Race/Ethnicity and Geographic Disparities in Learning-Adjusted Years of Schooling in the United States, 2009-2016

Hunter Wade York

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Public Health

University of Washington

2020

Committee:

Emmanuela Gakidou

Laura Dwyer-Lindgren

Program Authorized to Offer Degree:

Department of Global Health

University of Washington

Abstract

Race/Ethnicity and Geographic Disparities in Learning-Adjusted Years of Schooling in the United States, 2009-2016

Hunter Wade York

Chair of the Supervisory Committee:

Emmanuela Gakidou

Health Metrics Sciences

Learning-adjusted years of schooling (LAYS) is a metric of education that tracks both schooling completion and schooling quality. It has been used to rank the development of countries, especially as it relates to human capital. Though it can also be used to track smaller units of geography, until now it has found no such use. In this paper, I create metrics of years of schooling, learning quality, and learning-adjusted years of schooling for all school districts in the United States, school years 2009-2010 through 2015-2016. These metrics are further disaggregated by race/ethnicity, grade, and subject area where applicable. I also aggregate metrics to create state rankings and race/ethnicity comparisons in digestible formats. Most states have increased their average LAYS across the timespan observed in this paper. However, I find that large disparities in learning persist, across all states and between races/ethnicities, while disparities in school completion are comparatively smaller. Additionally, I find that Black-White gaps are growing in most states, while Hispanic-White gaps are staying relatively constant. This descriptive dataset is well-suited for analyses of policy impact at the local level, state- and county-level comparisons when aggregated, and for analyzing spatial disparities in other key social dimensions in future studies.

Contents

Introduction

Educational attainment is a key component of human capital. It is most often quantified as the years of schooling one has received over the course of one's lifetime, and it is commonly used in social science research and health research as a proxy for socioeconomic status. It is also commonly cited in aggregate as an indicator of the development of a geographic unit, from the community level[1] up to the nation-state level[2,3]. This metric, while accurately capturing time spent within educational institutions, does little to track the most commonly cited benefit of education: learned skills. The distributions of these skills and of educational attainment are likely unequal across time, space, race, and class. Despite a common intuition amongst academics and policymakers that learning outcomes are an important topic[4,5], few researchers have created comparable metrics of learning that can span borders. Additionally, no researchers have incorporated dropout rates into spatial metrics of educational quality. In this paper, I estimate years of schooling (YOS), learning quality, and learning-adjusted years of schooling (LAYS) at the school district level for all United States school districts years 2009-2016, disaggregated by four races and ethnicities using publically available data.

Quantifying Learning

The measurement of learning outcomes has lagged significantly behind tracking of the educational status of populations. This is primarily due to the available data from testing instruments being non-standard across time and space. In the United States, for example, testing is conducted at the state level due to the decentralization of our schools, leading to geographic incomparability of learning outcomes across state lines. Additionally, from one year to the next there is no guarantee that metrics of success will be comparable, leading to temporal biases in the comparability of metrics.

Internationally standardized assessments of learning also exist that have attempted to surmount obstacles posed by differences in language and curricula. Tests like the Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS), and the Progress in International Reading Literacy Study (PIRLS) are examples of this and have been used to create internationally comparable metrics of learning (see [6] for an example). However, these tests are rarely offered to all students in a country. Instead, select districts or schools are chosen, resulting in a lack of geographic granularity in their final data.

Domestically in the United States, there are some inter-state tests that are able to surmount the problem of comparability between geographic units. College preparedness tests such as the SAT, ACT, and PSAT, all are standardized across state lines, and temporal changes in testing instruments are dampened[7]. While these tests could present an opportunity for researchers to analyze US-based disparities in learning outcomes, they are not taken by all pupils[8] and are costly, introducing further compositional biases in their samples[9].

The National Assessment of Education Progress[10], an initiative mandated by Congress under the National Center for Education Statistics, a division of the US Department of Education, is an exception to this rule. This test is offered biannually in all states and serves as a gold standard data source to which more granular data are standardized in the primary dataset used in this analysis.

Even if comprehensive data on testing outcomes were available, there remains the question of the most appropriate indicator of learning to use. Commonly cited metrics such as "percent achieving proficiency" are relatively uninformative without a rigorous explanation of proficiency standards. Even with standardized proficiency standards, appropriate metrics of learning are not straightforward. The Stanford Educational Data Archive (SEDA)[11] uses a cohort-based measure that tracks improvements of a cohort as they mature across grades, claiming that this tracks school effectiveness and therefore learning. They also generate average test scores, intended to reflect educational opportunity benchmarked against a national average. While they only brand the former as a metric of learning, I see both indicators as tracking learning, though in importantly distinct ways. Average performance is an absolute metric, which, though anchored by definition to a national average, allows for the opportunity to compare across time and space. This contrasts to the cohort-based learning metric which is instead anchored on a cohort's achievement at the youngest grade observed. The bulk of this paper refers to learning metrics in age- and period-space, such that most metrics reflect the average learning outcomes for a given district-grade-subgroup-subject combination in a given year, compared to the national average for that year. In later metrics, for display purposes, these numbers are collapsed to encompass all grades and both subjects.

To date, SEDA is the only other source of metrics of learning spatially-precise beneath the state level, and my estimates are identical to theirs where data are complete. A key difference in this paper and the estimates produced by SEDA is the use of imputation for missing data points that exist both in time series for districts with incomplete data and in districts with no data whatsoever.

Quantifying Years of Schooling

Though simpler to quantify accurately than its learning counterpart, quantifying exact years of schooling also poses methodological challenges, especially for school-age populations. In order to accurately capture the amount of schooling achieved by an individual, it is necessary to know exactly when a pupil quits schooling. This is not captured by common questions of educational qualifications on surveys and censuses, since dropouts do not leave the schooling system with a qualification. For example, qualifications such as high school diplomas, bachelor's degrees, and associate's degrees would fail to capture students who never finished high school or who prematurely left college before earning a degree, resulting in an imprecise measurement of years of schooling. While improved data with more granular questions on schooling could be used for a measure of years of schooling (see[11] for a discussion on this topic), there is a significant lag in decennial census data collection and publication. Additionally, the census is not well equipped to capture granularity in the age period needed (school-age children) without access to microdata, since their most common age aggregations do not perfectly line up with this demographic group. Additionally, data necessary for calculating years of schooling is not captured by questionnaires completed by school administrators that ask for school dropout rates, since it is impossible to know the dropout pattern across grades, a necessary piece of information for calculating years of schooling.

