$$(\ln)\mathbf{SCHOOL} = b_0 + b_1 State_i + b_2 LaborMarket_{ij} + b_3 CWI_{ij} + b_4 \mathbf{FINANCE}_{ij} + b_5 PopulationDensity_{ij} + b_6 Enrollment_{ij} + b_7 \mathbf{INDICATORS}_{ij} + b_8 Scale_{ij} + b_9 Poverty_{ij} + b_{10} SchlType_{ij} + b_{11} \mathbf{DATABASE}_{ij} + e$$

ALBERT SHANKER INSTITUTE

RUTGERS Graduate School of Education

STATE INDICATORS DATABASE

USER'S GUIDE AND CODEBOOK

VERSION 3 (RELEASED 2021)

This School Finance Indicators Database (SFID) system is a collection of data and analysis measuring the adequacy and fairness of K-12 education finance and resources. The purpose is to provide a single source of data for policymakers, the public, and researchers working in the fields of education finance and economics.

This **State Indicators Database** is the primary public product of the SFID. It contains a set of state level equity, spending, salary, staffing, and contextual measures for each state from 1993 to 2018 (not all variables are available in all years). These indicators are generated in large part using data from our District Indicators Database, available separately by request, which includes indicators constructed using American Community Survey Data, the Stanford Education Data Archive, the Schools and Staffing Survey, and other sources described herein. This dataset and its documentation are presented in a manner that is accessible to both researchers and non-researchers.

The full dataset, as well as reports, state profiles, and briefs using the data, are available at: http://schoolfinancedata.org.

SECTIONS

1. **Data use agreement**
2. **Data sources**: a list of data sources used in the School Finance Indicators Database
3. **Guide to variables**: a non-technical description of the variables included in each of 13 types of state indicators, and how they might be used and interpreted
4. **List of variables**: a list of all variables in the State Indicators Database, the years in which they are available, special notes, and technical details
5. **Changes to the dataset**: a record of significant changes since the original 2019 release

This project has in the past been supported by a grant from the William T. Grant Foundation. More information is available at:
http://wtgrantfoundation.org/browse-grants#/grant/183939

RUTGERS Graduate School of Education

ALBERT SHANKER INSTITUTE

DATA USE AGREEMENT

The School Finance Indicators Database, as well as the contents of this report, are the sole property of the authors. Public use of the datasets and results is encouraged, with proper attribution. Any alternative use of the data, models, or methods of the SFID must be approved by the authors.

You agree not to use the data sets for commercial advantage, or in the course of for-profit activities. Commercial entities wishing to use this Service should contact Rutgers University's Graduate School of Education at this link: http://gse.rutgers.edu/about/contact

You agree that you will not use these data to identify or to otherwise infringe the privacy or confidentiality rights of individuals.

THE DATA SETS ARE PROVIDED "AS IS" AND THE AUTHORS, RUTGERS UNIVERSITY, THE ALBERT SHANKER INSTITUTE, THE WILLIAM T. GRANT FOUNDATION, AND ALL OTHER ASSOCIATED PARTIES MAKE NO REPRESENTATIONS AND EXTEND NO WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED. THE ABOVE PARTIES SHALL NOT BE LIABLE FOR ANY CLAIMS OR DAMAGES WITH RESPECT TO ANY LOSS OR OTHER CLAIM BY YOU OR ANY THIRD PARTY ON ACCOUNT OF, OR ARISING FROM THE USE OF THE DATA SETS.

You agree that this Agreement and any dispute arising under it is governed by the laws of the State of New Jersey of the United States of America, applicable to agreements negotiated, executed, and performed within New Jersey.

You agree to acknowledge "*Rutgers Graduate School of Education/Albert Shanker Institute: School Finance Indicators Database*" as the source of these data. In publications, please cite the data as:

> Baker, Bruce D., Di Carlo, Matthew, Srikanth, Ajay, and Weber, Mark A. 2021. *Albert Shanker Institute/Rutgers Graduate School of Education: School Finance Indicators Database.* Retrieved from: http://www.schoolfinancedata.org.

Subject to your compliance with the terms and conditions set forth in this Agreement, Rutgers University and the Albert Shanker Institute grant you a revocable, non-exclusive, non- transferable right to access and make use of the Data Sets.

DATA SOURCES

The external data sources used in our system are:

- Decennial Housing and Income Data
- Decennial District and County Population Density, 2000 & 2010
- American Community Survey
- Education Comparable Wage Index (ECWI)
- U.S. Census Bureau Small Area Income and Poverty Estimates (SAIPE)
- F33 School District Fiscal Data, reduced dataset
- F33 School District Fiscal Data, full dataset
- NCES Common Core of Data: local education agency (LEA) level
- NCES Common Core of Data: school level (aggregated to LEA level)
- Stanford Education Data Archive (SEDA)
- NCES Schools and Staffing Survey
- Bureau of Economic Analysis
- U.S. Census Bureau Annual Survey of State and Local Finances

The data drawn from many of these sources, and details about their use, are available in the data and documentation for our District Indicators Database. The district-level data are used to construct most of the indicators in this state-level database, as described below.

GUIDE TO VARIABLES

The variables in the State Indicators Database are divided into 13 categories. In this section, we briefly describe the types of variables within each category and how they might be interpreted and used. The complete list of variables in each section, and technical details about the methods used to generate these measures, can be found in the next section ("List of Variables").

CATEGORY 1: GEOGRAPHY

The identification variables in this category indicate state (full name and two letter abbreviation), state FIPS code, Census region (four categories), and Census division (nine categories).

