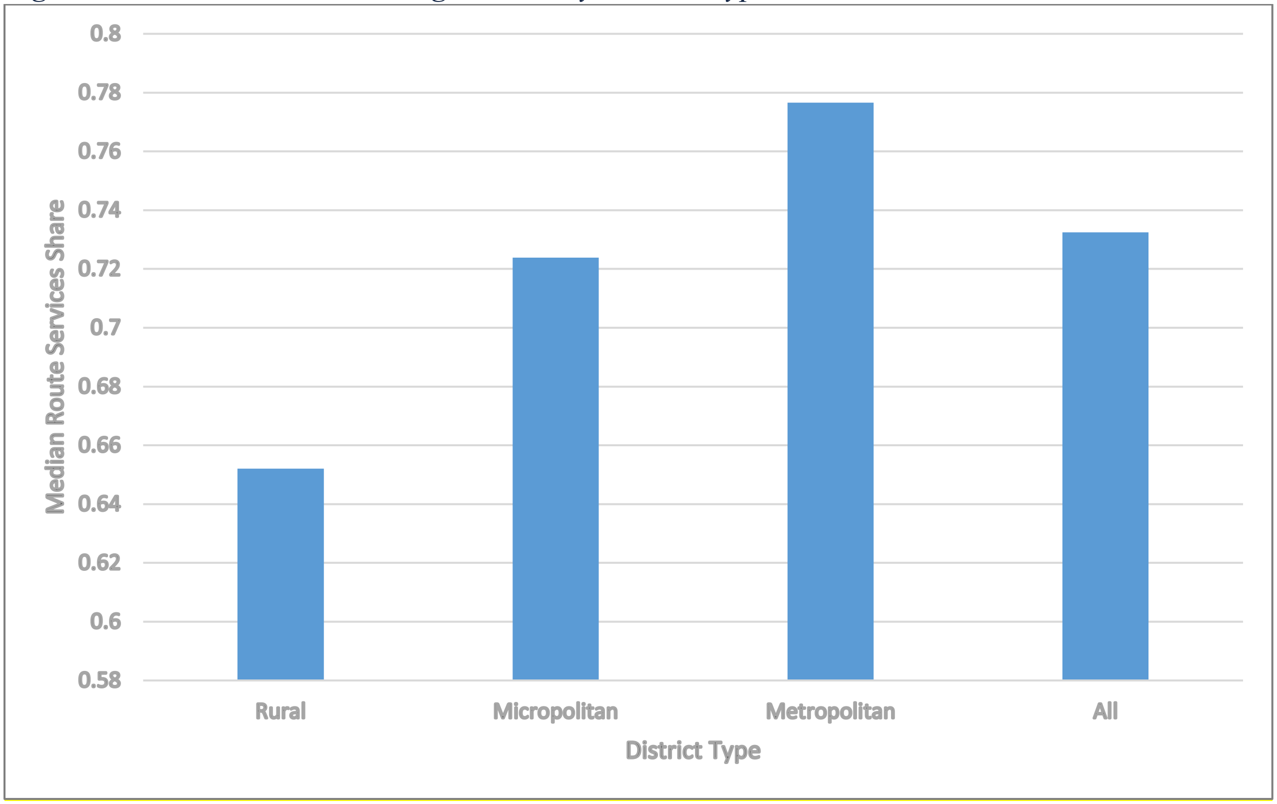
The common perspective of the principal role for the yellow bus system is home to school transport. For the majority of school districts, the largest percentage of bus miles are, indeed, route services miles. The median of the distribution of route service bus mile shares across school districts is 73%. As shown in the frequency distribution graph below, however, the route share does vary considerably. Alternatively stated, the variation in the share of bus miles devoted to extracurricular/cocurricular and other trips is nontrivial.

Figure 3-1: Route Services Mileage Share across Districts, 2018–19



The median proportion of total bus miles that are associated with route services varies across the three district groups—Rural, Micropolitan and Metropolitan. As shown in the bar graph in Figure 3-2 below, the median of the route share miles for the rural districts is considerably lower than the share for both micropolitan (by over 7 percentage points) and metropolitan (by over 12 percentage points). The average extracurricular trip lengths (rather than number of trips) are likely greater for rural districts than for metropolitan and micropolitan districts.

Figure 3-2: Route Services Mileage Shares by District Type, 2018–19



Source: Texas Education Agency (2019).

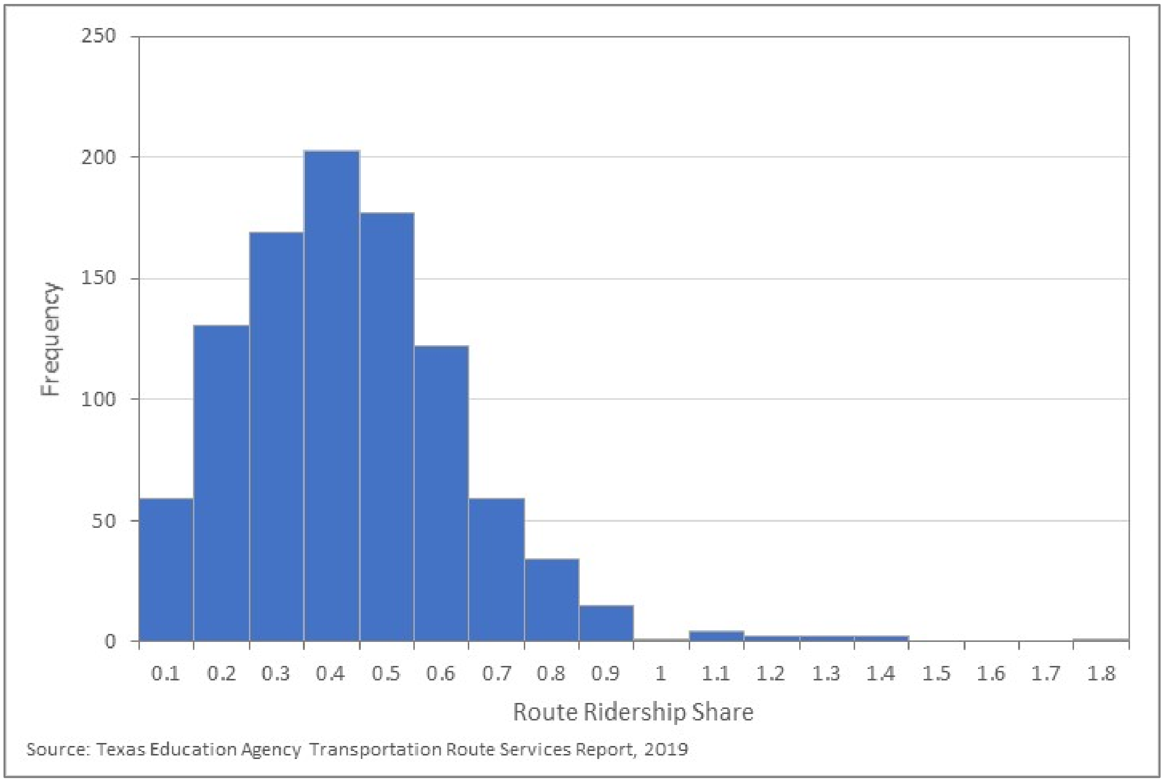
Choices about Travel Modes

Students travel to school by four different modes: school buses, personal vehicles, walking/biking, or public transit. According to the National Household Travel Survey (2017), roughly a third of children ages 5–17 travel to school on a school bus. The majority (roughly 54%) travel to school in personal vehicles.

In Texas, approximately 32% of the students ride from home to school on the school bus. The eligible ridership share roughly matches the national share. There is, however, considerable variation in ridership share across the school districts in Texas. The frequency distribution for rider share is shown in Figure 3-3 below. The median of the distribution is 35.2%. The rider share is 24% at the 25th percentile and 50% at the 75th percentile.

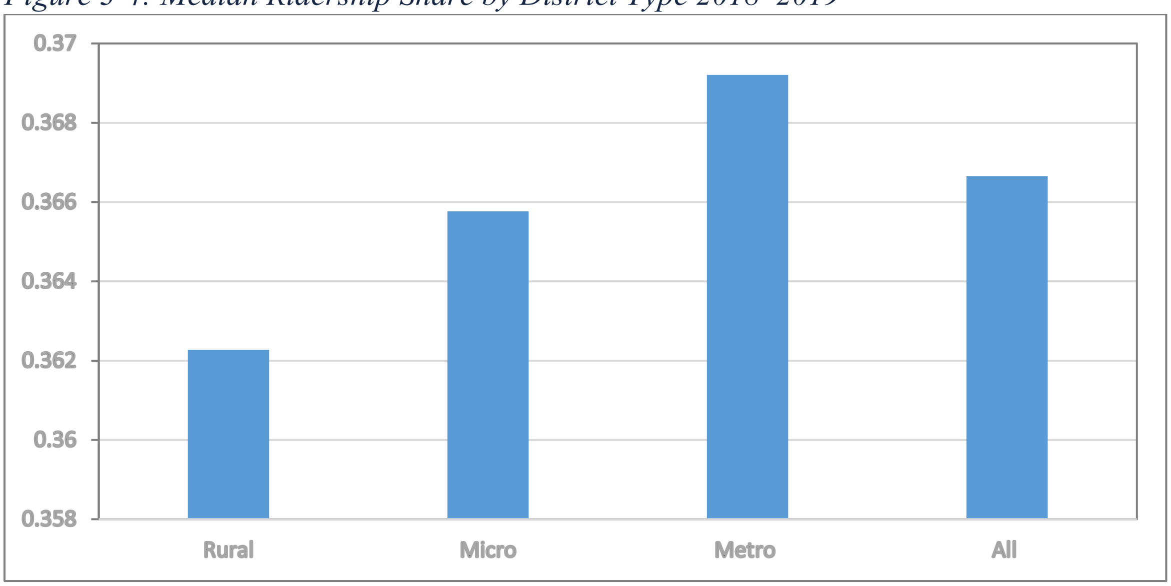
We would note that the reported eligible ridership share is a downward-biased measure of home to school bus ridership. Districts are allowed to pick up and transport non-eligible students who live inside of the two-mile "funding radius" for their campus. These are not necessarily free riders, as districts can charge parents a fee for the seats.

Figure 3-3: Route Ridership Share Distribution across Districts, 2018–19



The ridership share distributions for rural, micro, and metro school districts are quite similar, and the medians of the district share distributions differ by less than 1 percentage point.

Figure 3-4: Median Ridership Share by District Type 2018–2019



Source: Authors' calculations, 2019; Texas Education Agency Transportation Route Service.

The Transportation Funding Allotment, 2014–15 through 2018–19

School districts receive funding support for their transportation expenditures in the form of a transportation allotment. The size of these allotments is determined by statute and by the current biennium's General Appropriations Act. Across the country, states have different formulae for reimbursement of public school transportation expenditures. The three major funding categories are mileage-based reimbursement, cost-based reimbursement, and per student reimbursement (often with a cost factor adjustment). Texas utilizes a mileage-based funding formula approach.

The Texas transportation allotment for a district is equal to the sum of the funding for each of the four categories of the district's route services: regular, special, CTE, and private. The funding per category is determined by multiplying the total eligible mileage for the category by the per-mile rate for the category.

During our 2015–2019 study period, the calculation of funding for regular route services is the most complex. The eligible mileage includes both two-or-more-mile-only miles and hazardous-traffic/high risk of violence area miles, but the funding for hazardous-traffic and high risk of violence area mileage is capped at 10 percent of the total funding for two-or-more-mile service. The per-mile funding rate varies according to the district's effective linear density. The effective linear density is calculated as the average number of riders (who live two or more miles from campus) per mile for the school year. The mapping between district linear density and the district funding rate per mile comes from the following schedule, which was established by the Texas Legislature in 1984. The funding rate is increasing in rider density. The linear density funding model has been replaced by a flat rate per mile model under House Bill 3. We consider some of the implications of that model change later in this report.

Table 3-1: Rate per Mile for Linear Density Groups, 2018–19

Linear Density	Rate per Mile of Approved Route
2.400 or above	$1.43
1.650 – 2.399	$1.25
1.150 – 1.649	$1.11
0.900 – 1.149	$0.97
0.650 – 0.899	$0.88
0.400 – 0.649	$0.79
Up to 0.399	$0.68

Source: Texas Education Agency Transportation Route Services Report, 2019.

We display the distribution of districts by linear density category in Table 3-2 below. The distribution is right-skewed, with 40% of the districts in the two lowest allotment rate categories and less than 10% of the districts in the two highest linear density/rate allotment bins.

Table 3-2: Distribution of Districts by Linear Density Group, 2018–2019

Linear Density Group	Allotment Per Mile of Approved Route	Number of Districts Per Group	Percentage of Districts Per Group
2.40 and above	$1.24	31	3%
1.65 to 2.40	$1.15	59	6%
1.15 to 1.65	$1.07	145	15%
0.90 to 1.15	$0.99	147	15%
0.65 to 0.90	$0.89	214	22%
0.40 to 0.65	$0.79	193	20%
Up to 0.40	$0.68	200	20%

The regular route funding allotment is then calculated by multiplying a district's total eligible regular-route-service mileage by its linear-density based per-mile rate for regular route services.

The funding for special route services is calculated by multiplying the district's total eligible special route mileage by the lesser of a State-determined rate per mile ($1.08 in 2018) or the expenditure per mile for regular route services for the preceding year.

The funding for CTE route services is calculated by multiplying the district's total eligible CTE-route-service mileage by its expenditure per mile for regular route services for the preceding year.

The funding for private route services is calculated by multiplying the district's total eligible private-route-service mileage by a State-determined rate per mile ($0.25 in 2018), with a maximum reimbursement cap ($816/eligible student rider in 2018).

We have obtained data on the FSP Transportation Allotments by district for our sample period 2014–15 through 2018–19. As shown in Table 3-3 below, the State increased transportation allotment spending by a little over 11% during the period. The majority, over 60%, of the allotment spending is for the Regular Program, although that share declined over the period. The Special Program allotments and Career and Technical Education allotments increased significantly during the period. The Special Program allotments increased by nearly $14 million (over 14% increase) and the Career and Technical Allotments increased by almost $16 million (greater than 75% increase).

Table 3-3: Transportation Allotment by Program, 2015–2019

Program	2014–15	2015–16	2016–17	2017–18	2018–19
Regular	$212,118,720	$214,435,712	$215,277,936	$213,341,056	$218,867,808
Special	$97,394,616	$99,426,192	$100,249,232	$101,651,264	$111,314,768
Career and Technical Education	$21,145,720	$26,394,934	$29,466,602	$31,481,640	$37,106,500
Private	$92,546	$99,079	$69,949	$104,542	$86,324

As shown in Table 3-4 below, for our sample of 980 TPS districts, the Transportation Allotment totaled $367.38 million in 2019 and funded 23% of the Total district expenditures. The percentage of expenditures funded decreased from 27% to 23% over the five-year period.

Table 3-4: FSP Transportation Allotment and Local Expenditures for Public School Transportation in Millions, School Years 2014–2015 to 2018–2019

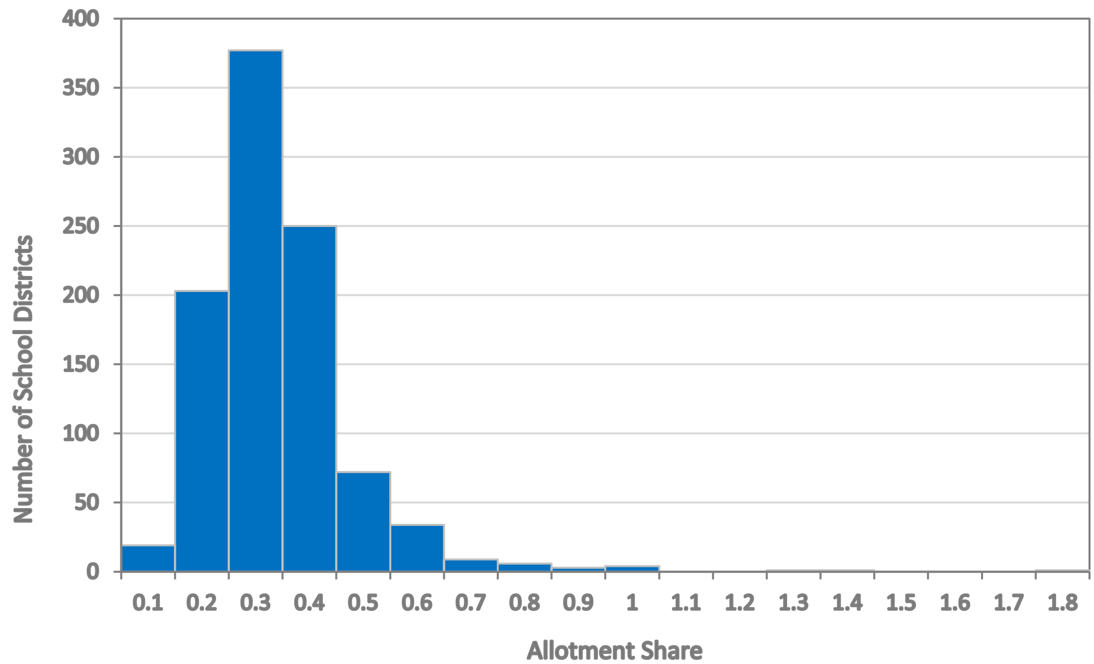
School Year	FSP Transportation Allotment Entitlement	ISD Transportation Expenditures	FSP Transportation Allotment as a Percentage of Expenditures
2014–15	$330.75	$1,233.47	27%
2015–16	$340.36	$1,277.58	27%
2016–17	$345.06	$1,348.82	26%
2017–18	$346.58	$1,442.53	24%
2018–19	$367.38	$1,565.73	23%

Source: Texas Education Agency's PEIMS actual financial data. PEIMS annual transportation expenditures are calculated as the total reported expenditures in PEIMS actual financial data, Function 34, object codes 6100–6499 for each ISD.

When initially rolled out, under the linear density funding model the State provided 70 to 80 percent of total transportation costs. According to a 2013 Legislative Budget Board Brief, The Foundation School Program Transportation Allotment funded between 27% and 30% of total ISD and charter school annual expenditures for transportation during school year period 2006–07 through 2010–11. The percentage of expenditures funded by the FSP transportation allotment fell throughout the last decade.

We can also use these data to study the distribution of allotments and allotment funding shares across districts. We show the frequency distribution of the allotment funding shares for 2019 in Figure 3-5 below. The median district funding share is 26.3%. There is a concentration of share values in the middle of the distribution, but there is considerable variation across districts. The 10% percentile district funding share is 15.4% and the 90th percentile district funding share is 44%.

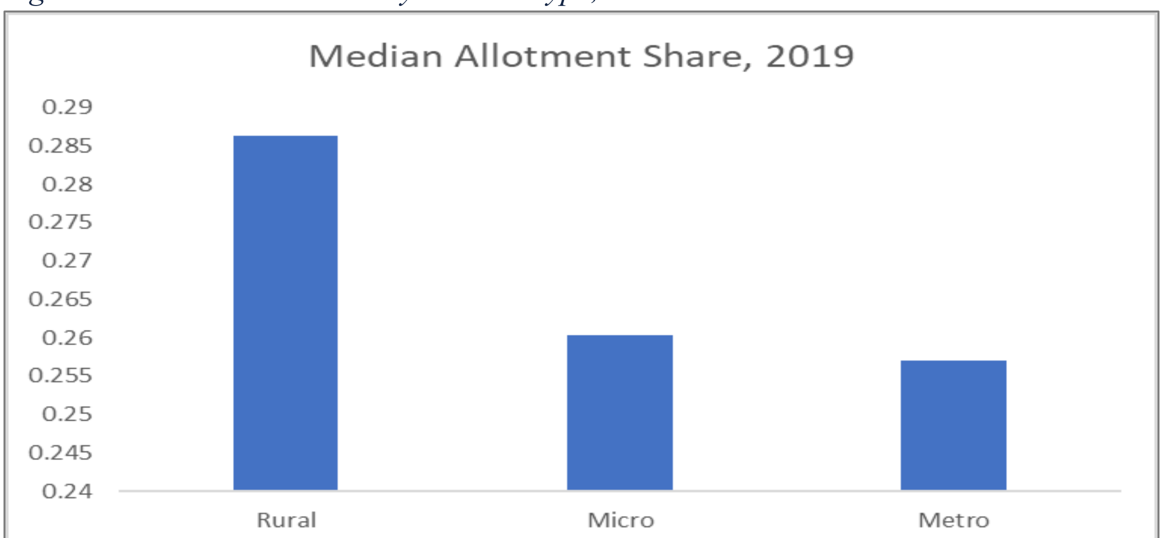
Figure 3-5: Transportation Allotment Funding Shares across Districts: 2018–19



Source: Texas Education Agency Division of School Funding and Public Education Information Management System, 2019.

As a group, rural districts receive a higher transportation funding share than do metropolitan and micropolitan districts. As shown in Figure 3-6 below, the median district allotment share among rural districts is 3 percentage points higher than the metro district median and 2.5 percentage points higher than the micro district median share value.

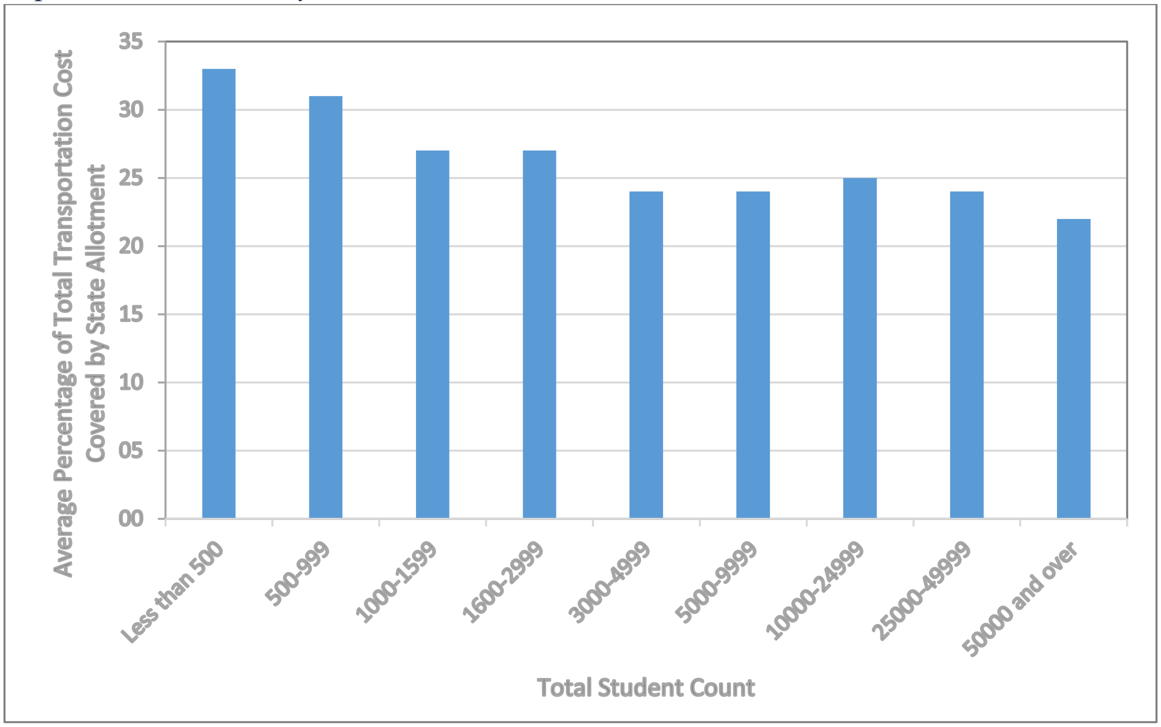
Figure 3-6: Allotment Shares by District Type, 2018–19



Source: Authors' calculations, 2019; Texas Education Agency Division of School Funding and Public Education Information Management System, 2019.

The average percentage of transportation expenditures covered by the transportation allotment also differs by district size as measured by enrollment. We show the relationship between the average allotment/expenditure coverage by nine district enrollment size categories in Figure 3-7 below. For districts with fewer than 500 students, the transportation allotment covered an average of 33 percent of district expenditure. The allotment share falls fairly steadily across the size categories, with the lowest average percentage coverage at 22% for the largest enrollment district group.

Figure 3-7: Districts by Total Student Count Average Percentage of Total Transportation Expenditure Covered by State Allotment, 2018–2019



Source: Authors' calculations, 2019.

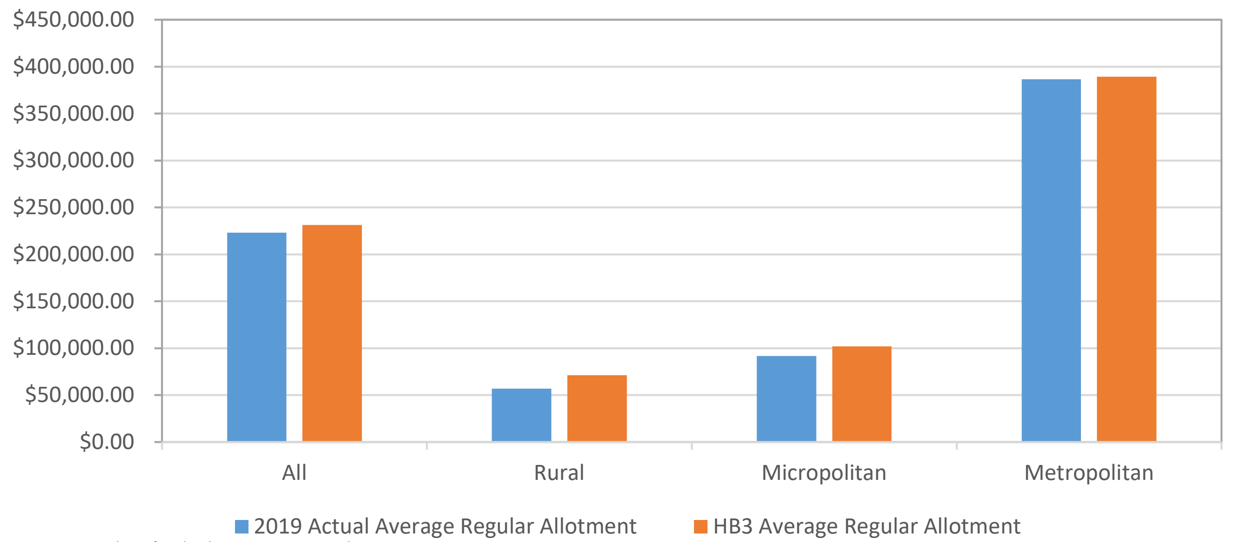
The Transportation Funding Allotment after HB3

During the 2015–2019 time period, the Regular program allotment was determined using the linear density-based formula described earlier in this section. The formula for determining the regular transportation allotment was amended under House Bill 3 (HB 3) in June 2019. Under HB 3, the regular program allotment will be determined based on a flat rate per mile to be set by the Legislature in the General Appropriations Act (GAA). The rate adopted for 2020–21 under the current GAA is $1 per mile.

The switch to the $1 per mile rate will change the aggregate level of regular program funding support as well as the distribution of support across districts. We can get a sense of the direction of impact on funding by simulating the distribution of regular program allotments across our sample of traditional public districts that would have been obtained had the $1 per mile rate been in place for the 2018–19 allotment determination. In 2019, the State actual total regular allotment spending was $218,867,808. We calculate that the State total regular allotment under the HB3 rate would have been $226,630,000. In the aggregate, the State subsidy for school regular program transportation would have been increased by just under $8 million, an increase in spending of around 3.5%. There will be redistributive effects associated with a switch to the uniform $1 per mile rate as well. The 219 districts who were funded at more than a $1 per mile under the pre-HB 3 rate schedule would realize a decrease in their funding allotment (holding miles fixed at the 2019 values). The remaining 761 districts would receive an increase in allotment funding. The average impact across rural, metro, and micro district types is shown in Figure 3-8 below.

The Rural districts would have received the biggest average increase, just over $14K per district. The Micropolitan district average increase would have been a little over $10K per district, while the Metropolitan districts would have enjoyed the smallest average funding gain of a little over $2800 per district.

Figure 3-8: Average Transportation Allotment by District Type, 2018–19



Source: Authors' calculations; Texas Education Agency Transportation Operations Report, 2019.

Although the total regular program allotment funding would be higher, the percentage of total regular program expenditures covered remains modest. The portion of transportation expenditures not funded by the transportation allotment must be covered out of general revenues. The strong crowd-out of uncovered transportation expenditures for instructional expenditures does create strong incentives for efficiency in transportation operations. The magnitude of the crowd-out is, however, potentially higher in Texas than in several other states. In New York, for example, the state contribution covers over half of the district cost of transportation. Increased funding for school transportation in Texas will free up general district revenues for investment in educational enrichment activities and initiatives. Given the current low proportion of regular program expenditures funded by the regular transportation allotment, there is room for substantial increases in transportation allotment funding without significantly weakening the efficiency incentives for transportation operations. The marginal returns to developing strategies to reduce student transportation costs per mile remain intact—a dollar saved on transportation is a dollar available for instruction.

District Transportation Expenditures

School districts in Texas report produce an annual Transportation Operations Report (TOR) that provides information on the expenditures (direct and indirect) incurred providing transportation services for the school year. Expenditures are reported under five object code expenditure categories: Salaries and Benefits, Purchased and Contracted Services, Supplies and Materials, Annual Depreciation and Other Operating Expenses, and Debt Service. The expenditures are subdivided into expenditures attributable to transporting regular-program students and expenditures attributable to transporting special-program students (with both excluding students who were provided private route services).

School districts in Texas also report their expenditures on student transportation in their annual district financial report. In order to maintain consistency with our estimation of instructional costs, we will use district financial data on transportation outlays from the Public Education Information Management System (PEIMS). Within the TEA reporting system, we will use the data on Student (Pupil) Transportation Expenditures that is recorded under function code 34 in PEIMS.

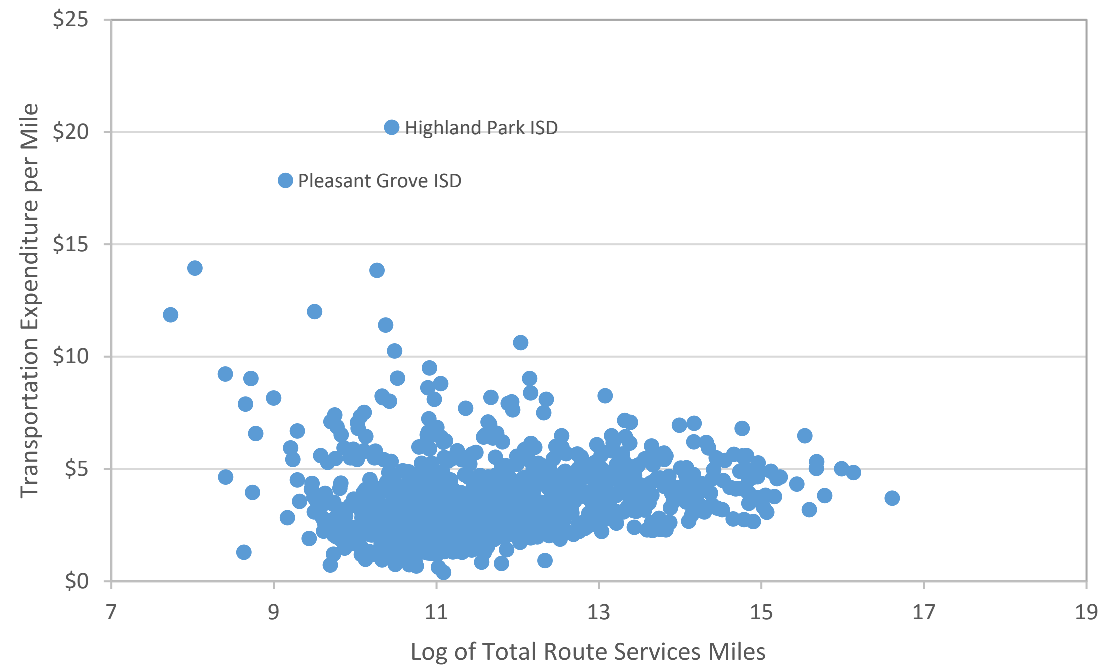
The Texas Education Agency Financial Accounting System Reporting Guide instructs districts that "Your district must use function code 34, Student Transportation, to account for only the cost of transporting students to and from school for the regular instructional day." Districts are allowed to use district buses to transport students for other purposes, e.g., extracurricular events and class field trips. The transportation expenditures on these alternative uses should, however, be allocated to the relevant alternative function codes, e.g., Extracurricular Activities (function code 36) and Instructional (function code 11).

The key point here is that the function code 34 expenditures are intended to capture only the outlays associated with the Route Services supplied by the school districts. This point has two implications for our analysis. First, we need to identify and measure our cost variables in the context of a Route Services cost function. Second, we need to be careful when referencing data from the Transportation Operating Reports. For example, the TOR operating cost data are aggregate costs that include expenditures incurred from non-Route services, e.g., extracurricular transportation of students.

A key issue in many school policy discussions is the relationship between district cost and district size or scale. Texas is a state with huge variation in district size. An important empirical question is whether or not there are expenditure advantages/expenditure economies to size. In the school transportation context, the two obvious candidates for measuring operating size would be bus miles produced or district student rider demand serviced. We can provide some rough suggestive evidence of the transportation expenditure-size relationship by plotting expenditure per mile against total bus miles and expenditure/rider against total number of riders by district using PEIMS data for 2019.

Figure 3-9 below shows expenditure per mile against the log of total miles (for route miles only). The scatterplot reveals a general slightly increasing relationship between average expenditure per mile and number of bus miles supplied for most of the district routes.

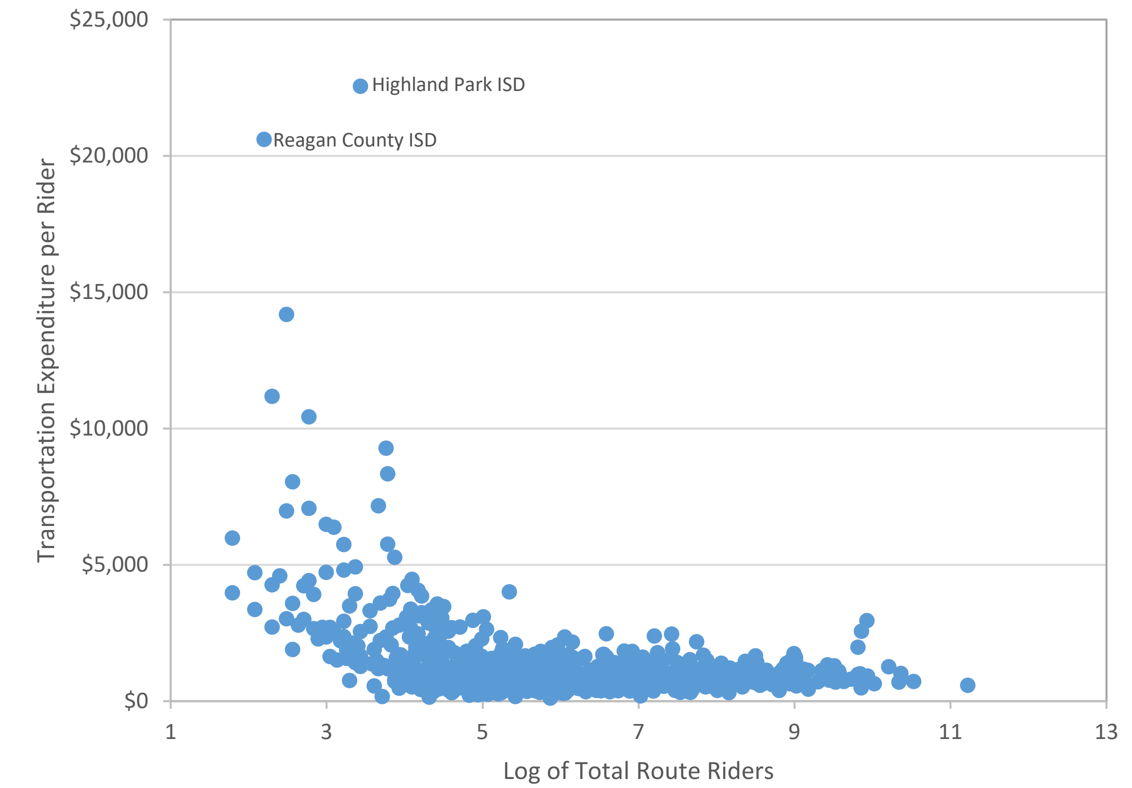
Figure 3-9: Transportation Expenditure/Route Mile against Log of Total Route Services Miles, 2018–19



Source: Authors' calculations; Texas Education Agendu Transportation Rouse Services Report, 2019.

Figure 3-10 displays a plot of expenditure/rider against total (log) riders. The scatterplot reveals some decline in the relationship between average expenditure per rider and number of route riders served by the district transportation system for a few very small ridership districts. This suggests the potential existence of cost economies with respect to ridership size, but these cost reductions occur among districts with an average number of riders per day between 1 and 300 riders. The expenditures per rider are fairly constant beyond 300 riders/day up to around 5,000 riders per day. They decline slightly again for ridership greater than 5,000, indicating the possibility of additional economies of scale available to very high rider demand districts.

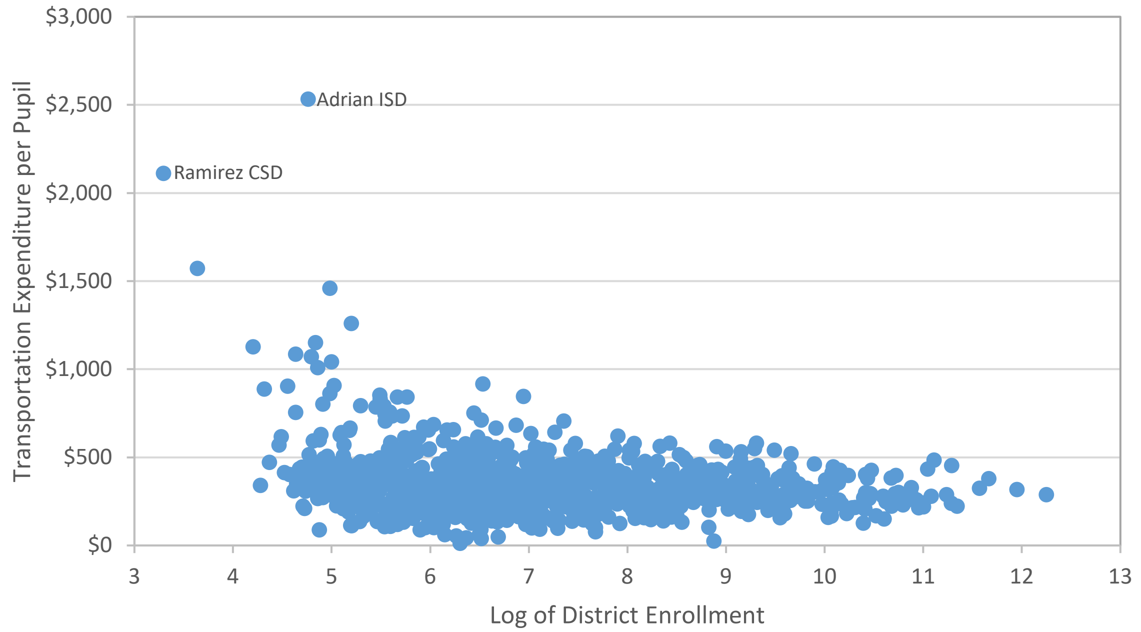
Figure 3-10: Transportation Expenditure/Rider against Log of Number of Route Riders, 2018–19



Source: Texas Education Agency Transportation Operations Report & Public Education Information Management.

The potential ridership size economies story in Figure 3-10 matches up reasonably well with the possible transportation cost economies to enrollment size discussion in Chapter 2. The average transportation expenditure per pupil in Table 2-1 showed declining average costs for districts up to size 1,000, fairly constant average expenditures for districts with enrollment between 1,000 and 10,000, and an additional decline for the largest districts with over 10,000 students. With an average ridership share of around 31 percent, there is a general consistency to the rider and enrollment potential size economy evidence. The consistent relationship between the expenditure/rider to number of riders and expenditure/pupil to number of pupils in the raw data can be seen by comparing Figure 3-10 above to Figure 3-11 below.

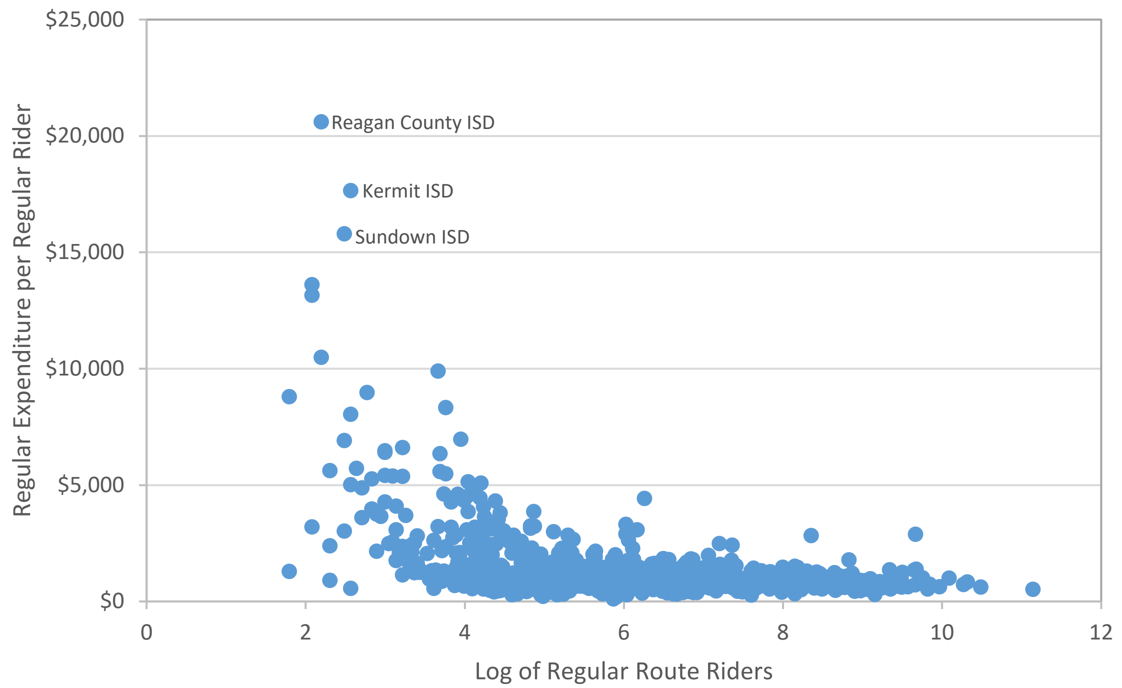
Figure 3-11: Expenditure/Pupil against Log of Number of Pupils, 2018–19



Source: Texas Education Agency Transportation Operations Report & Public Education Information Management System, 2019.