Instead, registration systems with the capacity to track pupils across school systems are necessary to create both individual- and aggregate-level statistics of years of schooling completed. Unfortunately, in the US such a system does not exist. High school stock data poses a crude solution to this conundrum. Similar to a vital registration system that can track pupils across school systems, high school stock data tracks the number of pupils enrolled in any district annually by grade and other demographic characteristics. Unlike a true vital registration system, it is impossible to track pupils that move between districts and out of the schooling system altogether. Therefore, it is necessary to control for migration when using stock data to track years of schooling completed.

At the time of the drafting of this document, there have been no attempts to explicitly measure years of schooling attained during high school alone, which allows for metrics that are significantly less lagged than other metrics of years of schooling such as those produce by Barro & Lee[3], which are only really relevant after the majority of a cohort has completed schooling. Additionally, no other researchers have explicitly measured dropout rates or retention statistics at the school district level, though a journalist has compiled statistics reported on state websites in an online report in 2015[12], though dropout rates are an imperfect proxy for years of schooling, as previously stated.

Evidence for a Composite Indicator of Learning-Adjusted Years of Schooling

While it is necessary to also analyze both learning and years of schooling in isolation, there is a benefit to investigating the intersection of these two indicators. Just as many inequalities in America are not experienced in isolation, it is likely that schooling and learning deprivation are correlated. Composite indicators such as LAYS bring this double disadvantage to light. There is also the possibility that they are inversely correlated. This could be the case if school districts are forced to choose between two funding buckets—one devoted to maintaining school retention and one devoted to increasing learning outcomes—thereby imposing a reverse correlation between these outcomes. Again, composite measures are key to investigating how these two indicators interact. To capture both the component pieces of these complex interactions and their compounding effects, I therefore present all three metrics of schooling and learning throughout this paper: year of schooling (YOS), mean achievement, and learning-adjusted years of schooling (LAYS).

Decomposing Data by Geography and Race/Ethnicity

Ultimately schools are governed and financed at the school district level. Despite districts falling under state management, large inequities persist[13], and with over 13,500 school districts as of the 2010 US Census[14], there is a clear utility for geospatial indicators of education outcomes. While aggregate measures of learning can be useful for ranking larger geographic units, only the most spatially granular estimates of school characteristics can allow for targeted interventions.

Just as it is critical to decompose LAYS by their component pieces and to highlight geographic disparities at the finest resolution possible, it is paramount that schooling outcomes be evaluated separately by race/ethnicity, especially in the United States where race/ethnicity are essential components of social class and lived experiences.

Publicly available data, however, must not be identifiable at the individual level, leading to the masking of data referring to small numbers of students. This leads to an unfortunate scarcity of data for non-White students in many school districts where minority students number in the single digits within each grade. This problem is especially prevalent amongst Native, Asian, and Hispanic populations. Such small numbers can introduce noise in data due to both missingness and lack of precision and must be treated appropriately. Due to limitations in data quality and availability, I therefore only disaggregate these estimates into five race/ethnicity groups: Black, Hispanic, Asian, Native American/Alaska Native, and White. Because race and ethnicity are assessed separately in the learning data, The Hispanic grouping is not mutually exclusive with other groupings. All reported races are, however, mutually exclusive.

Investigating Associations between LAYS and Common Social Indicators

The achievement gap in the United States has grown significantly over the past 25 years[15]. Theories explaining this increase in inequality often implicate rising income inequality, though researchers are increasingly less concerned with family-level traits as they seek to highlight the salience of policy decisions that have led to widening inequalities. While the learning gap is likely not solely driven by individual-level decisions, aggregate indicators of individual traits such as the percent of the population in poverty, average socioeconomic status, and median income all serve as useful means of investigating why and how inequalities persist across race and class lines.

Roscigno and colleagues lay out an interpretive framework for understanding spatial inequalities in school testing data[16]. They outline how lack of resources both at the institutional and family level inhibit educational success. Central to their argument is the fact that the most poorly performing schools serve the poorest and most underresourced families. While it may seem tempting to interpret relationships between social indicators and educational outcomes as causal and directional, it is important to recognize how social capital interacts with geography in order to produce spatial disparities with large social gradients and how social capital. To that end, the associations we produce between schooling outcomes and community-level indicators are merely for descriptive purposes, and they should be used primarily to highlight how social indicators interact with race/ethnicity to manifest as inequalities in educational outcomes.

Methods

Learning

Data Inputs

Data were sourced from the Stanford Education Data Archive (SEDA)[17], itself a compendium of school test data sourced from EDFacts[18], a data system housed in the US Department of Education. This data system collects each states' standardized testing data, as mandated by law. SEDA standardizes state tests such that they are comparable across state lines and across time, since proficiency standards are not consistent across testing instruments. In brief, SEDA standardizes every state's testing instruments such that their proficiency standards line up with the National Assessment of Educational Progress (NAEP) scale. States with higher cutoffs for proficiency are placed higher on the NAEP achievement scale, while states with less strict cutoffs are placed lower on the NAEP scale[19]. SEDA does not impute data, and many districts fail to report testing numbers due to a variety of reasons including student privacy, missed tests, and other unforeseen obstacles. Therefore, the availability of a complete assessment of the status of learning competency in the US is lacking.

SEDA data are disaggregated by grade (3-8), school year (2009/10-2015/16), subject (math and English/language arts), and race/ethnicity (hereafter abbreviated as subgroup for readability). Subgroup-location-years with small numbers (less than 20 students) are masked to protect student privacy, resulting in a patchwork of data across the US.