CATEGORY 2: FISCAL EFFORT

These are two indicators measuring direct state K-12 expenditures as a proportion of state economic capacity, with the latter represented by either gross state product (GSP) or aggregate personal income (thus generating two separate variables). Effort gauges how much states spend on K-12 education as a proportion of their capacity to raise revenue for public services. In other words, how much does each spend on education as a proportion of its total "economic pie?" States that spend a larger share of their pies can be viewed as exhibiting more "effort" toward their schools.

The two versions of the variable — i.e., spending as a proportion of GSP and spending as a proportion of aggregate personal income — are highly correlated, which means if one is relatively high or low in a given state the other will also tend to be relatively high or low.

Whereas most of the other measures in the SID focus on levels and allocation of school funding, effort is about state funding as a proportion of *potential* state funding. Users might, for instance, compare state effort with other SID variables, such as adequacy (described below). Policymakers in states with inadequate funding and low effort levels are making a deliberate choice to underfund their schools.

Note, however, that states with larger economies, such as New York and California, can put forth less effort than states with smaller economies, such as Mississippi and Alabama, but still raise the same amount of revenue. Of particular concern, then, are low capacity states in which funding is low but effort is high, as these states' smaller economies make it difficult to raise revenue for schools.

CATEGORY 3: PREDICTED REVENUE/SPENDING AND FAIRNESS

This group of indicators are revenue and spending adjusted for student and district characteristics. Specifically, the variables in this section represent predicted revenue or spending per pupil, within a given state in a given year, at four different (U.S. Census) district child poverty levels (0, 10, 20, and 30 percent), controlling for population density, district size, and labor market costs.

These statistical controls account for differences in factors, such as district size and the cost of hiring teachers and other personnel, that affect the cost of providing a given level of educational quality. Without these adjustments, comparisons are highly problematic. Districts with higher costs of living, for instance, have to spend more on salaries than lower cost of living districts; any differences in raw, unadjusted spending between these districts will at least partly reflect the cost of living differences rather than the generosity or adequacy of funding levels. Adjusted values, in contrast, permit more valid comparisons of revenue and spending, within and between states, since they allow one to compare *similar districts* within and between those states. For example, do high poverty (30 percent) districts in one state spend more than similar high poverty districts in another state? This is one way to assess adequacy — i.e., against the standard of other states' adjusted levels.

There are sets of four variables (predicted values at 0, 10, 20, and 30 percent district poverty) for each of the following revenue/spending types: per pupil federal revenue, per pupil state revenue, per pupil local revenue, per pupil state and local revenue, and per pupil current spending.

For each of these types, there is a fifth variable: a progressivity (or "fairness") measure ("substantial progressivity") that compares revenue or spending between high poverty (30 percent) districts and zero poverty districts. These variables are calculated by simply dividing the 30 percent poverty by the 0 percent poverty measure. For example, do high poverty districts in one state receive/spend more resources than otherwise similar zero poverty districts (0 percent) in the same state (values above 1 indicate progressive funding - i.e., high poverty districts receive more funding than zero poverty districts). These progressivity variables are important because districts serving higher poverty student populations must spend more to provide the same quality of education as districts serving lower poverty populations.

Users can also calculate substantial progressivity ratios between any two of the poverty levels (e.g., whether low and medium poverty districts [10 or 20 percent] spend/receive more resources than the lowest poverty [0 percent] districts in the same state).

Finally, the SID includes an additional progressivity measure, "systematic progressivity." Rather than comparing adjusted revenue/spending between districts at different poverty levels ("substantial progressivity"), this measure represents the consistency of the relationship (correlation) between labor market-centered district revenue and district poverty within each state. In states where higher poverty districts tend to receive more revenue than lower poverty districts, this number will be positive. Negative numbers indicate the opposite — i.e., higher poverty districts actually tend to receive less revenue, and vice versa. Note, however, that while substantial and systematic progressivity are fairly strongly correlated, states with positive systematic progressivity values might be substantially regressive (i.e., fairness ratios less than one), and vice-versa.

CATEGORY 4: ADEQUACY RELATIVE TO COMMON OUTCOME GOALS

The variables in this section are our primary measures of the adequacy of states' educational spending. In short, they compare how much states actually spend with how much they *would have to spend* to achieve a common "benchmark" goal — i.e., national average test scores.

All these estimates come from the National Education Cost Model (NECM), which is part of our system. The NECM defines adequacy in terms of testing outcomes, despite the fact that test scores are highly imperfect and incomplete performance measures. The purpose of testing outcomes in the NECM, rather, is to serve as a common outcome goal or benchmark, which is a requirement of any adequacy indicator. Moreover, it bears mentioning that the benchmark goal in this case — national average test scores — may be a rather modest goal, particularly for lower poverty states and districts. If a given state is spending above our estimated adequate levels, this does not mean that the state is "overspending" in any absolute sense, only that it is spending more than required to meet this common outcome goal. Adequacy is a relative concept.

That said, the estimates of actual and required spending (per pupil) are provided statewide, and also separately by district poverty quintile (that is, for the 20 percent lowest poverty districts, 20-40 percent, 40-60 percent, and so on). For example, how much would a given state have to spend (per pupil) for its highest poverty districts (top 20 percent) to achieve national average testing outcomes, and how does this compare to how much that state *actually* spends on these districts (i.e., is spending adequate to achieve average test scores)? Users can also make this comparison across an entire state, though bear in mind that, in some states, actual statewide per pupil spending is greater than required spending solely because actual spending in the lowest poverty districts is much higher than the estimated required levels, even though actual spending is below adequate levels in higher poverty districts (often far below adequate).

Although these measures are designed primarily to assess adequacy on a state-by-state basis, users can also compare adequacy between states, with a couple of caveats. First, poverty quintiles are defined state-by-state, which means, for example, that the highest poverty districts in one state may be more or less poor than the highest poverty districts in a different state. This will affect the estimates of required spending, since districts serving larger proportions of disadvantaged students will have to spend more than their counterparts in other states serving lower proportions of these students. Second, the NECM calculates required (adequate) spending levels in terms of the estimated amounts that would be necessary to achieve national average test scores. This means that a state with test scores far below the U.S. average will have higher required levels than a different state with scores closer to the average, even if those states are similar in terms of the students they serve, labor market costs, etc. Adequate spending in one state may not be adequate in a different state.