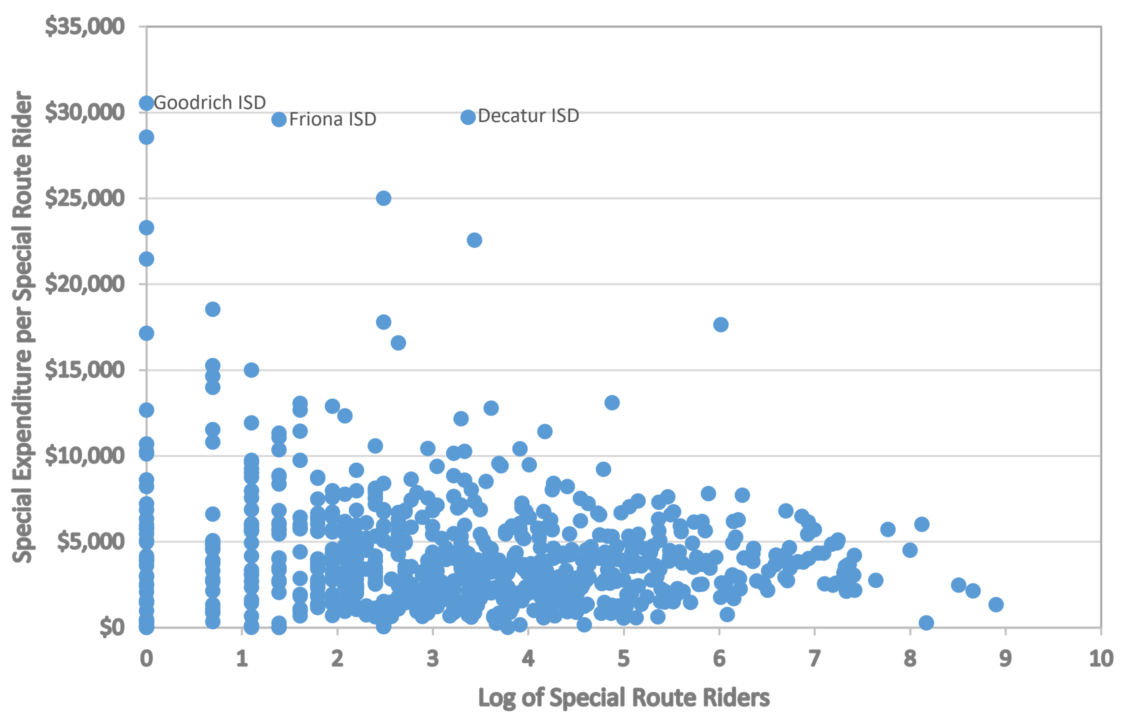
A recent public policy report on school transportation from the Urban Institute (Chingos and Blagg 2017) argues, based upon published evidence, that "Transportation expenditures can also be driven by changes in policy on the provision of transportation for students with special needs, which are governed by federal law as well as state and local policy" (op.cit. p.4). School districts in Texas are required to provide separate reporting of regular program miles and expenditures from special program miles and expenditures in their Transportation Route Reports and Transportation Operating Reports. As noted earlier, we are using the PEIMS financial report transportation expenditure data, and those data do not distinguish between regular and special program expenditures. We estimated the regular and special program transportation PEIMS expenditures by assuming that the program expenditure shares were proportional to the route mile shares. So, if a district's Function 34 PEIMS expenditures were $1 million, and if the 90% of the district route miles were regular program miles and 10% were special program miles, we assumed that regular program expenditures were $900K and that special program expenditures were $100K. We then divided the estimated program expenditures by the number of program route riders to generate separate estimates of the expenditure per rider for regular and special program riders for each district. We display the distribution of the expenditures per rider by program type in Figures 3-12 and 3-13 below.

Figure 3-12: Regular Program Expenditure per Regular Program Rider against Log of Total Regular Route Riders, 2018–19



Source: Authors' calculations; Texas Education Agency Transportation Operations Report, 2019; PEIMS.

Figure 3-13: Special Program Expenditure per Special Program Rider against Log of Special Route Riders, 2018–19

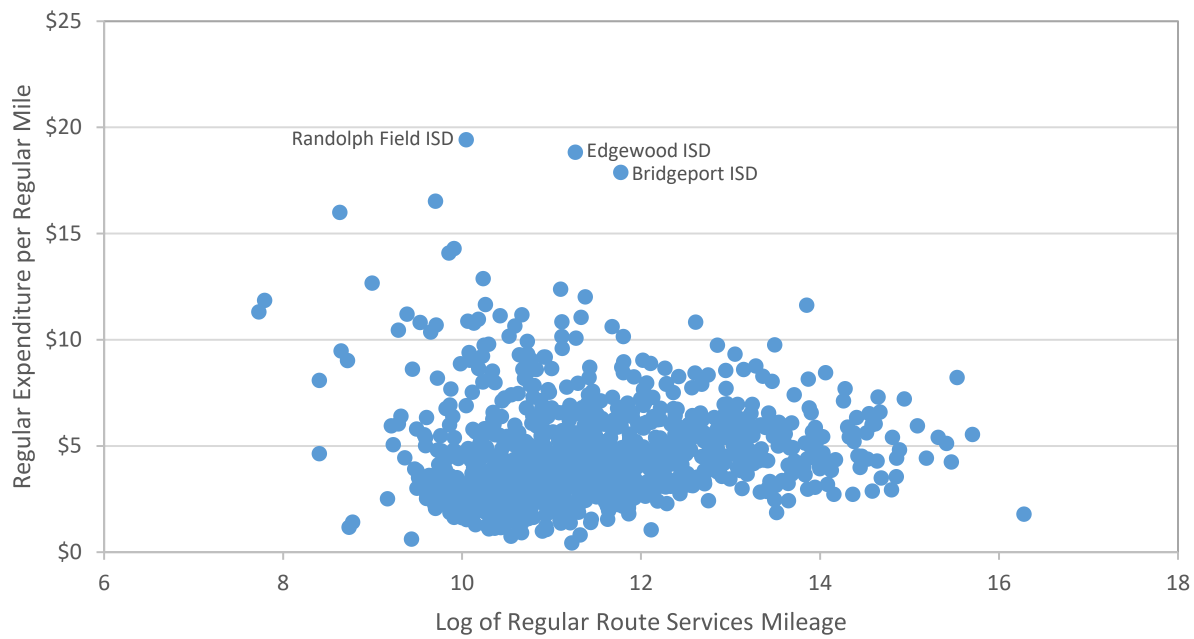


Source: Authors' calculations; Texas Education Agency Transportation Operations Report, 2019; PEIMS.

The median district special program expenditure per rider ($3598) is more than three times the median district regular program expenditure per rider ($987).

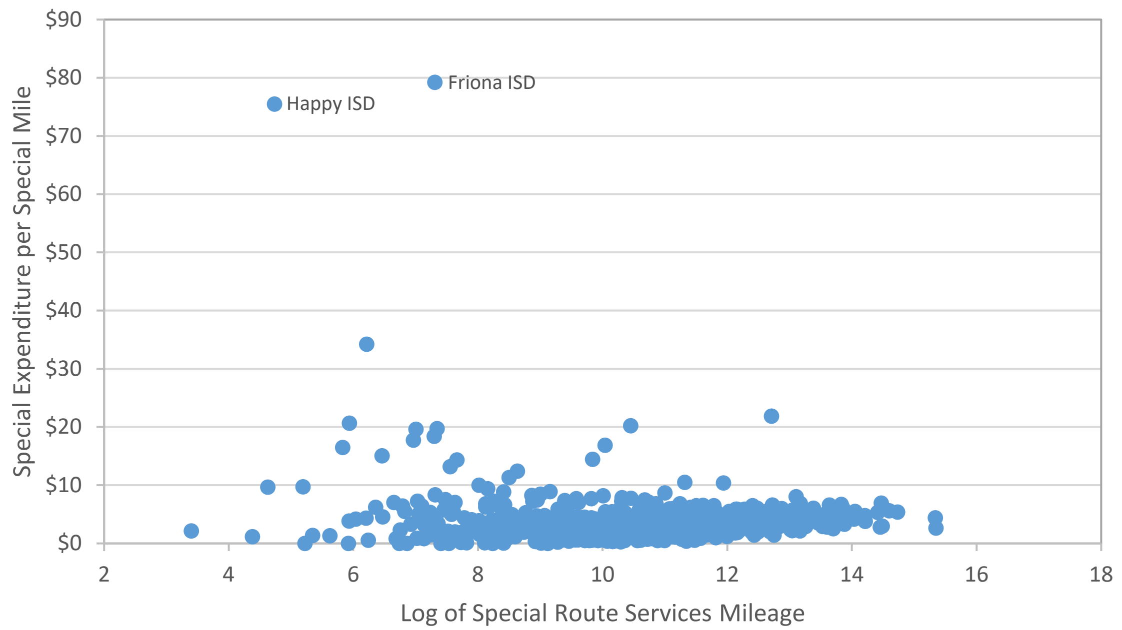
We can also look for evidence of differences in the expenditures associated with producing regular program and special program miles. We display the distributions of expenditures per mile against miles for the two program types in Figures 3-14 and 3-15.

Figure 3-14: Regular Program Expenditure per Mile against Log of Regular Program Miles, 2018–19



Source: Authors' calculations; Texas Education Agency Transportation Operations Report; PEIMS. 2019.

Figure 3-15: Special Program Expenditure per Mile against Log of Special Program Miles, 2018–19



Source: Texas Education Agency Transportation Operations Report, 2019.

The median expenditure per mile for regular program miles is $2.78 and for special program miles is $2.02. The mean regular-special expenditures per mile differential is $0.49. There are 233 districts that have 0 special program miles. There are also several extremely high per mile special program expenditures.

The cost of producing transportation services can also vary by the density of demand for bus seats. We can take a simple correlative look at the cost-density relationship by plotting district expenditure per mile and expenditure/pupil against district population/density.

Figure 3-16 graphs expenditure per mile against population density. There is an upward-sloping relationship, which is consistent with the linear density rate per mile schedule used for calculating the regular route transportation allotment.

Figure 3-17 graphs expenditure per rider against population density. Here there is a concentration of very high expenditure per rider values associated with extremely low population density districts. On both graphs there is an outlier observation for one high cost and high-population-density district, Highland Park ISD.

Figure 3-16: Expenditure per mile against Log of Population Density, 2018–19

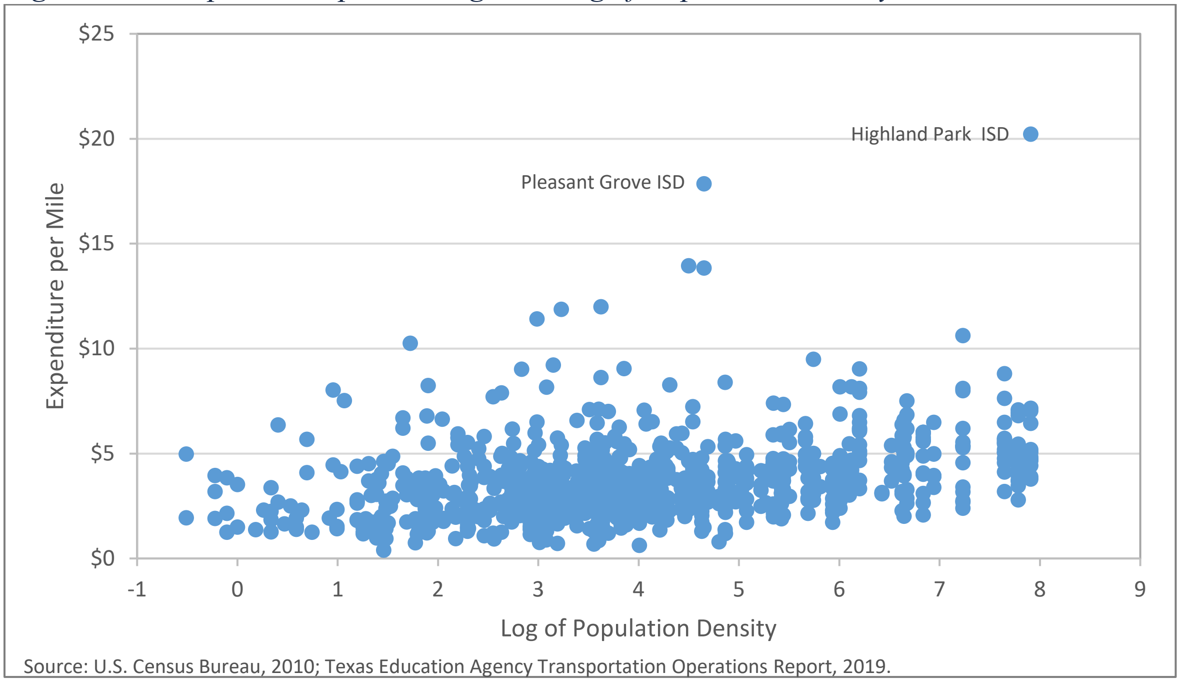
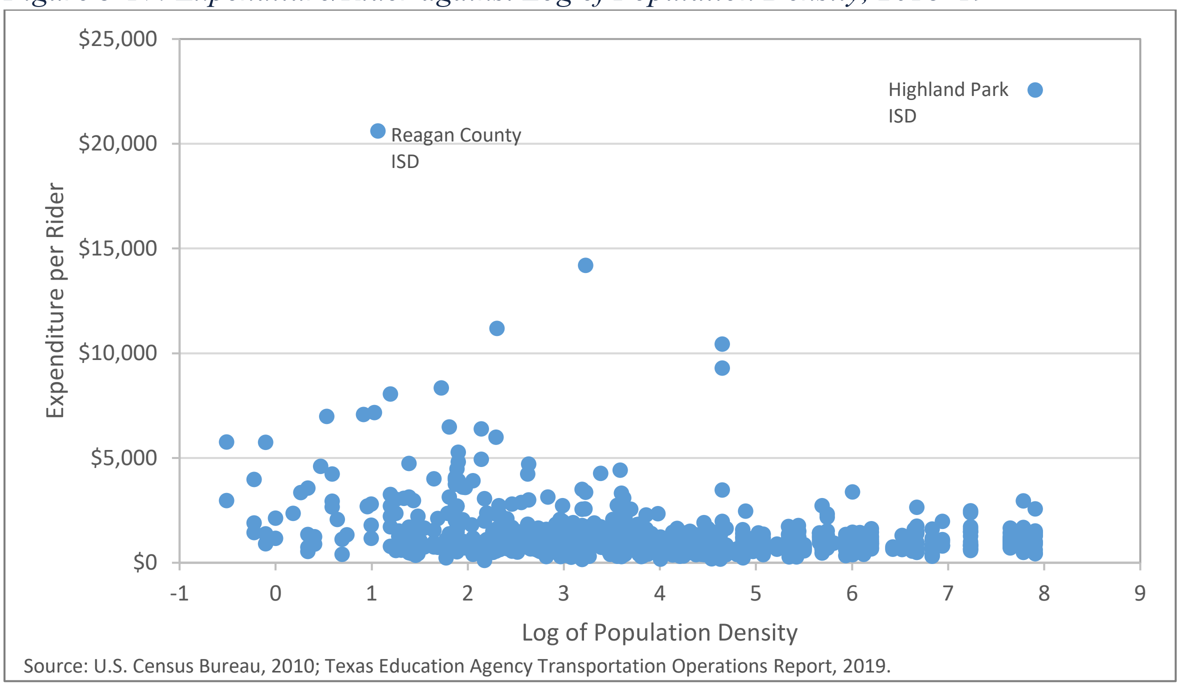
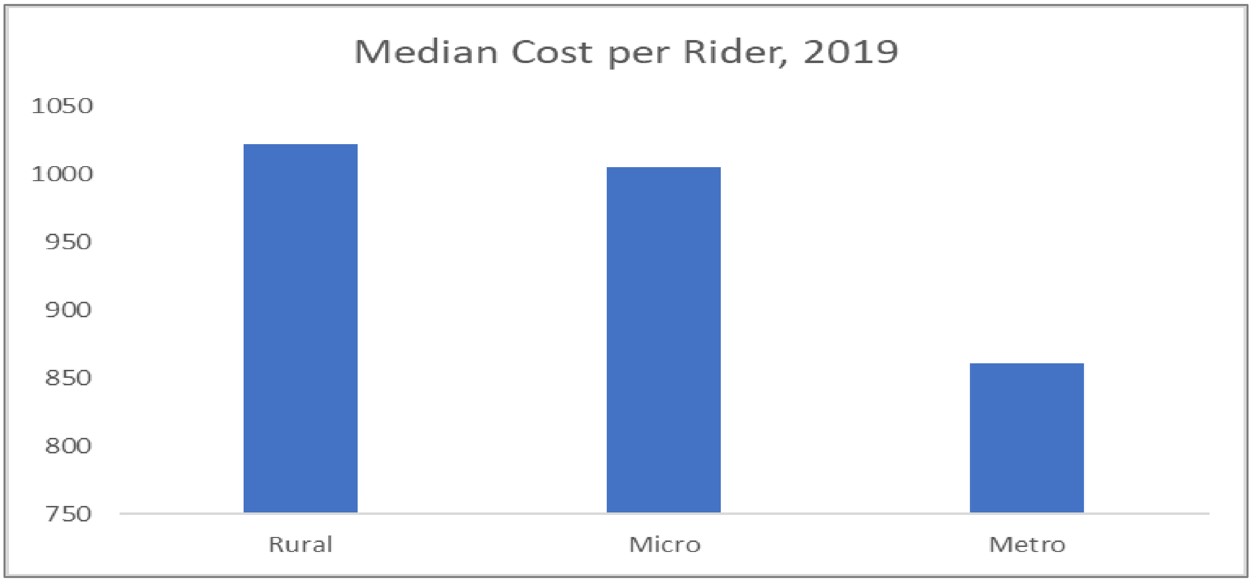


Figure 3-17: Expenditure/Rider against Log of Population Density, 2018–19



Given the relationship between rider density and expenditures per rider shown above, it is not surprising that average (and median) expenditures per rider are higher among rural districts than among metro districts. The average expenditures per rider for micropolitan districts sits between the average values for rurals and metros, and the median expenditure/rider for micro districts is very similar to the rural median.

Figure 3-18: Expenditure per Rider by District Type, 2018–19

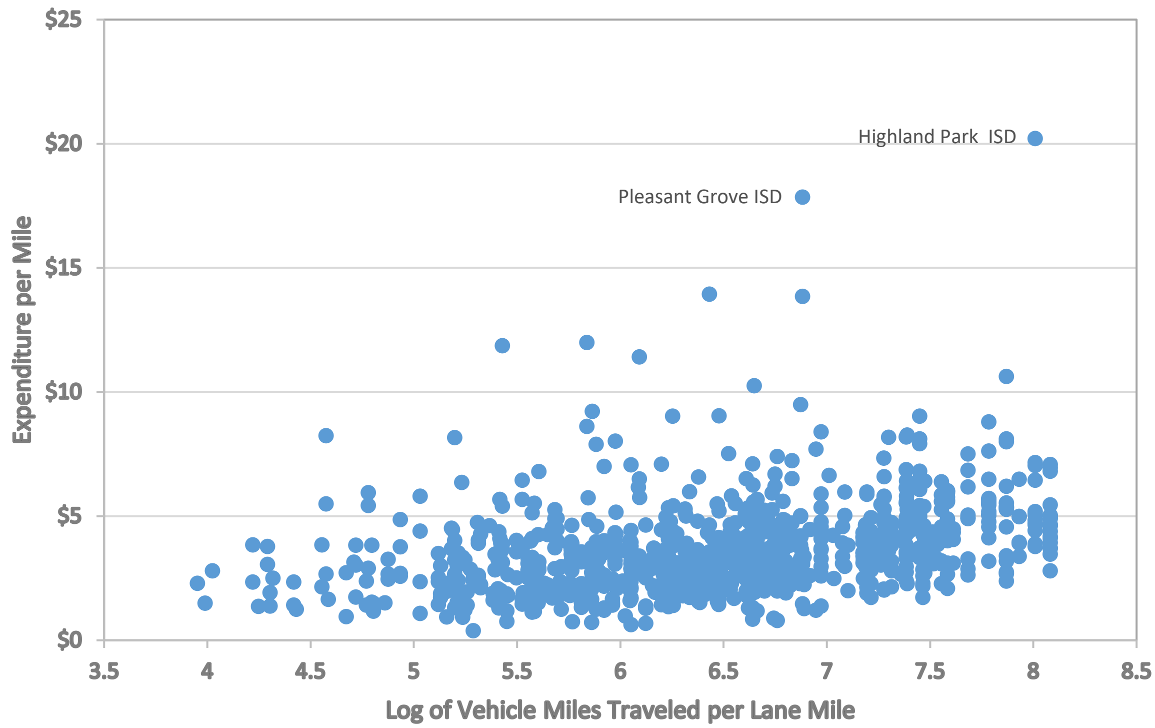


Source: Authors' calculations, 2019; Texas Education Agency Transportation Operations Report, 2019.

The school bus system utilizes the publicly-provided road infrastructure to transport students within the district (home to school trips) and across districts (extracurricular trips). The quality of the public road infrastructure is expected to impact the cost to the school district of providing student transportation. We can take a preliminary look at the relationship between school transportation expenditures and road system quality using data provided by the Texas A&M Transportation Institute. Figures 3-19 and 3-20 plot route expenditure per mile and route expenditure per rider against a measure of roadway utilization (vehicle miles traveled per lane mile, VMTperLM). Higher values indicate greater road system congestion.

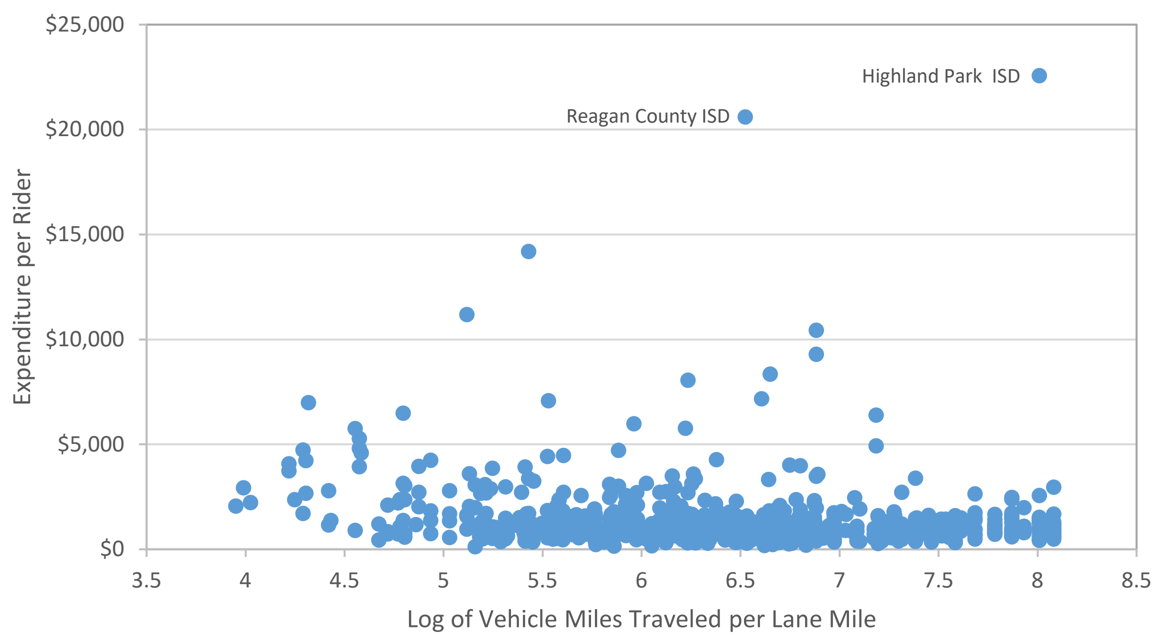
Figure 3-19 indicates that the expenditure per mile for school buses is positively related to the measure of total roadway utilization/congestion. This relationship matches a priori expectations. Figure 3-20 shows the expenditure per rider as being slightly higher in districts with very low roadway utilization rates.

Figure 3-19: Expenditure per Mile against Log of Vehicle Miles Traveled per Lane Mile, 2018–19



Source: Authors' calculations; Texas Education Agency Transportation Operations Report, 2019; Texas Transportation Institute, 2019.

Figure 3-20: Expenditure per Rider against Log of Vehicle Miles Traveled per Lane Mile, 2018–19



Source: Authors' calculations; Texas Education Agency Transportation Operations Report, 2019; Texas Transportation Institute, 2019.

It is important to note that, although the series of figures above provide an interesting first look at several relationships in the data, these are all just descriptive correlations. They do not represent information on causal relationships. The cost function model will allow us to estimate the marginal causal effects of the multiple factors that impact district costs of transporting their students.

Student Transportation Cost Function Model

The underlying assumption of our cost function analysis is that school districts produce transportation outputs—quantity and quality—using a production process (a technology) that combines input factors that are purchased (for example, bus drivers and fuel) with environmental input factors that are not purchased (for example, the system of roads in the district). Thus, school district transportation costs are a function of the outputs produced, the prices of inputs, and other features that influence the production of transportation services, such as the spatial distribution of student riders and the school transportation regulatory environment.

By definition, the cost function represents the minimum cost of producing outputs, given prices of input choices and given various exogenous environmental conditions. (For more on the cost function methodology, see Appendix E.) There exists a sizeable literature which finds that school districts do not all operate in an efficient, cost-minimizing fashion in producing educational outcomes within schools, and that the degree of inefficiency varies considerably across districts. It seems reasonable to assume that school districts may also not operate efficiently in production transportation services. We will use a stochastic frontier cost function model to address this potential for inefficiency.

An important modeling decision for any researcher is the choice of functional form for the cost function. Hutchinson and Pratt (1998, 2007) adopt a translog specification for their school transportation studies. In the bus transit literature, the translog cost function is also the go-to specification in many of the published studies. The translog cost function is a local second-order approximation to an arbitrary cost function. Thus, up to a second order, the translog can serve to approximate any number of possible cost function specifications. Therefore, as in Chapter 2, a translog cost function was used here.

Key components of any cost function analysis are the units of analysis, the measure of expenditures, outcome measures, input prices and environmental factors. We discuss each in turn.

Unit of Analysis

Texas is a district-run transportation system, and districts control all elements of school transportation production. Our cost function is developed and estimated at the district level of aggregation. We will estimate the cost function using a pooled sample of a cross-section of traditional public school districts for the five-year time period 2015–2019.

Expenditures

The dependent variable in our analysis is district variable operating expenditures on student transportation. We exclude depreciation, debt service payments, and capital purchase outlays. We use the PEIMS financial report expenditure data, which reports district expenditures on student transportation under Function Code 34. As pointed out earlier, student transportation expenditures under Function 34 are limited to Route transportation services only. Thus, we are not considering the costs of extracurricular and other non-Route transportation services that districts provide for their students.

Outputs

The definition of outputs is critical to any cost function analysis. In the case of school bus transportation, there are two relevant measures of output. One measure is bus miles. The second output measure is number of student trips. As discussed in our review of the literature on bus transit, bus miles is a vehicle-based or technical output measure. The number of student trips is a passenger-based or demand-based output measure. Cost is, fundamentally, a producer or supplier type concept. In the school transportation context, school districts are suppliers of bus miles. Districts combine labor (bus drivers, fleet maintenance staff, professional staff), materials (utilities services, oil and fuel) and capital (bus fleet, bus barn, road infrastructure) to produce bus miles. The costs that are reported in the PEIMS Student Transportation Function are the expenditures on all of the purchased inputs used to supply Route service bus miles.

Bus miles are, however, best considered an intermediate output. The bus miles are used to transport district students to and from home for the regular instructional day. The student passenger trips or the student passenger miles are the final output.

As argued by Berechman and Guiliano (1984), bus miles represents a measure of output capacity, while passenger trips represent the intensity of utilization of that capacity. When measured with respect to the supply of bus miles, economies of size then measure the change in total cost with respect to a change in capacity. When demand for trips is used, economies of size measures the change in total cost with respect to the density of utilization of capacity.

In our analysis, we will estimate a cost function with both bus miles and student passenger trips as output measures. This approach is similar to that used by Tauchen, Fravel, and Gilbert (1983) in a study of the US intercity bus industry, by Windle (1988) in a study of the US urban bus transit industry, and by Harmatuck (2005) in his cost function study of Midwest bus transit systems. The school transportation cost function privatization studies by Hutchinson and Pratt (1998, 2007) also include bus miles and student trips as their measures of output. A preferred measure would be passenger miles, as used by Windle (1988) in his study of the US urban bus transit industry. Average trip lengths will vary across districts, and this will captured in the passenger mile measure. We do not, however, have data on passenger miles for school districts in Texas. We will include bus miles directly as an output, and include passenger trips measured as riders per mile. This specification matches the most recent published school bus cost specification by Thompson (2011). This specification also measures output in the two dimensions, total miles and linear rider density, which were used in the transportation allotment formula in place during the 2015–19 time period. Given the restrictions on PEIMS Student Transportation Expenditures noted above, we restrict the bus miles to Route miles and the student passenger trips to Route trips.

One issue in defining our outputs is the potential need to separate regular student transportation and special student transportation. The current system of reporting transportation mileage and costs in the district Transportation Route Reports and Transportation Operation Reports maintains this division. If the resource requirements for transporting special-program students are different than for transporting regular-program students, then treating special program outputs as distinct from regular program outputs is appropriate. The PEIMS transportation cost data do not provide a separate cost accounting for regular program and special program student transportation expenditures. We will, therefore, use total bus miles and total student passenger trips as our output measures. We will also include the percentage of riders classified as special program riders and the percentage of miles classified as special program miles to help control for potential differences in the costs of transporting special program students. This is similar to the approach taken by Tauchen, Fravel, and Gilbert (1983) to explore potential differences in costs for different types of municipal bus transit services.

There is also an output measurement issue for student passenger trips. The average number of eligible, i.e., eligible for State funding, students riding to and from school per day is reported in the Transportation Route Services Report. School districts can, however, choose to make district bus service available to non-eligible regular program students. A district can charge a fee for busing the non-eligible students. We do not have data on the average number of non-eligible Route riders. Our measure of the average number of regular Route riders is, therefore, an undercount for some districts. We do not know the number of districts involved nor the magnitude of the undercounting. We may be able to mitigate the potential ridership measurement error problem by instrumenting for the number of riders, using total student enrollment as the instrumental variable.

A more general issue for our outputs is the potential for endogeneity bias. In all of the school and transit bus literature, passengers and miles are treated as being exogenous variables in the transportation cost estimation. In the theory of cost functions, the output is also an exogenous variable. But in the data generating process, both riders and miles could be viewed as choice variables. Potential endogeneity concerns are common in the instructional educational cost function literature. We will follow Gronberg et al. (2015) and Gronberg, Jansen and Taylor (2017) and utilize a control function approach to address the potential endogeneity of these output variables.

Ideally, we would include measures of output quality. The two obvious candidates are trip time and safety. Unfortunately, we do not have data on these two important dimensions of trip quality. Trip time will, however, be related to the number of bus routes. All else equal, increasing the number of routes will decrease the average time kids spend on the bus. We do not have a measure of the number of routes that a district uses to transport its students to and from school. The number of buses may, however, serve as a sound proxy for this missing output-type of bus network quality measure in our estimation. The Transportation Operations Report provides data on the number and type of buses and other student transport vehicles as well as information on the vintage of the rolling stock of buses. We will include the number of buses as an output quality measure and include the percentage of newer buses as a quality characteristic of that output.

Input Prices

Labor costs are the lion's share of bus system operating costs. For the set of districts that run their own student transportation operation, i.e., do not contract out their transportation services, salary and benefit costs are around 80% of their variable operating costs. Variations in the price that districts must pay to hire their transportation employees is thus expected to be a key driver in variations in transportation costs across districts. Given the anticipated pivotal role of labor price differences in understanding operating cost differences, we devote special attention to constructing a sound measure of transportation labor price variation. Our wage index approach is spelled out earlier in Chapter 1 of this report.

Fuel costs are a second important component of annual transportation costs. We will include a measure of district diesel fuel prices to account for variation in this crucial input price. Unfortunately, we do not have data on the prices that districts are paying to fuel up their buses and other student transport vehicles. We purchased a dataset from Oil Price Information Services (OPIS) of average annual diesel fuel prices by county in Texas for the 2015–19 time period. These data are collected by OPIS on a daily basis from a sample of reporting suppliers. The OPIS data are retail prices, which include federal and state diesel taxes. In Texas, during the sample period the state tax on diesel was 20 cents/gallon and the federal tax was 24.4 cents/gallon. Since school districts are exempt from these taxes, we subtracted 44.4 cents from the average county diesel fuel prices reported in the OPIS data.

We recognize that these tax-adjusted retail prices are imperfect measures of the actual prices paid by districts. We expect that districts purchase fuel under a variety of contracting arrangements with fuel suppliers, many/most of which are with wholesale suppliers. Contract lengths likely vary, and districts may band together to negotiate better contracting terms. We are assuming that the underlying exogenous market conditions that generate the quite persistent retail diesel price differences across counties in Texas, e.g., transportation/distribution cost differentials, lead to matching exogenous variation in the wholesale contract prices at which districts actually transact.

Environmental Factors

It is important to account for exogenous factors that impact the transportation cost decision environment. From a cost function estimation perspective, it is important to control for these environmental differences in order to generate valid estimates of the impacts of outputs and input prices on transportation costs from the cross-sectional observations across districts. From a funding policy perspective, it is important to have estimates of the role of differences in these "uncontrollable" factors in determining observed differences in costs across districts.

The decision problem facing school transportation suppliers includes fundamental spatial dimensions that make the problem inherently complex. One key spatial element is the locational distribution of students in the district. The best routing strategies for school transportation suppliers will depend upon the density of student riders in their districts. We will include district population density as an environmental control variable in our cost function analysis. This consideration is consistent with the linear density approach to adjusting rate per mile in the funding formula for regular route services.

The routing choices for school transportation will also depend upon characteristics of the road network system in the district. As discussed above, we will include total vehicle miles per lane mile as a measure of road network quality for each district. The vehicle miles per lane mile is included as a measure of roadway utilization (a congestion proxy). This measure was provided for use in our study by the Texas A&M Transportation Institute. The measure was developed at the county level. Many school districts overlap more than one county boundary. We assign each district the road measure for the district's primary county.

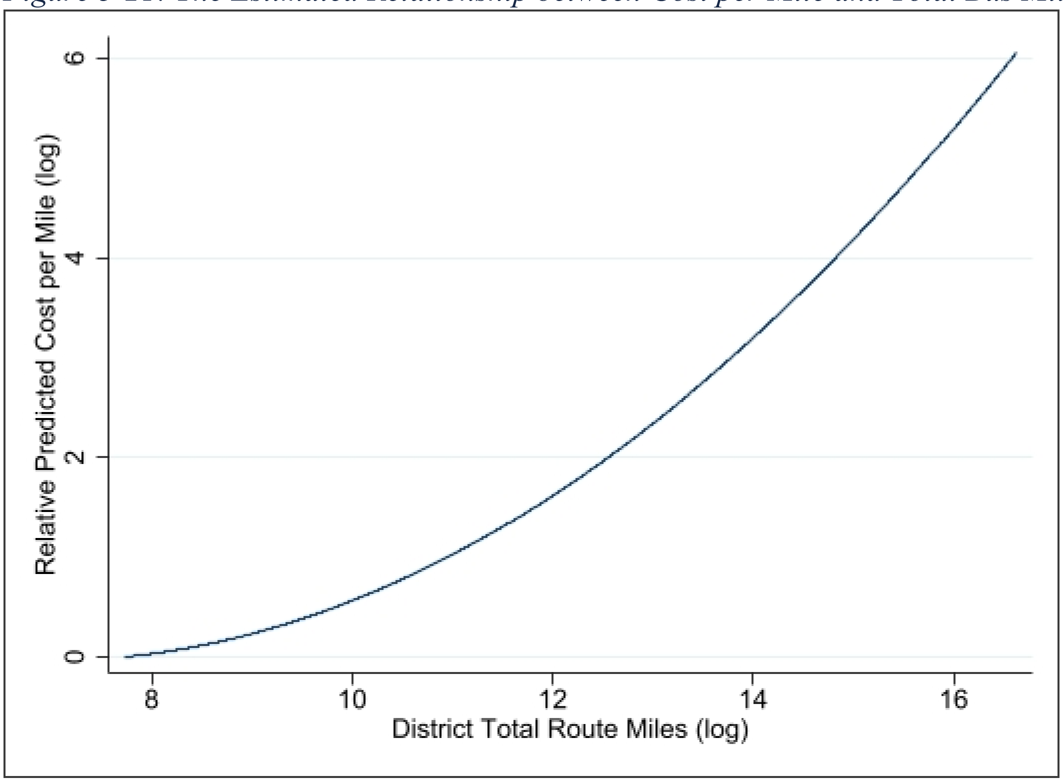
Estimation Results

As discussed in Appendix G, the cost function analysis provides a quite reasonable picture of the supply of route student transportation services. Costs are increasing in outputs and in input prices.

Differences in the density of the distribution of potential riders and in the congestion features of the district road system also impact the costs of hauling kids from home to school. The negative relationship between density and cost matches the finding in Hutchinson and Pratt (1999) from their study of school transportation costs in Tennessee, but is opposite of their finding for school transportation costs in Louisiana (2007).

Figure 3-21 graphs the impact of changes in log miles on predicted cost. The figure is generated by varying miles, while holding all other cost variables at their sample mean values. The slope of the graph is the marginal effect, and the shape of the graph indicates that there are diseconomies in producing bus miles, holding riders per mile and bus capital constant. Since this exercise holds the number of buses, which proxies the number of bus routes, fixed, the increased miles here are generated by increased miles/bus, i.e., either longer average routes or additional bus runs per route.

Figure 3-21: The Estimated Relationship between Cost per Mile and Total Bus Miles



Source: Authors' calculations from Appendix G.

As an alternative exercise, we can ask what the effect of scaling up all dimensions of the bus supply process, i.e., increasing the number of buses/routes proportional to the increase in miles (and riders). The estimated first-order marginal effect at the mean of a 1% increase in route miles, riders, and buses/routes is a 0.0048% increase in cost per mile. This is a very small per mile cost effect, implying nearly constant returns to scale (when evaluated at sample means).

The analysis finds evidence that districts are spending more than what would be expected if they were operating efficiently. The extent of the estimated inefficiency is, however, quite modest. The average degree of estimated inefficiency is only 5.6%, i.e., the average spending only exceeds the estimated minimum frontier cost by 5.6%. The estimated inefficiency is only 10% at the 5th percentile. The relatively strong efficiency estimates are not altogether surprising. A potential benefit of the low level of state support for school transportation is the strong incentive for districts to run their transportation operations efficiently. It is also the case that the technology for producing bus miles is relatively straightforward and well-known.

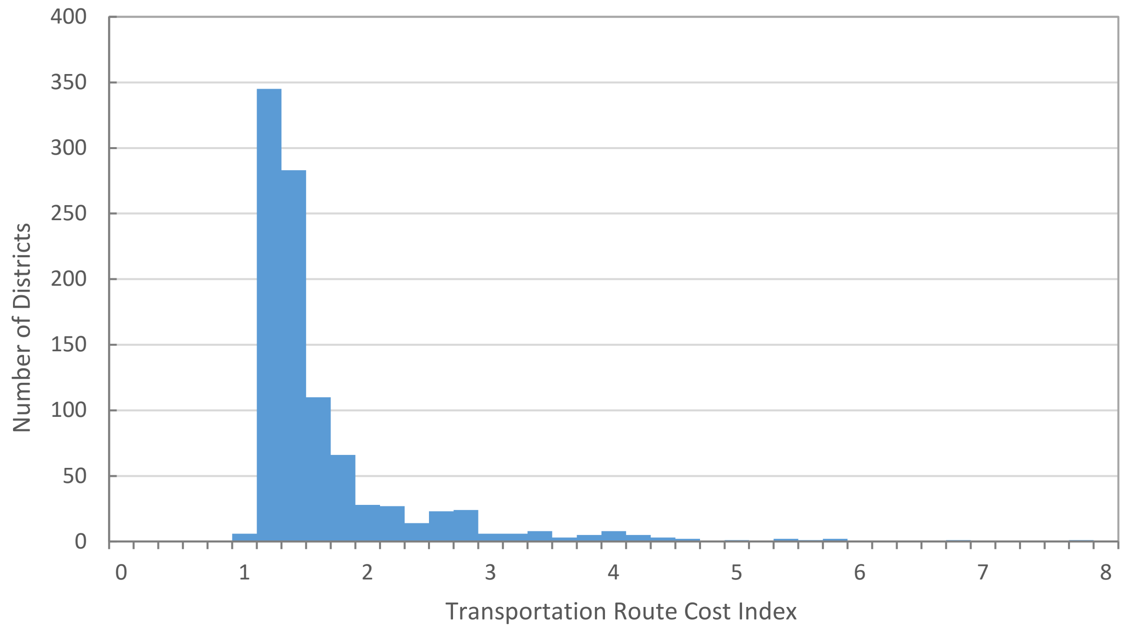
The Transportation Route Cost Index

Once the transportation cost function has been estimated, transportation cost indices can be generated. These cost indices indicate how much more or less it costs to produce bus miles in Houston than in Hutto. The development of the indices involves several steps. For the three output measures in our cost function, the objective is to evaluate districts at common output levels. This effectively holds these cost factors constant across school districts. The usual convention in constructing such indices is to use the state sample average values as the common output levels. For the other cost factors, which are treated as uncontrollable, the cost model is evaluated at the actual value for these variables in each district.