Supplementary data were compiled for imputation and aggregation from the Common Core of Data (CCD)[20], which provides both fiscal and non-fiscal data at the school district level. Additionally, demographic data on race/ethnicity breakdowns were sourced from the American Community Survey (ACS)[21] and the 2010 US Census, both of which were obtained using Tidycensus[22].

Data on high school completion rates were sourced from the Common Core of Data non-fiscal school characteristics data for years 2007-2018. Enrollment numbers, disaggregated by race/ethnicity, sex, year, and grade, are available for each district. When analyzed for cohorts, these numbers can be used to create dropout statistics to track high school completion over time.

Table 1. Data Availability by School District and Race/Ethnicity

Race/Ethnicity	Population	Median Income	Testing Outcomes	Teacher Salaries	High School Retention	Total Districts
Asian	100.00%	64.50%	47.99%	61.52%	61.82%	568
Black	100.00%	64.41%	53.96%	61.60%	61.52%	1,901
Hispanic	100.00%	64.06%	50.67%	61.30%	63.15%	2,170
Native	100.00%	59.79%	44.81%	46.51%	57.43%	79
White	98.43%	64.02%	53.01%	58.79%	61.30%	9,423
All	98.67%	62.87%	50.15%	58.44%	58.25%	11,234

Percentages are measured such that the denominator is the intersection of all possible combinations of school districts, grades, subjects, and school-years. Therefore, the number of school districts with some data is larger than what is reported, which is the percent of all possible data at the most granular level that are complete.

Creating Retention Rates from School District Stock Data

To create high school dropout statistics that are comparable across state boundaries, I used the Common Core of Data non-fiscal student registry data. These data do not account for exogenous forces such as migration and mortality, so students may enter and leave a cohort as time progresses. While this is impossible to fully account for given the available data, I attempt to control for migration by calculating the retention rates in grades 10-12 as the change in a population's cohort over time less the average change in population for younger cohorts grades 1-6. Additionally, I incorporate a temporal smoother that averages the dropout rates in the proceeding and following years in order to allow for the analysis of school districts with very small populations that would otherwise prove difficult to analyze.

My method for deriving the probability of advancement from *grade* n to *grade* $n+1$ for *years i = 2009 to i = 2015* can be expressed via the following formula for grades 1 to 11:

$$ {}_1p_{grade}^{subgroup,year,district} = \frac{1}{3}\sum_{i=year-1}^{year+1}\left(N_{grade+1}^{subgroup,,district,year_i} - N_{grade}^{subgroup,,district,year_i}\right) $$

Which, for grades 9-11, is standardized by subtracting the average probability of advancement for grades 1-6:

$$ {}_1p_{grade}^{subgroup,year,dist.,standardized} = {}_1p_{grade}^{subgroup,year,district} - \frac{1}{6}\sum_{n=1}^{6} p_{grade=i}^{subgroup,year,district} $$

Where:

- ${}_1p_{grade}^{subgroup,year,district}$ is the probability of surviving from *grade* to the following grade,
- $N_{grade}^{subgroup,year,district_i}$ is the population reported in CCD for the subgroup in question at year *i* and the grade in question.

Supplementary Figure 1 displays the effect that the above standardization has an aggregate retention rates.

Imputing Missing Data

For aggregation purposes and for purposes of visualization, data were imputed when missing. This ensured that compositional biases did not affect aggregate estimates composed from incomplete time series. For example, school districts showing an upward trend in 2009-2012 but missing data for 2013-2016 would have an 8-year aggregated mean performance that was biased downwards without imputation, assuming increases in performance would have continued from 2013 onwards. Additionally, counties with missing data in certain school districts will be biased in their aggregated numbers if missingness is not at random spatially (if low-performing districts are missing more data than high performing districts, for example). Data completeness and summary characteristics are described in Table 1.

I use a two-step imputation process for mean achievement in learning outcomes at the district level, average teacher salary per student taught, and high school retention rates data (Supplementary Figure 1). First, I implemented a simple linear imputation for school-district/grade/subject/subgroup combinations with missing years of data. If a time-series was over half complete (i.e., five or more years out of eight total years of data), a linear model was fit to impute missing data. This model took the functional form of $Y \sim \mathrm{B}_0 + \mathrm{B}_1 * Year$ where $Year$ is a continuous variable indexing the year within the complete time series (2009 is 1, 2010 is 2, etc.). This model was run separately for every school-district/subject/subgroup/grade combination and fitted coefficients were used to predict missing values, though in the case of teacher salary data, estimates are not differentiated by grade.

To impute the remaining data in school-district/grade/subject/subgroup combinations with less than 5 years of data, I implemented a random forest algorithm that allows missingness in inputs as well as outputs to create a complete series of outputs. This approach was used because even with a covariates assembled from a variety of sources (CCD, ACS, and the US Census), a complete set of covariates was not available for all school districts. Because the time frames are nonstandard across these data sources (CCD is yearly, ACS covers 5-year spans, and the US Census is decennial), estimates are not available for all school districts as boundaries change across time. Additionally, there are limitations in data availability due to concerns such as data masking and data mismanagement.

The algorithm I implemented, MissForest[23], allows for mixed-type data (continuous and discrete covariates) and is non-parametric. It also allows for flexibility in linearity, directionality, and interactions between predictor and response variables. Due to computational constraints, I ran the algorithm separately by state and race/ethnicity.

For each race/ethnicity, complementary data for another, more data-rich subgroup (the all students population), were used as covariates alongside median income and average teacher salary per student. The data was formatted such that the algorithm estimates each school-district/year separately, allowing for estimates across grades and subgroups to influence estimates where otherwise missing.