Also included in this category of SID indicators are two additional variables, again statewide and by poverty quintile. The first is enrollment. The second is a measure of the gap between “target” (national average) test scores and actual test scores (expressed in

standard deviations). This allows users to compare spending adequacy with actual testing outcomes. Insofar as the NECM calculates adequate spending based on testing outcomes, states and poverty quintiles with more adequate spending will also tend to have better testing outcomes, and states and poverty quintiles with larger gaps will require more spending to achieve the target outcomes. But users might, for instance, be interested in identifying states that are exceptions to that aggregate relationship, or in comparing these testing outcome gaps to other measures in the SID.

One final note: unlike all other variables in the SID, adequacy estimates are only available for a single year (the latest year in the dataset). This is because the NECM uses multiple previous years of data to calculate required spending for the latest year.

CATEGORY 5: PREDICTED STAFFING RATIOS AND FAIRNESS

The measures in this category are similar to the predicted revenue/spending variables in category 3. In this case, however, the measures are of teachers per student (i.e., staffing ratios), adjusted for student and district characteristics. As such, this is a measure of how states and districts *spend* resources, vis-à-vis similar districts in other states. Education is a labor-intensive enterprise — compensation accounts for a larger share of school budgets than any other type of expenditure.

As is the case with the revenue/fairness indicators, the variables in this section represent average staffing ratios, within a given state, at four different Census district poverty levels (0, 10, 20, and 30 percent), controlling for population density, district size, and labor market costs.

Once again, these statistical controls account for differences in factors, such as district size and labor market costs, that affect the cost of providing a given level of educational quality (in this case, the cost of hiring and retaining additional teachers). They therefore allow for more valid comparisons of staffing ratios, within and between states, since they mean you are comparing *similar districts* within and between those states. For example, do high poverty (30 percent) districts in one state employ, on average, more teachers per student than similar high poverty districts in another state?

These variables are also used to calculate progressivity/fairness indicators – i.e., whether high poverty districts in one state have higher/lower staffing ratios than otherwise similar low poverty districts in the same state. The dataset includes a variable comparing staffing ratios between the highest and lowest poverty districts (ratio of 30 to 0 percent poverty) but users can calculate ratios between any two of the poverty levels.

CATEGORY 6: EARLY CHILDHOOD EDUCATION COVERAGE AND GAPS

This small group of variables is focused on how many of states' young children (3 and 4 year olds) are enrolled in early childhood education programs, and how these enrollment rates vary by poverty. There are two variables representing the number of young children in each state and the number of low-income young children in each state, and

two variables measuring the number of young children and young low-income children who are enrolled in early education programs.

These two sets of variables are then used to calculate the percentage of all young children and the percentage of young low-income children who are enrolled. Finally, the SID includes a variable measuring the ratio of the low income enrollment rate to the overall enrollment rate (lower values indicate larger enrollment gaps between low income children and all children).

Like staffing ratios (category 5), this too is a measure of how states and districts spend their resources, in this case the degree to which they invest in access to early childhood education. Moreover, the comparison of enrollment rates between all young children and low income young children is a rough measure of the equality of access to early childhood education programs (or lack thereof).

CATEGORY 7: TEACHER/NON-TEACHER WAGE COMPETITIVENESS

These indicators compare teacher wages with those of similar non-teachers in each state, by teacher age. For example, do young teachers in one state earn more than their young counterparts in other states, all else being equal?

This resource allocation measure is important because, put simply, teachers are important, and how teachers are paid, while far from the only factor that matters, does influence the quality of applicants into the profession, and their retention. Note that these estimates include both public and private school teachers, though the inclusion of the latter does not affect results dramatically, as private school teachers constitute only a small share of the teacher workforce, and public/private teacher pay gaps tend not to vary widely between states.

These comparisons of teacher with non-teacher wages are derived from models that control for various factors that affect wages, such as education, hours, and age. These controls allow for better comparisons within and between states. The SID includes variables measuring both teacher and similar non-teacher wages for four age ranges (25-30, 31-40, 41-50, and 51-60), and there are four variables that calculate the teacher/non-teacher ratio at specific ages (25, 35, 45, and 55). The breakdown by age, which is a rough proxy for teacher experience, is important because the teacher/non-teacher wage gap varies by how long teachers stay in the profession.

CATEGORY 8: PREDICTED CLASS SIZE AND FAIRNESS

This group of variables measures average class size by district poverty (poverty is calculated as a percentage of average district poverty within the districts' labor markets), calculated separately for departmentalized and self-contained classes, controlling for labor market (i.e., class sizes compared between classes in schools in the same labor market).

This too is a measure of how states spend their resources. For example, do relatively high poverty districts (160% of their labor market's district average) in one state have larger classes than comparable districts in another state? The focus on district poverty stems from the fact that more affluent districts tend to have more resources to hire additional teachers, and thus offer smaller classes. Moreover, because these models use teacher-level data (from the Schools and Staffing Survey), the estimates represent actual reported class sizes (predicted at different poverty levels), rather than approximations based on aggregate staffing ratios.

These variables are also used to calculate progressivity/fairness indicators – i.e., whether relatively high poverty districts in one state, all else being equal, have higher class sizes than otherwise similar low poverty districts in the same state (with, to reiterate, poverty defined relative to the labor market in which districts are located). The dataset includes a variable comparing only the highest/lowest ratios (ratio of 160 to 60 percent of the average district poverty rate in the labor market), but users can calculate ratios between any two of the poverty levels.