We estimated the cost index for each district by dividing the predicted spending level for each district by the minimum predicted spending level among the sample population of districts. The index values so generated provide a measure of the cost in a district due to its uncontrollable factors relative to the cost in a district with the most cost-favorable characteristics for the uncontrollable factors. For example, an index value of 1.5 indicates that a district is predicted to require 50 percent more dollars per mile to achieve the standardized output levels than the minimum cost district. This normalization implies that the base allotment would only be adjusted upwards. This choice of normalization is motivated by our observation above that Texas currently funds a relatively small share of school transportation costs and that the funding share has fallen significantly over time. Other normalizations are, of course, possible. For example, the reference cost level could be the predicted cost of producing the standardized outputs for a district with the average values of the uncontrollable cost factors. The index values would then range from below one to above one, and some districts would have their regular allotments reduced relative to their allotments under the uniform base rate alternative.

Figure 3-22 illustrates the frequency distribution of the normalized-to-minimum cost index, labeled the Transportation Route Cost Index or TRCI. The graph illustrates the range of the resulting cost index, which is from 1.00 to 7.80. The median of the TRCI is 1.29, so half of the district values of costs per mile, given their district-specific uncontrollable factors, were between 100 percent and 129 percent of the minimum estimated district cost per mile of achieving the standardized level of educational outputs. The TRCI distribution is rather heavily skewed, with a long right tail of districts with TRCI values greater than 2.00. Still, these extreme values of the TRCI distribution are outliers. More than 85% percent of districts have an index value less than 2.00, and only five percent of the values are more than 3.00.

Figure 3-22: Normalized-to-Minimum Transportation Route Cost Index



Source: Authors' calculations.

Table 3-5 provides an additional perspective on the TRCI. As the table illustrates, the average TRCI is highest in rural districts and higher in micropolitan districts than in metropolitan districts. Small districts have higher average TRCI values. Transportation systems in districts located in sparsely populated counties had higher TRCIs than districts in more populous counties.

Table 3-5: The Transportation Route Cost Index, by Location and School District Type, 2019–20

	Number of Districts	Mean	Minimum	Maximum
Metropolitan	475	1.23	1.01	3.05
Micropolitan	190	1.52	1.00	5.75
Rural	315	2.03	1.10	7.80
Very Sparsely Populated County	152	2.84	1.28	7.80
Sparsely Populated County	106	1.70	1.16	2.50
Other County	722	1.25	1.00	2.00
Small District	603	1.73	1.00	7.80
Midsized District	198	1.32	1.00	3.80
Large District	179	1.18	1.00	1.78
Highest Poverty Quintile	195	1.61	1.00	6.68
Lowest Poverty Quintile	196	1.46	1.00	7.80

Source: Authors' calculations from Appendix G

Conclusions

This chapter develops and estimates a model of the costs to school districts of transporting students to and from school. The model is grounded in the academic literature on bus transportation, both school bus services and municipal transit services. The cost of producing school bus miles depends on the number of miles the buses are covering, the cost of bus mile inputs (such as labor prices and fuel prices), the number and spatial distribution of student riders, and upon the environment in which the bus miles are being produced (such as features of the road infrastructure). Cost is an efficiency-based concept, and the observed expenditures on school transportation may deviate from the minimum cost of delivering the level of services provided due to inefficient decision-making practices. We adopt a stochastic frontier translog cost function approach to identifying the role of the various cost factors, including factors that are inherently outside of a district's control in determining district transportation costs while also estimating the degree to which observed district transportation expenditures deviate from best practice minimum costs.

Our cost function estimates well organize the data and provide plausible characterizations of the role of various cost factors in determining the variation across districts in the cost per mile of transporting students to and from school. The characterization of the efficiency—or inefficiency—of the provision of transportation services across districts also seems plausible.

Chapter 4: Strategies to Address Geographic Cost Differences

The analyses in Chapters 1 through 3 demonstrate clearly that there are large geographic differences in the cost of providing educational and transportation services in Texas. Those differences arise from a lack of population density and economies of scale in rural Texas, higher labor costs in urban Texas, and district-by-district differences in uncontrollable cost factors like student need.

The cost function analyses also generated cost indices that could be used to adjust the Foundation School Program (FSP) and the transportation allotment for those differences.

Uncontrollable Cost Adjustment Using the ECI

The ECI could be used as an adjustment for the basic allotment in the FSP. If the legislature chose to go that route, the ECI would replace the small and midsized allotments, the compensatory education allotments, the dyslexia allotment and the special education allotments, as well as the bilingual/ESL allotment in Tier I of the funding formula. Because those allotments largely define weighted average daily attendance (WADA), the ECI would also largely redefine WADA in Tier II of the FSP.

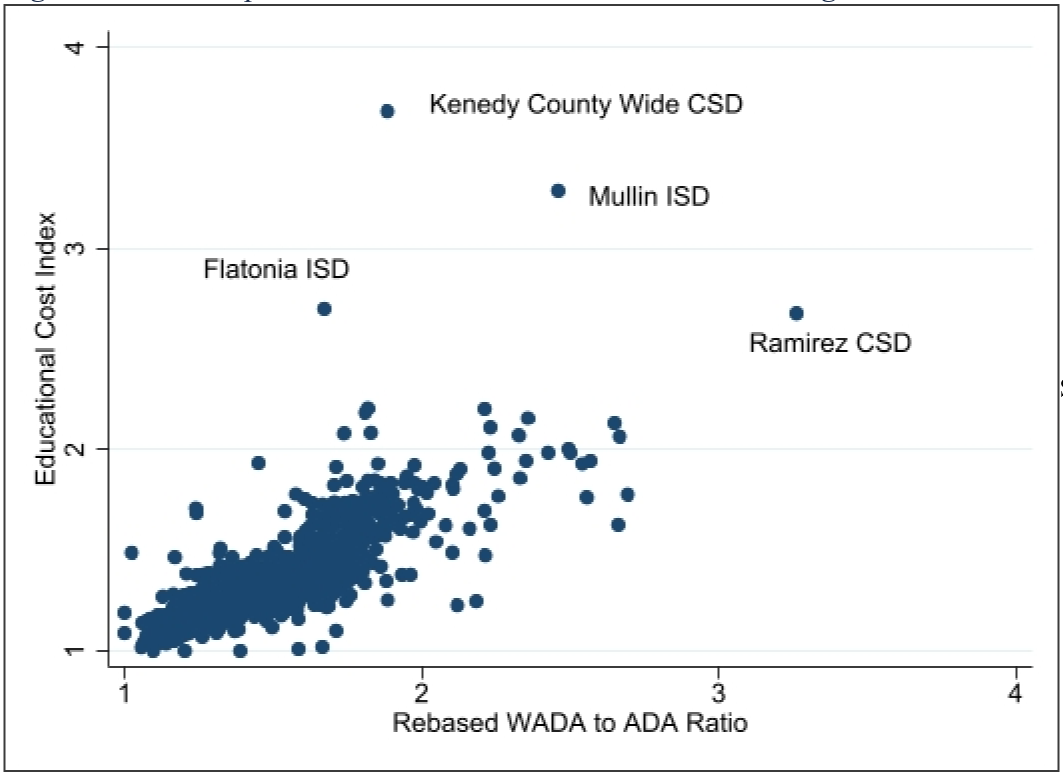
The impact on the FSP allocations is best understood by comparing the ECI to the weights implied by WADA. A district's WADA ratio (i.e., WADA divided by ADA) is calculated by dividing the sum of the school district's Subchapter B & C allotments (i.e., the regular program allotments, small and mid-size allotment, special education adjusted allotment, dyslexia allotment, compensatory education allotment, bilingual education allotment, career and technology allotment, early education allotment and a number of smaller allotments) by the district's basic allotment. In 2019–20, that WADA ratio ranged from 1.12 to 11.26, indicating that one district in the state (San Vicente ISD with its ADA of 10.8 and WADA of 121.1) received nearly 11 times as much allotment funding per pupil as the district at the other end of the spectrum (Highland Park ISD with its ADA of 6549.9 and WADA of 7334.8).

Rebasing the WADA ratio to start at 1.00 allows for a direct comparison with the ECI.[41] Excluding the extreme outliers in the WADA ratio, the ECI and the WADA ratio were reasonably well correlated.[42] (See Figure 4-1.) The median ECI (1.30) was lower than the median rebased WADA ratio (1.50), so adopting the ECI as a replacement for most of the allocations that lead to the WADA ratio would imply a substantial, offsetting increase in the basic allotment. As a general rule, geographic adjustments using the ECI rather than the WADA ratio would reduce the differential for district size, but lead to higher relative funding for districts in counties with very low population density and districts in counties with high labor costs.

[41] Rebasing the WADA ratio means dividing the WADA ratio by the lowest value (in this case, 1.12).

[42] There were six districts with a WADA to ADA ratio greater than 4 in 2019–20. In addition to San Vicente ISD, those outlier district were Dell City ISD, Divide ISD, Doss Consolidated CSD, Marathon ISD, and Valentine ISD. The largest of these six districts had an ADA of 50.1. The correlation between the ECI and the rebased WADA ratio was 0.779 excluding these seven districts, and 0.785 including them.

Figure 4-1: A Comparison between the ECI and the WADA Weights, 2019–20



Source: Authors' calculations; PEIMS 2019.

Uncontrollable Cost Adjustment Using the TRCI

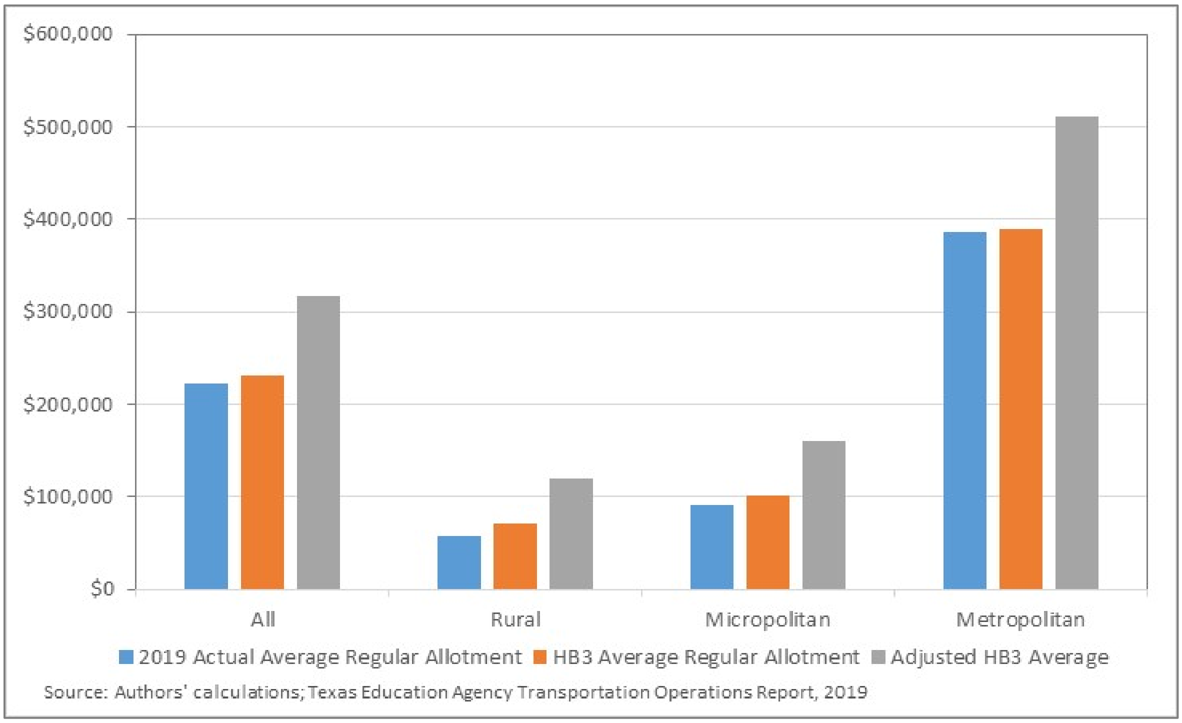
Similarly, the TRCI could be used to adjust the transportation allotment for geographic differences in the cost of education. During the 2014–15 through 2018–19 time period, the regular program allotment was determined using a linear density-based formula that provided a higher rate per mile for districts with a larger number of riders per mile. The formula for determining the regular transportation allotment was amended under House Bill 3 (HB 3) in June 2019. Under HB 3, the regular program allotment will be based on a flat rate per mile to be set by the Legislature in the General Appropriations Act (GAA). The rate adopted for 2020–21 under the current GAA is $1 per mile.

The TRCI could be used directly to adjust the base allotment rate per mile for the estimated differential costs associated with the different uncontrollable cost environments facing district transportation planners. A district with an estimated TRCI of 1.29 (i.e., the median district) would be assigned a regular program allotment rate of 1.29 times the base allotment rate per mile. Assuming no change in the base allotment, then at the median the current HB3 rate of $1 per mile would be increased to $1.29/ mile.

We illustrate the potential application of the TRCI for adjusting the base transportation allotment rate and identify the potential funding implications. For each district in our sample, we multiply the district TRCI times the $1 per mile base allotment to generate adjusted allotment rates by district. For our application exercise, we assume that the TRCI is capped at 2.0 for allotment funding purposes. The maximum adjusted allotment will be $2.00/ mile. We then calculate the adjusted regular program allotments for 2018–19 for our sample districts. The State funding for the regular program transportation allotment would be $276,216,992. The cost factor adjustments would increase State transportation regular allotment spending by $49,586,992 relative to funding at the flat $1 per mile HB3 rate, a 21.88 % spending increase.

We show the distributive effect of the increased transportation funding under the TRCI adjusted funding approach in Figure 4-2 below.

Figure 4-2: Allotment Comparison: HB3, TRCI Adjusted HB3 and Actual Average Regular Allotment



The cost index adjustment would provide a substantial boost to the average allotments for districts of all three types. The Metropolitan districts would benefit the most from use of the TRCI-adjusted allotment rates relative to the flat rate. The average regular allotment funding would be bumped up by almost $62,000 per Metropolitan district.

Of course, the legislature could also use its discretion to make a revenue-neutral, downward adjustment to the base allotment per mile. A base allotment of $0.82 per mile multiplied by the TRCI would have the same predicted impact on the state's total transportation allotment as the flat $1 per mile under HB3. Under a TRCI-driven model, relatively more transportation funding would flow to the districts where uncontrollable transportation costs are higher.

Long-term Implementation of the ECI and TRCI

One key to successful long-term implementation of either the ECI or the TRCI would be the development of a strategy for regularly updating the indices. Although many of the factors that drive differences in the costs of education and transportation are unlikely to change over time, other factors—such as wage levels outside of education, fuel costs and student enrollments—are sensitive to changing economic and socioeconomic conditions. To ensure that the cost indices are functioning as intended, the ECI and TRCI should be updated regularly, either by using the estimated cost models to generate new cost predictions corresponding to new values for the various cost factors, or by re-estimating the cost models themselves.

Uncontrollable Cost Adjustment Using Individual Cost Factor Adjustments

While it would be straightforward and analytically sound to use the ECI and TRCI as black-box cost adjustments, the legislature may instead choose to use the information provided herein to refine key components of the two funding models. For example:

The Compensatory Program Allotments:

HB3 instructed the Commissioner of Education to develop new measures of student socioeconomic status. The new measures were to be based on the demographics of the Census block where each educationally disadvantaged student resides. The funding formula weight was increased by 2.5 percentage points for economically disadvantaged students who live in Census blocks that were the least disadvantaged; and increased by 7.5 percentage points for economically disadvantaged students who live in severely disadvantaged Census blocks.

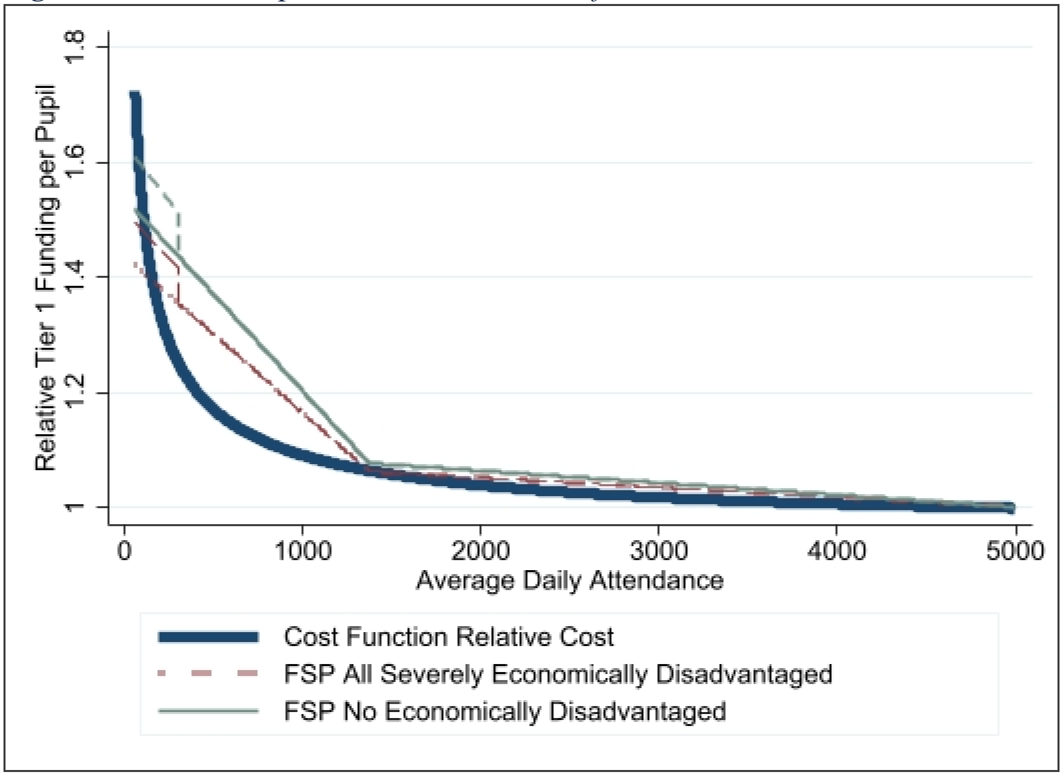
The cost function estimates suggest that the concentration of poverty in the school –not just the concentration of poverty in the student's neighborhood—can have a significant impact on the cost of education. The higher the percentage of economically disadvantaged students, the higher the increase in cost associated with an additional disadvantaged student. The legislature may wish to consider adding an intensity adjustment, perhaps modeled after the concentration grants or the targeted assistance grants that are part of the Title 1 program (Baker et al. 2015).

The Small and Midsized Allotments:

HB3 replaced the small and midsized adjustments in the funding formula with small and midsized allotments. This change treated the scale adjustments in a manner analogous to the allotment for compensatory education. "Instead of flowing funds to small and mid-size districts as an adjustment that occurs before other funding adjustments, the funding now flows as an allotment under Tier I at the same time as other funding adjustments, such as the compensatory education allotment and the bilingual allotment" (TEA 2019). As a result, the small and midsized adjustments no longer have a multiplicative effect on the other allotments, such as the compensatory or bilingual program allotments. This change reduced the funding differential for small and midsized districts.

The cost function estimates suggest that the small and midsized allotments still overstate the relationship between school district size and the cost of education for all but the smallest districts. Figure 4-3 compares the small and midsized allotments expected under HB3 (as a percentage of the funding for an otherwise identical district that was not eligible for the size adjustments) to those implied by the cost function analysis. (The dashed lines at the far left indicate the supplemental allotment provided to districts with fewer than 300 students when the district is the only one in the county.) There are two alternatives for the FSP—one in which all the students are economically disadvantaged and live in a severely disadvantaged Census block, and one in which none of the students are economically disadvantaged. As the figure illustrates, the cost function estimates indicate that a district with 300 students costs 25% more per pupil to operate than a school district with 5,000 students, whereas the funding formula provides an additional 35% to 44% per pupil, depending on the percentage of economically disadvantaged students. The gap between the FSP and the cost function-based estimates is even wider for districts with between 300 and 1,000 students.

Figure 4-3: Two Perspectives on Economies of Scale



Source: Authors' calculations from the FSP and Appendix F.

The Cost of Education Index:

HB3 removed the Cost of Education Index (CEI) that had been part of the FSP since 1991. The CEI was a labor cost index that ranged from 1.02 to 1.20, indicating that the cost of labor was 18% higher in some parts of the state than it was in others. The CEI adjusted both Tier I and Tier II of the FSP, although the adjustment to Tier II was only half as large as the adjustment to Tier I, by formula. Every school district received some adjustment under the formula to compensate for uncontrollable variations in labor costs.

While the CEI was clearly outdated, Chapter 1 of this report provides evidence that significant regional differences in labor cost persist, and offers a ready-made replacement for legislative consideration, namely the TCI. The TCI ranged from 1.00 to 1.37, indicating that the regional differences in labor cost are even wider now than they were 25 years ago, when the CEI was developed.

Because the non-labor components of a school district's budget are unlikely to have the same geographic pattern as labor costs, the legislature may wish to embed the Texas TCI or the ACS-CWIFT in a regional cost index. As discussed in Taylor (2015) a regional cost index can be constructed as a weighted average of the various price indices.

$$Regional\ Cost\ Index = P_1S_1 + P_2S_2 + P_3S_3 \ldots P_nS_n$$

where P_1 is the relative price a district must pay for its most important input and S_1 is the share of the budget devoted to that input, P_2 is the relative price for a second type of input and S_2 is the share of the budget devoted to that second input, and so on. Prior to HB3, the CEI was operationalized into Tier I of the funding formula as if it were embedded in a regional cost index with a labor weight of 0.71 (Taylor, 2015b).

On average over the period from 2014–15 through 2018–19, Texas school districts allocated 72 percent of their current operating expenditures to salaries and benefits for teachers and other professionals, 14 percent of their spending to support personnel and the remaining 14 percent of their spending to non-labor expenses.[43] Assuming that the prices of non-labor inputs do not vary geographically, a regional cost index for Texas could be constructed as:

$$RCI_i = TCI_i \cdot 0.72 + APCI_i \cdot 0.14 + 0.14,$$

where TCI_i is the TCI for school district i and $APCI_i$ is the APCI for school district i. (Of course, the ACS-CWIFT and/or HS-CWI could be used as alternative price indices in the construction of the geographic cost of education index.) The resulting regional cost index for 2020 would range from 1.00 to 1.29, with a median of 1.10. Such a regional cost index could then be multiplied by the basic allotment in Tier I of the funding formula, and (unlike the CEI) be fully incorporated into the calculation of WADA.

Should the legislature choose to adopt the Texas TCI or APCI (either individually or as part of a regional cost index) it would be prudent to also adopt a process by which the indices would be updated, so that the indices would continue to perform as intended when economic conditions changed. As with the ECI and TRCI, the Texas TCI and APCI could be updated either by using the estimated models to generate new indices based on updated values of the uncontrollable labor cost factors, or by re-estimating the wage models themselves. A hybrid approach, with annual updates to the uncontrollable cost factors and periodic updates to the analysis could be a particularly useful strategy (and could be easily implemented by the TEA or some other state agency).

A commitment to regular studies—as opposed to regular updates—is likely to be less effective, as illustrated by previous experiences with TCIs in Texas and Wyoming. Texas did not updated the CEI for 25 years, despite multiple studies demonstrating that the geographic pattern of labor costs had changed (e.g., Alexander et al. 2000; Taylor 2004 & 2015a) and Wyoming has not updated the Hedonic Wage Index used in its school funding formula for more than 15 years, despite a similar series of legislatively commissioned studies also demonstrating changes in the geographic pattern of labor costs (e.g., Taylor 2010, 2015c & 2020).

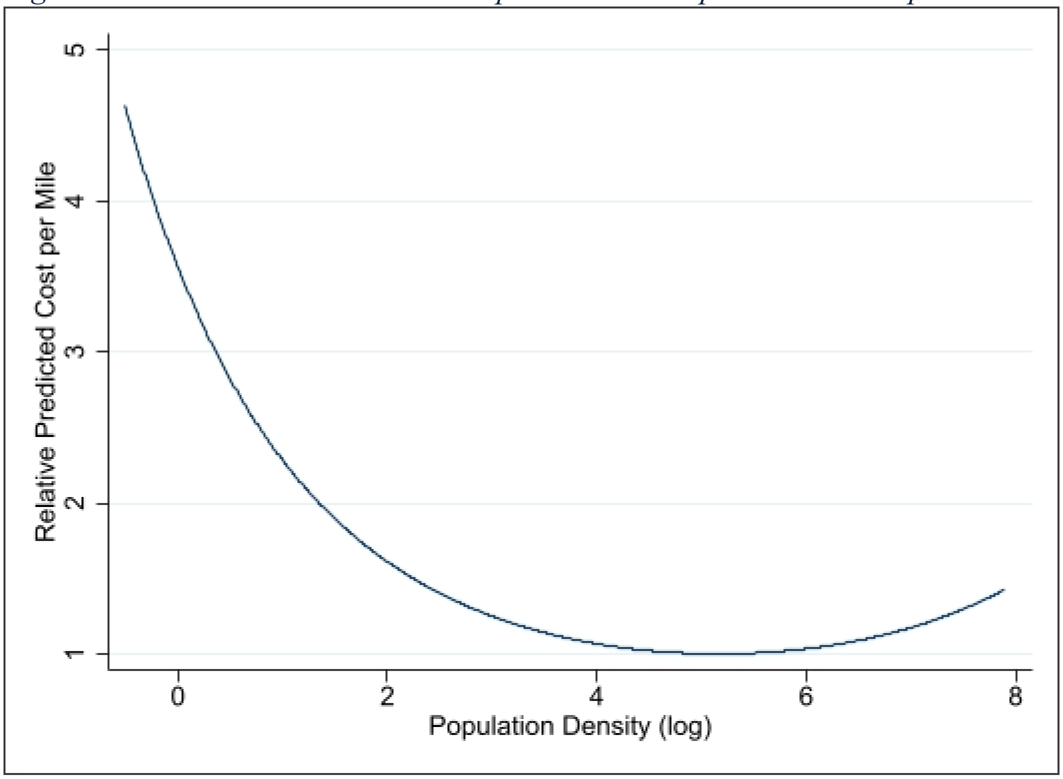
[43] The budget shares were calculated for school district operating expenditures, excluding food and transportation services as in the calculation of the ECI.

Transportation Cost Adjustments:

Just as the educational cost function can provide useful insights into the design of funding formula adjustments, the transportation cost function can provide useful insights into appropriate adjustments to state's transportation allotments.

For each of the uncontrollable cost factors in the transportation cost model, one can estimate the percentage change in cost per mile due to a one percent change in value of the cost factor, holding the value of all other cost factors fixed (at the statewide mean). For example, Figure 4-4 graphs the relationship between the log of population density and cost per mile (relative to the minimum predicted cost per mile). As the figure illustrates, holding all other cost factors constant, the cost per mile was sharply higher for districts in very sparsely populated parts of the state than it was for districts in more densely populated areas.

Figure 4-4: The Estimated Relationship between Cost per Mile and Population Density



Source: Authors' calculations from Appendix G.

The cost function analysis suggested that three uncontrollable, transportation cost factors were particularly important—fuel costs, labor costs and population density. Therefore we used the estimated marginal effects to generate a set of cost adjustment factors for each of these three key cost factors.

To simplify implementation, we divided the cost factor data into quartiles for each cost factor, and then calculated the appropriate adjustment for the median of the quartile (relative to the minimum input price) to yield a predicted percentage increase in cost per mile. (See Appendix G.) We then treated the predicted percentage cost increase as a transportation allotment cost adjustment factor for all districts in that quartile.

For example, the first quartile of the diesel fuel price in our 2018–19 sample ranged from the sample minimum of \$2.106 per gallon to \$2.253 per gallon. The median quartile fuel price of \$2.229 per gallon is 5.8% higher than the minimum price. That 5.8% higher fuel price is estimated to increase cost per mile by \$0.027, so the fuel price adjustment factor assigned to districts in the first quartile is 0.03. This fuel price adjustment would increase the regular program allotment rate to \$1.03 for these districts. We repeated this process for the other three fuel price quartiles, and ended up with four fuel price adjustment factors, one for each of the quartile fuel price ranges. We applied the same process to generate four quartile adjustment factors for the transportation labor wage index, the other key exogenous input price cost factor. The input price adjustment factors are given in Tables 4-1 and 4-2.

Table 4-1: Transportation Regular Program Allotment Rate Adjustment Factor for Fuel Index

Quartile	Value Range	Adjustment Factor
First Quartile	2.106 – 2.253	0.03
Second Quartile	2.254 – 2.296	0.04
Third Quartile	2.297 – 2.345	0.05
Fourth Quartile	2.346 – 2.921	0.07

Table 4-2: Transportation Regular Program Allotment Rate Adjustment Factor for Wage Index

Quartile	Value Range	Adjustment Factor
First Quartile	1 – 1.068	0.02
Second Quartile	1.069 – 1.108	0.03
Third Quartile	1.109 – 1.166	0.05
Fourth Quartile	1.167 – 1.363	0.07

For population density, the key environmental cost factor, we modified the adjustment factor generating process slightly. Population density in Texas ranged widely, from a minimum density of 0.6 persons per square mile to a maximum of 2,718 persons per square mile.[44] The population density distribution is also highly skewed, with over half of the densities below 50, and a 75th percentile value of only 209.5. Furthermore, the estimated marginal effect of population density is negative, so higher population density represents a more advantageous cost environment. Therefore, we subdivided the top quartile into two parts—those above and those below the 90th percentile of population density (i.e., a population density of 769.9 persons per square mile)—and used that 90th percentile as the reference point for cost adjustments. This approach assigned a zero density adjustment factor to all districts in the top decile of the population density distribution. The density adjustment factors are given in Table 4-3.

[44] Population density was measured at the c

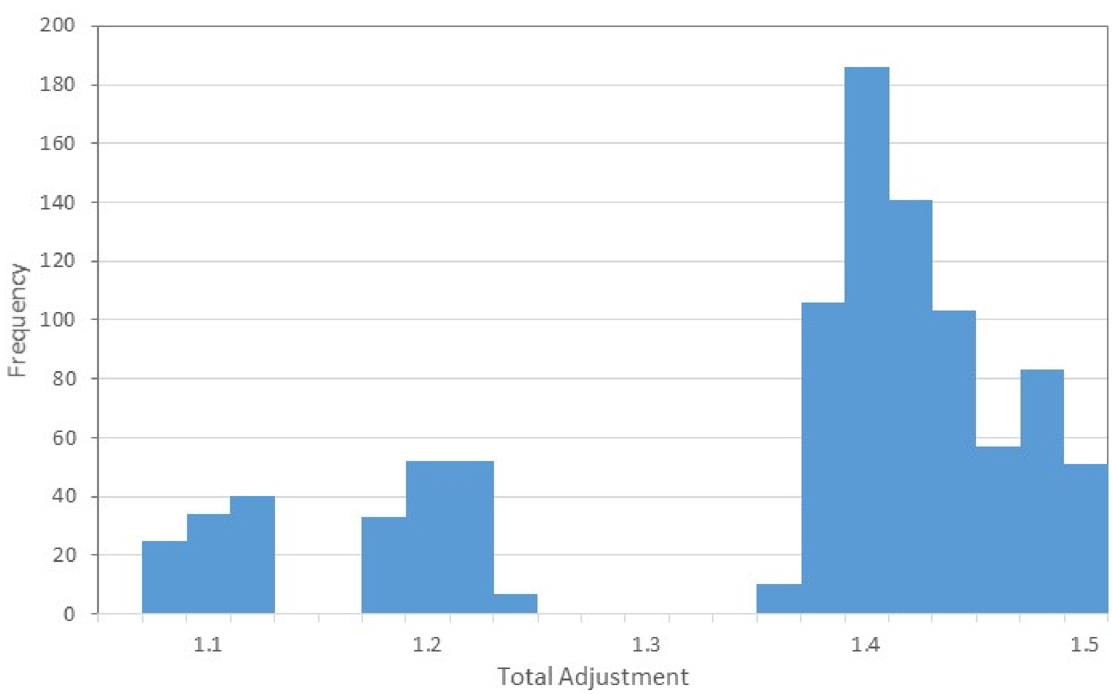
Table 4-3: Transportation Regular Program Allotment Rate Adjustment Factor for Population Density

Quartile	Value Range	Adjustment Factor
First Quartile	0.6 – 19.4	0.35
Second Quartile	19.5 – 47.2	0.33
Third Quartile	47.3 – 209.5	0.31
Fourth Quartile, Up to 769.9	209.6 – 769.9	0.13
Fourth Quartile, 770 and above	770 – 2,718	0.00

A district could be in the first quartile for the fuel index, the second for the wage and the third for population density (or vice versa). The total input cost factor adjustment to the basic transportation allotment would simply be the sum of the three adjustment factors for each district.

The histogram for the distribution of the cost factor adjusted regular program transportation allotment rates across our sample districts is given in Figure 4-5 below. The three groupings are driven by the jumps in the density adjustment—all of the districts with cost-factor adjustments above 1.3 are in the first three quartiles of the density adjustment. The median adjustment is 1.40, or 140% per mile.

Figure 4-5: Cost Factor Adjusted Regular Program Transportation Allotment Rates

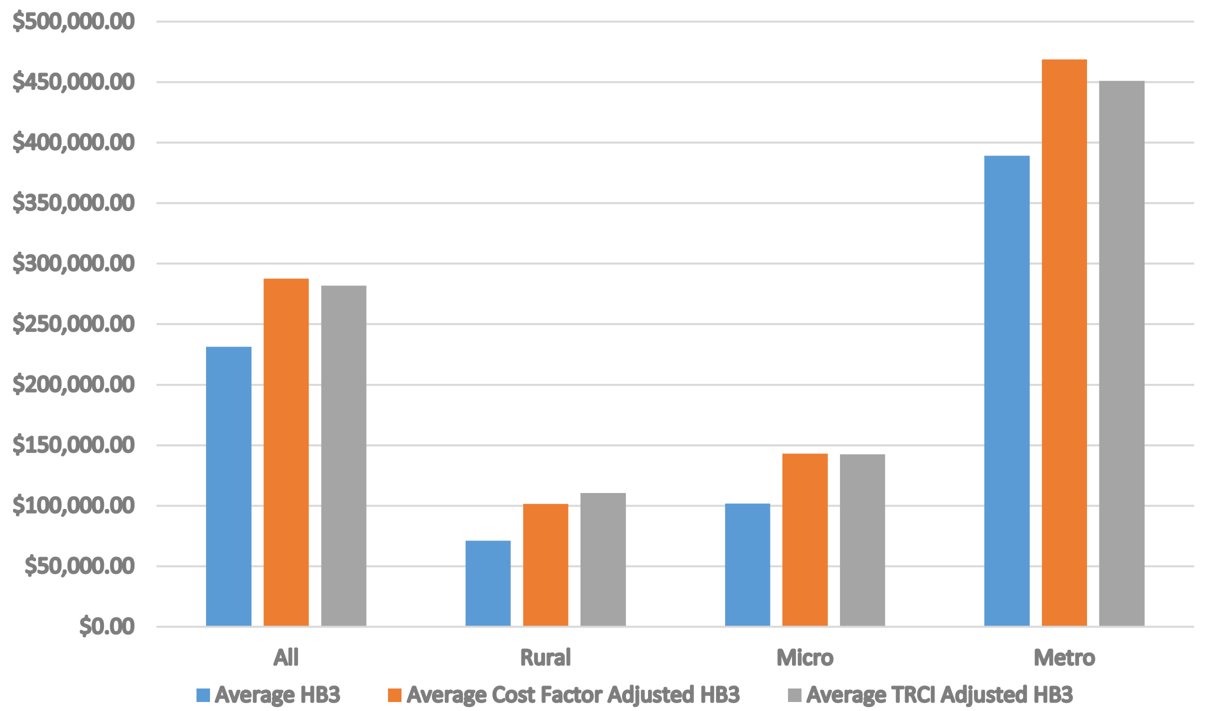


Source: Authors' calculations; Texas Education Agency Transportation Report, 2019; U.S. Census Bureau, 2010

If we use these per mile allotment rates to calculate the regular program allotments based upon 2018–19 regular mileage, the State funding for the regular program transportation allotment would be $281,873,792. The cost factor adjustments would increase State transportation allotment spending by $55,243,792 relative to funding at the flat $1 per mile HB3 rate. This is more than a $5.5 million dollar increase from the projected allotment spending that our Transportation Cost Index adjustment would have generated.

The distributional effects of replacing the flat allotment rate with the cost factor adjusted rates by district type are shown in Figure 4-6 below. The average cost factor adjusted allotments are less generous than the TRCI adjusted allotments for rural districts and more generous than the TRCI adjusted allotments for metropolitan counties. The average allotments for districts in micropolitan counties are almost identical for the two alternative adjustment experiments.

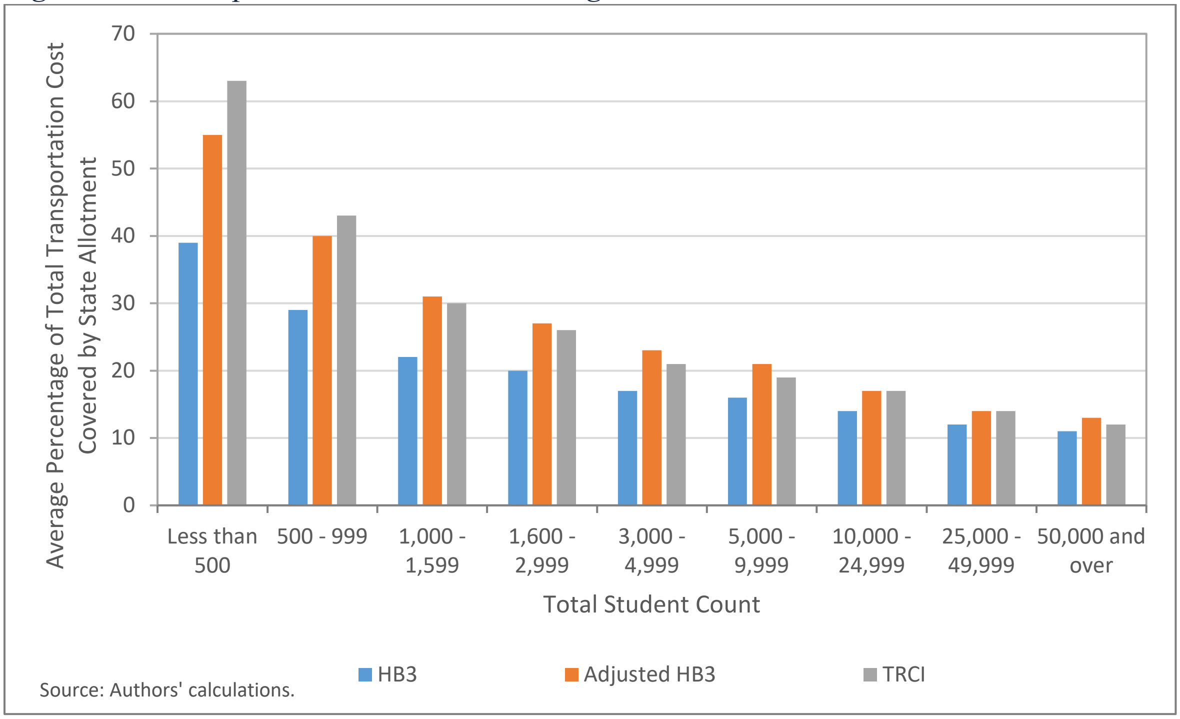
Figure 4-6: Distributional Effects of Replacing the Flat Allotment Rate with the Cost Factor Adjusted Rates



Source: Authors' calculations; Texas Education Agency Division of School Finance, 2019.

We can also calculate the transportation allotment funding shares that would be generated under our two alternative allotment rate cost-adjustment simulations. We report those generated funding shares in Figure 4-7. There are two main takeaways from this figure. First, the average gain in transportation allotment funding share under the two cost adjustment approaches is decreasing in district size. The smallest districts would realize the largest average jump in funding share if the regular program allotment rate were adjusted for differences in uncontrollable costs under adjustment simulations. Second, the average percentage of transportation costs covered by the cost-adjusted allotments would remain well below 50% for the majority of districts. The incentives for districts to seek cost-reducing strategies for their transportation operations would remain strong.

Figure 4-7: Transportation Allotment Funding Share Simulation



Conclusions

We demonstrate how the estimated cost functions could be meaningfully used in the context of the FSP and the school district transportation allotment. Our basic approach was to use cost function analysis to examine geographic variations in the costs of education and transportation due to factors beyond the control of school districts. The generated Transportation Route Cost Index (TRCI) captures the relative cost of supplying the same transportation outputs in a tough input environment relative to a more favorable input environment; the Educational Cost Index (ECI) similarly captures the relative cost of supplying the same educational outputs in various educational environments. Both the ECI and the TRCI could be used directly to enhance the Texas school funding models by modifying key allotments in the funding formulas. However, since the ECI and TRCI are complicated functions of the cost factors, we also provide more simplified, i.e., more linearized, and alternative cost index approaches. Either approach offers a way to adjust the transportation cost allotment and the FSP for relevant uncontrollable cost factors.

References

Alexander, Celeste, Timothy Gronberg, Dennis Jansen, Harrison Keller, Lori Taylor, and Philip Uri Treisman. 2000. *A Study of Uncontrollable Variations in the Costs of Texas Public Education*. A summary report prepared for The 77th Texas Legislature.

Alliance for Excellent Education. 2004. *Tapping the Potential: Retaining and Developing High-Quality New Teachers*. Retrieved December 9, 2020 from https://all4ed.org/reports-factsheets/tapping-the-potential-retaining-and-developing-high-quality-new-teachers/.

Andrews, Matthew, William Duncombe, and John Yinger. 2002. "Revisiting Economies of Size in American Education: Are We Any Closer to a Consensus?" *Economics of Education Review* 21, no. 3 (June): 245–262. doi: 10.1016/S0272-7757(01)00006-1.

Baker, Bruce D., Lori L. Taylor, and Arnold Vedlitz. 2008. *Adequacy Estimates and the Implications of Common Standards for the Cost of Instruction*. National Academy of Science, National Research Council.

Battese, G. E., and T. J. Coelli. 1995. "A model for technical inefficiency effects in a stochastic frontier production function for panel data." *Empirical Economics* 20: 325–332. doi: 10.1007/BF01205442.