Calculating Learning-Adjusted Years of Schooling (LAYS) and Years of Schooling

Learning-adjusted years of schooling (LAYS) are calculated under the assumption that learning data are representative (i.e., if all children grades 3-8 are on average 1 year behind, they will continue to be 1 year behind for the grades to come). Additionally, I assume that no children drop out before the beginning of 10th grade. Because compulsory education laws extend to ages 15-18, with an average of 16, I assume that each child receives at least 10 years of education. One should note that if a child repeats a grade, they will still accrue 10 years of education, even if they do not complete 10 grades of education. Children are also unable to accrue more than 13 years of schooling (representing completion of kindergarten through 12th grade), even if they repeat a grade. YOS are therefore bounded by 10 and 13, inclusive. The formula to calculate Years of Schooling (YOS) is as follows:

$$YOS^{subgroup,year,district} = 10 + {}_1p_{grade\ 9}^{subgroup,year,district} + \prod_{n=grade\ 9}^{grade\ 10} {}_1p_n^{subgroup,year,district} + \prod_{n=grade\ 9}^{grade\ 11} {}_1p_n^{subgroup,year,district}$$

Learning-adjusted years of schooling (LAYS) are derived by adding the learning components:

$$LAYS^{subgroup,year,district} = YOS^{subgroup,year,district} + Learning^{subgroup,year,district}$$

Where:

- $Learning^{subgroup,year,district}$ is the average performance across grades on standardized testing, measured in years ahead or behind grade level as compared to the national average.

Assessing Imputation Validity

Performance of both imputation methods—linear imputation and random forest imputation—was assessed in tandem. I used several out-of-sample predictive validity tasks wherein 10% of the data were removed, the imputation model was run, and estimates were compared to withheld data. Specifically, a 10% sample was withheld from each subgroup/state combination, separately for learning outcomes and for high school retention data. Each of the tasks was performed with three folds and summarized across holdouts to assess model performance.

Comparing LAYS with Community-Level Social Indicators

Due to the overwhelming number of school districts and the irregularity of the population density by race/ethnicity geographically, we randomly selected estimates of 250 district-years for each race/ethnicity as a means of visualizing associations between community-level social indicators and indicators of educational outcomes. Social indicators were sourced from SEDA's dataset, which uses American Community Survey 5-year datasets from 2006-2010 through 2012-2016[19]. Additionally, we sourced median income and average teacher salary per student from the US 2010 Census and the Common Core of Data, respectively, as previously described. These data are shown for all race/ethnicities while SEDA covariates are only provided for White, Black, and Hispanic students.

Results

Learning-Adjusted Years of Schooling, Years of Schooling, and Mean Achievement

Estimates varied from 6.88 LAYS in Oglala Lakota County School, SD, in 2015 to 16.97 LAYS in Mountain Lakes Borough School District, NJ, in 2016 (Figure 1a). The spatial distribution of mean achievement and learning-adjusted years of schooling (LAYS) follows other, similar patterns of deprivation observed in social sciences and public health. Notably, the Deep South fares worse than the northeast, and borderlands are worse off than the US interior. Years of schooling is not as geographically variable, though some counties with large American Indian/Alaska Native communities stand out, such as those in Yakima County, Washington; Oglala Lakota County, South Dakota; and the Four Corners Area.

State-level aggregations place Massachusetts at the top of US states in terms of LAYS with 13.88 LAYS, and Washington, DC, is last with 10.53 LAYS (Figure 1b). Disparities in LAYS are largely driven by mean achievement, which ranges from +1.13 grade levels in Massachusetts to -1.53 grade levels in Washington, DC. However, years of schooling is also quite variable, ranging from 12.94 YOS in Utah to 12.06 in Washington, DC.

In nearly all states, Black students have the poorest learning outcomes (Figure 2). In some states with large native population, such as Alaska and Arizona, American Indian/Alaska Native students fare worse than their Black counterparts. Years of schooling show different patterns of disparities, with Hispanic students, Black students, and Native students receiving the fewest years of schooling in varying states, indicating a less rigid hierarchy of school completion. Whereas learning disparities may be the result of lack of community investment in school quality or the ability of families to invest in learning outside of school, the disparities in years of schooling may due in larger part to labor forces and external pressures to begin earning money as a worker sooner.

White and Asian students leave school with the best learning outcomes and the highest completion rates. Asian students in nearly every state outpace White students academically, while their time spent in the schooling system is about even. While the relative positions of each race/ethnicity across states is relatively structured, the gaps between races/ethnicities is less standard. As previously discussed, White-Black and White-Hispanic gaps in LAYS are variable by state, and the gaps in YOS and mean achievement are even more complex.

Figure 1a.