Unfortunately, these variables are only available in the SID every four years due to data availability, and the latest year in which they are available is 2012, as the survey necessary to estimate these models was redesigned after 2012, and the latest data are not yet available (we will soon be assessing whether or not we will be able to estimate comparable models using the redesigned dataset).

CATEGORY 9: TEACHER SALARY COMPETITIVENESS AND FAIRNESS BY POVERTY

Rather than assessing competitiveness by comparing teachers with non-teachers (category 7), these measures assess competitiveness by comparing the salaries of teachers in districts with different poverty levels *to those of similar teachers* working in the same labor market, controlling for other factors that affect earnings, such as degree and teaching experience. Note that these models include public school teachers only.

Like the class size estimates (category 8), there are variables at different district poverty levels, with poverty again expressed as a percentage of average district poverty within the labor market. For example, how do the salaries of teachers working in districts with high relative poverty rates (160 percent of the labor market average) compare to the salaries of similar teachers in the same area (regardless of district poverty)?

This is important because higher poverty districts tend have more trouble than lower poverty districts recruiting and retaining teachers, and the former also tend to have less funding with which to pay teachers.

These indicators are also used to calculate a progressivity/fairness measure – i.e., the ratio of the competitiveness of teacher salaries in high poverty districts (160% of labor market) to that of teacher salaries in lower poverty districts (60% of the labor market).

The SID includes a variable comparing only the highest/lowest ratios (ratio of 160 to 60 percent of the labor market poverty rate), but users can calculate ratios between any two of the poverty levels.

Unfortunately, these variables are only available in the SID every four years due to data availability, and the latest year in which they are available is 2012, as the survey necessary to estimate these models was redesigned after 2012, and the latest data are not yet available (we will soon be assessing whether or not we will be able to estimate comparable models using the redesigned dataset).

CATEGORY 10: FAMILY INCOME ABOVE/BELOW FRL CUTPOINTS

This small group of contextual variables reports average income of families with incomes above and below two important income thresholds, both of which are based on the federal poverty line. The two thresholds are 130 percent of the federal poverty line (i.e., income no more than 30 percent higher than the poverty line) and 185 percent of the federal poverty line. The former (130 percent) is the eligibility cutpoint for free school lunch, and the latter (185 percent) is the cutpoint for reduced-price lunch.

Free and reduced-price lunch eligibility rates are widely used in education research as proxies for student poverty, but they are highly imperfect proxies since the actual incomes of families above or below these thresholds can vary widely between states (and districts and schools). In other words, the families below the 130 or 185 percent threshold in one state may have far lower incomes, on average, than families below the threshold in a different state. And, conversely, the incomes of families *above* the thresholds may vary as well.

Accordingly, the indicators in this group report the average income of families above and below the two thresholds, as well as the ratios of the averages of below to above for each threshold. Lower ratios indicate larger gaps — more inequality — in income between families above and below the subsidized lunch eligibility thresholds.

CATEGORY 11: COVERAGE AND CHARTER SCHOOL MARKET SHARE

The two simple contextual variables in this category are both "coverage" variables. The first is the proportion of each state's 6- to 16-year-olds enrolled in public schools and the second is the statewide share of public school students enrolled in charter schools.

Both of these coverage measures can have important implications for school finance. For example, charter school proliferation affects revenue (which is based on enrollment), and it may cause other finance-related complications related to school building utilization, transportation costs, and other areas.

CATEGORY 12: PUBLIC/NON-PUBLIC SCHOOL FAMILY INCOME GAPS

Like the measures in category 10, this set of three variables compares average incomes between two groups of families in each state: those with children attending public schools and those with children attending non-public (i.e., private) schools. These two

variables are also used to calculate a ratio of average income among public school families to that of non-public school families (ratios closer to 1 indicate smaller gaps in income between public and non-public school families).

If, for instance, private school families tend to be much more affluent than public school families, this may reflect greater demographic sorting of students into sectors, which in turn may affect school funding. In addition, some states have enacted tax credits in an effort to shift more students to private schools.

CATEGORY 13: STATE FINANCE LITIGATION

These non-numeric variables list major state finance adequacy and equity court cases and outcomes in a given state and year. Users might, for example, examine the relationship between these cases and the other indicators in the database. Note that these variables have not been updated since 2013.

LIST OF VARIABLES AND METHODOLOGY

The table below provides a list of all variables in the State Indicators Database, along with descriptions, notes, and the years in which they are available. Below the headers for each of the 13 categories are data sources and technical details. With the exception of the state, region, and the finance litigation variables (category 13), all variables are continuous (numeric).

Variable	Description	Years available	Notes
year	Year of Data	All years	Year refers to the spring of the school year. For example, 2018 refers to the 2017-2018 school year (or the 2018 fiscal year).

1. GEOGRAPHY			
stabbr	State abbreviation	All years	
state_name	State	All years	
statefip	State FIPS code	All years	Can be used to merge this dataset with other state-level datasets.
region4	Census region	All years	U.S. Census Bureau regional codes (four categories)
region9	Census division	All years	U.S. Census Bureau division codes (nine categories)

2. EFFORT

Data source(s): District Indicators Database; U.S. Bureau of Economic Analysis

Effort is calculated by dividing total state and local expenditures (direct to education) by Gross State Product (effort) or aggregate state personal income (inc_effort) in a given state and year.

NOTE: *Effort is not calculated for the District of Columbia in all years, and for Vermont in 2018.*

effort	Fiscal effort (% GSP)	1997-2000; 2002; 2004-2018	Gross State Product data from the U.S. Bureau of Economic Analysis.
inc_effort	Fiscal effort (% personal income)	1997-2000; 2002; 2004-2018	Personal income data from the U.S. Bureau of Economic Analysis.