Belfield, Clive R., and Henry M. Levin, 2002. "The Effects of Competition between Schools on Educational Outcomes: A Review for the United States." *Review of Educational Research* 72, no.2 (Summer): 279–341. doi: 10.3102/00346543072002279

Berechman, Joseph. 1983. "Costs, Economies of Scale and Factor Demand in Bus Transit: An Analysis." *Journal of Transport Economics and Policy* 17, no. 1 (January): 7–24. https://www.jstor.org/stable/20052664.

Berechman, Joseph, and Genevieve Giuliano. 1984. "Analysis of the Cost Structure of an Urban Bus Transit Property." *Transportation Research Part B: Methodological* 18(4–5): 277–87. doi: 10.1016/0191-2615(84)90011-0.

Berechman, Joseph, and Genevieve Giuliano. 1985. "Economies of Scale in Bus Transit: A Review of Concepts and Evidence." *Transportation* 12 (May): 313–332. doi: 10.1007/BF00165470.

Billingsley, Bonnie S. 1993. "Teacher Retention and Attrition in Special and General Education: A Critical Review of the Literature." *The Journal of Special Education* 27(2): 137–174. doi: 10.1177/002246699302700202

Bogler, Ronit. 2001. "The Influence of Leadership Style on Teacher Job Satisfaction." *Educational Administration Quarterly* 37, no. 5 (December): 662–683. doi: 10.1177/00131610121969460.

Borman, Geoffrey D. and N. Maritza Dowling. 2008. "Teacher Attrition and Retention: A Meta-Analytic and Narrative Review of the Research." *Review of Educational Research* 78(3): 367–409. doi: 10.3102/0034654308321455

Button, K.J. and O'Donnell, K.J., 1985. "An Examination of the Cost Structures Associated with Providing Urban Bus Services in Britain." *Scottish Journal of Political Economy*, 32, no. 1 (February): 67–81. doi: 10.1111/j.1467-9485.1985.tb00788.x.

Carr, Matthew. 2011. "The Impact of Ohio's EdChoice on Traditional Public School Performance." *Cato Journal* 31, no. 2 (July): 257–84. http://object.cato.org/sites/cato.org/files/serials/files/cato-journal/2011/5/cj31n2-5.pdf.

Caves, Douglas W., and Laurits R. Christensen. 1988. "The Importance of Economies of Scale, Capacity Utilization, and Density in Explaining Inter-Industry Differences in Productivity Growth." *Logistics and Transportation Review* 24 (January): *3–32*.

Chingos, Matthew and Kristin Blagg. 2017. *Student Transportation and Educational Access: How Students Get to School in Denver, Detroit, New Orleans, New York City, and Washington, DC*. Washington, DC: Urban Institute. Retrieved from https://www.urban.org/research/publication/student-transportation-and-educational-access.

De Borger, Bruno, Kristiaan Kerstens, and Álvaro Costa. 2002. "Public Transit Performance: What Does One Learn from Frontier Studies?" *Transport Reviews* 22 (1): 1–38. doi: 10.1080/01441640010020313.

De Rus, Ginés. 1990. "Public Transport Demand Elasticities in Spain." *Journal of Transport Economics and Policy* 24, no.2 (May): 189–201. https://www.jstor.org/stable/20052917.

Dodson III, Marvin, and Thomas A. Garrett. 2004. "Inefficient Education Spending in Public School Districts: A Case for Consolidation?" *Contemporary Economic Policy* 22, no. 2 (April): 270–280. doi: 10.1093/cep/byh019.

Downes, Thomas A., and Thomas F. Pogue. 1994. Adjusting School Aid Formulas for the Higher Cost of Educating Disadvantaged Students. *National Tax Journal* 47 (1): 89–110.

Duncombe, William, Jerry Miner, and John Ruggiero. 1995. "Potential Cost Savings from School District Consolidation: A Case Study of New York." *Economics of Education Review* 14, no. 3 (September): 265–284. doi: 10.1016/0272-7757(94)00011-T.

Duncombe, William, and John Yinger. 1997. "Why Is It so Hard to Help Central City Schools?" Journal of Policy Analysis and Management, 16, no. 1 (Winter): 85–113. doi: 10.1002/(SICI)1520-6688(199724)16:1<85::AID-PAM5>3.0.CO;2-E

Duncombe, William, and John Yinger. 2005a. "How Much More Does a Disadvantaged Student Cost?" *Economics of Education Review* 24, no. 5 (October): 513–532. doi: 10.1016/j.econedurev.2004.07.015.

Duncombe, William, and John Yinger. 2005b. *Estimating the Costs of Meeting Student Performance Outcomes Adopted by the Kansas State Board of Education*. A study prepared for the Kansas Division of Legislative Post Audit.

Duncombe, William, and John Yinger. 2011. "Are Education Cost Functions Ready for Prime Time? An Examination of Their Validity and Reliability." *Peabody Journal of Education* 86(1): 28–57. doi: 10.1080/0161956X.2011.539954.

Egalite, Anna J. 2016. "The Competitive Effects of the Louisiana Scholarship Program on Public School Performance." Louisiana Scholarship Program Evaluation Report #4. Fayetteville: Univ. of Arkansas, School Choice Demonstration Project; New Orleans: Tulane Univ., Education Research Alliance. doi: 10.2139/ssrn.2739783.

Eom, Tae Ho, William Duncombe, Phuong Nguyen-Hoang, and John Yinger. 2014. "The Unintended Consequences of Property Tax Relief: New York's STAR Program." *Education Finance and Policy* 9, no. 4 (Fall): 446–480. doi: 10.1162/edfp_a_00143.

Erichsen, Kristen and John Reynolds. 2020. “Public School Accountability, Workplace Culture, and Teacher Morale.” *Social Science Research*, 85 (January): 102347. doi: 10.1016/j.ssresearch.2019.102347

Fazioli, R., M. Filippini, and P. Prioni. 1993. “Cost-Structure and Efficiency of Local Public Transport: The Case of Emilia Romagna Bus Companies.” *International Journal of Transport Economics* 20, no.3 (October): 305–324. https://www.jstor.org/stable/42747453.

Figlio, David, and Cassandra M.D. Hart. 2014. “Competitive Effects of Means-Tested School Vouchers.” *American Economic Journal: Applied Economics* 6, no. 1 (January): 133–56. doi: 10.1257/app.6.1.133.

Figlio, David, and Krzysztof Karbownik. 2016. “Evaluation of Ohio’s EdChoice Scholarship Program: Selection, Competition, and Performance Effects.” Thomas B. Fordham Institute, July 7, 2016. Retrieved from https://fordhaminstitute.org/ohio/research/evaluation-ohios-edchoice-scholarship-program-selection-competition-and-performance.

Goldhaber, Dan. 1999. “An Alternative Measure of Inflation in Teacher Salaries.” In *Selected Papers in School Finance, 1997–99*, NCES 1999–334, edited by William J. Fowler Jr., 29–54. Washington, DC: US Department of Education, National Center for Education Statistics.

Goldhaber, Dan, Katherine Destler, and Daniel Player. 2010. “Teacher Labor Markets and the Perils of Using Hedonics to Estimate Compensating Differentials in the Public Sector.” *Economics of Education Review* 29, no. 1 (February): 1–17. doi: 10.1016/j.econedurev.2009.07.010.

Golebiewski, Julie Anna. 2011. “An Overview of the Literature Measuring Education Cost Differentials.” *Peabody Journal of Education* 86 (1): 84–112.

Gronberg, Timothy J., Dennis W. Jansen, Mustafa U. Karakaplan, and Lori Taylor. 2015. “School District Consolidation: Market Concentration and the Scale-Efficiency Tradeoff.” *Southern Economic Journal* 82, no. 2 (October): 580–597. doi: 10.1002/soej.12029.

Gronberg, Timothy J., Dennis W. Jansen, and Lori L. Taylor. 2011a. "The Adequacy of Educational Cost Functions: Lessons from Texas." Peabody Journal of Education 86 (1): 3–27. doi: 10.1080/0161956X.2011.539953.

Gronberg, Timothy J., Dennis W. Jansen, and Lori L. Taylor. 2011b. “The Impact of Facilities on the Cost of Education.” *National Tax Journal* 64 (1): 193–218. doi: dx.doi.org/10.17310/ntj.2011.1.08.

Gronberg, Timothy J., Dennis W. Jansen, and Lori L. Taylor. 2012. “The Relative Efficiency of Charter Schools: A Cost Frontier Approach.” *Economics of Education Review* 31, no. 2 (April): 302–317. doi: 10.1016/j.econedurev.2011.07.001.

Gronberg, Timothy J., Dennis W. Jansen, and Lori L. Taylor. 2017. “Are Charters the Best Alternative? A Cost Frontier Analysis of Alternative Education Campuses in Texas.” *Southern Economic Journal* 83, no. 1 (January):721–743. doi: 10.1002/soej.12173.

Gronberg, Timothy J., Dennis W. Jansen, Lori L. Taylor, and Kevin Booker. 2004. *School Outcomes and School Costs: The Cost Function Approach*. Austin, TX: Texas Joint Select Committee on Public School Finance. Retrieved from http://bush.tamu.edu/research/faculty/TXSchoolFinance/.

Gronberg, Timothy J., Dennis W. Jansen, Lori L. Taylor, and Kevin Booker. 2005. *School Outcomes and School Costs: A Technical Supplement*. Austin, TX: Texas Joint Select Committee on Public School Finance. Retrieved from http://bush.tamu.edu/research/faculty/TXSchoolFinance/.

Grosskopf, Shawna, Kathy J. Hayes, and Lori L. Taylor. 2014. "Efficiency in Education: Research and Implications." *Applied Economic Perspectives and Policy* 36, no. 2 (June): 175–210. doi: 10.1093/aepp/ppu007.

Grosskopf, Shawna, Kathy J. Hayes, Lori L. Taylor, and William L. Weber. 1999. "Anticipating the Consequences of School Reform: A New use of DEA." *Management Science* 45, no. 4 (April):608–620. doi: 10.1287/mnsc.45.4.608.

Grosskopf, Shawna, Kathy J. Hayes, Lori L. Taylor, and William L. Weber. 2001. "On the Determinants of School District Efficiency: Competition and Monitoring." *Journal of Urban Economics* 49, no. 3 (May):453–478. doi: 10.1006/juec.2000.2201.

Grosskopf, Shawna, Kathy J. Hayes, Lori L. Taylor, and William L. Weber. 2013. "Centralized or Decentralized Control of School Resources? A Network Model." *Journal of Productivity Analysis* 43 (2): 139–150. doi: 10.1007/s11123-013-0379-2.

Hanley, Paul F. 2007. "Transportation Cost Changes with Statewide School District Consolidation." *Socio-Economic Planning Sciences* 41, no. 2 (June): 163–179. doi: 10.1016/j.seps.2005.10.004.

Hanushek, Eric A. 1999. "Adjusting for differences in the costs of educational inputs." In *Selected Papers in School Finance, 1997–99*, NCES 1999–334, edited by William J. Fowler Jr., 13–28. Washington, DC: US Department of Education. National Center for Education Statistics.

Harmatuck, Donald J. 2005. "Cost Functions and Efficiency Estimates of Midwest Bus Transit Systems." *Transportation Research Record: Journal of the Transportation Research Board* 1932 (1): 43–53. doi: 10.1177/0361198105193200106.

Hoxby, Caroline M. 2003. "School Choice and School Productivity." *The Economics of School Choice*, edited by Caroline M. Hoxby, 287–341. Chicago: Univ. Of Chicago Press. Retrieved from http://www.nber.org/chapters/c10091.pdf.

Hutchinson, E. Bruce, and Leila J. Pratt. 1999. "The Comparative Cost of Privatized Public School Transportation in Tennessee." *Policy Studies Journal* 27, no. 3 (August): 446–458. doi: 10.1111/j.1541-0072.1999.tb01979.x.

Hutchinson, E. Bruce, and Leila J. Pratt. 2007. "Is contracting out government services the great panacea? Evidence from Public School Transportation in Louisiana." *Journal of Private Enterprise* 23, no. 1 (Fall): 67–83.

Ingersoll, Richard M. 2001. "Teacher Turnover and Teacher Shortages: An Organizational Analysis." *American Educational Research Journal* 38, no. 3 (Fall): 499–534. doi: 10.3102/00028312038003499

Ingersoll, Richard M., and Thomas M. Smith. 2003. The Wrong Solution to the Teacher Shortage. *Educational Leadership* 60, no. 8 (May): 30–33.

Imazeki, Jennifer. 2008. "Assessing the Costs of Adequacy in California Public Schools: A Cost Function Approach." *Education Finance and Policy* 3, no. 1 (Winter): 90–108. doi: 10.1162/edfp.2008.3.1.90

Imazeki, Jennifer, and Andrew Reschovsky. 2004. "Is No Child Left Behind an Un (or Under) Funded Federal Mandate? Evidence from Texas," *National Tax Journal* 57. no.3 (September): 571–588. https://www.jstor.org/stable/41790231.

Imazeki, Jennifer, and Andrew Reschovsky. 2006. “Does No Child Left Behind Place a Fiscal Burden on States? Evidence from Texas.” *Education Finance and Policy* 1, no.2 (Spring): 217–246. doi: 10.1162/edfp.2006.1.2.217.

Jimenez-Castellanos, Oscar, and Amy Marcetti Topper. 2012. “The Cost of Providing an Adequate Education to English Language Learners: A Review of the Literature.” *Review of Educational Research* 82, no. 2 (June): 179–232. doi: 10.3102/0034654312449872.

Jinnai, Yusuke. 2014. “Direct and Indirect Impact of Charter Schools’ Entry on Traditional Public Schools: New Evidence from North Carolina.” Economics Letters 124, no.3 (September): 452–56. doi: 10.1016/j.econlet.2014.07.016.

Kang, Byung-Goo, and Kenneth V. Greene. 2002. “The Effects of Monitoring and Competition on Public Education Outputs: A Stochastic Frontier Approach.” *Public Finance Review* 30 (1): 3–26. doi: 10.1177/109114210203000101.

Karakaplan, Mustafa U., and Levant Kutlu. 2019. “School District Consolidation Policies: Endogenous Cost Inefficiency and Saving Reversals.” *Empirical Economics* 56, no. 5 (May): 1729–1768. doi: 10.1007/s00181-017-1398-z.

Karlaftis, Matthew, Patrick S. McCarthy, and Kumares C. Sinha. 1999. “System Size and Cost Structure of Transit Industry.” *Journal of Transportation Engineering* 125, no. 3 (May): 208–215. doi: 10.1061/(ASCE)0733-947X(1999)125:3(208).

LaFaive, Michael D., and James M. Hohman. 2015. *School Support Service Privatization: A Five-State Survey*. Mackinac Center for Public Policy. Retrieved from https://www.mackinac.org/s2015-06.

Lazarus, Sheryl S., and Gerard J. McCullough. 2005. “The Impact of Outsourcing on Efficiency in Rural and Nonrural School Districts: the Case of Pupil Transportation in Minnesota.” *Journal of Regional Analysis and Policy* 35 (1): 54–68.

Levaggi, Rosella. 1994. “Parametric and Non-Parametric Approach to Efficiency: The Case of Urban Transport in Italy.” *Studi Economici* 49: 67–88.

Loeb, Susanna, Linda Darling-Hammond, and John Luczak. 2005. “How Teaching Conditions Predict Teacher Turnover in California Schools.” *Peabody Journal of Education* 80(3): 44–70. doi: 10.1207/s15327930pje8003_4.

Matas, Anna, and José Luis Raymond. 1998. “Technical Characteristics and Efficiency of Urban Bus Companies: The Case of Spain.” *Transportation* 25, no. 3 (August): 243–263. doi: 10.1023/A:1005078830008.

McMahon, Walter W. 1996. “Intrastate Cost Adjustments.” In *Selected Papers in School Finance, 1994*, NCES 1996–068, edited by William J. Fowler Jr., 93–114. Washington, DC: US Department of Education. National Center for Education Statistics.

Millimet, Daniel L., and Trevor Collier. 2008. "Efficiency in Public Schools: Does Competition Matter?" *Journal of Econometrics* 145, no.1–2 (July): 134–157. doi: 10.1016/j.jeconom.2008.05.001.

Misra, Kaustav, Paul W. Grimes, and Kevin E. Rogers. 2012. “Does Competition Improve Public School Efficiency? A Spatial Analysis.” *Economics of Education Review* 31, no. 6 (December): 1177–1190. doi: 10.1016/j.econedurev.2012.08.001.

Nguyen, Tuan D., Lam D. Pham, Michael Crouch, and Matthew G. Springer. 2020. "The Correlates of Teacher Turnover: An Updated and expanded Meta-Analysis of the Literature." *Educational Research Review*, 31 (November): 100355.

Nweke, Winnifred C., Gerald M. Eads, Comfort Y. Afolabi, Cynthia E. Stephens, and F. D. Toth. 2006. *The impact of Teacher Workforce Retirement/Attrition on Personnel Replacement Cost*. Atlanta: Georgia Professional Standards Commission. https://www.gapsc.com/Commission/Research_Publication/downloads/Teacher_Workforce_Retirement_and_Turnover_Cost.pdf.

Pinkovitz, William H., Joseph Moskal, and Gary Green. 1997. "How Much Does Your Employee Turnover Cost?" *Small Business Forum* 14 (3): 70–71.

Reback, Randall. 2008. "Teaching to the rating: School Accountability and the Distribution of Student Achievement." *Journal of Public Economics* 92, no. 5–6 (June): 1394–1415. doi: 10.1016/j.jpubeco.2007.05.003.

Roseman, Edward. 1981. *Managing Employee Turnover: A Positive Approach*. New York, NY: AMACOM.

Rumberger, Russell W., and Patricia Gándara. 2008. "Resource Needs for Educating Linguistic Minority Students." In *Handbook of Research in Education Finance and Governance*, edited by Helen F. Ladd and Edward B. Fiske, 591–611.

Small, Kenneth A. 1992. *Urban Transportation Economics*. Chur Switzerland: Harwood Academic.

Smith, Thomas M., and Richard M. Ingersoll. 2004. "What Are the Effects of Induction and Mentoring on Beginning Teacher Turnover?" *American Educational Research Journal* 41 (3): 681–714. doi: 10.3102/00028312041003681.

Synar, E. A. 2010. Examination of the Financial Costs of Teacher Turnover in Mid-Sized Urban School Districts. (Doctoral dissertation, University of Oklahoma).Retrieved from ProQuest.

Tauchen, Helen, Frederic D. Fravel, and Gorman Gilbert. 1983. "Cost Structure of the Intercity Bus Industry." *Journal of Transport Economics and Policy, 17, no. 1 (January): 25–47*. https://www.jstor.org/stable/20052665.

Taylor, Lori L. 2000. "The Evidence on Government Competition." *Economic and Financial Policy Review* 2: 2–10. Federal Reserve Bank of Dallas.

Taylor, Lori L. 2004. "Adjusting for Geographic Variations in Teacher Compensation: Updating the Texas Cost-of-Education Index." Austin, TX: Joint Select Committee on Public School Finance.

Taylor, Lori L. 2010. "An Analysis of Wyoming Cost Adjustments." A report prepared for the Wyoming Select Committee on School Finance Recalibration.

Taylor, Lori L. 2015a. "Adjusting For Geographic Variations in Teacher Compensation: Updating The Texas CEI," Testimony before the House Public Education Committee.

Taylor, Lori L. 2015b. "When equality is not equity: Regional cost differences and the real allocation of educational resources." In Legal frontiers in education: Complex law issues for leaders, policymakers and policy implementers. Emerald Group Publishing Limited.

Taylor, Lori L. 2015c. "Options for Updating Wyoming's Regional Cost Adjustment," A report prepared for the Wyoming Select Committee on School Finance Recalibration

Taylor, Lori L. 2020. “A Vision for Regional Cost Adjustment: The 2020 Hedonic Wage Index,” A report submitted to The Select Committee on School Finance Recalibration, December 2020.

Taylor, Lori L., Timothy Gronberg, and Dennis Jansen. 2017. *Would School District Consolidation Lead to Cost Savings in Major Metropolitan Counties? A Cost Function Analysis,* A report prepared for the Texas Education Agency and the University of Texas at Dallas Education Research Center. Retrieved from https://tea.texas.gov/sites/default/files/TEC%2012%201013%20report%202017.pdf

Taylor, Lori L., Timothy J. Gronberg, Dennis W. Jansen, and Mustafa U. Karakaplan. 2014. *Anticipating the Consequences of School District Consolidation in Major Metropolitan Areas: A Simulation Based on Cost Function Analysis.* A report prepared for the Texas Education Agency and the University of Texas at Dallas Education Research Center. Retrieved from https://tea.texas.gov/sites/default/files/TEC%2012%201013%20Report_final_v2.pdf.

Taylor, Lori and Harrison Keller. 2003. “Competing Perspectives on the Cost of Education.” In *Developments in School Finance 2001–02,* NCES 2003–403, edited by William J. Fowler Jr., 111–126. Washington, DC: US Department of Education. National Center for Education Statistics.

Thompson, Owen. 2011. “The Estimated Cost Impact of Privatizing Student Transportation in Minnesota School Districts.” *Public Choice* 146, no.3–4 (March): 319–339. https://www.jstor.org/stable/41483632.

Tillman, Wanda R. and C. Justice Tillman. 2008. “And You Thought It Was the Apple: A Study of Job Satisfaction Among Teachers.” *Academy of Educational Leadership Journal*, 12(3): 1–18. Retrieved from https://www.abacademies.org/articles/aeljvol12no32008.pdf.

Texas Smart Schools. 2019. *Texas Smart Schools Methodology Technical Description*. Retrieved from https://txsmartschools.tamu.edu/pdf/Methodology/TXSmartSchools%20Methodology%20Technical%20Description%202019.pdf

Vincent, Jeffrey M., Carrie Makarewicz, Ruth Miller, Julia Ehrman and Deborah L. McKoy. 2014. *Beyond the Yellow Bus: Promising Practices for Maximizing Access to Opportunity Through Innovations in Student Transportation*. Berkeley, CA: Center for Cities + Schools, University of California. Retrieved from https://citiesandschools.berkeley.edu/uploads/CC+SYellowBus2014.pdf.

Viton, Philip A. 1981. “A Translog Cost Function for Urban Bus Transport.” *The Journal of Industrial Economics* 29, no. 3 (March):287–304. https://www.jstor.org/stable/2098305

Williams, Martin, and Ardeshir Dalal. 1981. “Estimation of the Elasticity of Factor Substitution in Urban Bus Transportation: A Cost Function Approach.” *Journal of Regional Science* 21 (2): 263–275. doi: 10.1111/j.1467-9787.1981.tb00698.x.

Williams, Martin and Carol Hall. 1981. “Returns to Scale in the United States Intercity Bus Industry.” *Regional Science and Urban Economics* 11, no. 4 (November): 573–584. doi: 10.1016/0166-0462(81)90039-9.

Windle, Robert J. 1984. “The Cost Structure of the U.S. Bus Transit Industry.” PhD Dissertation, University of Wisconsin.

Windle, Robert J. 1988. "Transit Policy and the Cost Structure of Urban Bus Transportation." In *Bus Deregulation and Privatization*, edited by John S. Dodgson and Neville Topham, 119–140. Aldershot: Avebury.

Zimmer, Timothy, Larry DeBoer, and Marilyn Hirth. 2009. "Examining Economies of Scale in School Consolidation: Assessment of Indiana School Districts." *Journal of Education Finance* 35, no. 2 (Fall): 103–127. doi: 10.1353/jef.0.0012.

Appendix A: Sources for Cost Adjustment Strategies for Other States

Alaska:

- Alaska Department of Education & Early Development (2019, January). Public School Funding Program Overview. https://education.alaska.gov/SchoolFinance/pdf/Funding-Overview.pdf
- Review of Alaska's School Funding Formula (2015, July). Prepared for the Alaska State Legislature by Augenblick, Palaich and Associates. https://lba.akleg.gov/download/publications/school2015.pdf

Colorado:

- Colorado Department of Education (2019, July). Financial Policies and Procedures Handbook. https://www.cde.state.co.us/cdefinance/fpphandbookfy19-20

Florida:

- Florida Department of Education. 2019–20 Funding for Florida Schools. http://www.fldoe.org/core/fileparse.php/7507/urlt/Fefpdist.pdf
- Dewey, J. (2019). 2018 Florida Price Level Index. https://www.bdb.org/clientuploads/Research/0-2020_Data/Consumer%20Price%20Index/2018fpli.pdf

Maine:

- Maine Department of Education (2018, July). Maine School Funding Model. https://www.maine.gov/doe/sites/maine.gov.doe/files/inline-files/History_Overview_School_Funding_Model_July2018_revMay2019_JLA_1.pdf
- Sloan, James E. and Amy Johnson (2019, November). Review of Geographic Cost Adjustment Component in the Essential Programs and Services Model. Maine Education Policy Research Institute, University of Southern Maine Office. https://www.maine.gov/doe/sites/maine.gov.doe/files/inline-files/EPS%20Regional%20Adjustment_11.27.19.pdf

Maryland:

- Maryland State Department of Education (2018, January). http://www.marylandpublicschools.org/about/Documents/DBS/BudgetRes/2018/FY19StateAidPreliminaryDraftCalculations011918.pdf
- Geographic Cost of Education Adjustment for Maryland (2015, November). Prepared for The Maryland State Department of Education by Jennifer Imazeki, Picus Odden & Associates. http://marylandpublicschools.org/Documents/adequacystudy/APA-POA-GCEI-Report-Rev-11232015.pdf

Massachusetts:

- Massachusetts Department of Elementary and Secondary Education (2018, September). School Finance: Chapter 70 Program, FY19 Chapter 70 Aid and Required Contribution

Calculations. http://www.doe.mass.edu/finance/chapter70/fy2019/chapter-19-whitepaper.docx

Missouri:

- Missouri Department of Elementary & Secondary Education. Dollar Value Modifier by Year. https://dese.mo.gov/sites/default/files/sf-Complete-DVM-List.pdf

New Jersey:

- State of New Jersey Department of Education. Geographic Cost Adjustment (GCA) Update FY2014. https://www.nj.gov/education/sff/gca2014.pdf found at https://www.nj.gov/education/sff/

New York:

- New York State Education Department. 2018–19 State Aid Handbook. http://www.oms.nysed.gov/faru/PDFDocuments/Primer2018-19.pdf found at https://stateaid.nysed.gov/generalinfo/

Virginia:

- Virginia Department of Education (2018, September). Overview of Virginia K–12 Funding Formulas and Formula Approaches to Recognize Student Need. http://www.doe.virginia.gov/boe/meetings/2018/work-session/09-sep/overview-of-virginia-k12-funding-formulas.pptx
- Lou, Cary and Kristin Blagg (2018, December). School District Funding in Virginia: Computing the Effects of Changes to the Standards of Quality Funding Formula. Urban Institute. https://www.urban.org/research/publication/school-district-funding-virginia

Washington:

- Washington State Legislature, Senate Ways and Means Committee. 2020 Citizen's Guide to K–12 Finance. http://leg.wa.gov/LIC/Documents/EducationAndInformation/Citizens%20Guide%20to%20K-12%20Finance.pdf
- Washington Office of Superintendent of Public Instruction (2017). EHB 2242 Guidance. https://www.k12.wa.us/policy-funding/school-apportionment/instructions-and-tools/ehb-2242-guidance

Wyoming:

- Wyoming Legislature (2020, March). State of Wyoming School Foundation Program Flow Chart. https://wyoleg.gov/docs/SchoolFinance/SchoolFoundationBlockGrantFlowChart.pdf

Appendix B: Variable Definitions and Sources

Variable Name	Definition and Source
Teacher Experience	Years of teaching experience (PEIMS)
First Year Teacher	A variable that takes on the value of 1 if the teacher has zero years of teaching experience, and zero otherwise (PEIMS)
No Degree	A variable that takes on the value of 1 if the individual does not have at least a bachelor's degree, and zero otherwise (PEIMS)
Master's Degree	A variable that takes on the value of 1 if the individual holds a master's degree, and zero otherwise (PEIMS)
PhD or EDD	A variable that takes on the value of 1 if the individual holds a doctoral degree, and zero otherwise (PEIMS)
Assigned Multiple Campuses	A variable that takes on the value of 1 if the individual teacher was assigned to work at multiple campuses, and zero otherwise (PEIMS)
New Hire Indicator	A variable that takes on the value of 1 if it is the teacher's first year of service in the district, and zero otherwise (PEIMS)
Subject Matter Assignment	A variable that takes on the value of 1 for the subject taught by the teacher (e.g., elementary, math, science), and zero otherwise (PEIMS)
Grade-level Assignment	A variable that takes on the value of 1 for the grade-level taught by the teacher (e.g., elementary, pre-K), and zero otherwise (PEIMS)
Campus Type	A variable that takes on the value of 1 for school type (e.g., elementary, middle), and zero otherwise (TAPR)
Department Head	A variable that takes on the value of 1 if a person is a department head, and zero otherwise (PEIMS)
Administrator	A variable that takes on the value of 1 if a person is a school or district administrator, and zero otherwise (PEIMS)
Support Staff	A variable that takes on the value of 1 if a person is support staff, and zero otherwise (PEIMS)
Percentage EverELL	Percentage of students in the school ever considered Limited English proficient or English learners (PEIMS data housed at the Educational Research Center at the University of Texas at Dallas)

Variable Name	Definition and Source
Social Security for All	A variable that takes on the value of 1 if a district's teachers participate in the social security system, and zero otherwise (Texas Classroom Teachers Association and authors' analysis of PEIMS financial data)
Social Security for Some or All	A variable that takes on the value of 1 if at least some of the district's employees participate in the social security system, and zero otherwise (Texas Classroom Teachers Association and authors' analysis of PEIMS financial data)
ACS-CWIFT	Prevailing wage for college graduates in the county (NCES)
Fair Market Rent Index	Fair market rent for a two-bedroom apartment in the county, as a percentage of the state average fair market rent (US Department of Housing and Urban Development)
Midsized District	A variable that takes on the value of 1 if district enrollment is between 1,600 and 5,000 students, and zero otherwise (TAPR)
Large District	A variable that takes on the value of 1 if district enrollment is greater than 5,000 students, and zero otherwise (TAPR)
Distance to Nearest EPP	Distance in miles from campus to the nearest accredited educator preparation program (EPP) (Authors' calculation from data provided by NCES and TEA)
Distance to Nearest Metropolitan Area	Distance in miles from campus to the center of the nearest metropolitan area (Authors' calculation from data provided by NCES and the US Census Bureau)
Central Metropolitan County	A variable that takes on the value of 1 if at least half of the population resides within areas of 10,000 or more population, and zero otherwise (US Census Bureau)
Outlying Metropolitan County	A variable that takes on the value of 1 if the county meets Census Bureau requirements of commuting to or from central counties, and zero otherwise (US Census Bureau)
Micropolitan County	A variable that takes on the value of 1 if a county has a central city with a population of between 10,000 and 50,000 residents, and zero otherwise (US Census Bureau)

Variable Name	Definition and Source
Rural County	A variable that takes on the value of 1 if a county is not considered a metropolitan or micropolitan area, and zero otherwise (US Census Bureau)
Sparsely Populated County	A variable that takes on the value of 1 if county population per square mile in 2020 was less than or equal to 15 and greater than 10, and zero otherwise (US Census Bureau)
Very Sparsely Populated county	A variable that takes on the value of 1 if county population per square mile in 2010 was less than or equal to 10, and zero otherwise (US census Bureau)
Unemployment rate	County unemployment rate (US Bureau of Labor Statistics)
Cooling Degree Days	The average number of cooling degree days per year during the 30-year period from 1981–2010 at the three weather stations that are closest to campus (authors' calculations from data provided by NCES and NOAA)
Heating Degree Days	The average number of heating degree days per year during the 30-year period from 1981–2010 at the three weather stations that are closest to campus (authors' calculations from data provided by NCES and NOAA)
Female	A variable that takes on the value of 1 if the individual is female, and zero otherwise (PEIMS)
Percent Day	The percentage of a standard district work day for which the employee is hired to work (PEIMS)
Days Employed	The actual number of at-work days within the school year that a person is scheduled to work in the district. This number does not include holidays, weekends, and any other days the employee is not scheduled to work (PEIMS)
Square miles > 400	A variable that takes on the value of 1 if the district has more than 400 square miles, and zero otherwise (TEA)
K–8 District	A variable that takes on the value of 1 if district does not serve high school grades, and zero otherwise (TAPR)
HS-CWI	The High School Comparable Wage Index (Texas Smart Schools)
Number of Potential Employers	A pupil-weighted average of the number of employers in the zip codes where campuses are located (authors' calculations from US Census data)

Variable Name	Definition and Source
Per-pupil operating expenditure	Actual current, per-pupil operating expenditures (PEIMS object codes 6100 through 6499), excluding food and student transportation expenditures, and adjusted for share services agreements.
Campus enrollment	Total number of students enrolled at the school (TAPR)
Average Conditional NCE	A normalized gain score indicator of student performance on the State of Texas Assessments of Academic Readiness (STAAR®) Grades 3–8 and end-of-course (EOC) exams (Authors' calculations from PEIMS data on individual students)
Teacher Cost Index	The predicted salary in the district for a teacher with a standard set of characteristics who was assigned to a standard campus, divided by a minimum predicted salary for that year (Authors' calculations from multiple data sources)
Auxiliary Personnel Cost Index	The predicted salary in the district for an auxiliary worker with a standard set of characteristics who was assigned to a standard campus, divided by a minimum predicted salary for that year (Authors' calculations from multiple data sources)
District enrollment	Total number of students enrolled in the district (TAPR)
% Economically disadvantaged	The percentage of students eligible for free or reduced-price lunch at a campus (TAPR)
% Special education	The percentage of students eligible for special education services at a campus (TAPR)
% High needs special education	The percentage of students with a special education classification other than speech-language difficulties or learning disabilities (Authors' calculations from PEIMS data)
First Tier Coastal County	A variable that takes on a value of 1 if a county is Aransas; Brazoria; Calhoun; Cameron; Chambers; Galveston; Jefferson; Kenedy; Kleberg; Matagorda; Nueces; Refugio; San Patricio; or Willacy County, and zero otherwise (Texas Department of Insurance)
Herfindahl Index	The sum of squared enrollment shares for metropolitan areas, micropolitan areas or rural counties (authors' calculations from TAPR data)
Share of spending unallocated	The share of current operating expenditures in a campus that were imputed on a per capita basis (Authors' calculations from PEIMS data)

Variable Name	Definition and Source
Number with test scores	Number of students who's test scores in reading or mathematics contribute to the calculation of the Conditional NCE score for a campus (Authors' calculations from PEIMS data)
Manufacturing establishments in zip	Number of establishments engaged in the mechanical, physical, or chemical transformation of materials into new products (U.S. Census Bureau)
Percent of county in places	Percentage of residents in a county residing in a Census-designated place in 2010 (U.S. Census Bureau)
Square Miles	Square miles in the school district (TEA)
Total Transportation Expenditures	A district's total reported current operating expenditures (PEIMS object codes 6100-6499) under PEIMS Function 34 (PEIMS)
Expenditures per Mile	A district's total expenditures under PEIMS Function 34 divided by their total route services mileage (Authors' calculations; PEIMS and Texas Education Agency Transportation Operations Report)
Riders per Mile	A district's total average daily ridership divided by their annual total route services mileage (Authors' calculations; Texas Education Agency Transportation Operations Report and Transportation Route Services Report)
Total Route Miles	The sum of a district's regular route services mileage and special route services mileage (Texas Education Agency Transportation Operations Report)
Diesel Price	Price of diesel in the county (Oil Price Information Services)
Population Density	A county's population in 2010 divided by its total land area (U.S. Census Bureau)
Congestion	Vehicle miles travelled per lane mile (Texas Transportation Institute)
Number of Special Riders	District special program average daily ridership (Texas Education Agency Transportation Route Services Report)
Number of Total Riders	District total of special program average daily ridership and regular program average daily ridership (Texas Education Agency Transportation Route Services Report)

Variable Name	Definition and Source
Percent Special Riders	Percentage of total district ridership that is special program riders (Authors' calculations; Texas Education Agency Transportation Route Services Report)
Number of Special Miles	District total special program route related services mileage (Texas Education Agency Transportation Operations Report)
Number of Regular Miles	District total regular program route related services mileage (Texas Education Agency Transportation Operations Report)
Percent Special Miles of Total Miles	Percentage of total district route mileage that is special route miles (Authors' calculations; Texas Education Agency Transportation Operations Report)
Total Vehicles	District total reported vehicles (Texas Education Agency Transportation Operations Report)
Number of Regular Service Buses Less than 5 Years Old	Number of district buses used for regular route services that are 0-5 years old (Texas Education Agency Transportation Operations Report)
Number of Special Service Buses Less than 5 Years Old	Number of district buses used for special route services that are 0-5 years old (Texas Education Agency Transportation Operations Report)
Number of Total Buses Less than 5 Years Old	Total buses within a school district that are 0-5 years old (Texas Education Agency Transportation Operations Report)
Percent of Buses Less than 5 Years Old	Percentage of total district buses that are 0-5 years old (Authors' calculations; Texas Education Agency Transportation Operations Report)

Appendix C: Coefficient Estimates and Robust Standard Errors for the Teacher Salary Model

Variables	Model 1	Std. Errors	Model 2	Std. Errors
Years of experience (log)	0.0544***	(0.00342)	0.0659***	(0.00130)
Years of experience (log), sq.	-0.0184***	(0.00204)	-0.0332***	(0.000736)
Years of experience (log), cubed	0.0148***	(0.000646)	0.0116***	(0.000125)
First year teacher	0.00948***	(0.00130)	0.0215***	(0.000670)
No degree	0.00767***	(0.00106)	-0.000687	(0.000541)
Master’s degree	0.0205***	(0.000411)	0.0241***	(0.000200)
PhD or EdD	0.0255***	(0.00232)	0.0309***	(0.00107)
New hire	-0.00430***	(0.000189)	-0.00547***	(0.000124)
Assigned multiple campuses	-0.00168***	(0.000481)	-0.000610**	(0.000271)
Elementary subjects	0.00386***	(0.000265)	0.00339***	(0.000202)
Language arts teacher	-0.000958***	(0.000226)	-0.00468***	(0.000153)
Math teacher	4.88e-06	(0.000260)	-0.000481***	(0.000168)
Science teacher	0.000151	(0.000256)	-0.000699***	(0.000170)
Social studies teacher	0.000419*	(0.000220)	0.00207***	(0.000153)
Health/PE teacher	0.0132***	(0.000381)	0.0221***	(0.000186)
Foreign language teacher	-0.00367***	(0.000510)	-0.00475***	(0.000332)
Fine arts teacher	0.00224***	(0.000441)	0.00559***	(0.000241)
Computer teacher	-0.00455***	(0.000630)	-0.0112***	(0.000376)
Vocational/technical teacher	-0.00287***	(0.000559)	-0.00711***	(0.000292)
Special subject teacher	0.00323***	(0.000501)	0.00156***	(0.000340)
Tested grade or subject teacher	-0.000465**	(0.000227)	-0.00272***	(0.000149)
Assigned non-graded students	-0.00107***	(0.000275)	-0.000604***	(0.000172)
Assigned elementary students	-0.00587***	(0.000450)	-0.00925***	(0.000326)
Assigned secondary students	0.0239***	(0.00880)	0.0181***	(0.00525)
Assigned pre-K students	0.000274	(0.000636)	0.000214	(0.000455)
Assigned kindergarten students	0.00288***	(0.000469)	0.00459***	(0.000350)
Elementary school campus	0.00982***	(0.00200)	0.0122***	(0.000730)
Middle school campus	0.0198***	(0.00197)	0.0281***	(0.000726)
High school campus	0.0227***	(0.00198)	0.0337***	(0.000730)
Large High School Campus	0.0333***	(0.00196)	0.0494***	(0.000731)
Department head	0.0126***	(0.00155)	0.0156***	(0.00122)

Variables	Model 1	Std. Errors	Model 2	Std. Errors
Administrator	0.0413***	(0.00880)	0.0856***	(0.00245)
Support staff	0.00579***	(0.00223)	0.00875***	(0.00141)
Social Security for All	0.0520***	(0.00167)	0.0364***	(0.000506)
Percentage EverELL students	0.0560***	(0.00200)	0.0893***	(0.000621)
Unemployment rate	-0.00474***	(0.000175)	-0.00534***	(9.26e-05)
Fair Market Rent Index	0.236***	(0.0138)	0.613***	(0.00748)
ACS-CWIFT	0.591***	(0.0145)	1.069***	(0.00798)
ACS-CWIFT*Rent Index	-0.252***	(0.0142)	-0.653***	(0.00788)
Distance to the center of the nearest metropolitan area (log)	-0.00582***	(0.00140)	-0.00833***	(0.000533)
Distance to metro, squared	0.00334***	(0.000341)	0.00403***	(0.000120)
Distance to nearest EPP	0.0167***	(0.000899)	0.0243***	(0.000350)
Distance to EPP, squared	-0.00623***	(0.000271)	-0.00772***	(9.44e-05)
Cooling Degree days (log)	0.0824***	(0.00321)	0.0602***	(0.000684)
2014–2015	-0.0615***	(0.00232)	-0.111***	(0.000226)
2015–2016	-0.0532***	(0.00185)	-0.0938***	(0.000165)
2016–2017	-0.0467***	(0.00140)	-0.0775***	(0.000173)
2017–2018	-0.0414***	(0.000930)	-0.0619***	(0.000135)
2018–2019	-0.0376***	(0.000469)	-0.0477***	(0.000104)
Central metropolitan County	0.00993***	(0.00103)	0.0109***	(0.000319)
Micropolitan county	0.0149***	(0.00167)	0.0178***	(0.000502)
Sparsely populated county	-0.0581***	(0.00675)	-0.0606***	(0.00185)
Very Sparsely populated county	-0.0503***	(0.00449)	-0.0610***	(0.00126)
Rural: Years of experience (log)	-0.162***	(0.00745)	-0.218***	(0.00358)
Rural: Years of experience (log), sq.	0.110***	(0.00439)	0.147***	(0.00197)
Rural: Years of experience (log), cubed	-0.0180***	(0.000791)	-0.0244***	(0.000327)
Rural: First year teacher	-0.0748***	(0.00358)	-0.102***	(0.00187)
Rural: No degree	-0.0110**	(0.00449)	-0.0128***	(0.00185)
Rural: Master's degree	-0.00685***	(0.00151)	-0.00704***	(0.000566)
Rural: PhD or EdD	-0.00884	(0.0119)	-0.00538	(0.00406)
Rural: New hire	0.00196***	(0.000587)	0.00656***	(0.000328)
Rural: Assigned multiple campuses	0.00219*	(0.00117)	0.00375***	(0.000583)
Rural: Elementary subjects	-0.00411***	(0.000836)	-0.00416***	(0.000531)

Variables	Model 1	Std. Errors	Model 2	Std. Errors
Rural: Language arts teacher	-0.00306***	(0.000724)	-0.00376***	(0.000382)
Rural: Math teacher	-0.00155*	(0.000840)	-0.00199***	(0.000417)
Rural: Science teacher	-0.000217	(0.000836)	-0.00112***	(0.000428)
Rural: Social studies teacher	-0.000346	(0.000732)	-0.000565	(0.000385)
Rural: Health/PE teacher	0.0104***	(0.00120)	0.0150***	(0.000492)
Rural: Foreign language teacher	0.00340	(0.00208)	0.00100	(0.000992)
Rural: Fine arts teacher	-0.00639***	(0.00119)	-0.00864***	(0.000570)
Rural: Computer teacher	-0.000540	(0.00139)	0.00398***	(0.000817)
Rural: Vocational/technical teacher	-0.000199	(0.00132)	0.00162***	(0.000601)
Rural: Special subject teacher	-0.00501***	(0.00177)	-0.00425***	(0.000981)
Rural: Tested grade or subject teacher	-0.000279	(0.000801)	0.000529	(0.000392)
Rural: Assigned non-graded students	0.00356***	(0.000777)	0.00405***	(0.000418)
Rural: Assigned elementary students	-0.000647	(0.00145)	0.00140*	(0.000769)
Rural: Assigned secondary students	-0.00214	(0.0195)	-0.00237	(0.00991)
Rural: Assigned pre-K students	0.000827	(0.00196)	-0.00454***	(0.00110)
Rural: Assigned kindergarten students	-0.000419	(0.00157)	-0.00372***	(0.000884)
Rural: Elementary school campus	-0.00527*	(0.00276)	-0.00640***	(0.00100)
Rural: Middle school campus	-0.000756	(0.00284)	0.000936	(0.00103)
Rural: High school campus	0.000829	(0.00281)	0.00450***	(0.00101)
Rural: Large High School Campus	-0.00928**	(0.00401)	-0.00537***	(0.00186)
Rural: Department head	0.00998	(0.0124)	0.00994	(0.00681)
Rural: Administrator	0.00762	(0.0128)	0.00100	(0.00384)
Rural: Support staff	0.00172	(0.00689)	-0.00264	(0.00325)
Rural: Social Security	-0.0189***	(0.00656)	-0.00484**	(0.00196)
Rural: Percentage EverELL students	0.0661***	(0.00536)	0.0377***	(0.00153)
Rural: Unemployment rate	0.00410***	(0.000286)	0.00329***	(0.000139)
Rural: Rent Index	-0.156***	(0.0434)	-0.249***	(0.0253)
Rural: CWIFT	-0.551***	(0.0408)	-0.804***	(0.0240)
Rural: CWIFT*Rent Index	0.236***	(0.0507)	0.367***	(0.0299)
Rural: Distance to the center of the nearest metropolitan area (log)	0.169***	(0.0237)	0.142***	(0.00696)
Rural: Distance to Metro, squared	-0.0282***	(0.00313)	-0.0243***	(0.000911)

Variables	Model 1	Std. Errors	Model 2	Std. Errors
Rural: Distance to nearest EPP	-0.0675***	(0.00298)	-0.0791***	(0.000967)
Rural: Distance to EPP, squared	0.0179***	(0.000690)	0.0199***	(0.000214)
Rural: Cooling Degree days (log)	-0.0161***	(0.00406)	0.00727***	(0.00116)
Rural: 2014–2015	0.275***	(0.0655)	0.366***	(0.0255)
Rural: 2015–2016	0.274***	(0.0655)	0.365***	(0.0255)
Rural: 2016–2017	0.271***	(0.0655)	0.361***	(0.0255)
Rural: 2017–2018	0.269***	(0.0655)	0.358***	(0.0255)
Rural: 2018–2019	0.273***	(0.0655)	0.359***	(0.0255)
Rural: 2019–2020	0.312***	(0.0655)	0.396***	(0.0255)
Rural: Sparsely populated county	0.0595***	(0.00706)	0.0655***	(0.00195)
Rural: Very Sparsely populated county	0.0486***	(0.00501)	0.0615***	(0.00143)
Constant	9.458***	(0.0300)	9.297***	(0.00867)
Estimation technique	Teacher Fixed effects		AR Teacher Random effects	
Wooldridge test for autocorrelation F(1, 314545)			9536.594	
Number of Observations	1,809,066		1,809,066	
Number of Individuals	441,671		441,671	

Note: All variables labeled “Rural: x” represent the coefficients on the interaction between the non-metropolitan county indicator and the designated variable.