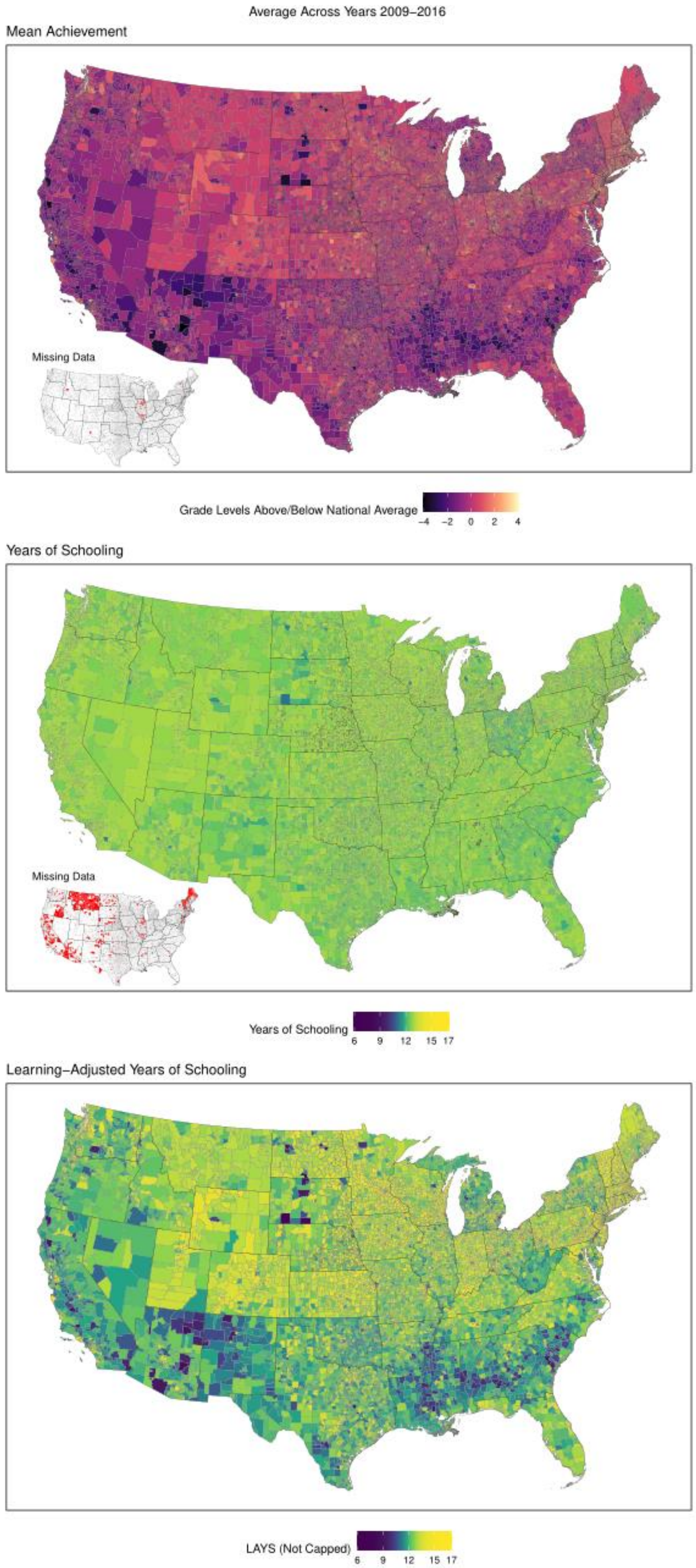


Figure 1b.

US States Ranked by Average Learning–Adjusted Years of Schooling Alongside Other Metrics of Schooling and Equity

US State/Territory, Ranked by Average LAYS 2009–2016	Average LAYS	Average YOS	Avg. Learning (Grades Ahead/Behind)	% Change LAYS 2009–2016	Avg. Black–White Gap (LAYS)	Average Hispanic–White Gap (LAYS)
Massachusetts	13.88	12.76	1.13	1.41		−2.71
Minnesota	13.65	12.87	0.78	0.73	−2.92	
Vermont	13.52	12.8	0.72	−0.52		
New Hampshire	13.51	12.58	0.93	1.95		
New Jersey	13.35	12.79	0.57	2.73	−2.81	−2.2
Washington	13.32	12.86	0.46	3.82		−2
Pennsylvania	13.29	12.77	0.52	0.18	−2.75	−2.62
Indiana	13.29	12.82	0.47	2.94	−2.21	−1.45
Wyoming	13.24	12.73	0.51	3.69		−1.61
Maine	13.23	12.83	0.39	−0.36		
Colorado	13.22	12.8	0.42	−2.72		−1.98
Virginia	13.22	12.71	0.51	2.22	−2.3	−1.65
North Dakota	13.18	12.78	0.39	0.1		
Utah	13.17	12.94	0.23	4.34		−2.36
Iowa	13.17	12.87	0.29	1.64		−1.89
Connecticut	13.16	12.72	0.44	0.11	−3.15	−3.19
Wisconsin	13.14	12.82	0.32	0.86		−2.29
Kansas	13.11	12.77	0.34	−0.62		−2.01
Nebraska	13.08	12.87	0.22	5.14		−2.04
Montana	13.04	12.69	0.35	−1.01		
New York	13.02	12.78	0.24	−1.64	−2.63	−1.92
Idaho	12.96	12.79	0.18	0.09		−1.99
Missouri	12.93	12.75	0.18	−1.02	−2.46	
Maryland	12.87	12.56	0.31	0.05	−2.98	−2.63
Ohio	12.86	12.49	0.37	1.58	−2.89	
Kentucky	12.83	12.72	0.11	0.85	−2.23	
South Dakota	12.82	12.63	0.19	−1.36		
Florida	12.82	12.68	0.14	2.24	−2.19	−1.07
North Carolina	12.8	12.6	0.21	0.79	−2.75	−2.3
Oregon	12.78	12.85	−0.07	0.69		−1.99
Rhode Island	12.73	12.65	0.08	3.46		−2.77
Michigan	12.7	12.7	0	0.55	−2.7	
Arizona	12.63	12.77	−0.15	3.5	−2.23	−2.02
Texas	12.56	12.57	−0.01	1.4	−2.31	−1.8
Delaware	12.55	12.42	0.13	−1.29	−2.6	−2.05
Arkansas	12.51	12.76	−0.25	−1.69	−2.32	−1.18
Illinois	12.49	12.68	−0.19	1.78	−2.67	−1.84
Tennessee	12.45	12.76	−0.31	6.67	−2.11	−1.55
Nevada	12.41	12.91	−0.5	1.87	−2.57	−1.77
Oklahoma	12.4	12.7	−0.3	2.63		−1.62
Georgia	12.21	12.4	−0.18	1.48	−2.49	−1.84
California	12.13	12.84	−0.71	2.76		−2.26
West Virginia	12.11	12.64	−0.53	2.25		
South Carolina	12.11	12.44	−0.33	0.36	−2.72	
Alabama	12.03	12.66	−0.63	1.07	−2.25	
Mississippi	11.74	12.63	−0.89	2.88	−1.91	
Louisiana	11.73	12.48	−0.76	0.85	−2.24	
New Mexico	11.59	12.49	−0.91	1.92		−1.84
District of Columbia	10.53	12.06	−1.53	13.12	−5.92	−5.11

Figure 1a. Average learning achievement outcomes, years of schooling, and learning-adjusted years of schooling for all districts in the United States (AK, HI, not shown). Inset maps indicate districts with no data.