3. PREDICTED REVENUE/SPENDING & FAIRNESS

Data source(s): District Indicators Database

Predicted values based on regression model using district level panel data. Dependent variable in the model below ($lnRESOURCES_{dy}$) can represent, for district *d* and year *y*, state revenue PP, local revenue PP, current expenditures PP, state & local combined revenue PP, or federal revenue PP, each yielding separate sets of predicted values by poverty level.

$$lnRESOURCES_{dy} = b_0 + b_1State + b_2PovRatedy + b_3State \times PovRate_{dy} + b_4ECWI_{dy} + b_5PopDens_{dy} + b_6\ PopDens_{dy} \times Enroll{<}100_{dy} + b_7PopDens_{dy} \times Enroll101to300_{dy} + b_8PopDens_{dy} \times Enroll301to600_{dy} + b9PopDens_{dy} \times Enroll601to1200_{dy} + b_{10}PopDens_{dy} \times Enroll1201to1500_{dy} + b_{11}K12District_{dy} + e_{dy}$$

Models weighted by district enrollment. Each year estimated separately. Predicted values for lnRESOURCES in a K–12 district with x% Census poverty (0/10/20/30), 2,000 or more students, in an average wage labor market (1.0 ECWI).

predicted_fedrevpp0_	Predicted federal revenue PP at 0% district poverty	1993-2018	
predicted_fedrevpp10_	Predicted federal revenue PP at 10% district poverty	1993-2018	
predicted_fedrevpp20_	Predicted federal revenue PP at 20% district poverty	1993-2018	

Variable	Description	Years available	Notes
predicted_fedrevpp30_	Predicted federal revenue PP at 30% district poverty	1993-2018	
fairness_fedrevpp	Federal revenue progressivity (30:0% poverty ratio)		
predicted_strevpp0_	Predicted state revenue PP at 0% district poverty	1993-2018	
predicted_strevpp10_	Predicted state revenue PP at 10% district poverty	1993-2018	
predicted_strevpp20_	Predicted state revenue PP at 20% district poverty	1993-2018	
predicted_strevpp30_	Predicted state revenue PP at 30% district poverty	1993-2018	
fairness_strevpp	State revenue progressivity (30:0% poverty ratio)	1993-2018	Users can also calculate 20:0 and 10:0 ratios using the variables above.
predicted_locrevpp0_	Predicted local revenue PP at 0% district poverty	1993-2018	
predicted_locrevpp10_	Predicted local revenue PP at 10% district poverty	1993-2018	
predicted_locrevpp20_	Predicted local revenue PP at 20% district poverty	1993-2018	
predicted_locrevpp30_	Predicted local revenue PP at 30% district poverty	1993-2018	
fairness_locrevpp	Local revenue progressivity (30:0% poverty ratio)	1993-2018	Users can also calculate 20:0 and 10:0 ratios using the variables above.
predicted_slocrev0_	Predicted state/local revenue PP at 0% district poverty	1993-2018	
predicted_slocrev10_	Predicted state/local revenue PP at 10% district poverty	1993-2018	
predicted_slocrev20_	Predicted state/local revenue PP at 20% district poverty	1993-2018	
predicted_slocrev30_	Predicted state/local revenue PP at 30% district poverty	1993-2018	
fairness	State and local revenue progressivity (30:0% poverty ratio)	1993-2018	Users can also calculate 20:0 and 10:0 ratios using the variables above.
syst_prog	Systematic progressivity	2007-2018	Does not use estimates from model above, but rather state and local revenue and Census poverty centered around the average of the labor market in which the district is located (in the District Indicators Database, the poverty and revenue variables are *ctr_perpov* and *ctr_slocrevpp,* respectively). The within-state/year correlations are weighted by enrollment. Estimates are not available for Vermont in 2018, due to irregularities in that state's data.
predicted_curexpp0_	Predicted current spending PP at 0% district poverty	1993-2018	Estimates are not available for Vermont in 2018, due to irregularities in that state's data.
predicted_curexpp10_	Predicted current spending PP at 10% district poverty	1993-2018	
predicted_curexpp20_	Predicted current spending PP at 20% district poverty	1993-2018	
predicted_curexpp30_	Predicted current spending PP at 30% district poverty	1993-2018	
fairness_curexpp	Spending progressivity (30:0% poverty ratio)	1993-2018	Users can also calculate 20:0 and 10:0 ratios using the variables above.