Appendix D: Coefficient Estimates and Robust Standard Errors for the Auxiliary Personnel Wage Model

Variables	Coefficients	Standard Errors
Years of experience (log)	-13.57***	(0.886)
Years of experience (log), sq.	4.176***	(0.241)
Years of experience (log), cubed	-0.415***	(0.0217)
Female	-6.075***	(1.379)
Female * years of experience	6.237***	(1.127)
Female * years of experience (log), sq.	-2.066***	(0.305)
Female * years of experience (log), cu.	0.220***	(0.0275)
Advanced Degree	0.0382***	(0.00212)
New hire	-0.0428***	(0.000635)
Percent Day	-0.00234***	(3.57e-05)
Days Employed (log)	0.170***	(0.00333)
Business/Finance Clerical	0.327***	(0.00292)
Campus Office/Clerical	0.131***	(0.00168)
Central Office/Clerical	0.257***	(0.00207)
Child Nutrition	-0.145***	(0.00157)
Human Resources	0.353***	(0.00442)
Information Technology Technicians	0.394***	(0.00304)
Campus Technology Specialist	0.287***	(0.00497)
Custodial	-0.182***	(0.00178)
Maintenance	0.0674***	(0.00213)
Plumber	0.246***	(0.00728)
Painter	0.117***	(0.00770)
HVAC	0.244***	(0.00483)
Electrician	0.263***	(0.00642)
Warehouse	0.0623***	(0.00470)
Safety/Security	0.205***	(0.00290)
Transportation	-	
Square Miles > 400	-0.0267***	(0.00170)
Large District	-0.0390***	(0.00161)

Variables	Coefficients	Standard Errors
Social Security for Some or All	0.0680***	(0.00201)
K–8 District	0.109***	(0.0118)
Unemployment rate	-0.00856***	(0.000516)
Rent Index	-0.501***	(0.0465)
HS-CWI	-0.159***	(0.0421)
HS-CWI*Rent Index	0.605***	(0.0461)
Distance to the center of the nearest metro area (log)	-0.0209***	(0.000949)
Heating Degree days (log)	0.0317***	(0.00371)
Cooling Degree days (log)	0.131***	(0.00722)
Number of potential employers (log)	0.00694***	(0.000535)
2017–2018	-0.0749***	(0.000558)
2018–2019	-0.0565***	(0.000373)
Metropolitan County	-0.0284***	(0.00313)
Central metropolitan County	0.0175***	(0.00188)
Micropolitan county	0.0119***	(0.00295)
Sparsely populated county	0.0247***	(0.00357)
Very Sparsely populated county	0.125***	(0.00407)
Constant	17.14***	(1.086)
Number of Observations	510,341	
Number of Individuals	223,509	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix E: A General Overview of the Cost Function Methodology

The idea of a cost function is widely used in the education literature, and has been discussed in Gronberg, Jansen and Taylor (2011a) and in Taylor et al (2014, 2017). A cost function is a function relating a measure of cost to the prices of various inputs (input prices) and to other specific factors that impact cost (environmental factors). In a market economy, productively efficient firms will minimize costs consistent with each level of output produced, given market prices for inputs and the measures of other environmental factors that impact cost. Efficient non-profit organizations will also minimize costs consistent with the output produced for given input prices, and again conditional on the environmental factors that impact their costs.

Two important concepts are the short run cost function and the long run cost function. The short run cost function applies when the period of time is too short for the firm or nonprofit organization to adjust its capital stock. The long run cost function applies with the time period is long enough for all changes to occur in the desired level of the capital stock (and all other inputs). In this study we modeled the short run cost function, the cost function that considers the capital stock to be relatively fixed.

In economics the concept of a production function is relatively straightforward and posits a model linking the use of inputs (e.g., labor and capital) with the production of output. Again, there can be a short run and a long run version of the production function. The cost function is built on the concept of a production function linking outputs to inputs, along with the idea that an organization must pay market prices for inputs, with the price and quantity of inputs determining the organizations cost of producing output. To produce more output requires more inputs and leads to higher costs. Organizations that are efficient will strive to produce every level of output at the lowest possible cost. The cost function specifies this relationship between a left hand side variable, cost, and a set of right hand side variables that include output levels, input price levels, and measures of the relevant environmental variables. An item of interest is the impact of changes in these right hand side or explanatory variables on the left hand side variable, cost. For instance, the impact on cost of rising output levels is often of interest, as this is related to the concept of economies of scale. A firm or organization with positive economies of scale will see its per-unit cost declining as output increases. Typically economists posit a U-shaped relationship between the level of output and per-unit costs, with per-unit costs falling as output increases up to some point, after which per-unit costs remain constant or even increase with further increases in output.

Firms minimize cost, but they are also hypothesized to maximize profits. Nonprofit organizations do not maximize profits by definition, but if efficient they would still seek to minimize costs. In the public sector the absence of profit-maximizing incentives makes greater the possibility of cost inefficiency, and this must be addressed in any analysis of educational cost.

The issue of inefficiency can be addressed by using what is called a stochastic frontier cost function. In this model a cost function is modified to include a regression error term that contains two parts. One is a standard two-sided error term that is included in any regression model to capture random positive or negative 'shocks' or disturbances that make the relationship between the left hand side variable, cost, not perfectly described by the left hand side function of output levels, input prices, and environmental variables. The second is a one-sided error that can only be positive, and this one-sided error moves the left hand side variable measuring cost away from the value predicted using the cost function in a positive direction. The idea is that the cost function – including the traditional two-sided error – is the measure of cost when an organization is operating efficiently, and the one-sided error is the additional 'cost' or, better, the additional spending over and above the efficient level, that an organization is doing. This additional spending is inefficient. The stochastic cost frontier approach allows the data to reveal the degree of cost inefficiency while identifying properties of the true cost function.

Another important issue is the functional form of the cost function itself, including the two error terms. For the cost function, the 'translog' specification is a flexible functional form that can serve as a second order approximation to an arbitrary function. The word translog stands for "transcendental logarithmic' and the name refers to the logarithmic function which is one of the transcendental functions in mathematics. The translog is popular because of its flexibility – elasticities (including elasticity of scale) can change with output and with factor proportions.

The statistical distribution of the two error terms in the stochastic frontier cost function must also be specified. It is common in regression models to assume the two-sided error term has a normal distribution (or possibly a closely related distribution). The one-sided error, the real difference with the stochastic frontier model, has been developed for a series of one-sided distributions, and here we use the distribution that is STATA's default, the exponential distribution.

When properly specified and estimated using stochastic frontier analysis (SFA), the educational cost function is a theoretically and statistically reliable method for estimating the relationship between cost and measures of scale. These have often been used to measure scale economies for the provision of educational services. This analysis uses SFA to estimate an educational cost function for Texas. It also uses SFA to study the provision of school transportation services by school districts

The standard stochastic frontier model starts with a cost function. A cost function – a cost frontier – specifies the minimum cost necessary to achieve certain outcomes with specified inputs and specified environmental factors. The cost function can be written in general form as:

(1) $C = C(Z \mid \beta) \cdot \exp(\varepsilon)$

where the left hand side variable C is cost, and on the right $C(Z \mid \beta)$ is a function with a set of variables Z and a set of coefficients β. This function is the non-stochastic part of the cost function or cost frontier. The set of variables Z would include variables affecting the frontier level of cost, which includes output levels, the prices of inputs into production, and environmental factors that impact cost and production. We write this as $Z = \{y_1,...,y_n; w_1,...,w_k, z_1,...,z_m\}$, where we have n output measures, k input prices, and m environmental factors including fixed factors of production. The second term on the right hand side contains exp(ε), where ε is a random noise component representing exogenous random shocks to the relationship. This cost frontier $C(Z \mid \beta)$ is the true deterministic neo-classical cost function, the object of discovery. The error term, ε, indicates random deviations from the cost frontier due to measurement error and unforeseen random changes in cost due to factors not modeled in the cost function.

Equation (1) with a symmetric two-sided error term ε presents a general form for the standard empirical cost function, including the modeled cost frontier and the allowance for random deviations from the cost frontier.

In the stochastic frontier approach, the cost function in (1) is regarded as a frontier, a minimum cost of attaining given outputs with given inputs including environmental factors. Spending may then deviate from this cost frontier, exceeding the minimum cost specified in the cost frontier. Thus the stochastic frontier approach starts with (1) and adds the assumption that spending exceeds the cost frontier due to random errors that can be (but are not required to be) interpreted as inefficiency. The stochastic frontier approach basically takes equation (1) and assumes that the random error, ε, consists of two parts, a standard two-sided random error that can be positive or negative and on average is zero, and a one-sided error that is always positive (or at least not negative). This one-sided error captures the idea that individual decision-making units – districts or campuses in the school context – can at best be on the cost frontier, plus or minus the value specified in the two-sided random error. The one-sided error captures the idea that a fully efficient decision-making units can at best be on the cost frontier, and if they are inefficient this is captured or modelled by the one-sided error. The larger the one-sided error, the further a decision-making unit is from the frontier, and hence the more inefficient is the decision-making unit.

To model this, equation (1) is altered to specify the error term, ε, as consisting of two components, v plus u. The two-sided error is v, and the one-sided error is u. Because inefficiency increases cost above the frontier (i.e., above the minimum possible cost), values of u are zero or greater.

The stochastic frontier cost function is given as:

(2) $$E = C(Z \mid \beta) \cdot \exp(v+u) = C(Z \mid \beta) \cdot \exp(v) \cdot \exp(u)$$

where E is actual or observed expenditures and spending and $C(Z \mid \beta)$ is the cost frontier as described above. The expression $C(Z \mid \beta) \cdot \exp(v)$ is the traditional stochastic cost function, and the full expression including the one-sided error term $\exp(u)$ is the stochastic frontier model. Cost efficiency is defined as $CE = \exp(-u) \leq 1$. The distance below 1 is a measure of inefficiency.

Cost frontier estimates indicate the cost of achieving certain outcomes after controlling for output levels, input prices, environmental factors, and random two-sided disturbances. Note that if we take the natural logarithm of both sides of equation (2), we get:

(3) $$\ln(E) = \ln(C(Z \mid \beta)) + v + u.$$

It remains to say something more about the deterministic part of this cost function. This analysis estimates a version of a translog frontier cost function. The dependent variable is a measure of operating expenditures, as we are estimating a short run cost function and as such abstract from capital purchases. The explanatory variables—the right-hand-side variables—include measures of outputs, measures of input prices, and measures of environmental factors. The translog specification is:

(4)

$$
\begin{aligned}
\ln(E) &= \beta_0 + \sum_{i=1}^{n} \beta_{y,i} \ln(y_i) + \sum_{i=1}^{k} \beta_{w,i} \ln(w_i) + \sum_{i=1}^{m} \beta_{z,i} \ln(z_i) \\
&+ \frac{1}{2}\sum_{i=1}^{n} \beta_{yy,ii} (\ln(y_i))^2 + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{yy,ij} \ln(y_i)\ln(y_j) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{k} \beta_{yw,ij} \ln(y_i)\ln(w_j) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \beta_{yz,ij} \ln(y_i)\ln(z_j) \\
&+ \frac{1}{2}\sum_{i=1}^{k} \beta_{ww,ii} (\ln(w_i))^2 + \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k} \beta_{ww,ij} \ln(w_i)\ln(w_j) + \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n} \beta_{wy,ij} \ln(w_i)\ln(y_j) + \frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{m} \beta_{wz,ij} \ln(w_i)\ln(z_j) \\
&+ \frac{1}{2}\sum_{i=1}^{m} \beta_{zz,ii} (\ln(z_i))^2 + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \beta_{zz,ij} \ln(z_i)\ln(z_j) + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{k} \beta_{zw,ij} \ln(z_i)\ln(w_j) + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n} \beta_{zy,ij} \ln(z_i)\ln(y_j) \\
&+ v + u.
\end{aligned}
$$

where usual symmetry restrictions ($\beta_{ij} = \beta_{ji}$) for all i,j and for all variables r,s from the set $\{y, w, z\}$. There are times when a translog is modified to include a cubic term, especially a cubic term for an output variable. For instance, cost functions estimated for educational service providers often include a cube of (log) enrollment, since enrollment values can be quite large for some districts and quite small for other districts. Further, variables in a translog must be positive, as the log of zero and the log of a negative number are undefined. This can be dealt with in a number of ways. One common procedure is to substitute a percentage deviation of a variable from one. For values of x close to 1, ln(x) is approximately the percentage deviation of x from 1.

Equation (4) nests the popular Cobb-Douglas as a special case. If the coefficients on all of the quadratic terms are set to zero a Cobb-Douglas function is the result. Equation (4) also nests the classical (non-frontier) linear regression specification of the translog, which is obtained if the one-side error term is restricted to be identically zero. Thus, the general specification used in this analysis allows researchers to test empirically for alternative specifications sometimes found in the literature.

Following Gronberg et al. (2013) the one-sided error (u) is modeled as a function of a Herfindahl index of school district competition. This Herfindahl index is based on the enrollment shares of districts within a given county. The Herfindahl index for a perfectly competitive market with an infinity of small firms is zero; the Herfindahl index for a monopoly market with only a single firm is one. Larger values of the Herfindahl index indicate lower levels of competition.

Equation (4) is hard to interpret, but we can derive ‘marginal effects’ of changes in the explanatory variables on cost. This provides useful interpretable output from the estimation of equation (4). These marginal effects are derivatives of the dependent variable in equation (4), *ln(E)*, with respect to an explanatory variable in equation (4), such as *ln(y)* or *ln(w)*. Moreover, because we are taking derivatives in this way, the result is a percentage change in *E* with respect to a change in *y*, or with respect to a change in *w*.

For example, the marginal effect of a change in $ln(y_1)$ is calculated as:

$$(5)\ me(y_1) = \frac{d\ln(E)}{d\ln(y_1)} = \beta_{y,1} + 2\beta_{yy,11}\ln(y_1) + \sum_{i=2}^{n}\beta_{yy,1i}\ln(y_i) + \sum_{j=1}^{k}\beta_{yw,1j}\ln(w_j) + \sum_{j=1}^{m}\beta_{yz,1j}\ln(z_j)$$

Again as examples, the marginal effects of a change in $ln(w_1)$ is presented in equation (6), and the marginal effect of a change in $ln(z_1)$ is presented in equation (7):

$$(6)\ me(w_1) = \frac{d\ln(E)}{d\ln(w_1)} = \beta_{w,1} + 2\beta_{ww,11}\ln(w_1) + \sum_{i=2}^{n}\beta_{ww,1i}\ln(w_i) + \sum_{j=1}^{k}\beta_{wy,1j}\ln(y_j) + \sum_{j=1}^{m}\beta_{wz,1j}\ln(z_j)$$

$$(7)\ me(z_1) = \frac{d\ln(E)}{d\ln(z_1)} = \beta_{z,1} + 2\beta_{zz,11}\ln(z_1) + \sum_{i=2}^{n}\beta_{zz,1i}\ln(z_i) + \sum_{j=1}^{k}\beta_{zy,1j}\ln(y_j) + \sum_{j=1}^{m}\beta_{zw,1j}\ln(w_j)$$

Appendix F: The Educational Cost Function

Cost functions are widely studied in evaluating the cost of education and in evaluating school funding formulas including possible economies of scale and the cost implications of other differences across educational units. Obviously, economies of scale have been a large interest to educational funding agencies. Other interests include the cost of education difference by level of school (high schools versus elementary schools, for example), by student needs (low income students, English language learners), by geography (enrollment density or sparsity, remoteness), by teacher labor market characteristics as they influence the price of hiring a teacher, by other labor market characteristics that influence the price of hiring other workers, as well as a host of other matters. Cost functions, in principle, are well able to handle all of these issues. This analysis follows Taylor et al. (2014, 2017) by using stochastic frontier analysis (SFA) to estimate an educational cost function for Texas.

Background

This analysis draws on three key strands of scholarly literature. The first is the literature on the estimation of educational cost functions in the context of economies of scale. The second is the literature on educational cost functions in the context of measuring the impact of student need. The third is the literature on the link between competition and school district efficiency.

Cost Function Analysis and Economies of Scale in Education

Cost function analysis is particularly well suited to examinations of educational economies of scale, frequently in the context of school district consolidation. For example, Dodson and Garrett (2004) estimated a cost function for Arkansas school districts and found per-pupil cost savings of at least 19% from consolidating four small, rural districts into a county-level district. Duncombe, Miner and Ruggiero (1995) simulated the consolidation of New York school districts with fewer than 500 students and found large potential cost savings. Zimmer, DeBoer and Hirth (2009) also found large potential gains from their simulated consolidation of smaller (i.e., fewer than 1,000 pupils) districts in Indiana. Gronberg et al. (2015) simulated consolidation to the county level throughout Texas and found that consolidation would reduce per-pupil costs in many rural Texas counties, but raise per-pupil costs in most metropolitan counties. Taylor et al. (2014) simulated county-level consolidation in Texas five largest counties, and concluded that such consolidations would lead to dysfunctionally large districts and an increase in educational costs. Karakaplan and Kutlu (2019) simulated the consolidation of very small (<100 student) school districts in California, and concluded that the cost savings due to economies of scale would be more than offset by increased inefficiency due to the loss of competition.

Further evidence comes from cost function estimates of the relationship between school district size and the cost of education. Andrews, Duncombe and Yinger (2002) surveyed 10 cost studies that were published between 1985 and 1999, and concluded that per-pupil cost was very high for school districts with fewer than 500 students, lowest for school districts in the 2,000 to 5,000 student range, and somewhat higher for school districts with more than 5,000 students.

More recent cost-function analyses have reached similar conclusions about the high cost of operating small districts, but offer contradictory findings about the least-cost district configuration.

For example, Imazeki and Reschovsky (2006) found that most of the savings from economies of scale were realized by the time the district reaches 10,000 students, but that costs continued to decline with size until enrollments reached approximately 85,000. Gronberg, Jansen, and Taylor (2011a) and Eom et al. (2014) found that costs continued to decline with size for even the largest districts.

Most recently, researchers have examined economies of scale for district size while controlling for campus size. In such analyses, there are substantial economies of scale at the campus level, and district-wide economies of scale are found to be exhausted at much lower levels of total enrollment. Again, using Texas data, Gronberg, Jansen, and Taylor (2012) found that the economies of scale were fully exhausted when district enrollment reached 1,200 students, Taylor et al. (2014) found that economies of scale were fully exhausted when district enrollment reached 3,200, and Taylor et al. (2017) found that economies of scale were fully exhausted when district enrollment reached 7,700. Gronberg, Jansen, and Taylor (2017) compared alternative education campuses operated by traditional public school districts with those operated by open-enrollment charter schools, and concluded that the cost of alternative education fell with campus size but rose with district size once district size reached 570.

Cost Function Analyses and Student Need

In addition to economies of scale, researchers have also used cost function analyses to explore the additional costs associated with variations in student need. As discussed in Golebiewski (2011), Rumberger and Gandara (2008) and Baker, Taylor and Vedlitz (2008), cost function estimates of the .cost associated with student poverty gaps varied widely. Some of the studies they surveyed found that no additional funding would be needed (Downes and Pogue 1994) while other studies suggested that economically disadvantaged students require more than twice the funding of students who are not disadvantaged (Duncombe and Yinger 2005a). As a general rule, the highest cost estimates come from analyses of New York and the lowest cost estimates come from analyses of more rural states such as Arkansas, Arizona, Kansas and Texas.

Many cost-function researchers have also estimated the additional funding needed to achieve the same level of performance with English language learners (ELL) as with students who are already proficient in English. Recent reviews of the literature include Jimenez-Castellanos and Topper (2012), Golebiewski (2011) and Rumberger and Gandara (2008). They all found that the estimated range of costs is even wider for ELL students than for economically disadvantage students. For example, Duncombe and Yinger (2005b) estimated that the cost of serving an ELL student in Kansas was a statistically significant, but tiny, 0.14 percent higher than the cost of serving a student who was not ELL. At the other end of the spectrum, Duncombe and Yinger (1997) estimated that the cost of serving an ELL student in New York was four times the cost of serving a student who was not ELL. Taylor et al (2014) and Taylor, Gronberg and Jansen (2017) found that in Texas the cost of serving a student who had ever been identified as ELL was between 9 percent and 13.5 percent higher than the cost of serving a student who had never been identified as ELL.

Imazeki (2008) found in her analysis of California that the marginal cost of serving ELL student who were not Spanish speakers was four times greater than the marginal cost of serving Spanish-speaking ELL students.

More generally, a lack of economies of scale could make it more costly per student to provide bilingual education in some states or districts. A number of researchers, including Downes and Pogue (1994), Imazeki and Reschovsky (2004, 2006) and Gronberg et al. (2015) have found significant nonlinearities in the cost of serving ELL students.

A large literature has developed regarding the cost of serving special education students. Recent reviews of the literature include Golebiewski (2011) who notes that there is little consensus as to how to measure the extent of student disabilities, and even less consensus regarding the associated costs. A number of researchers have found that costs were systematically higher for students with more profound disabilities. For example, Gronberg et al. (2004) and Imazeki and Reschovsky (2004) found that the cost of serving students with speech and learning disabilities were much lower than the costs of serving other special education students although they were still significantly higher than the costs of serving students in regular education programs.

Competition and Efficiency in Education

Although the evidence is not uniform, researchers have generally found that a lack of choice among educational providers reduces the efficiency of the public school system. Much of the work has been done in Texas. For example, Grosskopf, Hayes, Taylor, and Weber (1999), Grosskopf et al. (2001), Gronberg et al. (2015) and Taylor et al. (2014) found that Texas school districts were less efficient (i.e., got less educational bang for the buck) when they were located in metropolitan areas with less choice.

Competition has also been found to effect school district efficiency in other states. Misra, Grimes, and Rogers (2012) found that elementary and secondary schools in Mississippi were more efficient in urban areas where competition from private schools was higher. Kang and Greene (2002) analyzed school districts in New York and concluded that efficiency was lower in counties with less competition. Hoxby (2003) studied school districts in Michigan and found less efficiency in school markets with less charter school competition. A recent paper by Jinnai (2014) finds a positive effect of charter school entry on student achievement in overlapping/matched grades in neighboring traditional public schools in North Carolina. Millimet and Collier (2008) and Karakaplan and Kutlu (2019) reached similar conclusions about the relationship between competition and school district efficiency in the states of Illinois and California, respectively.

Recent evidence suggests that another form of school choice—vouchers—leads to positive competitive effects. An important paper by Figlio and Hart (2014) found evidence that an increase in the competitiveness of private schools due to the introduction of a means-tested voucher program in Florida led to a modest increase in public school student performance. Figlio and Karbownik (2016) and Carr (2011) found similar results for scholarship voucher programs in Ohio, while Egalite (2016) found similar competitive effects in Louisiana.

The Estimation

The data for this analysis come from administrative files and public records of the Texas Education Agency (TEA), the Education Research Center at the University of Texas at Dallas, the National Center for Education Statistics (NCES), the US Bureau of Labor Statistics (BLS), the US Department of Housing and Urban Development (HUD), the U.S Census Bureau. The analysis covers the five-year period from 2014–15 through 2018–19.

The unit of analysis is the standard accountability campus in all traditional public school districts. Alternative Education Accountability (AEA) campuses (e.g., juvenile justice campuses, disciplinary education campuses, residential campuses and all other alternative education campuses) have been excluded because they are subject to different accountability requirements and may have different cost structures than other campuses (TEA 2014). Because they may have a different education technology that will not be available to traditional school districts, open-enrollment charter schools have also been excluded from the cost function analysis (although they are included in the measure of educational competition). Virtual campuses and campuses that lack reliable data on student performance (such as elementary education campuses that serve no students in tested grades, or very small campuses) have also been excluded.

Table F-1 provides means and standard deviations for the variables use in this analysis. Enrollment (both campus and district), the teacher cost index, the auxiliary personnel cost index and miles to the metro center enter the stochastic frontier regression in logs, while variables already in percentages and the indicator variables are not logged before entering the stochastic frontier regression.

Because school quality is generally thought of as a choice variable for school district administrators, the possible endogeneity of school quality indicators is a common concern for researchers estimating educational cost functions. (For example, see the discussion in Duncombe and Yinger (2005a, 2011); Imazeki and Reschovsky (2004); or Gronberg, Jansen, and Taylor. (2011a).) Campus size is also plausibly under the control of the school district—at least in the longer run. After all, larger school districts choose whether to have four 600 student high schools or two 1,200 student high schools. Even smaller districts choose whether to have an elementary school that serves kindergarten through sixth grade and a second school for seventh and eighth graders, or to have a larger, single school that serves kindergarten through eighth grade. This analysis follows Gronberg et al. (2015) and Gronberg, Jansen and Taylor (2017) and treats both school quality and campus size as potentially endogenous, using a control function approach.

Table F-1: Descriptive Statistics for Campuses, 2014–2015 to 2018–2019

Variables	**Mean**	**Std. Dev.**	**Minimum**	**Maximum**
Per-pupil operating expenditure	9,064	1,930	4,326	32,498
Campus enrollment	695.356	528.305	28	5,098
Average Conditional NCE	0.503	0.045	0.242	0.771
Miles to the metro center	22.948	19.443	1	190.554
Teacher Cost Index	1.234	0.095	1	1.392
Auxiliary Personnel Cost Index	1.149	0.065	1	1.371
District enrollment (log)	9.378	1.868	4.159	12.280
% Economically disadvantaged	0.605	0.259	0.000	1.000
% Ever English Language Learner	0.268	0.224	0.000	0.978
% Special education	0.090	0.030	0.000	0.292
% high needs special education	0.484	0.081	0.034	0.756
Middle School campus	0.226	0.418	0.000	1.000
High School campus	0.172	0.378	0.000	1.000
Multi-grade campus	0.033	0.177	0.000	1.000
K–8 district	0.006	0.077	0.000	1.000
Micropolitan county	0.084	0.277	0.000	1.000
Metropolitan county	0.813	0.390	0.000	1.000
Sparsely populated county	0.037	0.190	0.000	1.000
Very sparsely populated county	0.045	0.208	0.000	1.000
First Tier Coastal County	0.087	0.282	0.000	1.000
Herfindahl Index (log)	-2.019	0.884	-2.985	0.000
Share of spending unallocated	0.192	0.066	0.000	0.750
Number with test scores (log)	5.402	0.856	3.219	7.796
Manufacturing establishments in zip	18.878	21.726	0.000	378.000
Percent of county in places	0.719	0.213	0.000	0.997
Square Miles (log)	4.949	0.970	1.635	8.490
Number of observations				34,502

Note: Open-enrollment charter, virtual school, alternative education, juvenile justice and disciplinary justice campuses have been excluded, as have all campuses with fewer than 25 students for whom conditional normal curve equivalent (NCE) scores could be calculated. Sources: Academic Excellence Indicator System (AEIS) 2011–12; Texas Academic Performance Reports (TAPR) 2014–15 through 2018–19; Public Education Information Management System (PEIMS); National Center for Education Statistics (NCES).

The dependent variable used in the analysis is the log of actual current, per-pupil operating expenditures, excluding food and student transportation expenditures. As in Imazeki and Reschovsky (2006), Gronberg, Jansen, and Taylor (2011b) or Gronberg, Jansen, and Taylor (2017), food service expenditures have been excluded on the grounds that they are unlikely to be explained by the same factors that explain student performance, and therefore that they add unnecessary noise to the analysis. Transportation expenditures have been excluded on similar grounds.

All expenditures data have been adjusted to account for school districts that serve as a fiscal agent for another school district or group of districts.[45] Fiscal agents collect funds from member districts in a shared service agreement, and make purchases or pay salaries with those shared funds on behalf of the member districts. As a result, the spending of fiscal agents is artificially inflated while the spending by member districts is artificially suppressed. However, fiscal agents report annually to TEA about the amounts they spent on behalf of their member districts. These distribution data have been used to allocate spending by fiscal agents to their member districts on a proportional basis.[46]

Because not all school district expenditures are allocated to the campus level, and the share of allocated expenditures varies from district to district, researchers have distributed unallocated school district expenditures to the campuses on a per pupil basis. Thus, for example, if Little Elementary serves 20% of the students in its district, it is presumed to be responsible for 20% of the unallocated spending. While other allocation strategies are possible, this is the most common in the literature (e.g., Gronberg, Jansen, and Taylor, 2012; Grosskopf et al., 2013).

In the end, some schools still had anomalous spending patterns. The researchers excluded as unreliable any school where per-pupil expenditures exceeded $33,000 or were less than $3,500.

Outputs

As noted above, the independent variables measuring education output include both a quantity dimension of output—enrollment—and a quality dimension. Quantity is measured as the number of students in fall enrollment at the campus. The campus enrollment variable ranges from 29 to 5,098 with a mean of 696.

The quality measure captures differences in student performance. The measure is a normalized gain score indicator of student performance on the State of Texas Assessments of Academic Readiness (STAAR®) Grades 3–8 and end-of-course (EOC) exams. Although schools clearly produce unmeasured outcomes that may be uncorrelated with mathematics and reading test scores, and standardized tests may not measure the acquisition of all important higher-order skills, these are performance measures for which districts are held accountable by the state, and the most common measures of school district output in the literature (e.g., Gronberg, Jansen, and Taylor,

[45] For more on the allocation procedure, see Texas Smart Schools (2019)

[46] Due to data limitations, spending by fiscal agents could not be allocated back to specific campuses within member districts.

2011a, 2011b, 2017 or Imazeki and Reschovsky, 2006). Therefore, they are reasonable output measures for cost analysis.

STAAR® Grades 3–8 and EOC scores can be difficult to compare across grades, years or testing regimes. Therefore, this analysis relies on normalized (or equivalently, standardized) test scores. The normalization follows Reback (2008) and yields gain score measures of student performance that are not biased by typical patterns of reversion to the mean.[47]

The calculation of normalized gain scores proceeds in three steps. First, transform the scores of individual students into conditional z-scores. Denote the test scores for student (i), grade (g), and time or year (t), as S_{igt}, and measure each student's performance relative to others with same prior score in the subject as:

$$Y_{igt} = \frac{S_{igt} - E(S_{igt}|S_{i,g-1,t-1})}{[E(S_{igt}^2|S_{i,g-1,t-1}) - E((S_{igt}|S_{i,g-1,t-1})^2]^{.5}}$$

For example, consider all Grade 6 students who had a raw score of 30 on the prior year's Grade 5 STAAR®-Mathematics. For this subgroup of students with a Grade 5 score of 30, calculate the mean and standard deviations of the Grade 6 scores for STAAR®-Mathematics. The mean is the expected score in Grade 6 ($E(S_{igt}|S_{i,g-1,t-1})$) for someone with a Grade 5 score of 30; the standard deviation is the denominator in equation (1). Thus, the variable Y_{ijgt} measures individual deviations from the expected score, adjusted for the variance in those expected scores. This is a type of z-score. Transforming individual STAAR® scores into z-scores in this way allows researchers to aggregate across different grade levels, test subjects and test regimes despite the differences in the content or scaling of the various tests. It also provides a common frame of reference for incorporating the scores of students who, for example, took the STAAR®-Mathematics in Grade 7, but the Algebra 1 EOC in Grade 8.[48]

Second, calculate the average conditional z-score (i.e., the average Y_{igt}) across all required mathematics and reading tests for all of the students attending each school.[49] An average conditional z-score of 1 indicates that, on average, the students at Little Elementary scored one standard deviation above the expected score for students with their prior test performance. An average conditional z-score of -1 indicates that, on average, the students scored one standard deviation below expectations.

Finally, for ease of interpretation, transform the z-scores into conditional normal curve equivalent (NCE) scores. NCE scores (defined as 50+21.06*z) are a monotonic transformation of z-scores that are commonly used in the education literature and can be interpreted as percentile ranks.[50] A

---

[47] All students in the state, not just those in CBSAs were included in the calculation of standardized scores.

[48] Y_{igt} for this population is calculated by taking the mean and standard deviations of the Algebra 1 EOC scores among all of the students who took the Algebra 1 EOC and shared a common score on the prior year's STAAR®-Mathematics.

[49] Only students in the accountability subset (i.e., students who attended the same campus in the fall of the academic year as they did in the spring) are included in the campus average.

[50] Technically, this interpretation only holds if the scores are normally distributed. Given the large number of students tested each year in Texas, normality is a reasonable assumption.

conditional NCE score of 50 indicates that (on average) the students performed exactly as expected given their prior test performance; and a conditional NCE score of 90 indicates that (on average) they performed as well or better than 90% of their peers.

For estimation purposes, the conditional NCE scores are expressed as percentages. As Table F-1 documents, the campus-level average conditional NCE score had a mean of 0.50 with a minimum of 0.24 and a maximum of 0.77.

Input Prices

This analysis uses the new Texas TCI as the salary index for profession staff (i.e., teachers, administrators and professional staff) and the APCI as the wage index for other staff (i.e., auxiliary personnel and instructional aides). Both indices indicate that the price of labor is at least 36% higher in the highest cost locations than in the lowest cost locations. However, the correlation between the two indices is only 0.5776, suggesting that the local conditions that made a given school district particularly attractive to teachers are not always the same as the local conditions that make a district particularly attractive to auxiliary personnel.

In an ideal situation, the estimated cost function would include direct measures of local prices for instructional equipment and classroom materials. Such data are, unfortunately, not available to researchers. However, prices for pencils, paper, computers, and other instructional materials are largely set in a competitive market (and therefore unlikely to vary across schools), and prices for nonprofessional labor or building rents are largely a function of school location. Therefore, the cost analysis includes a measure of the distance from the campus to the center of the nearest metropolitan area.[51] This variable had an average value of 23 miles, a minimum of 1 mile, and a maximum of 191 miles, indicating the rather large distances sometimes involved in Texas.

Other Environmental Factors

The model includes indicators for a variety of environmental factors that influence district cost but which are not purchased inputs. A major environmental factor in this study is district enrollment. District enrollment averaged 37,294 students, with a minimum of 64 and a maximum of 215,408. Given the large range on this important environmental variable and the potential for unusually large districts to have undue influence on the estimation results, the cube of district enrollment and indicators for the two largest districts (Dallas ISD and Houston ISD) were added to the specification.

To capture variations in costs that derive from variations in student needs, the cost function included the percentages of students in each campus who were identified as special education or economically disadvantaged students. The models also included the percentage of students in each

[51] Miles to the center of the metropolitan area for each campus was calculated as-the-crow-flies using latitude and longitude information. The latitude and longitude of metro centers come from the US Census Bureau. Where available, latitude and longitude information for campuses are taken from the NCES' Common Core Database. The remaining campuses are assigned latitudes and longitudes according to the zip codes at their street address.

campus who were ever identified as English Language Learners (ELL) [52] and the percentage of special education students in the district with high needs.[53] (Following Gronberg et al. 2005, high needs special education students are those with a classification other than speech-language difficulties or learning disabilities.) The latter was measured at the district level rather than the campus level because privacy concerns led to excessive censoring when the variables were measured at the campus level.

To allow for the possibility that the education technology differs according to the grade level of the school, the cost model includes indicators for school type (middle school, high school and multi-grade schools). Fixed effects for year control for inflation and other time trends in Texas education.

The Texas Department of Insurance designates 14 Texas counties along the gulf coast as potential windstorm catastrophe areas.[54] Districts in those First Tier Coastal Counties (and in the cities of Morgan's Point, La Porte, Shoreacres, Pasadena and Seabrook) have elevated risk of damage from a hurricane or tropical storm, and therefore face higher costs to purchase insurance or self-insure. To capture geographic cost differences arising from difference in insurance risk, this analysis included an indicator for whether or not the district was in a designated catastrophe area.

The K–8 indicator takes on the value of one if the school district does not operate any high school grades, and zero otherwise. It has been included because the restricted grade range of a K–8 school district may allow it to specialize in ways not available to districts of similar size attempting to serve the full range of grades.[55]

Population density and metropolitan status are factors that constrain district choices about campus size and could influence other aspects of the educational technology. For example, districts in sparsely populated counties cannot take advantage of the school-level economies of scale available to other districts of similar size because their populations are so dispersed. Instead, such districts must operate smaller schools than other districts, which drives up costs. In addition, districts in metropolitan areas may incur costs (such as school security costs) that are not incurred by districts in other parts of the state. Therefore, this analysis included indicators for whether or not the district was located in a metropolitan or micropolitan county (as defined by the US Census Bureau) and indicators for whether or not the district was located in a sparsely or very sparsely populated

[52] Students who perform well on the English/Language Arts tests are no longer considered ELL, making the percentage ELL endogenous and introducing potential estimation problems. Therefore, each student's complete academic history was used to identify those students who have been categorized as ELL, at some point during their experience in Texas (Ever-ELL). While only 18.5% of students statewide were identified as ELL in 2015–16, nearly 30% of the students could be identified as Ever ELL.

[53] Following Gronberg et al. 2005, high needs special education students are those with a classification other than speech-language difficulties or learning disabilities. Where the share of students with speech-language difficulties or learning disabilities was censored (due to privacy concerns) the researchers presumed that all of the special education students were high needs students.

[54] The First Tier Counties are: Aransas, Brazoria, Calhoun, Cameron, Chambers, Galveston, Jefferson, Kenedy, Kleberg, Matagorda, Nueces, Refugio, San Patricio, and Willacy.

[55] The one traditional public school district that does not serve elementary grades, South Texas ISD, has been excluded from the analysis.

county.[56] (Note that these categories are not mutually exclusive: a county can be both sparsely populated and located in a metropolitan area.)