Figure 1b. Rankings of states by LAYS, also showing average years of schooling, average learning, % change in LAYS, and Black- and Hispanic-White gaps in LAYS. (States with less than a 10% population of Black/Hispanic students are not shown.

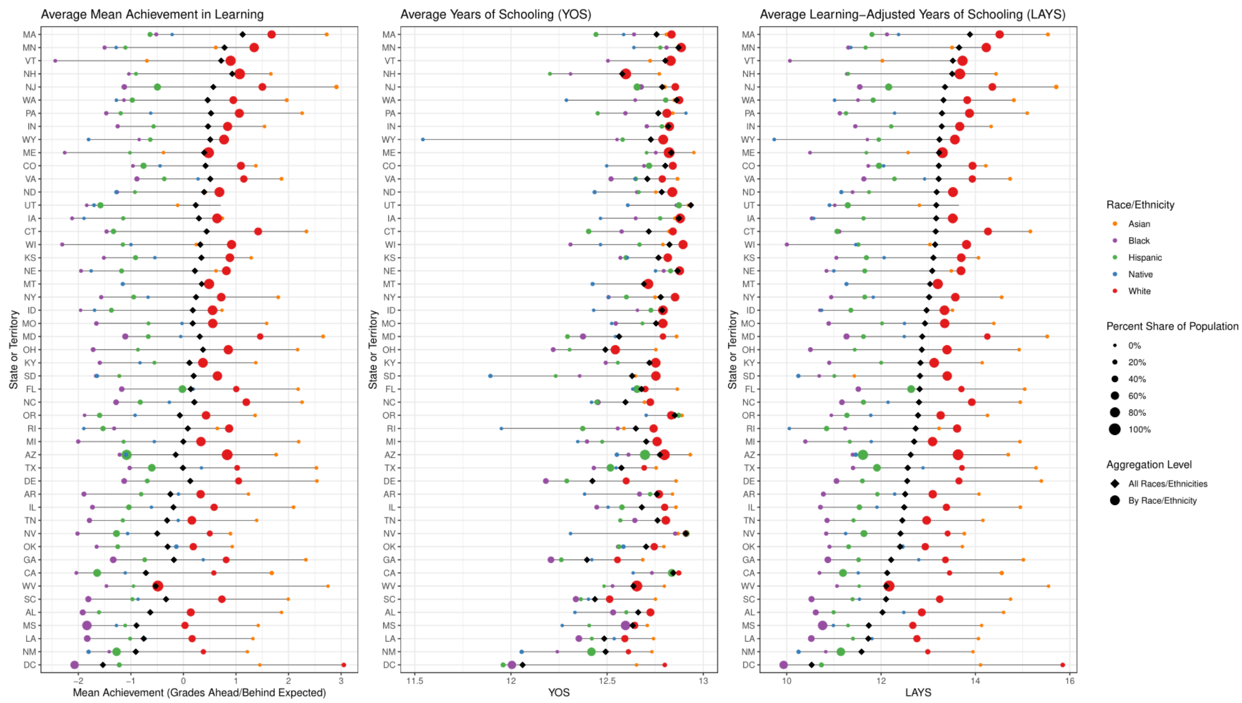


Figure 2. Average metrics of educational outcomes, disaggregated by race/ethnicity and US state/territory.

Associations of Learning-Adjusted Years of Schooling with Other Social Indicators

The distribution of learning-adjusted years of schooling at the district/race level is correlated with a number of other social indicators with a large social gradient (Figure 3). One interesting exception to this observation is the less strong correlation between average spending on teachers' salaries, standardized by the number of students in a district, and LAYS. While other indicators such as the poverty rate, the unemployment rate, and median income show clear correlations with LAYS, this indicator does not, which could indicate an inefficacy of policies directed at proximate determinants of learning at the school level that do not also tackle distal determinants of overall wellbeing.

This is consistent with the fact that teacher salaries are most variable between states, while variation in LAYS is larger between districts within states than it is between states themselves. This is due to legislation that sets teacher salary schedules at the state level, while allowing some variation between districts[24]. Additionally, this metric is sensitive to both the teacher/student ratio in a school district and to teacher pay, making it a less than ideal indicator for measuring teacher compensation or the differential spending between districts on students.

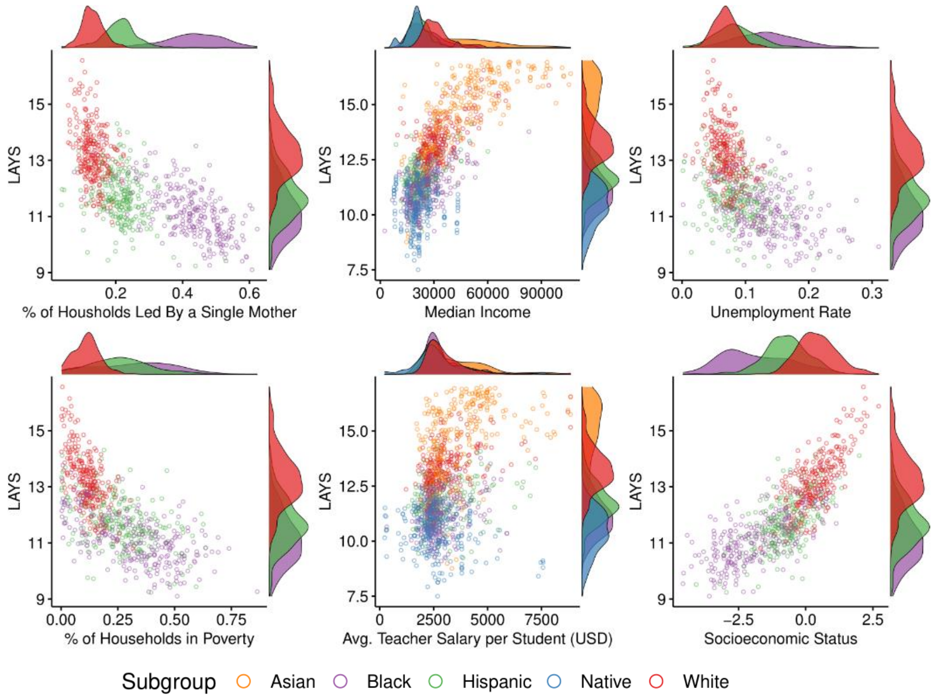


Figure 3. A 250-district sample was taken for each race/ethnicity and compared to six common social indicators. Social indicator data is missing for Asians for all indicators except median income. Average teacher salary per student is not race-/ethnicity-specific.