Variable	Description	Years available	Notes
4. ADEQUACY RELATIVE TO COMMON OUTCOME GOALS			
Data source(s): District Indicators Database; nationally-normed outcome measures from the Stanford Education Data Archive (SEDA) Estimates from the **National Education Cost Model** (NECM). The NECM uses nationally-normed outcomes (test scores) from the SEDA to model how much state and local spending would be required in order to achieve national average test scores from the previous year, by state-specific poverty quintile (in the list below, these are the five variables NECM_PREDCOST_Q1-Q5) and statewide (NECM_PREDCOST_STATE). These estimates can then be compared to actual spending, again by state-specific poverty quintile (NECM_PPCSTOT_Q1-Q5) and statewide (NECM_PPCSTOT_STATE). For each poverty quintile and statewide, we also include variables measuring the gaps in testing outcomes, expressed in standard deviations (NECM_OUTCOMEGAP_Q1-Q5 and NECM_OUTCOMEGAP_STATE), enrollment (NECM_ENROLL_Q1-Q5 and NECM_ENROLL_STATE), and the funding gap (NECM_FUNDINGGAP_Q1-Q5 and NECM_FUNDINGGAP_STATE), which is simply the difference ($ per pupil) between predicted required spending and current spending. Estimates for 2018 are from models that use SEDA and other data from 2015-2017. **NOTE**: *Variables in this section are not available for Hawaii (which contains only one school district) and Vermont (due to data irregularities in that state), and are only available for the highest poverty quintile in the District of Columbia.* For additional details on the methodology of the NECM, see: Baker, Bruce D., Weber, Mark, Srikanth, Ajay, Kim, Robert, and Atzbi, Michael. 2018. *The Real Shame of the Nation: The Causes and Consequences of Interstate Inequity in Public School Investments*. New Brunswick, NY: Rutgers University.			
necm_predcost_state	Required (adequate) spending PP - statewide	2018	
necm_ppcstot_state	Actual spending PP - statewide	2018	
necm_enroll_state	Enrollment - statewide	2018	Statewide enrollment may differ slightly from the sum of enrollment in the five quintiles in a few states.
necm_outcomegap_state	Test score gap b/w state and nat'l. avg. - statewide	2018	In standard deviations.
necm_fundinggap_state	Gap between actual and required spending PP - statewide	2018	This is the difference between the necm_ppcstot and necm_predcost variables.
necm_predcost_q1	Required (adequate) spending PP - lowest (Q1) poverty districts	2018	Poverty quintile 1 includes the 20 percent of districts in each state with the lowest Census poverty levels.
necm_ppcstot_q1	Actual spending PP - lowest (Q1) poverty districts	2018	
necm_enroll_q1	Enrollment - lowest (Q1) poverty districts	2018	
necm_outcomegap_q1	Test score gap b/w state and nat'l. avg. - lowest (Q1) poverty districts	2018	In standard deviations.
necm_fundinggap_q1	Gap between actual and required spending PP - lowest (Q1) poverty districts	2018	This is the difference between the necm_ppcstot and necm_predcost variables for this poverty quintile.
necm_predcost_q2	Required (adequate) spending PP - low (Q2) poverty districts	2018	
necm_ppcstot_q2	Actual spending PP - low (Q2) poverty districts	2018	
necm_enroll_q2	Enrollment - low (Q2) poverty districts	2018	
necm_outcomegap_q2	Test score gap b/w state and nat'l. avg. - low (Q2) poverty districts	2018	In standard deviations.
necm_fundinggap_q2	Gap between actual and required spending PP - low (Q2) poverty districts	2018	This is the difference between the necm_ppcstot and necm_predcost variables for this poverty quintile.
necm_predcost_q3	Required (adequate) spending PP - medium (Q3) poverty districts	2018	
necm_ppcstot_q3	Actual spending PP - medium (Q3) poverty districts	2018	
necm_enroll_q3	Enrollment - medium (Q3) poverty districts	2018	

Variable	Description	Years available	Notes
necm_outcomegap_q3	Test score gap b/w state and nat'l. avg. - medium (Q3) poverty districts	2018	In standard deviations.
necm_fundinggap_q3	Gap between actual and required spending PP - medium (Q3) poverty districts	2018	This is the difference between the necm_ppcstot and necm_predcost variables for this poverty quintile.
necm_predcost_q4	Required (adequate) spending PP - high (Q4) poverty districts	2018	
necm_ppcstot_q4	Actual spending PP - high (Q4) poverty districts	2018	
necm_enroll_q4	Enrollment - high (Q4) poverty districts	2018	
necm_outcomegap_q4	Test score gap b/w state and nat'l. avg. - high (Q4) poverty districts	2018	In standard deviations.
necm_fundinggap_q4	Gap between actual and required spending PP - high (Q4) poverty districts	2018	This is the difference between the necm_ppcstot and necm_predcost variables for this poverty quintile.
necm_predcost_q5	Required (adequate) spending PP - highest (Q5) poverty districts	2018	Poverty quintile 5 includes the 20 percent of districts in each state with the highest Census poverty levels.
necm_ppcstot_q5	Actual spending PP - highest (Q5) poverty districts	2018	
necm_enroll_q5	Enrollment - highest (Q5) poverty districts	2018	
necm_outcomegap_q5	Test score gap b/w state and nat'l. avg. - highest (Q5) poverty districts	2018	In standard deviations.
necm_fundinggap_q5	Gap between actual and required spending PP - highest (Q5) poverty districts	2018	This is the difference between the necm_ppcstot and necm_predcost variables for this poverty quintile.

5. PREDICTED STAFFING RATIOS & FAIRNESS

Data source(s): District Indicators Database

Predicted values based on regression model using district level panel data. Dependent variable in the model below is logged Staffing Ratio (lnSR).

$$lnSR_{dy} = b_0 + b_1State + b_2PovRate_{dy} + b_3State \times PovRate_{dy} + b_4ECWI_{dy} + b_5PopDens_{dy} + b_6\ PopDens_{dy} \times Enroll{<}100_{dy} + b_7PopDens_{dy} \times Enroll101to300_{dy} + b_8PopDensdy \times Enroll301to600_{dy} + b_9PopDens_{dy} \times Enroll601to1200_{dy} + b_{10}PopDens_{dy} \times Enroll1201to1500_{dy} + b_{11}K12District_{dy} + e_{dy}$$

Models weighted by district enrollment. Each year estimated separately. Predicted values for lnSR in a K–12 district with x% Census poverty (0/10/20/30), average density, 2,000 or more students, in an average wage labor market (1.0 ECWI).

predicted_tchph0_	Predicted teachers per 100 pupils at 0% district poverty	1994-2018	
predicted_tchph10_	Predicted teachers per 100 pupils at 10% district poverty	1994-2018	
predicted_tchph20_	Predicted teachers per 100 pupils at 20% district poverty	1994-2018	
predicted_tchph30_	Predicted teachers per 100 pupils at 30% district poverty	1994-2018	
fairness_tchph	Staffing (teachers per 100 pupils) progressivity (30:0% poverty ratio)	1994-2018	Users can also calculate 20:0 and 10:0 ratios using the variables above.