Efficiency Factors

The error terms for all frontier specifications depend on a number of factors that theory suggests may explain differences in school efficiency. Prior research has demonstrated that competition can reduce inefficiency in public education (e.g., Belfield and Levin 2002; Millimet and Collier 2008; Gronberg et al. 2015, Taylor et al. 2017). Therefore, the one-sided variance function is modeled as a function of the degree of educational competition in the district's metropolitan area, micropolitan area, or rural county.[57]

As is common in the literature, the degree of educational competition was measured with a Herfindahl index of enrollment concentration. A Herfindahl index (which is defined as the sum of the squared enrollment shares) increases as the level of enrollment concentration increases. A Herfindahl index of 1.00 indicates a metropolitan area, micropolitan area or rural county with a single local education agency (LEA); a Herfindahl index of 0.10 indicates a metropolitan area, micropolitan area, or rural county with 10 LEAs of equal size. Both traditional public school districts and open-enrollment charter schools are included in the calculation of enrollment concentration.

Heteroscedasticity in the two-sided error may also arise. To capture such a possibility, the two-sided variance is modeled as a function of the share of campus expenditures that was not specifically allocated to the campus by the district. This variable has been included because measurement error in the dependent variable (a common source of heteroscedasticity) is likely to be a function of the extent to which the dependent variable was imputed. Also included is the number of students who had a conditional NCE score. The second factor has been included because the larger the number of tested students, the smaller is the potential for measurement error in this key independent variable.

Instrumental Variables

The key to implementing the control function corrections is the identification of viable instruments for school quality and campus size. Human capital theory suggests that local labor market conditions can influence the demand for educational quality and the opportunity cost of staying in school so, as in Gronberg, Jansen and Taylor (2017), this analysis uses labor market conditions in the vicinity of the school site as instruments for the conditional NCE scores. The indicator of labor market conditions—the number of employers in the campus zip code that were manufacturers—came from the ZIP Business Patterns produced by the Census Bureau.

---

[56] A sparsely populated county has a population density of fewer than 20 persons per square mile; a very sparsely populated county has a population density of fewer than 10 persons per square mile.

[57] By assumption, the one-sided error term has an exponential distribution, which is the Stata default. Jenson (2005) finds that specifying a half-normal distribution for the inefficiency term generates more reliable estimates of technical efficiency than other assumptions about the distribution of inefficiency

This analysis uses three instruments for campus size. The first was the number of square miles in the school district. Campuses are likely to be smaller (all else equal) in districts with larger geographic footprints, where the time costs of transporting students to scale-efficient campuses could be prohibitive. The second was the share of the county population who lived in a city, town or other Census-designated place (CDP), which proved to be a particularly useful measure of population dispersion and highly correlated with school size.[58] The third was the interaction between the share of the county population who lived in a CDP and an indicator for rural counties.

Results

While the translog specification has the benefit of flexibility and generality compared to, say, the Cobb Douglas or other simple forms, the coefficient estimates from the translog specification are not readily interpretable. Most researchers present the change in cost arising from a small change in each explanatory variable, the so-called marginal effects. These marginal effects depend on the values of all the explanatory variables.[59]

Table F-2 indicates the marginal effects of a change in the various outputs, prices, and environmental variables on expenditures per pupil. For each explanatory variable, two entries are provided in each column. The first is the mean of the marginal effect of the variable in question, calculated for each data point in the sample. The second is the probability that all of the coefficients related to the variable in question (i.e., the direct effect and all interaction effects) are jointly zero.

The first column presents Model 1. It was estimated from campus level data treating NCE scores and campus enrollments as exogenous, as in Taylor et al. (2014).

The second column in Table F-2 presents a model in which campus size and student performance were both treated as endogenous using a control function correction. To implement the correction, the residual from a first stage regression of campus enrollment on the instruments and all of the exogenous explanatory variables was included as a regressor in this specification of the translog, as was the residual from a first-stage regression of the school quality measure (the average conditional NCE score) on the same set of instruments and exogenous variables.

As the table illustrates, the instruments met the necessary conditions for instrumental variables, being not only conceptually exogenous but also well correlated with campus enrollment and school quality. The first-stage F-statistics for the joint significance of the excluded instruments easily exceeded the benchmark threshold of 10.

[58] Census designated places (CDPs) are statistical geographic entities representing closely settled, unincorporated communities that are locally recognized and identified by name

[59] See Taylor et al. (2014) for details on the calculation of marginal effects.

Table F-2: Means of the Marginal Effects

Variable	Model 1	Model 2	Model 3
Campus Enrollment (log)	-0.163	-0.016	-0.035
Joint p-value	0.000	0.000	0.000
Average NCE	0.048	0.677	0.128
Joint p-value	0.000	0.001	0.000
Miles to Metro Center (log)	0.009	0.009	0.008
Joint p-value	0.000	0.000	0.000
Teacher Cost Index	0.732	0.652	0.656
Joint p-value	0.000	0.000	0.000
APCI	0.007	0.095	0.049
Joint p-value	0.000	0.000	0.000
District Enrollment (log)	0.003	-0.033	-0.028
Joint p-value	0.000	0.000	0.000
% Students Econ. Disadv.	0.125	0.240	0.181
Joint p-value	0.000	0.000	0.000
% EverELL	0.062	-0.026	0.010
Joint p-value	0.000	0.000	0.000
% Special Ed.	1.204	1.119	1.116
Joint p-value	0.000	0.000	0.000
% Special Ed. High Needs	0.144	0.126	0.121
Joint p-value	0.000	0.000	0.000
Middle School Campus	0.042	0.031	0.028
Joint p-value	0.000	0.000	0.000
High School Campus	0.206	0.127	0.132
Joint p-value	0.000	0.000	0.000
Multi-grade Campus	0.199	0.146	0.150
Joint p-value	0.000	0.000	0.000
K¬8 School district	-0.080	-0.106	-0.094
Joint p-value	0.000	0.000	0.000
Micropolitan County	-0.111	-0.139	-0.135
Joint p-value	0.000	0.000	0.000
Metropolitan County	-0.104	-0.132	-0.121
Joint p-value	0.000	0.000	0.000
Sparsely Populated County	-0.061	-0.026	-0.041
Joint p-value	0.000	0.000	0.000

Variable	Model 1	Model 2	Model 3
Very Sparsely Populated County	0.135	0.127	0.135
Joint p-value	0.000	0.000	0.000
First Tier Coastal County	-0.008	-0.016	-0.013
Joint p-value	0.000	0.000	0.000
School Size Residual		-0.148	-0.129
p-value		0.007	0.023
First Stage F-Statistic		26.090	23.535
School Quality Residual		-0.630	
p-value		0.142	
First-stage F-statistic		21.560	
One-sided error			
Herfindahl Index (log)	0.330	0.352	0.343
p-value	0.000	0.000	0000
Number of observations	34,502	34,502	34,502

Note: All models also include year fixed effects. P-values based on robust standard errors that were clustered by district and year.
Source: Authors' calculations from Table F-5.

The first-stage residuals for school quality were statistically insignificant at the 5-percent level in Model 2, but the first-stage residuals for school size were statistically significant. This suggested that test scores could safely be treated as exogenous, but school size should be treated as endogenous. Therefore, the third column in Table F-2 presents a model in which campus size was treated as endogenous but average NCE scores were treated as exogenous.[60] Because the first-stage residual for school size was statistically significant in Model 3 (which confirms the endogeneity of school size) Model 3 was the preferred specification.

The first variable listed in Table F-2 is the log of Campus Enrollment. Researchers calculated the marginal effect of an increase in campus enrollment for every sample data point and then averaged those estimates to yield the mean of the marginal effects. The table indicates that a 1% increase in campus enrollment had a mean marginal effect of -0.035, indicating that, on average, a 1% increase in campus enrollment was associated with a 0.035% decrease in cost per student.

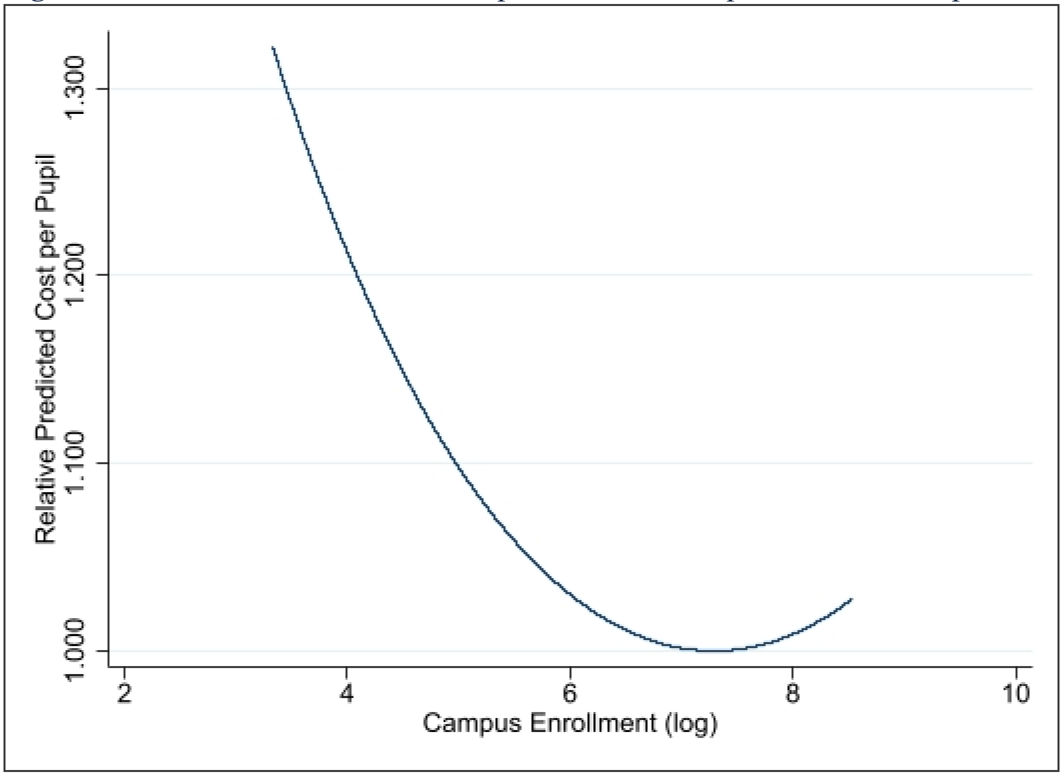
The joint p-value for the coefficients on campus enrollment and its interactions was zero to three decimal places, indicating that the coefficients on district enrollment in the cost function are jointly statistically significant at better than the 1-percent level.

Figure F-1 graphs the relationship between campus size and cost per student, holding all other variables at their sample mean values. As such, the figure illustrates the cost per pupil if every campus in a district of 11,823 students (the sample mean) had average demographics and the

[60] Thus, the residual from a first stage regression of campus enrollment on the instruments and all of the exogenous explanatory variables was included as a regressor in this specification of the translog

designated level of campus enrollment. For ease of interpretation, the predictions have been normalized so that a prediction of 1.30 indicates that costs are predicted to be 30% above the minimum predicted cost.

Figure F-1: The Estimated Relationship between Per-Pupil Cost and Campus Enrollment

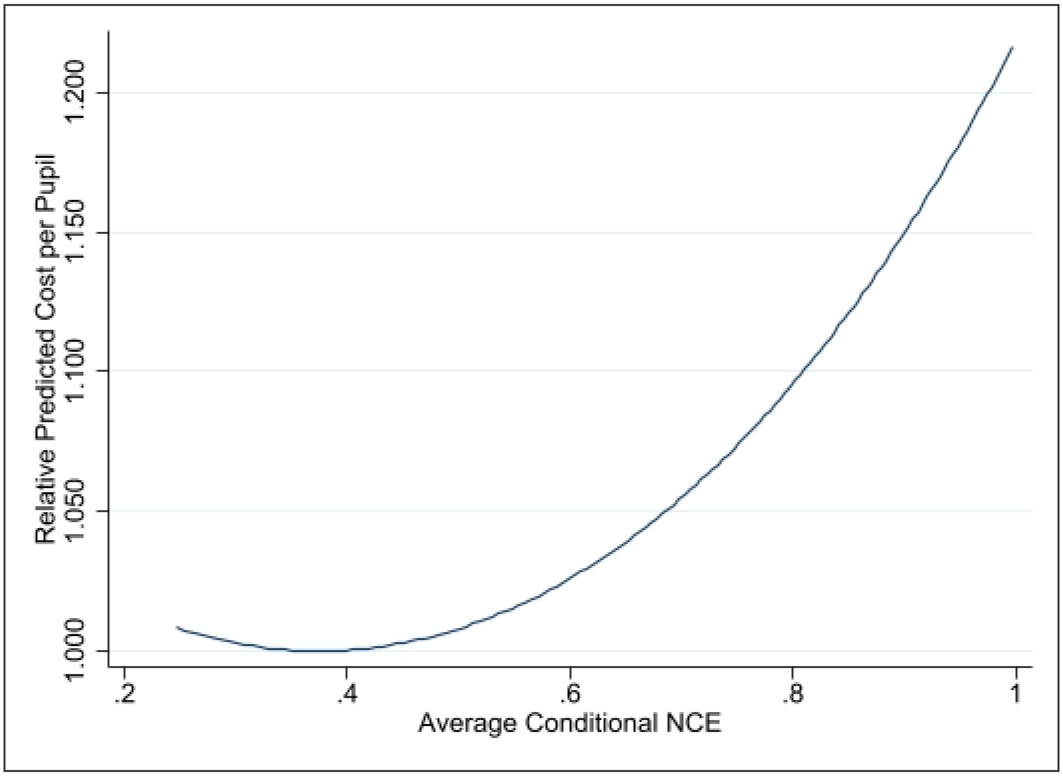


Source: Authors' calculations.

The slope of the graph is the marginal effect, and the shape of the graph in Figure F-1 indicates that, as a general rule and holding everything else constant, increases in campus size led to decreases in the cost of education. For example, the cost function indicated that all other things being equal, a 200-student campus cost 4% more to operate than a 400-student campus, which in turn costs 2.5% more to operate than an 800 student campus. Costs were minimized at a campus size of 1,500 students. However, the economies of scale at the campus level were largely exhausted once campus enrollment reached 1,000. The predicted cost per pupil for a campus of 1,000 students was 0.3% above the least-cost configuration, as was the predicted cost per pupil for a campus more than twice as large. The predicted cost of operating the largest campus in the sample (5,098 students) was 2.7% above the predicted cost of operating the least-cost campus configuration.

Figure F-2 presents a graph of how changes in campus average conditional NCE scores impact predicted cost. As the figure indicates, increases in educational quality were also associated with increases in the cost of education. As the Conditional NCE increased, the slope also increased, indicating the cost of producing additional academic gains was higher for campuses where gains were already high than it was for campuses where gains were relatively low.

Figure F-2: The Estimated Relationship between Per-Pupil Cost and the Average Conditional NCE Score



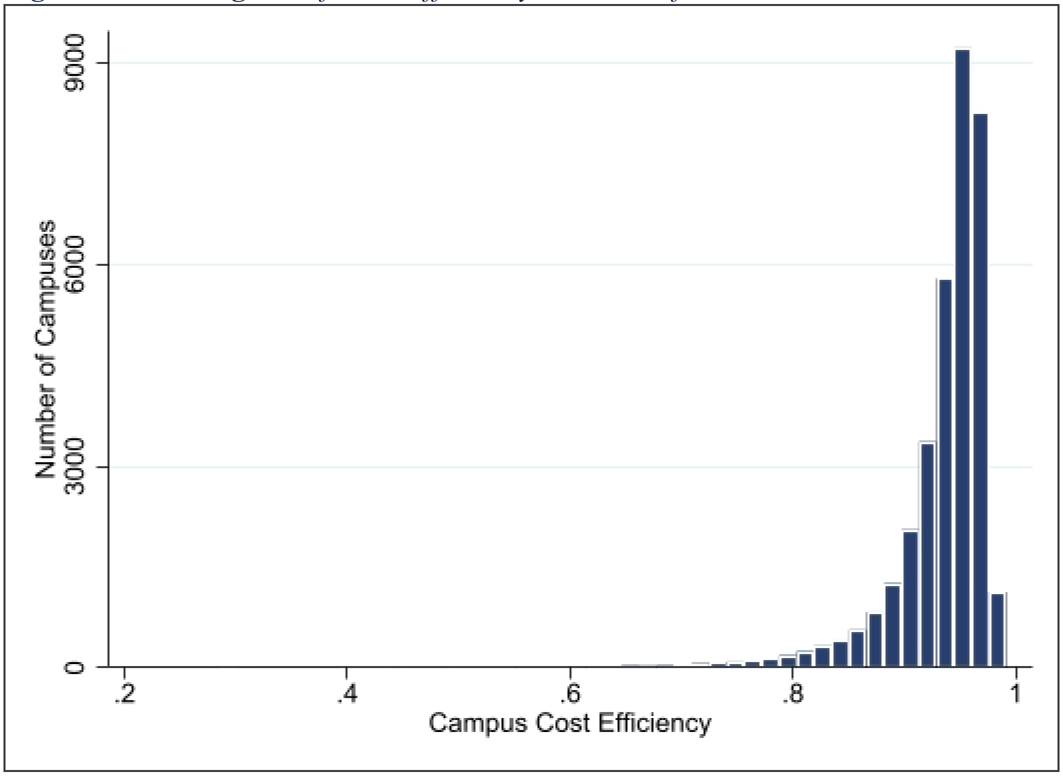
Source: Authors' calculations.

Efficiency Results

An important part of this study was the estimation of cost efficiency, or inefficiency. Figure F-3 graphs the distribution of cost efficiency for Model 3.[61] The average cost efficiency score was 0.93, indicating that campuses were producing 93% of their potential output, on average. Given that inefficiency in this context means unexplained expenditures, not necessarily waste, and that many campuses may have been producing outcomes that were not reflected in test scores, the average efficiency level was quite high. However, the minimum efficiency scores were well below 50%, suggesting that some campuses spent much more than could be explained by measured outcomes, input prices or student need. This higher the Herfindahl index, the higher the level of cost inefficiency, but competition alone explained only a fraction of the measured inefficiency. Sixty percent of the variation in cost efficiency came from differences within school districts.

[61] Cost efficiency was estimated following Battese and Coelli (1995).

Figure F-3: Histogram of Cost Efficiency Measures for Model 3



Source: Authors' calculations.

The Educational Cost index

Using the coefficient estimates from Model 3, one can predict how much each district must spend, each year, in order to produce a standard level of output, assuming it was making cost-minimizing choices about campus size. The Educational Cost Index (ECI) is the ratio of the predicted cost for the district, divided by the state minimum predicted cost.

As is customary in the literature, the level of output quality was set at the state average (or in other words a Conditional NCE score of 0.50). However, the level of output quantity (campus enrollment) to use in the construction of the ECI was not obvious.

Conceptually, the district should be expected to choose an average campus size that is as cost-effective as possible while still serving all of the students. Therefore, the research team used a grid search to identify the cost-minimizing average campus size. The team calculated the average campus size for each district under a variety of scenarios. The first scenario was constructed assuming only a single campus (so that campus enrollment equaled district enrollment); the second scenario was constructed assuming that campus enrollment was equal to district enrollment divided by two; the third scenario was constructed assuming that campus enrollment was equal to district enrollment divided by three; and so on until there were 300 scenarios.[62] Cost was predicted

[62] To avoid extrapolating outside of the experience of the data, scenarios that led to an average campus enrollment below 100 (roughly the first percentile of campus enrollment) or above 3000 (roughly the 99th percentile of campus enrollment) were

for each campus under each scenario, holding all cost factors except the two output measures at their observed values, setting the output quality at the state mean, and setting the campus enrollment for every campus in the district at the scenario average. The scenario with the lowest average predicted cost was identified for each district. Finally, the ECI for each district was defined as the average predicted cost per pupil under the least-cost scenario, divided by the state minimum average predicted cost per pupil. [63]

Robustness Checks

Table F-3 presents the marginal effects from two alternative specifications designed to examine the robustness of the educational cost function estimated as Model 3.

Table F-3: Means of the Marginal Effects from Alternative Specifications

Variable	Model 3	Model 3a	Model 3b
Campus Enrollment (log)	-0.035	-0.041	-0.038
Joint p-value	0.000	0.000	0.000
Average NCE	0.128	0.101	0.126
Joint p-value	0.000	0.000	0.000
Miles to Metro Center (log)	0.008	0.008	0.007
Joint p-value	0.000	0.000	0.000
Teacher Cost Index	0.656	0.685	0.667
Joint p-value	0.000	0.000	0.000
APCI	0.049	0.012	0.038
Joint p-value	0.000	0.000	0.000
District Enrollment (log)	-0.028	-0.027	-0.026
Joint p-value	0.000	0.000	0.000
% Students Econ. Disadv.	0.181	0.184	0.181
Joint p-value	0.000	0.000	0.000
% EverELL	0.010	0.013	0.009
Joint p-value	0.000	0.000	0.000
% Special Ed.	1.116	1.220	1.122

[63] The reference prediction used in the construction of the ECI is the prediction at the one-quarter percentile (so that only one quarter of one percent of the districts have a predicted wage below the reference wage). The ECI was set to 1.00 for the handful of districts with predicted wages below the reference wage. This approach ensures that the reference wage is not an extreme outlier

Variable	Model 3	Model 3a	Model 3b
Joint p-value	0.000	0.000	0.000
% Special Ed. High Needs	0.121	0.141	0.130
Joint p-value	0.000	0.000	0.000
Middle School Campus	0.028	0.030	0.027
Joint p-value	0.000	0.000	0.000
High School Campus	0.132	0.135	0.133
Joint p-value	0.000	0.000	0.000
Multi-grade Campus	0.150	0.226	0.150
Joint p-value	0.000	0.000	0.000
K¬8 School district	-0.094	-0.080	.
Joint p-value	0.000	0.000	.
Micropolitan County	-0.135	-0.131	-0.167
Joint p-value	0.000	0.000	0.000
Metropolitan County	-0.121	-0.126	-0.151
Joint p-value	0.000	0.000	0.000
Sparsely Populated County	-0.041	-0.047	-0.034
Joint p-value	0.000	0.000	0.000
Very Sparsely Populated County	0.135	0.148	0.081
Joint p-value	0.000	0.000	0.000
First Tier Coastal County	-0.013	-0.015	0.012
Joint p-value	0.000	0.000	0.000
Number of observations	34,502	32,149	34,298

Note: All models also include year fixed effects. P-values based on robust standard errors that were clustered by district and year.
Source: Authors' calculations from Table F-5.

The first column in Table F-3 replicates the marginal effects from Model 3 (the preferred specification). The second column in Table F-3 presents a version of model 3 wherein the state's two largest districts—Dallas ISD and Houston ISD—were excluded from the sample. At the table illustrates, the model was largely insensitive to the inclusion or exclusion of the state's two largest districts.

The final column in Table F-3 presents a version of model 3 that restricts the sample to only school districts that serve the full spectrum of grade levels. As such, K–8 districts have been excluded from Model 3b. Once again, the marginal effects indicate that the model was robust to this change in the estimation sample—with one exception. When K–8 districts were excluded from the sample, the mean marginal effect on cost from being in a First Tier Coastal County was positive, not negative. This pattern appears to arise from a slightly smaller coefficient on the interaction between district enrollment and the coastal county indicator under this specification.

The estimated ECI from both of the alternative specifications were very highly correlated with the baseline ECI. As Table F-4 illustrates, the correlation between the baseline ECI and the ECI from model 3a (the model excluding DISD and HISD) was 0.985. The correlation between the baseline ECI and the ECI from model 3b (the model excluding the K–8 districts) was lower (0.955). However, when attention was restricted only to districts that served the full spectrum of grades (i.e., the districts included in the estimation of model 3b) the correlation between the two ECIs was 0.9994). This pattern implies that including the K–8 districts in the estimation did not bias the index values for districts that served the full spectrum of grades, but did yield different index values for the K–8 districts themselves.

Table F-4: Pearson Correlation Coefficients for Alternative Model ECIs, 2018–19

Variable	Correlation with ECI Model 3, All Districts	Correlation with ECI Model 3, excluding DISD and HISD	Correlation with ECI Model 3, excluding K8 districts
ECI Model 3	1.0000	1.0000	1.0000
ECI Model 3a	0.9856	0.9856	0.9773
ECI Model 3b	0.9554	0.9554	0.9994

Table F-5 presents the estimated coefficients and robust standard errors from each of the model specifications.

Table F-5: Coefficient Estimates and Standard Errors from Alternative Specifications

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
District Enrollment	-0.0689 (0.098)	-0.7471*** (0.268)	-0.6853** (0.286)	-0.7041** (0.286)	-0.5821** (0.239)
District Enrollment, squared	0.0147 (0.012)	0.0790*** (0.026)	0.0720*** (0.027)	0.0766*** (0.027)	0.0601*** (0.022)
District Enrollment, cubed	-0.0004 (0.000)	-0.0025*** (0.001)	-0.0023*** (0.001)	-0.0024*** (0.001)	-0.0018*** (0.001)
District Enrollment* Campus Enrollment	-0.0147*** (0.003)	-0.0148*** (0.003)	-0.0148*** (0.003)	-0.0186*** (0.003)	-0.0146*** (0.003)
District Enrollment*NCE	0.0110 (0.022)	0.0092 (0.022)	0.0349 (0.025)	-0.0158 (0.025)	0.0381 (0.025)
District Enrollment*Distance (log)	-0.0013 (0.003)	-0.0015 (0.003)	-0.0016 (0.003)	-0.0007 (0.003)	-0.0017 (0.003)
District Enrollment*TCI (log)	0.0110 (0.061)	-0.0027 (0.063)	0.0002 (0.062)	0.0399 (0.063)	-0.0183 (0.065)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
District Enrollment* APCI (log)	0.0596 (0.041)	0.1079** (0.044)	0.0906** (0.045)	0.0718 (0.046)	0.1191** (0.052)
District Enrollment*Pct Poor	0.0076 (0.008)	0.0316** (0.013)	0.0323** (0.015)	0.0373** (0.015)	0.0327** (0.014)
District Enrollment * Pct EverELL	0.0020 (0.008)	-0.0115 (0.010)	-0.0125 (0.010)	-0.0214** (0.011)	-0.0116 (0.010)
District Enrollment* Pct Special Ed.	0.2924*** (0.036)	0.2395*** (0.054)	0.2315*** (0.046)	0.2991*** (0.048)	0.2280*** (0.046)
District Enrollment * Pct Special Ed. High Needs	-0.0226 (0.028)	-0.0302 (0.028)	-0.0305 (0.028)	-0.0160 (0.029)	-0.0297 (0.028)
District Enrollment* Middle School	0.0021 (0.003)	-0.0173** (0.007)	-0.0157** (0.008)	-0.0137* (0.008)	-0.0161** (0.008)
District Enrollment* High School	-0.0328*** (0.004)	-0.0650*** (0.012)	-0.0607*** (0.012)	-0.0594*** (0.013)	-0.0614*** (0.012)
District Enrollment* Multigrade School	0.0066 (0.009)	0.0238** (0.011)	0.0220* (0.011)	0.0557*** (0.012)	0.0173* (0.010)
District Enrollment* K¬8 District	-0.0335 (0.055)	-0.0075 (0.055)	-0.0119 (0.055)	0.0005 (0.056)	
District Enrollment* Micropolitan County	-0.0482*** (0.011)	-0.0494*** (0.011)	-0.0477*** (0.011)	-0.0456*** (0.011)	-0.0446*** (0.011)
District Enrollment* Metropolitan County	-0.0139 (0.012)	-0.0126 (0.012)	-0.0123 (0.012)	-0.0153 (0.012)	-0.0070 (0.013)
District Enrollment* Sparsely Populated County	-0.0243** (0.012)	-0.0114 (0.013)	-0.0178 (0.012)	-0.0168 (0.012)	-0.0172 (0.012)
District Enrollment* Very Sparsely Populated County	0.0691*** (0.016)	0.0829*** (0.017)	0.0775*** (0.016)	0.0809*** (0.016)	0.0730*** (0.016)
District Enrollment* Coastal County	-0.0183*** (0.005)	-0.0189*** (0.005)	-0.0185*** (0.005)	-0.0193*** (0.005)	-0.0177*** (0.005)
Campus Enrollment	-0.2719*** (0.051)	-0.1228 (0.079)	-0.1426* (0.082)	-0.1605* (0.086)	-0.1407* (0.080)
Campus Enrollment * Campus Enrollment	0.0182*** (0.003)	0.0179*** (0.003)	0.0178*** (0.003)	0.0204*** (0.004)	0.0176*** (0.003)
Campus Enrollment * NCE	0.0717 (0.045)	0.0735 (0.045)	0.0767* (0.045)	0.1576*** (0.047)	0.0733 (0.045)
Campus Enrollment * Distance (log)	-0.0010 (0.003)	-0.0014 (0.004)	-0.0015 (0.004)	-0.0037 (0.004)	-0.0014 (0.004)
Campus Enrollment * TCI (log)	-0.0841 (0.085)	-0.0949 (0.085)	-0.0958 (0.085)	-0.2020** (0.083)	-0.0919 (0.085)
Campus Enrollment * APCI (log)	-0.1067* (0.063)	-0.1009 (0.062)	-0.0968 (0.063)	-0.0205 (0.063)	-0.1011 (0.063)
Campus Enrollment * Pct. Poor	0.0284* (0.015)	0.0295* (0.015)	0.0307** (0.015)	0.0149 (0.016)	0.0305* (0.016)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Campus Enrollment * Pct. EverELL	-0.0026 (0.018)	-0.0008 (0.018)	-0.0018 (0.018)	0.0255 (0.018)	-0.0029 (0.018)
Campus Enrollment * Pct. Special Ed.	-0.6430*** (0.075)	-0.6387*** (0.075)	-0.6412*** (0.075)	-0.6670*** (0.078)	-0.6403*** (0.075)
Campus Enrollment * Pct. Special Ed. High Needs	0.0530 (0.038)	0.0591 (0.038)	0.0583 (0.038)	0.0500 (0.039)	0.0590 (0.038)
Campus Enrollment * Middle School	0.0566*** (0.006)	0.0577*** (0.006)	0.0578*** (0.006)	0.0536*** (0.007)	0.0581*** (0.006)
Campus Enrollment * High School	0.0823*** (0.006)	0.0844*** (0.006)	0.0846*** (0.006)	0.0828*** (0.008)	0.0850*** (0.006)
Campus Enrollment * Multigrade School	0.0047 (0.017)	0.0065 (0.017)	0.0071 (0.017)	-0.0303* (0.017)	0.0067 (0.017)
Campus Enrollment * K8 District	0.0820 (0.062)	0.0795 (0.062)	0.0793 (0.062)	0.0621 (0.062)	
Campus Enrollment * Micropolitan County	0.0293** (0.012)	0.0302** (0.012)	0.0304** (0.012)	0.0267** (0.012)	0.0282** (0.012)
Campus Enrollment * Metropolitan County	0.0049 (0.014)	0.0048 (0.014)	0.0045 (0.014)	0.0088 (0.014)	0.0020 (0.014)
Campus Enrollment * Sparsely Populated County	0.0378** (0.015)	0.0381** (0.015)	0.0378** (0.015)	0.0349** (0.015)	0.0342** (0.015)
Campus Enrollment * Very Sparsely Populated County	-0.1020*** (0.018)	-0.1050*** (0.018)	-0.1048*** (0.018)	-0.1113*** (0.018)	-0.1088*** (0.018)
Campus Enrollment * Coastal County	-0.0001 (0.008)	0.0009 (0.008)	0.0011 (0.008)	0.0042 (0.008)	0.0006 (0.008)
NCE	0.5710 (0.387)	1.1621** (0.590)	-0.9953 (0.824)	-1.0025 (0.832)	-0.9929 (0.808)
NCE * NCE	-0.7151*** (0.236)	-0.7212*** (0.235)	0.5068 (0.619)	0.4377 (0.620)	0.5122 (0.596)
NCE * Distance (log)	-0.0599* (0.032)	-0.0558* (0.032)	0.0013 (0.043)	-0.0011 (0.046)	0.0001 (0.042)
NCE * TCI (log)	0.0370 (0.628)	0.1496 (0.622)	-0.1563 (0.639)	-0.1809 (0.681)	-0.1350 (0.635)
NCE * APCI (log)	-0.7745 (0.512)	-0.7904 (0.510)	-0.6015 (0.519)	-0.7398 (0.533)	-0.7099 (0.514)
NCE * Pct. Poor	0.0335 (0.110)	0.0390 (0.110)	0.3301** (0.167)	0.3020* (0.171)	0.3089* (0.161)
NCE * Pct. EverELL	-0.3290*** (0.110)	-0.3273*** (0.110)	-0.4770*** (0.126)	-0.4996*** (0.140)	-0.4670*** (0.125)
NCE * Pct. Special Ed.	0.1385 (0.614)	0.1551 (0.611)	-2.0961* (1.141)	-1.2843 (1.154)	-2.0710* (1.096)
NCE * Pct. Special Ed. High Needs	0.1819 (0.291)	0.2137 (0.292)	0.0927 (0.297)	0.0177 (0.287)	0.1466 (0.299)
NCE * Middle School	-0.0547 (0.049)	-0.0525 (0.049)	0.0833 (0.077)	0.0544 (0.078)	0.0861 (0.077)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
NCE * High School	-0.1760***	-0.1714***	0.0287	-0.0133	0.0303
	(0.060)	(0.060)	(0.107)	(0.111)	(0.106)
NCE * Multigrade School	-0.0835	-0.0784	-0.0832	-0.2417**	-0.0682
	(0.126)	(0.126)	(0.126)	(0.118)	(0.126)
NCE * K8 District	0.4113*	0.4365*	0.3711	0.3129	
	(0.234)	(0.234)	(0.235)	(0.233)	
NCE * Micropolitan County	-0.1213	-0.1253	-0.1156	-0.1069	-0.1331*
	(0.080)	(0.081)	(0.080)	(0.080)	(0.080)
NCE * Metropolitan County	-0.1115	-0.1175	-0.0413	0.0168	-0.0655
	(0.102)	(0.102)	(0.108)	(0.113)	(0.108)
NCE * Sparsely Populated County	-0.0144	-0.0187	-0.1232	-0.1274	-0.1265
	(0.114)	(0.114)	(0.124)	(0.124)	(0.124)
NCE * Very Sparsely Populated County	0.2011*	0.2067*	0.1368	0.1373	0.1398
	(0.120)	(0.120)	(0.124)	(0.126)	(0.122)
NCE * Coastal County	0.1023	0.0914	0.0439	0.0622	0.0331
	(0.064)	(0.064)	(0.067)	(0.069)	(0.067)
Distance (log)	-0.0232	0.0062	-0.0232	-0.0171	-0.0267
	(0.040)	(0.040)	(0.041)	(0.043)	(0.042)
Distance (log) * Distance (log)	0.0037	0.0039*	0.0030	0.0035	0.0031
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Distance (log) * TCI (log)	-0.0420	-0.0646	-0.0530	-0.0575	-0.0661
	(0.086)	(0.088)	(0.088)	(0.095)	(0.090)
Distance (log) * APCI (log)	0.0172	-0.0586	-0.0504	-0.0303	-0.0297
	(0.067)	(0.073)	(0.073)	(0.073)	(0.070)
Distance (log) * Pct. Poor	-0.0173	-0.0359**	-0.0244*	-0.0183	-0.0247*
	(0.013)	(0.015)	(0.013)	(0.015)	(0.013)
Distance (log) * Pct. EverELL	0.0669***	0.0823***	0.0731***	0.0583***	0.0757***
	(0.016)	(0.016)	(0.016)	(0.019)	(0.016)
Distance (log) * Pct. Special Ed.	0.2772***	0.2753***	0.2482***	0.2033***	0.2422***
	(0.061)	(0.065)	(0.063)	(0.073)	(0.063)
Distance (log) * Pct. Special Ed. High Needs	0.0252	-0.0084	-0.0007	-0.0081	0.0042
	(0.036)	(0.037)	(0.037)	(0.040)	(0.038)
Distance (log) * Middle School	-0.0120***	-0.0094*	-0.0063	-0.0056	-0.0064
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Distance (log) * High School	-0.0213***	-0.0233***	-0.0206***	-0.0159***	-0.0210***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Distance (log) * Multigrade School	-0.0150	0.0001	-0.0013	-0.0001	-0.0014
	(0.012)	(0.013)	(0.014)	(0.014)	(0.014)
Distance (log) * K8 District	0.1273***	0.1409***	0.1347***	0.1343***	
	(0.029)	(0.029)	(0.029)	(0.028)	
Distance (log) * Micropolitan County	0.0097	0.0135	0.0154	0.0173	0.0183
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Distance (log) * Metropolitan County	0.0251*	0.0355**	0.0308**	0.0351**	0.0324**
	(0.013)	(0.014)	(0.014)	(0.015)	(0.015)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Distance (log) * Sparsely Populated County	0.0678*** (0.015)	0.0468*** (0.017)	0.0541*** (0.016)	0.0559*** (0.016)	0.0541*** (0.016)
Distance (log) * Very Sparsely Populated County	0.0721*** (0.018)	0.0703*** (0.018)	0.0714*** (0.018)	0.0687*** (0.019)	0.0701*** (0.018)
Distance (log) * Coastal County	0.0138 (0.009)	0.0187** (0.009)	0.0168* (0.009)	0.0186** (0.009)	0.0175** (0.009)
TCI (log)	2.6040*** (0.811)	3.1512*** (0.861)	3.1393*** (0.874)	3.3227*** (0.894)	3.4718*** (0.925)
TCI (log) * TCI (log)	0.9609 (1.194)	-0.7609 (1.352)	-0.0853 (1.212)	0.2494 (1.254)	-0.0866 (1.206)
TCI (log) * APCI (log)	-1.2174 (1.476)	0.7570 (1.669)	0.1503 (1.536)	-0.3548 (1.508)	0.3361 (1.558)
TCI (log) * Pct. Poor	-0.6991** (0.287)	-0.9688*** (0.315)	-0.9673*** (0.317)	-0.7137** (0.336)	-1.0120*** (0.322)
TCI (log) * Pct. EverELL	0.1554 (0.281)	0.1796 (0.281)	0.1663 (0.280)	-0.0024 (0.321)	0.1853 (0.282)
TCI (log) * Pct. Special Ed.	-8.0215*** (1.290)	-5.8788*** (1.453)	-6.1493*** (1.481)	-6.1818*** (1.497)	-5.9001*** (1.506)
TCI (log) * Pct. Special Ed. High Needs	-0.2650 (0.697)	-0.0594 (0.696)	-0.1909 (0.698)	-0.1276 (0.707)	-0.4175 (0.703)
TCI (log) * Middle School	0.0496 (0.072)	-0.1114 (0.094)	-0.0654 (0.088)	-0.0683 (0.091)	-0.0857 (0.093)
TCI (log) * High School	-0.0967 (0.097)	-0.2052* (0.108)	-0.2006* (0.107)	-0.1765 (0.113)	-0.2222** (0.110)
TCI (log) * Multigrade School	-0.0568 (0.214)	-0.3898 (0.247)	-0.3164 (0.243)	-0.2575 (0.250)	-0.3466 (0.248)
TCI (log) * K8 District	-0.8558* (0.515)	-1.2220** (0.536)	-1.1065** (0.527)	-1.1154** (0.518)	
TCI (log) * Micropolitan County	-0.1592 (0.232)	-0.2897 (0.246)	-0.3066 (0.241)	-0.2848 (0.242)	-0.3086 (0.246)
TCI (log) * Metropolitan County	-0.4143 (0.300)	-0.4572 (0.314)	-0.4702 (0.305)	-0.5155 (0.315)	-0.5349* (0.313)
TCI (log) * Sparsely Populated County	0.4143 (0.298)	0.2515 (0.307)	0.3583 (0.295)	0.3260 (0.296)	0.3152 (0.296)
TCI (log) * Very Sparsely Populated County	-0.4001 (0.331)	-0.8821** (0.379)	-0.6732** (0.340)	-0.5699* (0.342)	-0.7822** (0.349)
TCI (log) * Coastal County	0.2547 (0.197)	0.5498** (0.218)	0.4661** (0.210)	0.3797* (0.212)	0.4669** (0.208)
APCI (log)	-1.4937** (0.656)	-1.8952*** (0.669)	-1.7748*** (0.668)	-2.0748*** (0.674)	-2.2732*** (0.720)
APCI (log) * APCI (log)	1.4413** (0.709)	1.0114 (0.734)	1.0944 (0.720)	0.9824 (0.715)	0.9430 (0.746)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
APCI (log) * Pct. Poor	0.0043	0.0839	0.0802	-0.0313	0.1201
	(0.222)	(0.227)	(0.227)	(0.228)	(0.229)
APCI (log) * Pct. EverELL	-0.2892	-0.3979	-0.4178	-0.3624	-0.4658*
	(0.259)	(0.267)	(0.268)	(0.282)	(0.275)
APCI (log) * Pct. Special Ed.	5.4284***	5.9781***	5.8186***	6.6709***	5.4547***
	(1.115)	(1.140)	(1.131)	(1.114)	(1.120)
APCI (log) * Pct. Special Ed. High Needs	3.1419***	2.8163***	2.8360***	2.8884***	3.1638***
	(0.528)	(0.541)	(0.543)	(0.544)	(0.538)
APCI (log) * Middle School	-0.1347**	0.0263	0.0007	-0.0083	0.0385
	(0.066)	(0.090)	(0.089)	(0.091)	(0.099)
APCI (log) * High School	-0.1768**	0.0819	0.0589	0.0153	0.1017
	(0.075)	(0.126)	(0.131)	(0.135)	(0.142)
APCI (log) * Multigrade School	-0.1091	0.2046	0.1401	-0.0284	0.2567
	(0.193)	(0.226)	(0.225)	(0.229)	(0.252)
APCI (log) * K8 District	0.1257	0.4861	0.4558	0.5589	
	(0.447)	(0.468)	(0.472)	(0.455)	
APCI (log) * Micropolitan County	0.0405	0.1155	0.0806	0.0796	0.1186
	(0.162)	(0.165)	(0.164)	(0.164)	(0.164)
APCI (log) * Metropolitan County	-0.0883	-0.1986	-0.2092	-0.1288	-0.1135
	(0.201)	(0.204)	(0.204)	(0.203)	(0.201)
APCI (log) * Sparsely Populated County	0.0160	0.0505	-0.0226	-0.0130	0.0890
	(0.205)	(0.214)	(0.207)	(0.207)	(0.206)
APCI (log) * Very Sparsely Populated County	0.4695*	0.7651**	0.6048**	0.6493**	0.7639***
	(0.276)	(0.312)	(0.281)	(0.278)	(0.295)
APCI (log) * Coastal County	-0.1168	-0.1956	-0.1875	-0.1562	-0.1887
	(0.130)	(0.132)	(0.132)	(0.133)	(0.130)
Pct Poor	-0.0673	-0.1535	-0.4305**	-0.4950**	-0.4127*
	(0.139)	(0.174)	(0.216)	(0.220)	(0.211)
Pct. Poor * Pct. Poor	0.2012***	0.1930***	0.2431***	0.2835***	0.2353***
	(0.031)	(0.040)	(0.035)	(0.036)	(0.035)
Pct. Poor * Pct. EverELL	-0.2033***	-0.2446***	-0.2838***	-0.3277***	-0.2662***
	(0.047)	(0.056)	(0.057)	(0.064)	(0.055)
Pct. Poor * Pct. Special Ed.	-0.4826**	-0.8721***	-1.0313***	-0.8343***	-1.0240***
	(0.196)	(0.257)	(0.315)	(0.324)	(0.311)
Pct. Poor * Pct. Special Ed. High Needs	-0.3094***	-0.2511**	-0.2717**	-0.2363**	-0.2571**
	(0.117)	(0.119)	(0.118)	(0.117)	(0.120)
Pct. Poor * Middle School	0.0306**	0.0904***	0.0972***	0.0962***	0.0950***
	(0.014)	(0.026)	(0.031)	(0.032)	(0.030)
Pct. Poor * High School	0.0086	0.1555**	0.1689**	0.1873**	0.1644**
	(0.021)	(0.065)	(0.073)	(0.074)	(0.070)
Pct. Poor * Multigrade School	-0.0313	0.0017	0.0141	0.0696	0.0143
	(0.041)	(0.046)	(0.046)	(0.047)	(0.047)
Pct. Poor * K8 District	0.1141	0.1136	0.1433	0.1817*	
	(0.097)	(0.102)	(0.099)	(0.093)	