Gains in Learning-Adjusted Years of Schooling and Its Components from 2009-2016

Figure 4a showcases how dividends in learning, years of schooling, and LAYS are not distributed evenly across states from 2009 to 2016. Most states showed improvements in LAYS over the 7-year window, with Washington, DC, improving the most at 13.12% and Colorado fairing the worst, decreasing 2.72%.

Additionally, gains are not equally distributed amongst races and ethnicities. Figure 4b showcases how increases in schooling metrics are not evenly felt by all students. Hispanic students and White students (note that these two groups do overlap in varying amounts by state) have seen nearly equivalent increases in LAYS from 2009 to 2016, as is indicated by most states falling near the line of equivalence. Indeed, in many states, gains among Hispanic students have outpaced gains among White students. This is not the case, however, with Black students as compared to their White counterparts. Black-White and Hispanic-White gaps are large across all states with significant populations. The Black-White gap is smallest in Mississippi, with White students earning on average 1.91 LAYS more than Black students, and the Hispanic-White gap is smallest in Florida, with Hispanic students earning 1.07 LAYS less than white students. Despite Black students obtaining fewer LAYS in every state than White students at the start of the 7-year period, gains in LAYS in the period 2009 to 2016 were lesser for Black students than White students in nearly every state. In fact, while all but four states with substantial Black populations (over 10%) saw gains in LAYS for White students, over half of all states saw decreases in LAYS for Black students. These results, showing both how changes in LAYS are variable by state and by race indicate an unfortunate reality, that even when gains in education outcomes are seen within a district or a state, that they are not distributed equitably throughout the state.

Figure 4a.

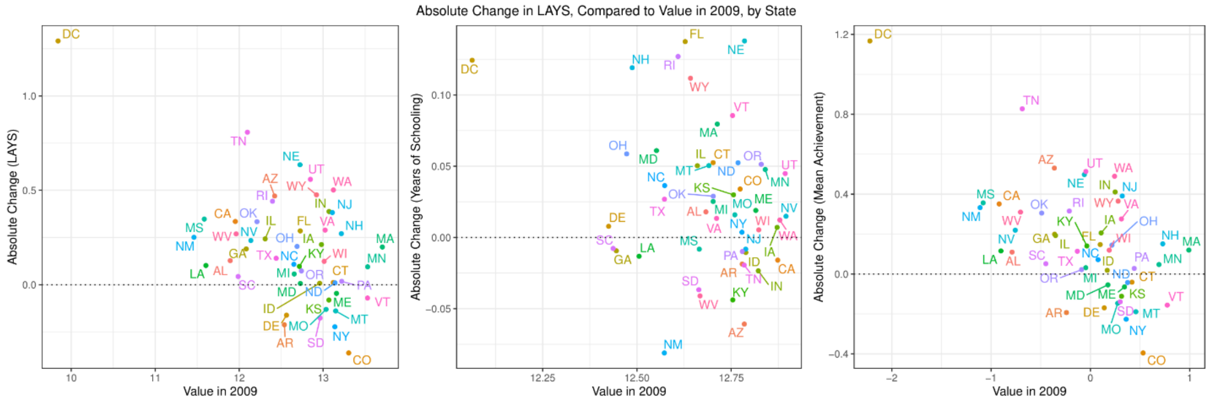


Figure 4b.

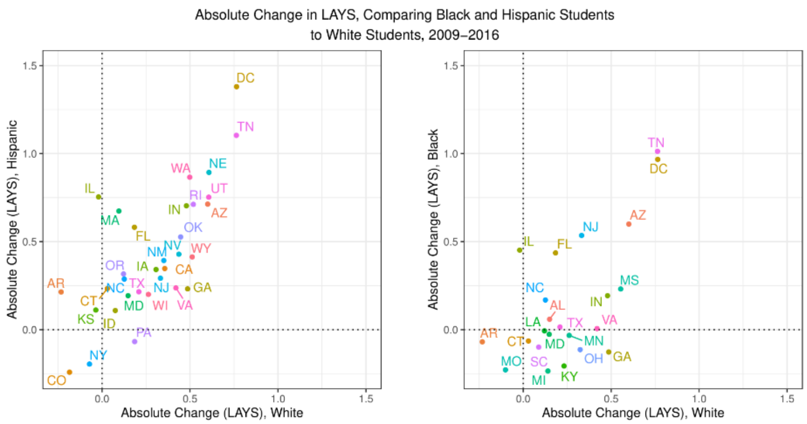


Figure 3a. Absolute change in LAYS, mean achievement, and years of schooling from year 2009 to 2016 compared to the value in 2009.

Figure 3b. Absolute change in LAYS comparing Hispanic and Black students to White students (only showing states with at least 10% of the school-age population being composed of Hispanic and Black students, respectively).

Imputation performance

Figure 5 shows performance of this models with respect to out-of-sample predictive validity, averaged across the three folds for both learning and high school retention data. Performance is consistently best for the aggregated "all" category and for White students, presumably due to a wealth of data. It follows that Asian and Native American and Alaska Native students are least accurate since they are the most data scarce categories. Overall, the average learning achievement and retention rates of all race/ethnicities were approximated well in the learning imputation task, with a median absolute error for all students being 0.343 grade-level equivalents. Given the noise in testing instruments across years, where a given cohort can perform quite erratically when measured across several grades, this is quite good performance.