Variable	Description	Years available	Notes
6. EARLY CHILDHOOD COVERAGE AND GAPS			
Data source(s): American Community Survey Tabulations by state and year weighted by sampling probability.			
count	Census count of all 3-4 year olds	2000-2018	
enrolled	Census count of 3-4 year olds enrolled in school	2000-2018	
lowinc_ec	Census count of low income 3-4 year olds	2000-2018	
lowinc_enrolled	Census count of low income 3-4 year olds enrolled in schools	2000-2018	
ec_enrollshare	Percent of all 3-4 year olds enrolled in school	2000-2018	
ec_lowinc_enrollshare	Percent of low income 3-4 years olds enrolled in school	2000-2018	
ec_enrollgapratio	Ratio of low income to all enrollment rate	2000-2018	

7. TEACHER/NON-TEACHER WAGE COMPETITIVENESS			
Data source(s): American Community Survey Based on regression model of wage income for teachers and non-teachers, weighted for sampling probability. Competitive wage ratio (INCWAGE), or predicted wage of elementary and secondary teachers divided by predicted wage of nonteachers working in the same state, with a master's degree, at specific ages: $\ln INCWAGE = b_0 + b_1 State + b_2 K12Teacher + b_3 State \times K\text{–}12Teacher + b_5 Age + b_6 Masters + b_7 lnhours + b_8 Weeks/Year + e$ Estimated separately for each year, and weighted by sampling probability weight (perwt). Includes only those individuals holding a bachelor's or master's degree. State represents a matrix of state dummy variables, K12Teacher is an indicator denoting that the occupation code is for teacher and the industry code is that for elementary/secondary education. Of particular interest is the estimated differential (b_3) between teachers and non-teachers' income from wages in each state (given the baseline difference b_2 between teacher and non-teacher income from wages). Age is a matrix of dummy variables for each age from 25 to 65. Typical hours per week are logged (lnhours) and weeks per year specified as a matrix of dummy indicators. Note that these models include both public and private school teachers.			
tchsalary25_30	Predicted teacher annual wage - age 25-30	2000-2018	
nontchsal25_30	Predicted non-teacher annual wage - age 25-30	2000-2018	
tchsalary31_40	Predicted teacher annual wage - age 31-40	2000-2018	
nontchsal31_40	Predicted non-teacher annual wage - age 31-40	2000-2018	
tchsalary41_50	Predicted teacher annual wage - age 41-50	2000-2018	
nontchsal41_50	Predicted non-teacher annual wage - age 41-50	2000-2018	
tchsalary51_60	Predicted teacher annual Wage - age 51-60	2000-2018	
nontchsal51_60	Predicted non-teacher annual Wage - age 51-60	2000-2018	
sal_parity25	Teacher/non-teacher annual wage ratio - age 25	2000-2018	
sal_parity35	Teacher/non-teacher annual wage ratio - age 35	2000-2018	
sal_parity45	Teacher/non-teacher annual wage ratio - age 45	2000-2018	
sal_parity55	Teacher/non-teacher annual wage ratio - age 55	2000-2018	

Variable	Description	Years available	Notes
8. PREDICTED CLASS SIZE RATIOS AND INCOME-BASED GAPS			
Data source(s): District Indicators Database; NCES Schools and Staffing Survey (SASS) in 1994, 2000, 2004, 2008, 2012 NCES Schools and Staffing Survey school level class size measures for individual teachers merged to district level panel. Regression model estimated to school level class size measures: $ClassSize_{tdy} = b_0 + b_1State + b_2POV(ctr)_{dy} + b_3State \times POV(ctr)_{dy} + b_4SecTch_{tdy} + b_5CBSA_{dy} + e_{tdy}$ For non-rural schools, where "t" is the individual teacher for whom class size is reported, CTR_POV is the labor market centered census poverty rate of the district and CBSA is the Core Based Statistical Area fixed effect, so as to compare class sizes across teachers in schools within the same labor market. SecTch is an indicator that a teacher is a secondary grades teacher. Regression weighted for sampling probability using balanced repeated replication (brr). Regression model used to generate predicted values of departmentalized and self-contained class sizes for a teacher working in a district at 60%, 80%, 100%, 120% 140% & 160% of the labor market average district Census poverty level.			
pred_csd_pov60	Predicted dept. class size - district at 60% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_csd_pov80	Predicted dept. class size - district at 80% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_csd_pov100	Predicted dept. class size - district at 100% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_csd_pov120	Predicted dept. class size - district at 120% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_csd_pov140	Predicted dept. class size - district at 140% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_csd_pov160	Predicted dept. class size - district at 160% of labor market poverty	1994; 2000; 2004; 2008; 2012	
csd_ratio	Dept. class size progressivity (160:60% LM poverty ratio)	1994; 2000; 2004; 2008; 2012	
pred_css_pov60	Predicted self-contained class size - district at 60% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_css_pov80	Predicted self-contained class size - district at 80% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_css_pov100	Predicted self-contained class size - district at 100% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_css_pov120	Predicted self-contained class size - district at 120% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_css_pov140	Predicted self-contained class size - district at 140% of labor market poverty	1994; 2000; 2004; 2008; 2012	
pred_css_pov160	Predicted self-contained class size - district at 160% of labor market poverty	1994; 2000; 2004; 2008; 2012	
css_ratio	Self-contained class size progressivity (160:60% LM poverty ratio)	1994; 2000; 2004; 2008; 2012	