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Pct. Poor * Micropolitan County	0.0097 (0.039)	0.0547 (0.043)	0.0532 (0.044)	0.0612 (0.043)	0.0574 (0.042)
Pct. Poor * Metropolitan County	0.1282*** (0.045)	0.1794*** (0.050)	0.1785*** (0.051)	0.1806*** (0.054)	0.1702*** (0.051)
Pct. Poor * Sparsely Populated County	0.0670 (0.048)	0.1274** (0.053)	0.0969** (0.049)	0.0970* (0.050)	0.0943* (0.051)
Pct. Poor * Very Sparsely Populated County	0.0293 (0.052)	0.0332 (0.053)	0.0191 (0.052)	0.0314 (0.053)	0.0204 (0.053)
Pct. Poor * Coastal County	0.0630*** (0.023)	0.0972*** (0.025)	0.0869*** (0.025)	0.0765*** (0.025)	0.0886*** (0.025)
Pct EverELL	-0.0035 (0.163)	0.0455 (0.168)	0.1827 (0.185)	0.1861 (0.205)	0.1388 (0.179)
Pct. EverELL * Pct. EverELL	-0.0068 (0.031)	0.1201** (0.053)	0.1055* (0.055)	0.1073* (0.060)	0.0965* (0.051)
Pct. EverELL * Pct. Special Ed.	0.8818*** (0.228)	1.4008*** (0.311)	1.4704*** (0.346)	1.5820*** (0.363)	1.4555*** (0.338)
Pct. EverELL * Pct. Special Ed. High Needs	0.0938 (0.118)	-0.0517 (0.136)	0.0231 (0.123)	0.0435 (0.128)	0.0582 (0.124)
Pct. EverELL * Middle School	-0.0222 (0.015)	-0.0366** (0.016)	-0.0467*** (0.018)	-0.0406** (0.018)	-0.0412** (0.017)
Pct. EverELL * High School	-0.0547** (0.021)	-0.1548*** (0.045)	-0.1588*** (0.049)	-0.1742*** (0.050)	-0.1504*** (0.046)
Pct. EverELL * Multigrade School	-0.0444 (0.048)	-0.1026* (0.054)	-0.1055* (0.055)	-0.1284** (0.055)	-0.0954* (0.054)
PCt. EverEll * K8 District	0.1485 (0.122)	0.1412 (0.123)	0.1724 (0.123)	0.1462 (0.124)	
Pct. EverELL * Micropolitan County	0.0333 (0.049)	-0.0280 (0.053)	-0.0078 (0.052)	-0.0049 (0.053)	-0.0149 (0.052)
Pct. EverELL * Metropolitan County	0.0997* (0.052)	0.0493 (0.054)	0.0604 (0.055)	0.0583 (0.058)	0.0712 (0.054)
Pct. EverELL * Sparsely Populated County	-0.2946*** (0.063)	-0.3612*** (0.068)	-0.3483*** (0.067)	-0.3426*** (0.068)	-0.3474*** (0.067)
Pct. EverELL * Very Sparsely Populated County	-0.0786 (0.058)	-0.1177** (0.060)	-0.1005* (0.059)	-0.0948 (0.062)	-0.0968* (0.058)
Pct. EverELL * Coastal County	-0.0515* (0.027)	-0.0742*** (0.028)	-0.0704** (0.028)	-0.0680** (0.029)	-0.0731*** (0.028)
Pct. Special Ed. * Pct. Special Ed.	-3.1725*** (0.716)	5.8269 (3.698)	4.7328 (3.576)	2.7658 (3.554)	4.6174 (3.444)
Pct. Special Ed. * Pct. Special Ed. High Needs	2.4106*** (0.521)	2.6095*** (0.524)	2.5991*** (0.523)	2.8851*** (0.519)	2.6573*** (0.533)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Pct. Special Ed. * Middle School	0.1090 (0.079)	-0.2959 (0.194)	-0.3079 (0.204)	-0.2586 (0.211)	-0.2937 (0.196)
Pct. Special Ed. * High School	0.8333*** (0.087)	-0.1933 (0.460)	-0.1634 (0.445)	-0.1288 (0.448)	-0.1344 (0.422)
Pct. Special Ed. * Multigrade School	0.7758*** (0.154)	0.4559* (0.259)	0.3479 (0.242)	0.4017 (0.264)	0.3773 (0.231)
Pct. Special Ed. * K8 District	-0.8451* (0.469)	-1.1621** (0.521)	-1.1806** (0.500)	-1.1200** (0.490)	
Pct. Special Ed. * Micropolitan County	0.9359*** (0.159)	0.9520*** (0.160)	0.9325*** (0.159)	0.9193*** (0.159)	0.8995*** (0.162)
Pct. Special Ed. * Metropolitan County	1.4194*** (0.194)	1.0196*** (0.238)	1.0565*** (0.244)	0.8315*** (0.248)	1.0417*** (0.238)
Pct. Special Ed. * Sparsely Populated County	-0.4133* (0.224)	-0.3905* (0.225)	-0.4081* (0.224)	-0.3330 (0.224)	-0.4318* (0.226)
Pct. Special Ed. * Very Sparsely Populated County	-1.5492*** (0.230)	-1.4245*** (0.237)	-1.3991*** (0.236)	-1.3250*** (0.239)	-1.3920*** (0.242)
Pct. Special Ed. * Coastal County	0.1247 (0.109)	0.0675 (0.114)	0.1150 (0.109)	-0.0009 (0.111)	0.1023 (0.109)
Pct Special Ed.	0.7131 (0.677)	-0.4445 (0.785)	1.1776* (0.710)	0.6526 (0.736)	1.2052* (0.707)
Pct. Special Ed. High Needs*Pct. Special Ed. High Needs	0.1102 (0.158)	0.1811 (0.161)	0.1438 (0.158)	0.2059 (0.152)	0.1446 (0.160)
Pct. Special Ed. High Needs * Middle School	-0.2474*** (0.034)	-0.3158*** (0.042)	-0.2950*** (0.041)	-0.2629*** (0.040)	-0.2967*** (0.041)
Pct. Special Ed. High Needs * High School	-0.2486*** (0.041)	-0.2951*** (0.045)	-0.2898*** (0.045)	-0.2588*** (0.045)	-0.2913*** (0.045)
Pct. Special Ed. High Needs * Multigrade School	-0.0517 (0.064)	-0.0584 (0.063)	-0.0590 (0.063)	-0.0033 (0.062)	-0.0635 (0.063)
Pct. Special Ed. High Needs * K8 District	-0.4325*** (0.115)	-0.3911*** (0.116)	-0.3917*** (0.116)	-0.3267*** (0.115)	
Pct. Special Ed. High Needs * Micropolitan County	0.0372 (0.048)	0.0241 (0.050)	0.0173 (0.049)	0.0176 (0.049)	0.0072 (0.049)
Pct. Special Ed. High Needs * Metropolitan County	-0.0076 (0.082)	-0.0321 (0.081)	-0.0144 (0.082)	-0.0432 (0.087)	-0.0097 (0.082)
Pct. Special Ed. High Needs * Sparsely Populated County	0.0932 (0.061)	0.1261** (0.062)	0.1085* (0.061)	0.1208** (0.061)	0.1122* (0.063)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Pct. Special Ed. High Needs * Very Sparsely Populated County	-0.3915*** (0.082)	-0.3648*** (0.083)	-0.3585*** (0.083)	-0.3493*** (0.084)	-0.4004*** (0.083)
Pct. Special Ed. High Needs * Coastal County	0.0314 (0.082)	-0.0217 (0.084)	-0.0188 (0.085)	-0.0349 (0.086)	-0.0339 (0.087)
Pct Special Ed. High Needs	-0.5633* (0.293)	-0.5043* (0.298)	-0.4282 (0.301)	-0.5564* (0.305)	-0.4951 (0.304)
Middle School * K8 District	0.1412*** (0.037)	0.1771*** (0.042)	0.1791*** (0.041)	0.1821*** (0.041)	
Middle School Micropolitan County	-0.0085 (0.009)	-0.0363** (0.015)	-0.0365** (0.015)	-0.0356** (0.015)	-0.0347** (0.015)
Middle School * Metropolitan County	-0.0187* (0.011)	-0.0176 (0.012)	-0.0204* (0.011)	-0.0190 (0.012)	-0.0204* (0.011)
Middle SchoolSparsely Populated County	-0.0195* (0.011)	-0.0064 (0.012)	-0.0124 (0.012)	-0.0101 (0.012)	-0.0142 (0.012)
Middle SchoolVery Sparsely Populated County	-0.0134 (0.014)	-0.0069 (0.014)	-0.0075 (0.014)	-0.0052 (0.014)	-0.0140 (0.014)
Middle SchoolCoastal County	0.0135** (0.006)	0.0277*** (0.010)	0.0179*** (0.007)	0.0162** (0.007)	0.0185*** (0.007)
Middle School	-0.1536*** (0.048)	0.0504 (0.090)	-0.0593 (0.065)	-0.0572 (0.067)	-0.0592 (0.065)
High School * Micropolitan County	-0.0350*** (0.011)	-0.0618*** (0.015)	-0.0583*** (0.015)	-0.0550*** (0.015)	-0.0552*** (0.015)
High School * Metropolitan County	-0.0284** (0.013)	-0.0457*** (0.015)	-0.0432*** (0.015)	-0.0344** (0.015)	-0.0437*** (0.014)
High School * Sparsely Populated County	-0.0051 (0.014)	0.0140 (0.016)	0.0060 (0.015)	0.0076 (0.015)	0.0031 (0.015)
High School * Very Sparsely Populated County	0.0509*** (0.014)	0.0563*** (0.015)	0.0487*** (0.014)	0.0539*** (0.015)	0.0422*** (0.015)
High School * Coastal County	0.0189** (0.008)	0.0038 (0.011)	0.0026 (0.011)	0.0000 (0.011)	0.0038 (0.011)
High School	0.2620*** (0.058)	0.5319*** (0.115)	0.3758*** (0.076)	0.3540*** (0.079)	0.3787*** (0.076)
Multigrade School * Micropolitan County	-0.0496*** (0.019)	-0.0765*** (0.022)	-0.0730*** (0.021)	-0.0612*** (0.021)	-0.0678*** (0.021)
Multigrade School * Metropolitan County	-0.0254 (0.026)	-0.0039 (0.027)	-0.0132 (0.026)	-0.0194 (0.027)	-0.0078 (0.027)
Multigrade School * Sparsely Populated County	-0.0038 (0.022)	0.0156 (0.024)	0.0045 (0.023)	0.0088 (0.022)	0.0026 (0.022)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Multigrade School * Very Sparsely Populated County	0.0696** (0.030)	0.0501 (0.032)	0.0468 (0.032)	0.0740** (0.032)	0.0325 (0.034)
Multigrade School * Coastal County *	0.0326 (0.038)	0.0332 (0.038)	0.0341 (0.038)	0.0243 (0.041)	0.0340 (0.038)
Multigrade School	0.2231* (0.125)	-0.0089 (0.152)	0.0194 (0.156)	0.0477 (0.157)	0.0415 (0.149)
K8 District * Micropolitan County	-0.0734 (0.046)	-0.0926** (0.047)	-0.0897* (0.047)	-0.0851* (0.047)	
K8 District * Metropolitan County	0.1686** (0.067)	0.2074*** (0.069)	0.1987*** (0.068)	0.1899*** (0.067)	
K8 District * Sparsely Populated County	-0.1465*** (0.053)	-0.1158** (0.055)	-0.1158** (0.055)	-0.1188** (0.055)	
K8 District * Very Sparsely Populated County	0.0822 (0.080)	0.0761 (0.083)	0.0646 (0.083)	0.0575 (0.081)	
K8 District * Coastal County	-0.0180 (0.052)	-0.0241 (0.053)	-0.0138 (0.052)	-0.0209 (0.053)	
K8 District	-0.6722** (0.282)	-0.9777*** (0.302)	-0.9116*** (0.300)	-0.9318*** (0.294)	
Micropolitan County * Sparsely Populated County	0.0189 (0.016)	0.0135 (0.016)	0.0107 (0.016)	0.0155 (0.016)	0.0083 (0.017)
Micropolitan County * Very Sparsely Populated County	0.0694*** (0.026)	0.0405 (0.029)	0.0416 (0.029)	0.0330 (0.029)	0.0421 (0.028)
Micropolitan County * Coastal County	0.4531*** (0.087)	0.4717*** (0.091)	0.4892*** (0.090)	0.4968*** (0.090)	0.0771 (0.071)
Micropolitan County	0.0675 (0.082)	0.0585 (0.083)	0.0420 (0.084)	0.0290 (0.083)	0.0312 (0.085)
Metropolitan County * Sparsely Populated County	-0.0178 (0.041)	-0.0121 (0.042)	-0.0127 (0.041)	-0.0136 (0.042)	-0.0092 (0.042)
Metropolitan County * Very Sparsely Populated County	0.0609* (0.034)	0.0713** (0.035)	0.0659** (0.033)	0.0516 (0.034)	0.0631* (0.034)
Metropolitan County * Coastal County	0.3849*** (0.091)	0.3882*** (0.094)	0.4094*** (0.093)	0.4258*** (0.092)	-0.0048 (0.067)
Metropolitan County	-0.1756 (0.112)	-0.1869* (0.113)	-0.2120* (0.116)	-0.2282* (0.118)	-0.2309* (0.123)
Sparsely Populated County	-0.2945*** (0.108)	-0.3387*** (0.111)	-0.2444** (0.111)	-0.2534** (0.111)	-0.2242** (0.111)
Very Sparsely Populated County	0.0804 (0.150)	-0.0051 (0.154)	0.0683 (0.150)	0.0587 (0.152)	0.1427 (0.151)

Variables	Model 1	Model 2	Model 3	Model 3a	Model 3b
Very Sparsely Populated County * Coastal County	0.4096***	0.4544***	0.4574***	0.4613***	0.0473
	(0.084)	(0.087)	(0.087)	(0.086)	(0.077)
Coastal County	-0.3911***	-0.4473***	-0.4209***	-0.4250***	
	(0.110)	(0.114)	(0.112)	(0.112)	
HISD	-0.0963***	-0.0667***	-0.0760***		-0.0827***
	(0.020)	(0.021)	(0.021)		(0.020)
DISD	0.0015	0.0136	0.0182		0.0114
	(0.023)	(0.023)	(0.023)		(0.022)
2014–15 School Year	-0.0792***	-0.0817***	-0.0819***	-0.0833***	-0.0814***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
2015–16 School Year	-0.0492***	-0.0510***	-0.0508***	-0.0547***	-0.0505***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)
2017–18 School Year	-0.0412***	-0.0417***	-0.0416***	-0.0427***	-0.0414***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
2018–19 School Year	-0.0141**	-0.0129**	-0.0133**	-0.0164***	-0.0131**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
School Size Residuals		-0.1476***	-0.1288**	-0.1275**	-0.1267**
		(0.055)	(0.057)	(0.057)	(0.054)
School Quality Residuals		-0.6303			
		(0.430)			
Constant	10.2232***	11.3182***	11.9656***	12.1393***	11.6861***
	(0.345)	(0.643)	(0.838)	(0.839)	(0.721)
One-sided error					
Herfindahl (log)	0.3304***	0.3523***	0.3430***	0.3621***	0.3362***
	(0.047)	(0.046)	(0.047)	(0.051)	(0.047)
Constant	-4.5997***	-4.5546***	-4.5718***	-4.5856***	-4.5885***
	(0.094)	(0.093)	(0.094)	(0.095)	(0.094)
Two-sided error					
Unallocated Share	3.3899***	3.3637***	3.3935***	3.3635***	3.5755***
	(0.404)	(0.404)	(0.403)	(0.399)	(0.403)
Number of students Tested (log)	-0.2638***	-0.2627***	-0.2601***	-0.2519***	-0.2513***
	(0.027)	(0.027)	(0.027)	(0.028)	(0.027)
Constant	-4.3869***	-4.3920***	-4.4122***	-4.4142***	-4.4991***
	(0.184)	(0.183)	(0.183)	(0.191)	(0.184)
Number of observations	34,502	34,502	34,502	32,149	34,298

Note: Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Appendix G: The Transportation Cost Function

The Literature

A school transportation policy monograph from the Center for Cities and Schools at the University California Berkeley (Vincent et al. 2014) reports that “Overall, student transport, and particularly the associated costs, is a grossly under-researched issue.” Consistent with this claim, we find that the academic journal literature on the economics of education has paid surprisingly little attention to the study of school transportation costs. There are two small school transportation research strands. One set of papers focuses upon the potential of cost advantages to district size and the second set of papers looks at the potential for cost savings from privatization of the school transportation function.

We find three published papers in the school cost economies literature that provide analyses of public school district transportation cost functions. Duncombe, Miner, and Ruggiero (1995) study school district costs for New York State using data from 1990. Dodson and Garrett (2004) use year 2000 data from Arkansas to study the determinants of variation in school costs across school districts. In the most recent study, Zimmer, DeBoer, and Hirth (2009) analyze a three-year panel (2004–2006) of cost data for school districts in Indiana.

For all three of these papers, the primary purpose is to develop and to estimate a cost function model of total operating costs for school districts. The outputs in the total cost function are measures of student performance, e.g., achievement test scores and dropout rates. The major input price is a measure of teacher salaries. Environmental factors, such as number of pupils and student/family characteristics, are also included. These papers are, therefore, part of the much larger literature on school cost functions that is summarized earlier in this report. The distinguishing feature of these three papers is that each one disaggregates total operating expenditures into its major subcomponents, including transportation, and then estimates separate cost functions for each subcomponent. The explanatory cost factors in the transportation cost function estimation are identical to those used in the total cost function estimation.

The policy focus of all three papers is consolidation, and the key findings from the transportation cost estimations are student enrollment scale results. In Duncombe, Miner, and Ruggiero (1995), the estimated relationship between transportation costs per pupil and enrollment is U-shaped, with a minimum at 1100 pupils. The per pupil transportation costs decline sharply by 25% from an enrollment of 50 to the minimum at 1100, then increase much more slowly beyond 1100. The scale results in Dodson and Garrett (2004) are similar to those in Duncombe, Miner, and Ruggiero (1995). The estimated transportation costs per pupil curve declines steeply to around 500 to 1000 students, then becomes essentially flat (so more L-shaped) over the reminder of the observed range of district sizes. Zimmer, DeBoer, and Hirth (2009) also finds an L-shaped relationship between transportation costs per pupil and enrollment over the enrollment range of the data. The cost-minimizing enrollment occurs at 27,510 students, although most of the decline in average costs is realized by an enrollment of 10,000 students.

There is a major shortcoming in the three papers referenced above. The transportation cost estimates are derived from empirical models that are, fundamentally, not developed as transportation cost function models. In Duncombe, Miner, and Ruggiero (1995) and Dodson and

Garrett (2004), the output measures are the number of students taught (quantity) and student achievement outcomes (quality), not transportation outputs. Only Zimmer, DeBoer, and Hirth (2009) includes a direct transportation output, bus miles, in addition to number of pupils and pupil achievement outputs. Since the total number of students may serve as a reasonable proxy for the number of student bus riders, there is some transportation cost function flavor to the results in these papers. However, neither Zimmer, DeBoer, and Hirth (2009) nor the other two papers included appropriate input prices for a transportation cost study. The only labor price is teacher salaries. For a school transportation cost analysis, the principal labor price is bus driver salaries. The transportation cost functions in the papers are poorly specified. Since the transportation cost function is not the primary object of interest, the weakness in specification is understandable, but it limits the usefulness of the findings for understanding the determinants of variation across school districts in the costs of transporting students.

The school bus privatization literature, on the other hand, does include papers that estimate credible school transportation cost function models. Lazarus and McCullough (2005) and estimate a log-linear model of variable school transportation costs using data from Minnesota school districts for school year 1999–2000. The model treats the number of pupils transported as the measure of transportation output. Input prices for bus drivers and for fuel and input controls for the number of miles of road in the district and the number school buses (separating small and large). The percentage of special transportation needs riders was also included as a regressor to allow for potential differential transportation costs for these rider types. The key additional control is a dummy variable for in-house or contract delivery of transportation services. They also explore the possibility of different cost functions for rural and nonrural districts. Thompson (2011) makes several refinements to the Lazarus and McCullough study. In particular, Thompson uses multiple years of Minnesota school district data rather than a single cross-section, a continuous measure of the degree of privatization, the addition of bus miles/student transported as a second output, and treatment of the potential endogeneity of the privatization variable/decision. The main conclusion of both of these papers is that contracting out did not lead to reductions in the cost of pupil transportation services relative to in house provision.

In a pair of papers, Hutchinson and Pratt (1999, 2007) explore the relative cost of contracting out versus in-house production of school bus transportation. The 1999 study uses data from Tennessee school districts while the 2007 study uses Louisiana school district data. The same basic empirical cost function model is used in both of these papers. The cost model assumes two outputs: the average number of students transported daily and the number of one-way bus miles driven, and two input prices: average annual bus driver salary and cost per gallon for fuel. Fixed inputs include the number of Type I and Type II buses and the district population density. A stronger, flexible translog functional form is assumed for the cost function specification. Again, the focus of these studies is on the comparative cost of in-house and contracted transportation institutions. Hutchinson and Pratt find that in-house was cheaper in Louisiana but that contracting out was cheaper in Tennessee.

Although the academic literature on school bus transportation functions is sparse, there is a robust academic literature on the costs of municipal bus transit. Most bus transit systems in the United States are publicly owned and operated. As with all regulated public industries, policy makers and government agencies are interested in understanding the underlying cost structure of firms in the industry in order to set industry pricing policies or to assess the potential cost advantages of increasing or decreasing the number of firm operators in a particular city. There is also considerable interest in the possibility of reducing operating costs by privatizing the transit system. Cost function studies can provide valuable information on scale economies, input price effects, and other cost factors that can be used to help shape bus transit policy.

The modern literature on transit system cost functions was largely developed in the 1980s and 1990s. The timing of this literature coincided with the development of flexible functional forms for the estimation of cost functions. The use of flexible forms, such as the transcendental logarithmic (or 'translog') cost model, allowed researchers to provide stronger econometric evidence on the key economic characteristics of the bus transit system that were of interest to policy makers and to transit system managers than could be obtained from more a priori restricted functional forms used in earlier studies.

As suggested above, a key element of any cost function study is the definition and measurement of the output for which the (minimum) costs are being determined. In a study of crude oil, the output measurement is straightforward—the number of barrels (quantity) of sweet light (quality) crude oil produced per day. For a study of transportation services, the choice of an appropriate and workable measure of output has been more controversial. As argued in an influential review article by Berechman and Giuliano (1985), in the case of bus transportation there are two relevant type of output measures. One is a vehicle-based or technical output measure, such as bus miles or bus-hours. The second is a passenger-based or demand-based output measure, such as passenger-trips or passenger-hours. Cost is, fundamentally, a producer or supplier type concept, and in the transportation context, transit operators assemble inputs to supply bus miles. Bus miles are, however, best thought of as an intermediate output. B miles are produced in order to transport passengers to destinations. The final output is the passenger trips, or the associated passenger miles, and the economic value is derived from this final output.

The cornerstone paper in the bus transit cost function literature, Viton (1981), uses bus miles as the measure of bus service output. Many of the bus system cost studies follow suit and use this supplier-related technical output level. Important examples include Williams and Dalal (1981) and Matas and Raymond (1998). Several studies, instead, have chosen to use the demand-related passenger miles as their output measure. Examples of this approach include Williams and Hall (1981), Button and O'Donnell (1985), and Berechman (1983). A third set of papers (e.g., Berechman and Giuliano1984; de Rus 1990; and Karlaftis, McCarthy and Sinha 1999) take a more agnostic stance and estimate two cost models, one with vehicles-miles as output and the second with passenger-miles as output. As suggested by Berechman and Giuliano (1984), a better approach would be to estimate a single model that included both output measures. In a well-cited review of the literature on economies of scale and economies of density in the transportation literature, Caves and Christensen (1988) cite an unpublished dissertation by Windle (1984) as providing the best evidence for bus transit cost economies. In a published article based upon his dissertation research, Windle (1988) reports the results of his two-output cost function specification, with bus route miles and passenger miles as the outputs. This is also the approach taken in a study of 68 Midwest bus transit systems by Harmatuck (2005). In a study of Class I intercity bus carriers in the United States, Tauchen, Fravel, and Gilbert (1983) include both bus miles and number of passengers per mile as output measures in their cost function.

In the public policy regulatory context, the key findings from bus cost function studies are evidence on the returns to scale characteristics of bus transit service production. According to Small (1992), cost functions that use vehicle-related outputs reveal increasing returns for small systems, constant returns for medium-sized systems, and mildly decreasing returns for large systems. Cost functions that use passenger-related outputs show increasing returns. These are, however, broadly generalized conclusions. There is considerable variation across the studies. From the studies that include both output types, Windle (1988) finds constant returns to bus miles and increasing returns to passenger miles. Harmatuck (2005) includes both types of output measures, finds constant returns to vehicle miles and slightly increasing returns to passenger trips. Tauchen, Fravel, and Gilbert (1983) finds constant returns to scale for bus miles and no significant cost effect of the number of passengers per mile.

All of the modern bus transit cost studies include measures of the price of labor and the price of fuel in their cost function. For bus transit firms, labor costs constitute around 75% of operating expenses (Harmatuck (2005)). Accounting for exogenous differences in wages of bus drivers (and other transport staff) is critically important to estimating bus transit cost functions. Similarly, the price of diesel fuel to run the buses is an important factor in determining bus transit operating costs.

In addition to labor and fuel, the third critical input to producing bus services is bus capital. The majority of the transit studies treat the rolling stock of bus capital as being fixed, and thus the cost function estimates are interpreted as short run bus variable operating cost functions. The number of buses is usually included as an explanatory variable. Some studies include average age of the buses as a measure of capital quality.

A second important public policy consideration is the efficiency of public transit operations. A common feature of public sector firms is a relatively weak incentive environment for realizing cost efficient service delivery. The absence of residual claimancy to the generation of profits and limited competition combine to make an assumption of cost minimizing behavior on the part of transit managers questionable. The hypothesis that bus transit firms may not minimize costs can be tested using stochastic cost frontier techniques. Harmatuck (2005) finds evidence of inefficiency ranging from very low (actual cost only 1% greater than predicted minimum cost) to very high (actual cost 30% greater than expected minimum cost) across city systems in the Midwest. Size did not appear to be a determinant of efficiency, but location (state) did show some systematic effects (e.g., most of the Wisconsin city transit systems were more cost efficient than were most of the Michigan city transit systems). In a review of a large set of (mostly 1990s) bus frontier studies, De Borger, Kerstens, and Costa (2002) report that "most studies report substantial remaining inefficiency among urban transit operators". Representative examples include Fazioli, Filippini, and Prioni (1993), with average cost efficiency per operator between 78% and 100% and Levaggi (1994), with average cost inefficiency per operator between 14% and 40%.

We would note that the Dodson and Garrett (2004) school transportation cost analysis employs a stochastic frontier cost function approach and the Duncombe, Miner, and Ruggiero (1995) school transportation cost paper includes a stochastic frontier cost function estimation in an appendix. Neither paper reports the estimated mean inefficiency for districts in their samples.

Hanley (2007) used linear programming techniques to explore the impact of school district consolidation on school bus routes and therefore school district transportation costs. He simulated the consolidation of school districts in Iowa up to a target enrollment of between 500 and 1,000 students and concluded that the increase in transportation costs would be large enough to offset at least half of the expected savings from administrative efficiencies.

Data

The data for this analysis come from administrative files and public records of the Texas Education Agency (TEA), the Education Research Center at the University of Texas at Dallas, the National Center for Education Statistics (NCES), the US Bureau of Labor Statistics (BLS), the US Department of Housing and Urban Development (HUD) and the U.S Census Bureau. The analysis covers the five year period from 2014–15 through 2018–19.

The unit of analysis is the district, as districts make transportation decisions for all campuses in the district. Alternative Education Accountability (AEA) campuses (e.g., juvenile justice campuses, disciplinary education campuses, residential campuses and all other alternative education campuses) have been excluded. Open-enrollment charter schools have also been excluded from the analysis. Virtual campuses and campuses that lack reliable data on student performance (such as elementary education campuses that serve no students in tested grades, or very small campuses) have also been excluded. The final sample includes 980 districts for each year between school years 2014–15 and 2018–19.

Table G-1 provides means and standard deviations for the variables used in this analysis. Total Transportation Expenditures is listed first, and Total Route Miles is listed fourth. We divide Total Transportation Expenditures by Total Route Miles to get Expenditures per Mile, which serves as our left hand side (or our dependent) variable after transforming to natural logarithm.

Table G-1: Descriptive Statistics for Transportation Cost Model AY 2015 –AY 2019

Variable	Mean	Std. Dev.	Minimum	Maximum
Total Transportation Expenditures	$1,411,742	4,920,305	$5,348	$66,600,000
Expenditures per Mile	$3.32	1.71	$0.23	$24.65
Riders per Mile	0.71	0.44	0.02	9.22
Total Route Miles	368,646.10	940,289.60	2,273	16,600,000
Diesel Price	$2.11	0.29	$1.55	$3.04
Wage Index	1.12	0.06	1	1.41
Population Density	243.73	508.93	0.6	2,718
Congestion	848.55	685.59	47	3,236
Number of Special Riders	117.16	423.60	0	7,335
Number of Total Riders	1,614.49	4,233.83	5	74,641
Percent Special Riders	4.78	9.25	0	100
Number of Special Miles	103,507.70	304,988.70	0	5,992,172
Number of Regular Miles	265,138.40	650,537	0	11,729,698
Percent Special Miles of Total Miles	18.04	16.55	0	100
Total Vehicles	42.59	88.93	1	1,133

Variable	Mean	Std. Dev.	Minimum	Maximum
Number of Regular Service Buses Less than 5 Years Old	10.23	23.00	0	388
Number of Special Service Buses Less than 5 Years Old	2.89	8.80	0	148
Number of Total Buses Less than 5 Years Old	13.11	30.75	0	470
Percent of Buses Less than 5 Years Old	27.55	0.18	0	1
Rural District Indicator	0.32	0.47	0	1
Micropolitan District Indicator	0.20	0.40	0	1

We have two output measures. The first is Riders per Mile, listed third in the table. Riders per Mile is calculated as Total Riders, listed 10th, divided by Total Route Miles, and it enters our regression after a natural logarithm transform. The second is Total Route Miles itself, which also enters our regression in natural logarithms.

We have two input price measures. The first is Diesel Prices, a per-gallon measure of diesel prices, listed fifth in the table. This variable enters our regression in natural logarithms. The second is our wage index for transportation and other auxiliary workers, APCI, discussed in Chapter 1. This variable enters our regression in natural logarithms.

We have seven environmental variables, plus a series of year fixed effects. The environmental variables include a measure of population density, Density, and a measure we think of as measuring congestion. We call it Congestion in the table. It is calculated by the Texas Transportation Institute and called vehicle miles per center lane mile. These two variables enter our regression in natural logs. We also calculate the percent of special riders as the ratio of Spec Riders to Total Riders and label it Pct Spec Riders. This enters our cost function regression as a percent. Similarly, we calculate the percent of special rider miles as the ratio of District Spec Miles to District Tot Miles, and label it Pct Spec Miles. This too enters our cost function as a percent. Another percent variable is the percent of new busses, busses less than five years old. For this, we add the total number of regular buses less than five years old to the number of special rider buses less than five years old to get the total number of buses less than five years old, and divide by the total number of district vehicles. This variable is labeled Pct Buses Less than 5 y.o., and it enters the regression as a percent as well.

The final variable we include is a measure of the capital stock for transportation. We use the variable Total Vehicles, and it enters our regression in natural logarithm. The capital stock may also serve as a proxy measure for the number of bus routes, which could be considered as an output-type variable.

Summary of Data with Sources:

Total transportation expenditures are collected from Texas Education Agency's Public Education Information Management System (PEIMS) Financial Actual Data files from 2014–2015 to 2018–2019. Total transportation expenditures were generated by taking the total of reported spending under PEIMS Function 34 (reported transportation expenditures).

Expenditures per mile are calculated by dividing total transport expenditures (from PEIMS Financial Actual Data files) by reported district total route services. Data for district total route services were collected from the Texas Education Agency's Transportation Operations Report (TOR). Total route services include regular and special route services (columns P and V of the TOR Excel spreadsheet).

Riders per mile are calculated by dividing a district's reported total average daily ridership by their annual total route services mileage. Route services mileage are collected from the TOR, as previously mentioned. Average daily ridership is collected from the Texas Education Agency's Transportation Route Report (TRR) and is the sum of reported regular program average daily ridership and special program average daily ridership (columns AC and AI of the TRR Excel spreadsheet).

Total route miles are generated from the TOR (columns P and V of the TOR Excel spreadsheet). They are the total of regular route services mileage and special route services mileage for each district.

Diesel prices were purchased from Oil Price Information Services (OPIS).

The **wage index** is discussed in Chapter 1.

Population density is calculated by dividing a county's population by its land area. Data on county population and land area in square miles were collected from the 2010 Census.

Congestion is vehicle miles travelled per lane mile. These figures were provided by TTI.

Ridership:

- District Special Riders: Special program average daily ridership from the TRR (column AI on the TRR Excel spreadsheet)
- District Total Riders: Total of special program average daily ridership and regular program average daily ridership as reported in the TRR (columns AC + AI on the TRR Excel spreadsheet)
- District Percent Special Riders: (Special program riders divided by total riders)*100

Mileage:

- District Special Miles: Special Mileage Summary—Route Related Services as reported in the TOR (column V of the TOR Excel spreadsheet)
- District Total Miles: Total of special program route related services and regular program route related services as reported in the TOR (columns P + V of the TOR Excel spreadsheet)
- District Percent Special Miles: (Special miles divided by total miles)*100

Vehicles: These all come from the TOR. The columns of the TOR Excel spreadsheet with the data are:

- District Buses Less than 5 y.o.: Column AZ
- District Special Buses Less than 5 y.o.: Column BE
- District Total Buses Less Than 5 y.o.: Columns AZ + BE
- District Percent Buses Less Than 5 y.o.: (Total buses less than 5 y.o./(total buses: columns BD + BI))*100

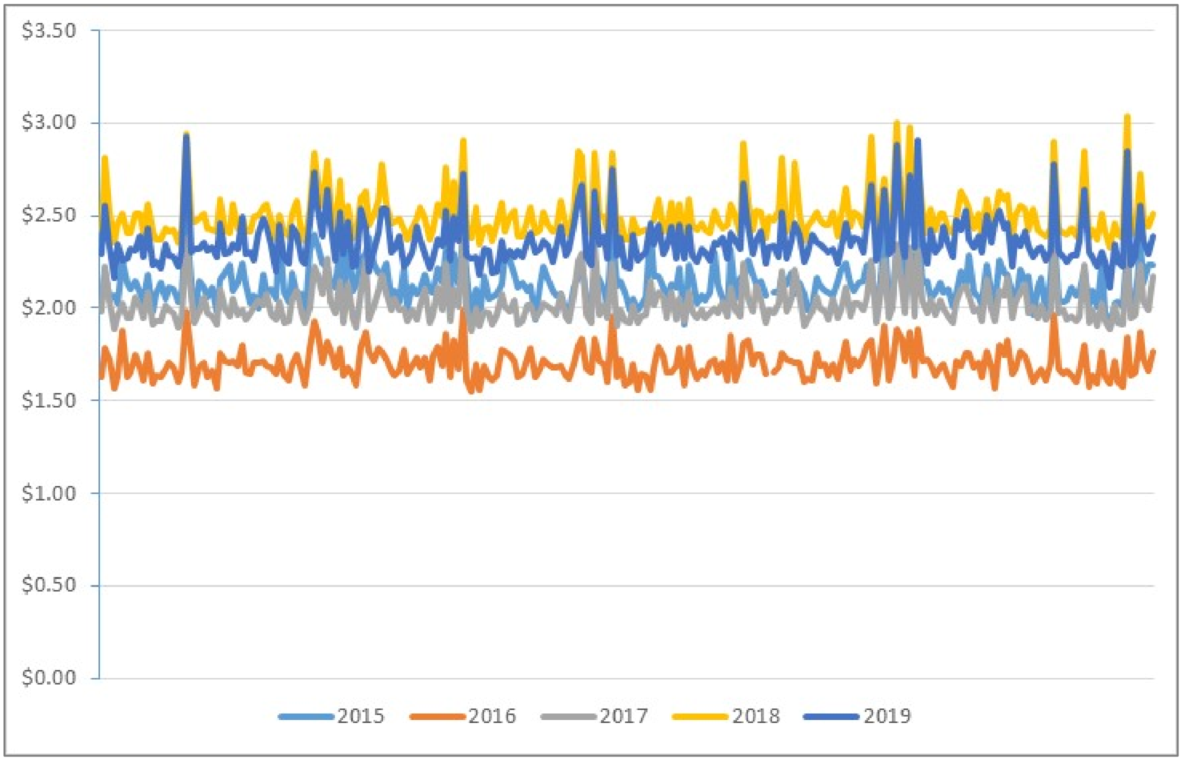
Rural and **micro** are both indicator variables discussed in the text.

Fuel Prices and West Texas Intermediate Prices

Our fuel price measure is a measure of diesel fuel prices by county. It seems clear that such prices are well outside the control of any school district. Fuel prices vary by geography—they vary systematically across counties—and they vary over time. The time variation is largely due, with some lags, to variation over time in the world price of oil. One important variable measuring the world price of oil is the price of West Texas Intermediate, or WTI. Variations in the price of WTI are a significant driver of oil prices at the county level over our sample period. In fact, a regression of our fuel price measure, in logs, on the price of WTI, in logs, has an r-square value of 78%. Adding county level fixed effects (county level indicators) increases that measure of fit to 88%. So, WTI explains a large majority of the movements in our county level diesel prices.

The graph below shows diesel prices for every county in Texas, for the years 2015–2019, where each year is represented by a different color. Notice the high correlation over the years in counties that have higher-than-typical diesel prices. Certain counties have high prices, for whatever reason, and those high prices persist over the years.

Figure G-1: Average Annual Diesel Prices by County: 2015–19



Source: Oil Price Information Services.

In addition, the price of WTI varies over time. WTI was highest in 2018, lowest in 2016. The following table shows the WTI price and the average of the county prices over time. WTI and average county diesel prices have varied closely over time.

Table G-2: Average Diesel Prices and West Texas Intermediate Prices: 2015–19

Diesel Price Types	2015	2016	2017	2018	2019
WTI (per gallon)	$1.16	$1.03	$1.21	$1.55	$1.36
Average Diesel Price (per gallon)	$2.13	$1.70	$2.02	$2.50	$2.36
Max Diesel Price	$2.49	$1.99	$2.40	$3.04	$2.92
90th Percentile Diesel Price	$2.29	$1.82	$2.18	$2.69	$2.52
75th Percentile Diesel Price	$2.19	$1.75	$2.06	$2.54	$2.41
Median Diesel Price	$2.11	$1.69	$1.99	$2.47	$2.33
25th Per	$2.06	$1.64	$1.95	$2.42	$2.28
Min Diesel Price	$1.90	$1.55	$1.88	$2.33	$2.11

The estimated marginal effect of fuel prices in our model is 0.1425, indicating that at the means of all other variables, a 10% increase in fuel prices would increase total transportation costs by 1.425%. The overall mean of diesel prices is $2.106 in our model, so a 10% increase would be an increase of $0.211 per gallon. Cost per mile averages $3.318 in our data, so an increase of 1.425 is an increase of $0.047 per mile.

The Two-Sided and One-Sided Error Terms—Heteroscedasticity and Efficiency

Heteroscedasticity in the two-sided error is literally a lack of a constant variance across the observations. It is often the case that the variance, or standard deviation, of the two-sided regression error may vary systematically with one or more of the explanatory variables in the regression. In the standard situation, this leads to problems with statistical inference, although the coefficient estimates are not impacted. In our model the two-sided error is related to the natural log of district enrollment and to the square of this variable. As enrollment increases, the variance of the two-sided error declines. Our coefficient estimates for the two sided error are that the standard deviation of the two sided error, SD(V), is given as:

$$SD(V) = -.9445067 + .0074094*dle - 0304401*(dle)^2.$$

Table G-3 presents values of SD(V) for the range of enrollment values from minimum to maximum in our sample. The SD(V) is largest for the smallest district, and declines steadily as district size grows.

Table G-3: Heteroscedasticity in Two-Sided Error

Variable	Minimum	Median	Mean	Maximum
Enrollment Values	27	1022	4904.124	215,408
Std. Dev. of V	0.5337	0.3081	0.2144	0.0665

The one-sided error is linked to efficiency, or its converse, inefficiency. We model the variance of the one-sided error as a function of educational competition, as measured by a Herfindahl index of enrollment concentration in the geographic area, where the geographic area was the CBSA. A Herfindahl index of 1.00 indicates a metropolitan area with a single local education agency (LEA); a Herfindahl index of 0.10 indicates a metropolitan area with 10 LEAs of equal size. Both traditional public school districts and open enrollment charter schools are included in the calculation of enrollment concentration. Table G-4 reports the mean value for the Herfindahl index in the sample is 0.24, with a minimum value of 0.05 and a maximum of 1.00. Our coefficient estimates for the one-sided are such that the standard deviation of the one sided error, SD(U), is given as:

$$SD(U) = -6.154216 - 0.4293874*lherf,$$

where lherf is the natural log of the Herfindahl index. Table G-4 presents values of SD(U) for values of our Herfindahl index from the lowest (.0505) to the highest (1.0) in our sample. As our Herfindahl index rises in value, as the educational services in an area become more concentrated, SD(U) decreases.

Table G-4: Inefficiency and the One-Sided Error

Variable	Minimum	Mean	Median	Maximum
Herfindahl Index	.0505	.2368	.2773	1.0000
Std. Dev. of U	0.0880	0.0627	0.0606	.0458

Endogeneity and Instrumental Variables

Our output measures are under the control of school districts. Districts set routes, determining the physical miles covered by each bus route, and arrange schedules of school opening times and bus routing, which impacts ridership and riders per mile. Parents and students have choices for transportation to schools, and districts can influence those choices in a number of ways.

Our methodology for dealing with endogeneity is to use instrumental variables in a control function approach. Basically the control function approach is a form of instrumental variables regression in which, in the first stage, a regression is run for each endogenous variable on a set of instruments. These instruments are variables that are correlated with the endogenous variable but that do not otherwise directly impact the dependent variable, which here is transportation cost per mile. In the control function approach, after the first stage regressions, the residual from each first stage regression is added as an additional explanatory variable in the main regression of interest, the stochastic frontier cost function regression. These residual variables both correct the final regression estimated coefficients and serve as a means of testing if the hypothesized endogenous variable is appropriately treated as endogenous.

Our instrumental variables are the log of district enrollment, the square of the log of district enrollment, the district enrollment in grade 9, the district enrollment in grade 10, the district enrollment in grade 11, the district enrollment in grade 12, and the log of square miles in a district. Table G-5 summarizes the first stage regressions.

Table G-5: Summary of First Stage Regression Results

Variable	Riders per Mile Coefficient	Riders per Mile p-value	Total Route Miles Coefficient	Total Route Miles p-value
Log of District Enrollment	.8186	0.000	-.4427	0.000
Log of District Enrollment, sq.	-.0455	0.000	.0419	0.000
Enrollment 9th	-.0002	0.042	.0002	0.012
Enrollment 10th	-.0000	0.819	-.0001	0.401
Enrollment 11th	.0001	0.630	-0002	0.277
Enrollment 12th	.00031	0.026	-.0001	0.271
Log of Square Miles in District	-.2878	0.000	.2357	0.000
F: test of instrument significance	132.11	0.000	122.66	.000
R-square of regression	.551		.927	

As is clear in the table, our instruments satisfy the requirement of entering significantly in the first stage, with very low probability values for individual hypothesis tests and for the tests of joint significance. The variables for square miles and for enrollment are especially strongly statistically significant in the individual hypotheses tests. Table G-8 reports the regression results for both first stages, plus a final column with results from the stochastic frontier regression. For the endogeneity issue, the most important information is summarized here in Table G-6. Both residual terms are strongly statistically significant, consistent with the notion that they are endogenous and that our first stage correction is appropriate.

Table G-6: Summary of Final Regression (SFA Cost Function) Results for Endogeneity

Variable	Coefficient	Standard Error	"t-stat"	p-value
First Stage Residual for Riders per Mile	-0.9314	.0361	25.78	0.000
First Stage Residual for Total Route Miles	-1.0602	.0453	23.41	0.000

Results

The translog specification has many benefits in terms of flexibility and generality compared to, say, the Cobb Douglas, but the coefficient estimates from the translog specification are not readily interpretable. There are quadratic terms and many interactions all of which impact the interpretation of how a change in an explanatory or right hand side variable will impact the dependent or left hand side variable. To interpret the coefficients researchers typically present the change in the dependent variable—here the change in cost per mile—that arises from a small change in an explanatory variable. These are called 'marginal effects.' Because the marginal effect of a change in any one explanatory variable will depend on the values of all the other explanatory variables, there is no unique marginal effect. That is, the impact of a change in any one explanatory variable depends on the values of all the other explanatory variables. So, to present marginal effects in a standard way, it is typical to calculate marginal effects based on the means of the explanatory variables. We will call these the marginal effects at the mean—the mean values of all the explanatory variables. We also calculate the marginal effects of each explanatory variables for all observations in our sample and then calculate the mean of those, and we call that the mean of the marginal effects. Both calculations provide a way to present the marginal effects in a standardized way.

Table G-7 indicates the marginal effects of a change in the various outputs, prices, and environmental variables on transportation expenditures per mile. For each explanatory variable there are two entries. First is the marginal effect at the mean—the marginal effect on per-pupil cost of a change in the explanatory variable in question, holding all other variables at their respective sample mean values. Second is the mean of the marginal effects—the mean of the marginal effect of the variable in question, calculated for each data point in the sample. A note at the bottom of the table indicates that the probability value is essentially zero for the null hypothesis for the variables in question that all the coefficients on the direct effect and all interaction effects are jointly zero. That is, all the variables are strongly statistically significant in the cost function.

The third variable listed in Table G-7 is the log of total route miles. The marginal effect of a change in total route miles calculated at the mean of all variables in the sample is 0.6198, indicating that an increase in total route miles of 1%, at the sample mean, will increase costs per mile by 0.6198%. To convey the magnitude of this effect, a 1% increase in route miles at the sample mean (about 122,000 miles) is an increase of about 1,220 miles. This causes a 0.6198% increase in transportation costs per mile. The mean of transportation costs per mile is about $3.32, so this is an increase of $0.0206 per mile, or about 2.1 cents per mile.

As mentioned, the mean of the marginal effects calculates the marginal effect of an increase in district enrollment for every sample data point and then averages those estimates to yield the mean of the marginal effects. Here a 1% increase in total route miles has the same marginal effect as reported above.

Table G-7: Marginal Effects at the Means

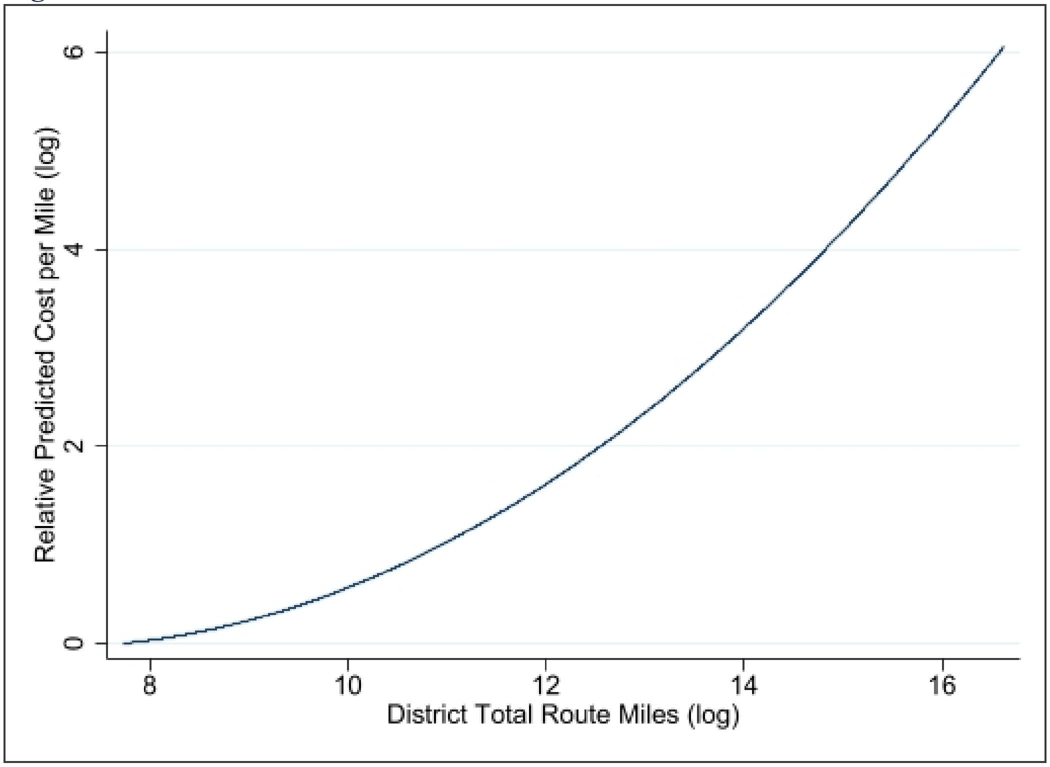
Variable	Marginal Effect at the Mean	Standard Error
Log of District Total Vehicles	-0.6691	0.0498
Log of Riders per Mile	1.0718	0.0342
Log of Total Route Miles	0.6198	0.0425
Log of Fuel Price	0.1533	0.1441
Log of APCI	0.1054	0.1055
Log of Population Density	-0.1068	0.0107
Log of Roadway Congestion	-0.1167	0.0178
Percent Special Program Riders	0.0283	0.0016
Percent Special Program Miles	-0.0054	0.0007
Percent District Buses < 5 y.o.	-0.0029	0.0008
Rural District	-0.0741	0.0336
Micropolitan District	0.0458	0.0289

Note: A hypothesis test for the joint p-value for the coefficients on each of the listed variables is zero to four decimal points, indicating that the coefficients on these variables in the cost function are, jointly, strongly statistically significant.

Figure G-2 graphs the impact of changes in log total route miles on predicted cost per mile (relative to the minimum predicted cost). The slope of the graph is the marginal effect, and the shape of the graph indicates that there are increasing costs. That is, transportation costs per mile are rising in the number of miles.

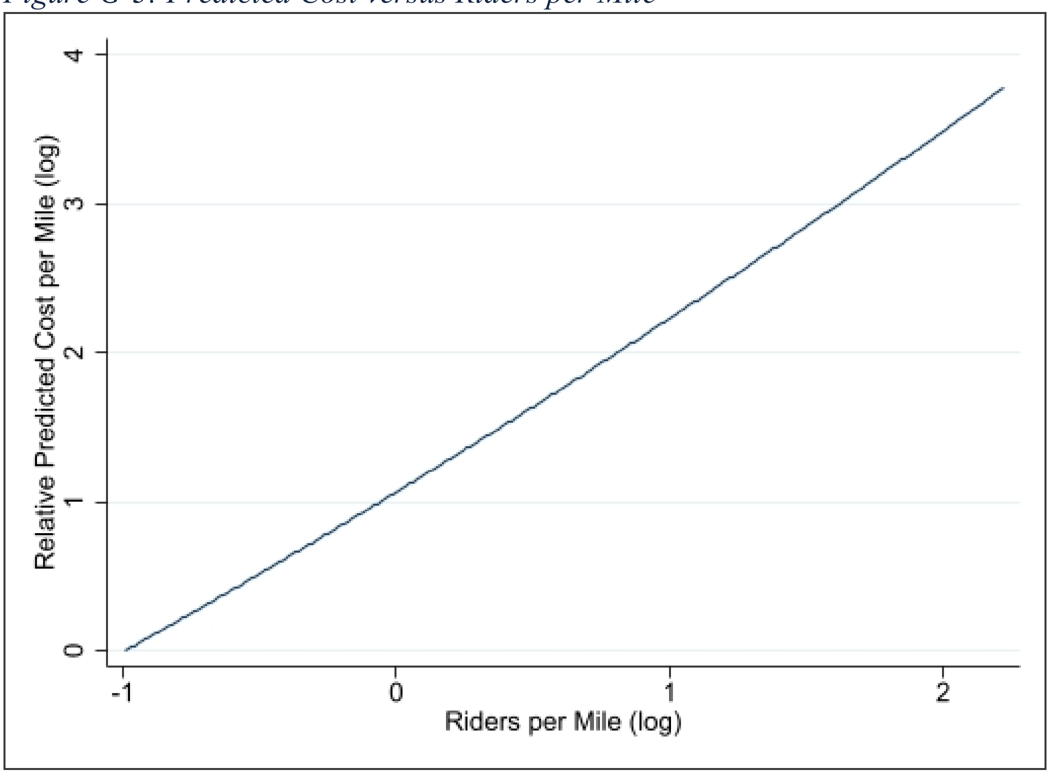
Figure G-3 graphs the impact of changes in log riders per mile on predicted cost per mile (relative to the minimum predicted cost). The slope of the graph is the marginal effect, and the shape indicates there are increasing costs. Transportation costs per mile are rising in the number of riders per mile. Moreover, the slope is fairly constant over the range of values for riders per mile. The marginal effect is about 1, actually 1.0718, indicating that a 1% change in riders per mile results in just over under a 1.1% increase in transportation costs per mile.

Figure G-2: Predicted Cost versus Total Route Miles



Source: Authors' calculations.

Figure G-3: Predicted Cost versus Riders per Mile

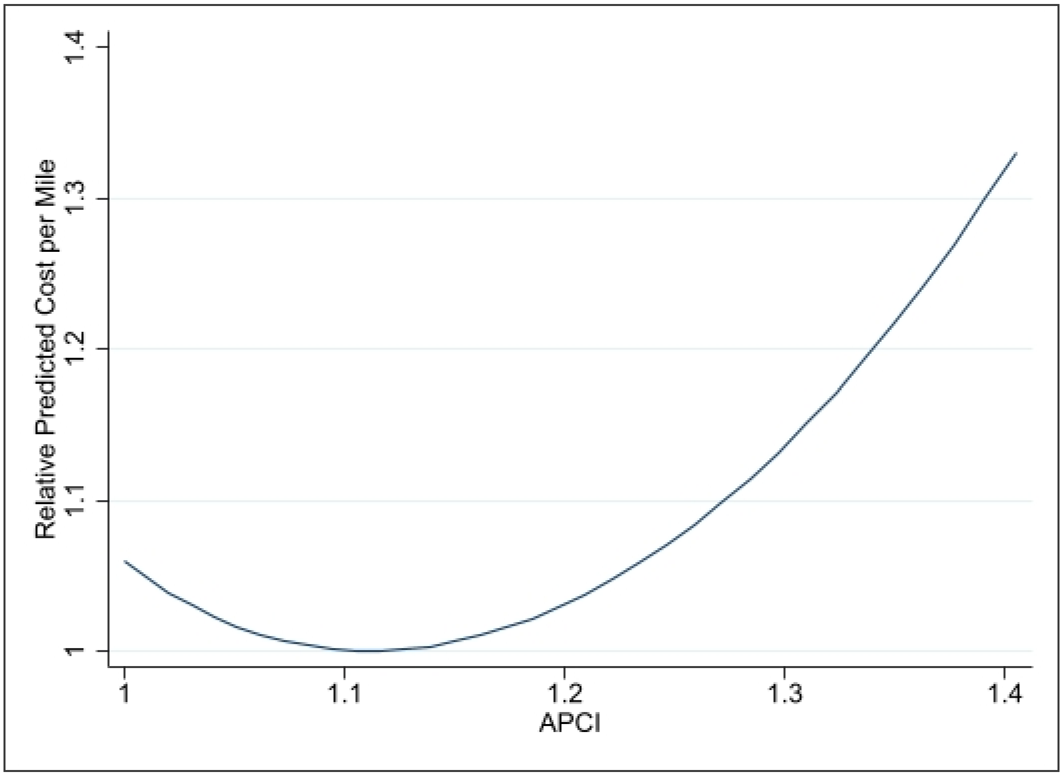


Source: Authors' calculations.

The auxiliary personnel cost index (ACPI) has a marginal effect at the mean of about 0.1054. On average, an increase in auxiliary personnel hiring cost of 1% results in a 0.1054% increase in per mile transportation costs, evaluated at the sample means. Labor costs are a large share of transportation spending, and it is expected that changes in labor costs will have a meaningful impact on transportation costs.

Figure G-4 graphs the impact of the ACPI on cost per mile (relative to the minimum predicted cost per mile) as the auxiliary personnel index ranges from 1.00 to 1.39 in the sample. As the figure illustrates, increases in personnel costs had a positive effect on cost per mile over much the relevant range, but the marginal effect was not well estimated at the very low end of the range.

Figure G-4: The Estimated Relationship between Cost per Mile and the Auxiliary Personnel Cost Index

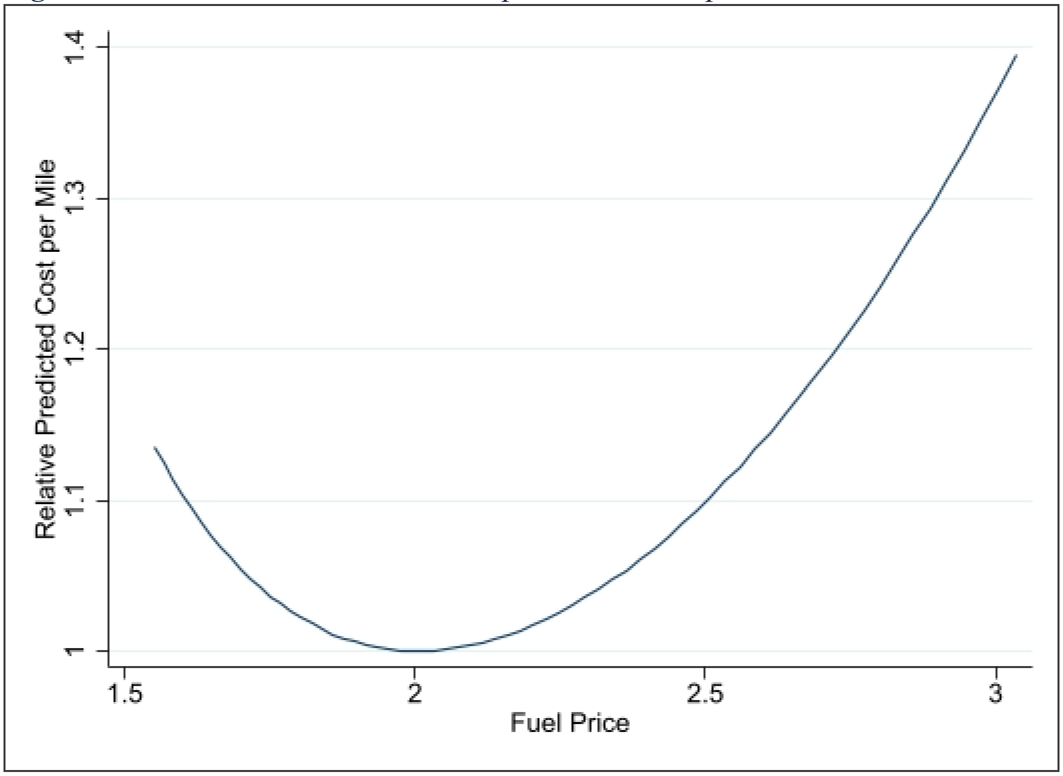


Source: Authors' calculations.

The fuel price has an estimated marginal effect at the mean of about 0.142. On average, an increase in fuel prices of 1% results in a 0.14% increase in per mile transportation costs, evaluated at the sample means. Fuel costs are a large share of transportation spending, and it is expected that changes in labor costs will have a significant impact on transportation costs.

Figure G-5 graphs the impact of the fuel price on cost per mile (relative to the minimum predicted cost per mile) as the fuel price varies from 1.55 to 3.04 in the sample. As the figure illustrates, increases in fuel costs had a positive effect on cost per mile over much the relevant range, but the marginal effect was not well estimated at the lower end of the range.

Figure G-5: The Estimated Relationship between Cost per Mile and District Fuel Price

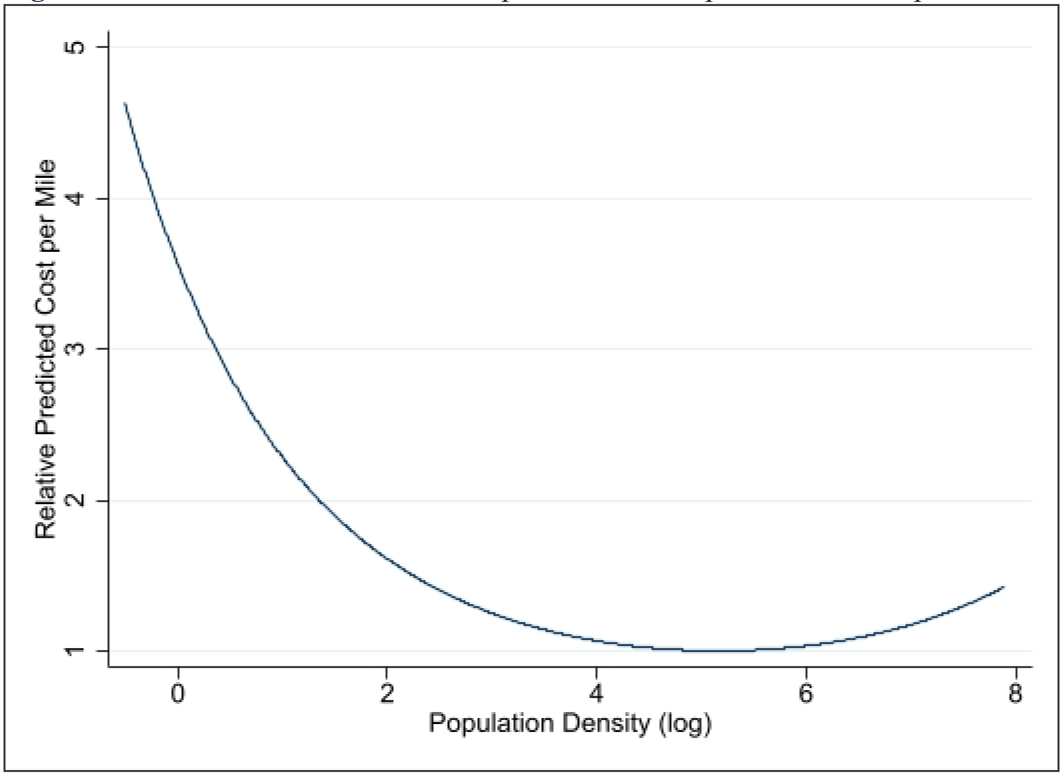


Source: Authors' calculations.

A 1% increase in population density for a district is associated, evaluated at the mean, with a decrease in district per mile student transportation costs of about 0.105%. Thus, the analysis indicates that for a campus with average characteristics, the per mile costs of supplying bus miles is decreasing in the geographic concentration of the student populations.

Figure G-6 graphs the relationship between the log of population density and cost per mile (relative to the minimum predicted cost per mile) as the population density varies from 0.6 to 2718 in the sample. As the figure illustrates, increases in population density had a negative effect on cost per mile over much the relevant range, but the marginal effect was positive at the extremely high densities at the top end of the range.

Figure G-6: The Estimated Relationship between Cost per Mile and Population Density

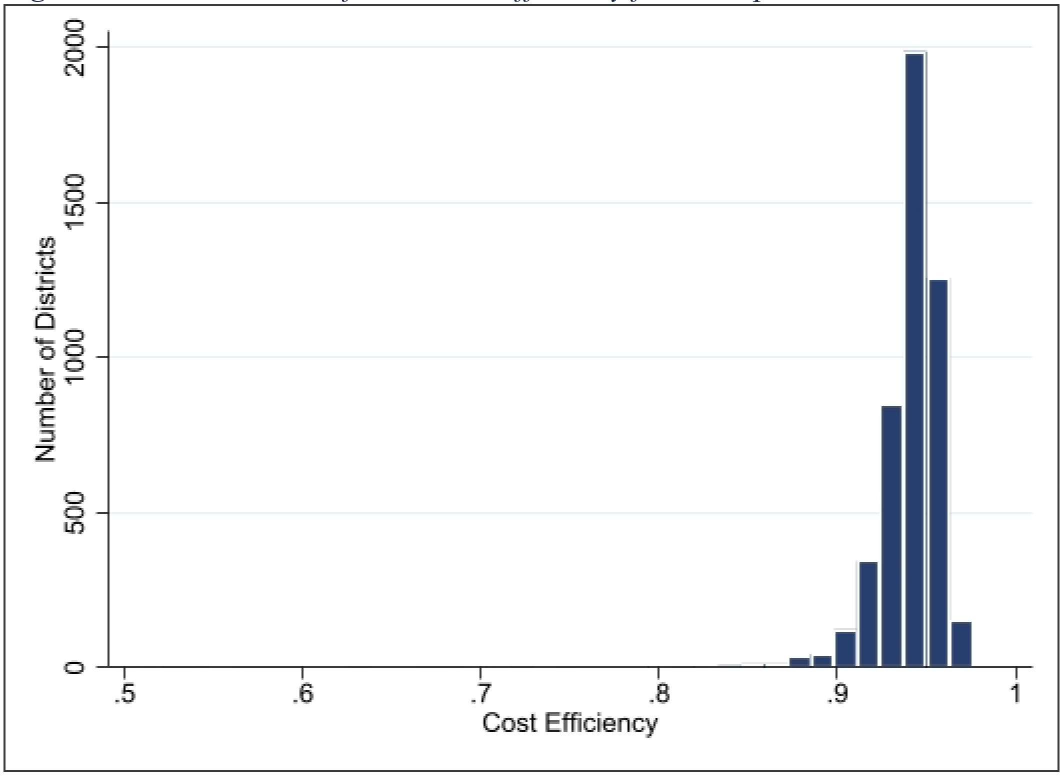


Source: Authors' calculations.

Efficiency Results

The one sided error is a measure of inefficiency. Alternatively, efficiency is measured as the complement of inefficiency, and is calculated as exp(-u), basically one minus percent inefficiency. A graph of efficiency is presented in Figure G-7, where 1 is 100% efficient. The mean efficiency is .9401, just above 94%. The lowest value in our sample is .5080, and the highest value is .9760. Clearly values are clustered near the median, .9443, and highly skewed. The first percentile is at .8686, and the tenth percentile .9206, indicating that almost all the districts had an efficiency value greater than 86% and a large majority over 92%.

Figure G-7: Distribution of Estimated Efficiency for Transportation Cost Function



Source: Authors' calculations.

One hypothesis would be that the lower the Herfindahl index, the more the competition leads school districts to pursue efficiency in the provision of services including transportation services. Certainly, this is what we found for educational services, although the impact on efficiency in transportation services is less clear.

In fact we find that the most concentrated districts are estimated to be the most efficient suppliers of student transportation services.

Calculating Adjustments Based on Marginal Effects

For the three principal exogenous input cost factors in our model—fuel price, labor price, and population density—we used the estimated marginal effect to generate a set of cost allotment adjustment factors. Using 2018–19 data, we first divide the input data into quartiles for each variable. For each quartile, we calculate the percentage difference between the median input price within the quartile and the minimum input price for the sample. We then multiply this percentage deviation of the mid-quartile price from the minimum price times the input's estimated marginal effect to yield a predicted percentage increase in cost per mile due to the higher fuel price. We then treat the predicted percentage cost increase as a transportation allotment cost adjustment factor for all districts in that fuel price quartile. For example, the first quartile of the diesel fuel price in our 2018–19 sample ranges from the sample minimum of \$2.106 per gallon to \$2.253 per gallon. The median quartile fuel price of \$2.229 per gallon is 5.8% higher than the minimum price. The marginal effect for fuel price is estimated to be \$0.0051 (for a 1% increase in fuel price), so the 5.8% higher fuel price is estimated to increase cost per mile by \$0.027. The fuel price adjustment factor assigned to districts in the first quartile is 0.03. This fuel price adjustment would increase the regular program allotment rate to \$1.03 for these districts. We repeat this process for the other three fuel price quartiles, and end up with four fuel price adjustment factors, one for each of the quartile fuel price ranges. We apply the same process to generate four quartile adjustment factors for the transportation labor wage index, the other key exogenous input price cost factor.

For population density, the key environmental cost factor, we modify the adjustment factor generating process slightly. Population density in Texas ranged widely, from a minimum density of 0.6 to a maximum of 2,718. The population density distribution is also highly skewed, with over half of the densities below 50, and a 75th percentile value of only 209.5. Also, the estimated marginal effect of population density is negative, so higher population density represents a more advantageous cost environment. We select the 90th percentile value of 769.9 as our reference density, and the percentage deviation of the median quartile density values from 769.9 times the estimated marginal cost per mile effect for population density (- \$0.0035 per mile) to generate our population density adjustment factors. This approach assigns a zero density adjustment factor to all districts in the top decile of the population density distribution.

Using 2018–19 data, we first divide the input data into quartiles for each variable. For each quartile, we calculated the percentage difference between the median input price within the quartile and the minimum input price for the sample. We then multiplied this percentage deviation of the mid-quartile price from the minimum price times the input's estimated marginal effect to yield a predicted percentage increase in cost per mile due to the higher fuel price. We then treated the predicted percentage cost increase as a transportation allotment cost adjustment factor for all districts in that fuel price quartile. For example, the first quartile of the diesel fuel price in our 2018–19 sample ranged from the sample minimum of \$2.106 per gallon to \$2.253 per gallon. The median quartile fuel price of \$2.229 per gallon is 5.8% higher than the minimum price. The marginal effect for fuel price is estimated to be \$0.0051 (for a 1% increase in fuel price), so the 5.8% higher fuel price is estimated to increase cost per mile by \$0.027. The fuel price adjustment factor assigned to districts in the first quartile is 0.03. This fuel price adjustment would increase the regular program allotment rate to \$1.03 for these districts. We repeat this process for the other three fuel price quartiles, and end up with four fuel price adjustment factors, one for each of the quartile fuel price ranges. We apply the same process to generate four quartile adjustment factors for the transportation labor wage index, the other key exogenous input price cost factor.

For population density, the key environmental cost factor, we modify the adjustment factor generating process slightly. Population density in Texas ranged widely, from a minimum density of 0.6 to a maximum of 2,718. The population density distribution is also highly skewed, with over half of the densities below 50, and a 75th percentile value of only 209.5. Also, the estimated marginal effect of population density is negative, so higher population density represents a more advantageous cost environment. We select the 90th percentile value of 769.9 as our reference density, and the percentage deviation of the median quartile density values from 769.9 times the estimated marginal cost per mile effect for population density (- \$0.0035 per mile) to generate our population density adjustment factors. This approach assigns a zero density adjustment factor to all districts in the top decile of the population density distribution.

Regression Results

Table G-8: Coefficient Estimates for First Stage and For Final Stage Transportation Cost Regressions

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
Log of District Total Vehicles	0.1058	0.4879***	0.316
	(0.149)	(0.119)	(0.362)
Log of District Total Vehicles, sq.	0.0116	0.0075	0.1150***
	(0.011)	(0.009)	(0.022)
Log of District Total Vehicles*Log of Riders per Mile			-0.0384
			(0.028)
Log of District Total Vehicles*Log of Total Route Miles			-0.1578***
			(0.036)
Log of District Total Vehicles*Log of Fuel Price	-0.0413	-0.0421	-0.0001
	(0.061)	(0.048)	(0.104)
Log of District Total Vehicles*Log of APCI	-0.2105	-0.8088***	1.4563***
	(0.157)	(0.125)	(0.294)
Log of District Total Vehicles*Log of Population Density	-0.0367***	0.0288***	0.0118
	(0.013)	(0.010)	(0.019)
Log of District Total Vehicles*Log of Roadway Congestion	0.0011	0.0590***	-0.036
	(0.026)	(0.021)	(0.036)
Log of District Total Vehicles*Percent Special Program Riders	-0.0012	0.0005	0.0072***
	(0.001)	(0.001)	(0.002)
Log of District Total Vehicles*Percent Special Program Miles	0.0027***	-0.0026***	0.0001
	(0.001)	(0.001)	(0.001)
Log of District Total Vehicles*Percent District Buses Less than 5 y.o.	-0.0003	-0.0002	0.0016**
	(0.000)	(0.000)	(0.001)
Log of District Total Vehicles*Rural District	0.0131	-0.0654**	0.2387***
	(0.034)	(0.027)	(0.053)
Log of District Total Vehicles*Micropolitan District	0.1101***	-0.0068	0.0685
	(0.029)	(0.023)	(0.053)
Log of Riders per Mile			0.7904***
			(0.219)

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
Log of Riders per Mile, sq.			0.0459***
			(0.012)
Log of Riders per Mile*Log of Total Route Miles			0.0585***
			(0.023)
Log of Riders per Mile*Log of Fuel Price			-0.0633
			(0.066)
Log of Riders per Mile*Log of APCI			-0.6333***
			(0.193)
Log of Riders per Mile*Log of Population Density			-0.0505***
			(0.011)
Log of Riders per Mile*Log of Roadway Congestion			0.0147
			(0.021)
Log of Riders per Mile*Percent Special Program Riders			-0.0083***
			(0.001)
Log of Riders per Mile*Percent Special Program Riders			0.0024***
			(0.001)
Log of Riders per Mile*Percent District Buses Less than 5 y.o.			0.0001
			(0.000)
Log of Riders per Mile*Rural District			-0.0263
			(0.034)
Log of Riders per Mile*Micropolitan District			-0.0398
			(0.033)
Log of Total Route Miles			-0.1395
			(0.31)
Log of Total Route Miles, sq.			0.0653***
			(0.016)
Log of Total Route Miles*Log of Fuel Price			0.0157
			(0.085)
Log of Total Route Miles*Log of APCI			-0.3656
			(0.245)
Log of Total Route Miles*Log of Population Density			-0.0281**
			(0.015)
Log of Total Route Miles*Log of Roadway Congestion			-0.0016

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
			(0.028)
Log of Total Route Miles*Percent Special Program Riders			-0.0077***
			(0.002)
Log of Total Route Miles*Percent Special Program Miles			0.0004
			(0.001)
Log of Total Route Miles*Percent District Buses Less than 5 y.o.			-0.0013**
			(0.001)
Log of Total Route Miles*Rural District			-0.0816**
			(0.041)
Log of Total Route Miles*Micropolitan District			-0.1047***
			(0.040)
Log of Fuel Price	2.0681*	0.4688	-2.6556***
	(1.231)	(0.980)	(1.029)
Log of Fuel Price, sq.	-0.4742	-0.9184	1.9390***
	(0.724)	(0.577)	(0.460)
Log of Fuel Price*Log of APCI	-1.5175	0.3856	1.3425**
	(0.935)	(0.745)	(0.594)
Log of District Total Vehicles	0.1058	0.4879***	0.316
	(0.149)	(0.119)	(0.362)
Log of District Total Vehicles, sq.	0.0116	0.0075	0.1150***
	(0.011)	(0.009)	(0.022)
Log of District Total Vehicles*Log of Riders per Mile			-0.0384
			(0.028)
Log of District Total Vehicles*Log of Total Route Miles			-0.1578***
			(0.036)
Log of District Total Vehicles*Log of Fuel Price	-0.0413	-0.0421	-0.0001
	(0.061)	(0.048)	(0.104)
Log of District Total Vehicles*Log of APCI	-0.2105	-0.8088***	1.4563***
	(0.157)	(0.125)	(0.294)
Log of District Total Vehicles*Log of Population Density	-0.0367***	0.0288***	0.0118
	(0.013)	(0.010)	(0.019)
Log of District Total Vehicles*Log of Roadway Congestion	0.0011	0.0590***	-0.036
	(0.026)	(0.021)	(0.036)

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
Log of District Total Vehicles*Percent Special Program Riders	-0.0012	0.0005	0.0072***
	(0.001)	(0.001)	(0.002)
Log of District Total Vehicles*Percent Special Program Miles	0.0027***	-0.0026***	0.0001
	(0.001)	(0.001)	(0.001)
Log of District Total Vehicles*Percent District Buses Less than 5 y.o.	-0.0003	-0.0002	0.0016**
	(0.000)	(0.000)	(0.001)
Log of District Total Vehicles*Rural District	0.0131	-0.0654**	0.2387***
	(0.034)	(0.027)	(0.053)
Log of District Total Vehicles*Micropolitan District	0.1101***	-0.0068	0.0685
	(0.029)	(0.023)	(0.053)
Log of Riders per Mile			0.7904***
			(0.219)
Log of Riders per Mile, sq.			0.0459***
			(0.012)
Log of Riders per Mile*Log of Total Route Miles			0.0585***
			(0.023)
Log of Riders per Mile*Log of Fuel Price			-0.0633
			(0.066)
Log of Riders per Mile*Log of APCI			-0.6333***
			(0.193)
Log of Riders per Mile*Log of Population Density			-0.0505***
			(0.011)
Log of Riders per Mile*Log of Roadway Congestion			0.0147
			(0.021)
Log of Riders per Mile*Percent Special Program Riders			-0.0083***
			(0.001)
Log of Riders per Mile*Percent Special Program Riders			0.0024***
			(0.001)
Log of Riders per Mile*Percent District Buses Less than 5 y.o.			0.0001
			(0.000)
Log of Riders per Mile*Rural District			-0.0263

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
			(0.034)
Log of Riders per Mile*Micropolitan District			-0.0398
			(0.033)
Log of Total Route Miles			-0.1395
			(0.31)
Log of Total Route Miles, sq.			0.0653***
			(0.016)
Log of Total Route Miles*Log of Fuel Price			0.0157
			(0.085)
Log of Total Route Miles*Log of APCI			-0.3656
			(0.245)
Log of Total Route Miles*Log of Population Density			-0.0281**
			(0.015)
Log of Total Route Miles*Log of Roadway Congestion			-0.0016
			(0.028)
Log of Total Route Miles*Percent Special Program Riders			-0.0077***
			(0.002)
Log of Total Route Miles*Percent Special Program Miles			0.0004
			(0.001)
Log of Total Route Miles*Percent District Buses Less than 5 y.o.			-0.0013**
			(0.001)
Log of Total Route Miles*Rural District			-0.0816**
			(0.041)
Log of Total Route Miles*Micropolitan District			-0.1047***
			(0.040)
Log of Fuel Price	2.0681*	0.4688	-2.6556***
	(1.231)	(0.980)	(1.029)
Log of Fuel Price, sq.	-0.4742	-0.9184	1.9390***
	(0.724)	(0.577)	(0.460)
Log of Fuel Price*Log of APCI	-1.5175	0.3856	1.3425**
	(0.935)	(0.745)	(0.594)
Log of District Total Vehicles	0.1058	0.4879***	0.316
	(0.149)	(0.119)	(0.362)
Log of District Total Vehicles, sq.	0.0116	0.0075	0.1150***
	(0.011)	(0.009)	(0.022)

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
Log of District Total Vehicles*Log of Riders per Mile			-0.0384
			(0.028)
Log of District Total Vehicles*Log of Total Route Miles			-0.1578***
			(0.036)
Log of District Total Vehicles*Log of Fuel Price	-0.0413	-0.0421	-0.0001
	(0.061)	(0.048)	(0.104)
Log of District Total Vehicles*Log of APCI	-0.2105	-0.8088***	1.4563***
	(0.157)	(0.125)	(0.294)
Log of District Total Vehicles*Log of Population Density	-0.0367***	0.0288***	0.0118
	(0.013)	(0.010)	(0.019)
Log of District Total Vehicles*Log of Roadway Congestion	0.0011	0.0590***	-0.036
	(0.026)	(0.021)	(0.036)
Log of District Total Vehicles*Percent Special Program Riders	-0.0012	0.0005	0.0072***
	(0.001)	(0.001)	(0.002)
Log of District Total Vehicles*Percent Special Program Miles	0.0027***	-0.0026***	0.0001
	(0.001)	(0.001)	(0.001)
Log of District Total Vehicles*Percent District Buses Less than 5 y.o.	-0.0003	-0.0002	0.0016**
	(0.000)	(0.000)	(0.001)
Log of District Total Vehicles*Rural District	0.0131	-0.0654**	0.2387***
	(0.034)	(0.027)	(0.053)
Log of District Total Vehicles*Micropolitan District	0.1101***	-0.0068	0.0685
	(0.029)	(0.023)	(0.053)
Log of Riders per Mile			0.7904***
			(0.219)
Log of Riders per Mile, sq.			0.0459***
			(0.012)
Log of Riders per Mile*Log of Total Route Miles			0.0585***
			(0.023)
Log of Riders per Mile*Log of Fuel Price			-0.0633
			(0.066)
Log of Riders per Mile*Log of APCI			-0.6333***

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
			(0.193)
Log of Riders per Mile*Log of Population Density			-0.0505***
			(0.011)
Log of Riders per Mile*Log of Roadway Congestion			0.0147
			(0.021)
Log of Riders per Mile*Percent Special Program Riders			-0.0083***
			(0.001)
Log of Riders per Mile*Percent Special Program Riders			0.0024***
			(0.001)
Log of Riders per Mile*Percent District Buses Less than 5 y.o.			0.0001
			(0.000)
Log of Riders per Mile*Rural District			-0.0263
			(0.034)
Log of Riders per Mile*Micropolitan District			-0.0398
			(0.033)
Log of Total Route Miles			-0.1395
			(0.31)
Log of Total Route Miles, sq.			0.0653***
			(0.016)
Log of Total Route Miles*Log of Fuel Price			0.0157
			(0.085)
Log of Total Route Miles*Log of APCI			-0.3656
			(0.245)
Log of Total Route Miles*Log of Population Density			-0.0281**
			(0.015)
Log of Total Route Miles*Log of Roadway Congestion			-0.0016
			(0.028)
Log of Total Route Miles*Percent Special Program Riders			-0.0077***
			(0.002)
Log of Total Route Miles*Percent Special Program Miles			0.0004
			(0.001)

Variables	First Stage Regression z1	First Stage Regression z2	Final Regression SFA
Log of Total Route Miles*Percent District Buses Less than 5 y.o.			-0.0013**
			(0.001)
Log of Total Route Miles*Rural District			-0.0816**
			(0.041)
Log of Total Route Miles*Micropolitan District			-0.1047***

Note: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.