Variable	Description	Years available	Notes
9. TEACHER SALARY COMPETITIVENESS BY DISTRICT POVERTY			
Data source(s): District Indicators Database; NCES Schools and Staffing Survey (SASS) in 1994, 2000, 2004, 2008, 2012 NCES Schools and Staffing Survey teacher level salary measures for individual teachers merged to district level panel. Construction of the salary competitiveness index involves a two-step process, the first of which uses a regression model to isolate salary variation at constant degree and experience among teachers in the same labor market: Regression model estimated to school level class size measures: $\ln Salary_{tdl} = b_0 + b_1 Experience_{tdl} + b_2 Degree_{tdl} + b_3 Labor\ Market_{tdl}$ where salary is the salary for teacher "t" in district "d" in labor market "l." And where the model includes each year of experience beyond year "0" as dummy variables, and masters, specialist and doctoral degree dummy variables, and finally, a matrix of labor market fixed effects, such that the model residuals are the difference in each individual teacher's salary from the labor market average for a teacher of the same degree and experience level (for full time classroom teachers). We express this residual as a ratio of the teacher's actual salary to the labor market average (predicted value). $Competitiveness\ Ratio_{tdl} = Actual_{tdl} / Predicted_{tdl}$ In the second step, we use another regression model to determine how the competitiveness of teacher salaries varies with respect to district poverty rates, similar to our class size models above: $Competitiveness\ Ratio_{tdl} = b_0 + b_1 State + b_2 POV(ctr)_{dy} + b_3 State \times POV(ctr)_{dy} + b_4 SecTch_{tdy} + b_7 CBSA + e_{tdy}$ For non-rural schools, where "t" is the individual teacher for whom the salary competitiveness ratio is calculated, CTR_POV is the labor market centered census poverty rate of the district and CBSA is the Core Based Statistical Area fixed effect. Regression model used to generate predicted values of salary competitiveness ratio for a teacher working in a district at 60%, 80%, 100%, 120% 140% & 160% of the labor market average district census poverty level. That is, are teacher salaries more competitive in lower or higher poverty settings?			
pred_salratio_pov60	Teacher salary competitiveness ratio (district poverty 60% of LM poverty)	1994; 2000; 2004; 2008; 2012	
pred_salratio_pov80	Teacher salary competitiveness ratio (district poverty 80% of LM poverty)	1994; 2000; 2004; 2008; 2012	
pred_salratio_pov100	Teacher salary competitiveness ratio (district poverty 100% of LM poverty)	1994; 2000; 2004; 2008; 2012	
pred_salratio_pov120	Teacher salary competitiveness ratio (district poverty 120% of LM poverty)	1994; 2000; 2004; 2008; 2012	
pred_salratio_pov140	Teacher salary competitiveness ratio (district poverty 140% of LM poverty)	1994; 2000; 2004; 2008; 2012	
pred_salratio_pov160	Teacher salary competitiveness ratio (district poverty 160% of LM poverty)	1994; 2000; 2004; 2008; 2012	
sal_ratio	Teacher salary competitiveness progressivity (160:60% LM poverty ratio)	1994; 2000; 2004; 2008; 2012	

10. FAMILY INCOME ABOVE/BELOW FRL CUTPOINTS			
Data source(s): American Community Survey Tabulations by state and year weighted by sampling probability.			
ftotinc_under185pov	Average income of families with incomes below 185% federal poverty line	1994; 1996; 1998; 2000-2018	185% of the federal poverty line is the typical cutoff point for eligibility for reduced-price school lunch.
ftotinc_over185pov	Average income of families with incomes above 185% federal poverty line	1994; 1996; 1998; 2000-2018	

Variable	Description	Years available	Notes
ftotinc_under130pov	Average income of families with incomes below 130% federal poverty line	1994; 1996; 1998; 2000-2018	130% of the federal poverty line is the typical cutoff point for eligibility for free school lunch.
ftotinc_over130pov	Average income of families with incomes abve 130% federal poverty line	1994; 1996; 1998; 2000-2018	
inc_gap185_ratio	Ratio of average family income below to above 185% federal poverty line	1994; 1996; 1998; 2000-2018	
inc_gap130_ratio	Ratio of average family income below to above 130% federal poverty line	1994; 1996; 1998; 2000-2018	

11. COVERAGE AND CHARTER MARKET SHARE			
Data source(s): American Community Survey (coverage); Common Core of Data Public School Universe Survey (state_chartershare) ACS tabulations by state and year weighted by sampling probability.			
coverage	Percent of 6-16 year olds enrolled in public school	2000-2018	
state_chartershare	Charter school coverage	1993-2018	

12. PUBLIC/NON-PUBLIC SCHOOL FAMILY INCOME GAPS			
Data source(s): American Community Survey Tabulations by state and year weighted by sampling probability.			
inc_pubsch	Average income of families with public school children	2000-2018	
inc_nonpubsch	Average income families with non-public school children	2000-2018	
pubprv_incratio	Ratio of income of public school to non-public school families	2000-2018	

13. STATE FINANCE LITIGATION			
Data source(s): Compiled by authors from various sources. Has not been updated since 2013.			
case	School finance litigation - case name	1993-2013	Major equity/adequacy state finance cases in a given state and year (where/when applicable)
citation	School finance litigation - case citation	1993-2013	
ruling	School finance litigation - high court ruling	1993-2013	Ruling by high court on major state finance cases – overturned or upheld
casetype	School finance litigation - equity/adequacy	1993-2013	Whether case addressed equity, adequacy, or both

CHANGES TO THE DATASET

This section provides a record of significant changes to the dataset since the initial release of the SFID in 2019.

2020 - no significant changes

2021

- Addition of five new variables to category 4: required spending statewide (necm_predcost_state); actual spending statewide (necm_ppcstot_state); outcome gaps statewide (necm_outcomegap_state); funding gap statewide (necm_fundinggap_state) and; enrollment statewide necm_enroll_state).
- New variable labels (revised for brevity and clarity)
- Reordered variables in category 3
- Removal of 2018 data for Vermont in category 2 (all variables), category 4 (all variables), and the current spending and systematic progressivity variables from category 2, due to irregularities in that state's spending data.