Retention data (the probability of advancing from one grade to the next) is also estimated well by this model, with the median absolute error for the "all" group being 0.012. Though this model performs well, the error is biased downwards, meaning I more often underestimate the probability of advancement. Given there is a sizeable population of well-off schools that do manage to advance all of their students from one grade to the next and the bulk of the density of the probability distribution is quite close to 1, this is an expected feature of this model.

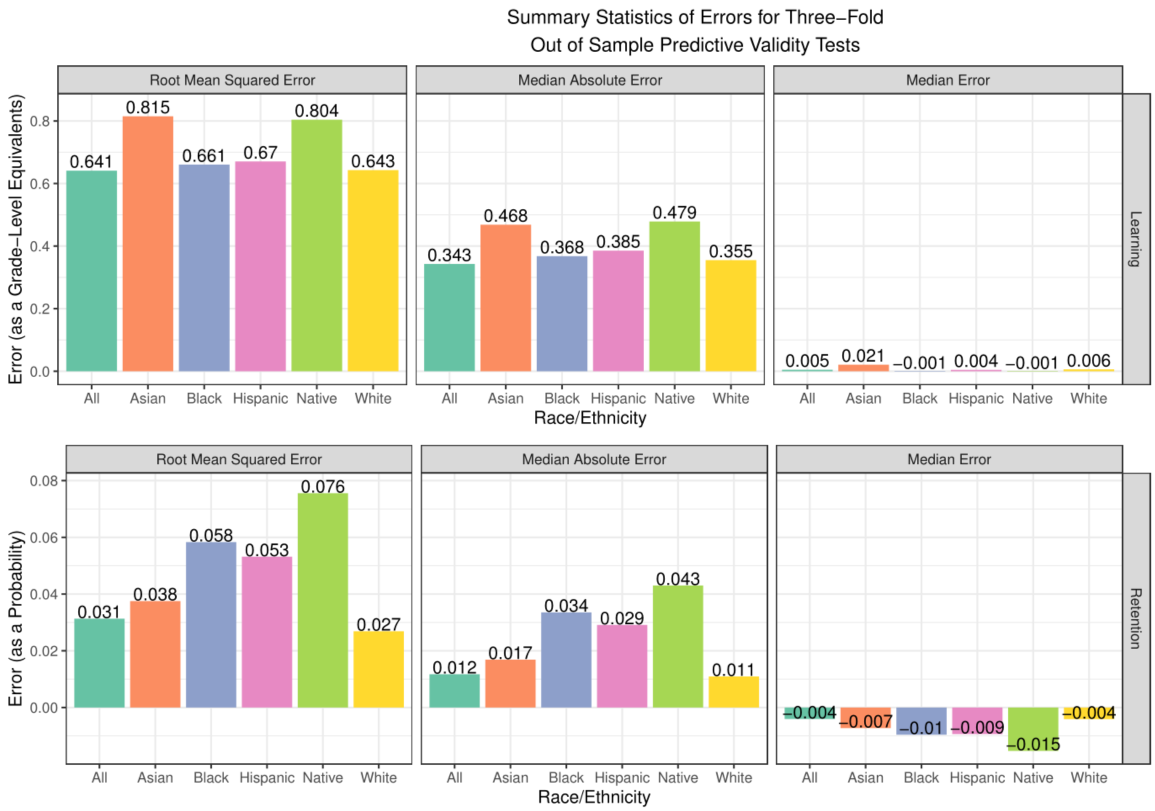


Figure 5. Summary error statistics represent three-fold out-of-sample predictive validity tasks wherein 10% of available data were withheld and predicted from the model.

Discussion

This paper has been the first to provide granular, geospatial estimates of learning-adjusted years of education (LAYS) across the United States. It is also the first study to assess years of schooling at the district level in the United States and to focus specifically on years of schooling accrued during elementary and secondary education. It complements efforts by the Stanford Educational Data Archive to measure learning quality across school districts during this time window by filling in gaps in their dataset left by missing and omitted poor-quality data using imputation methods. The estimates presented here are broken down by race/ethnicity and cover school years 2009-10 to 2015-16, and they have been aggregated to the state-level to calculate comparable rankings of educational quantity and quality over time.

Educational attainment and learning quality do not scale perfectly with each other. While some districts have succeeded in increasing both school retention and school quality, many districts fall short in one aspect or both. In the states that have excelled in both areas, students accrue upwards of 16 and 17 years of learning-adjusted education. In those districts that are doubly disadvantaged, pupils can only expect to accrue 6 to 7 years of learning-adjusted education.

The disparities in LAYS accrual are, like many things in the United States, stratified by race, space, and class. Corollaries between other common social indicators, such as those of poverty, single motherhood, and median income, are found with the distribution of LAYS. Importantly, school-level statistics such as average teacher salary per student do not seem to be as highly associated with the distribution of LAYS, and it is unclear the extent to which dividends in any of these social indicators lead to equitable advancement of learning and education across races/ethnicities.

The inequities of the relationships between social indicators and LAYS across race/ethnicities is not immediately apparent, but it is clear that there are structural forces preventing increases in LAYS across races/ethnicities at an even pace. For instance, from 2009-2016, in most states I observed Hispanic students to be increasing in LAYS in lockstep with White students, while Black students started behind and continue to stay behind, the Black-White gap continuing to grow in many states.

These patterns are likely not fixed across rural, suburban, and urban communities.[16] In fact, space has drastically different dimensions depending on context, with urban and rural school districts facing vastly different obstacles from connectivity, to funding, to labor availability, all mediated by spatial forces.[25] Whilst this dataset can be used to investigate these associations, proper policy experiments designed to test for causality are needed to fully understand how structural, family, and community forces manifest as disparities in educational outcomes.

Limitations

I present descriptive results and note that any associations remarked upon in this paper are not causal. While geographic granularity is key in interpreting underlying changes in schooling quantity and quality, these estimates are based on flawed aggregate reporting systems that can mask true individual-level results. The importance of individual-level data, for both learning and years of schooling, should be underscored in future studies, for it is impossible to know the movement of students between districts or how similar students fare in different districts with these data. Additionally, I assume learning to be static from grade 8 to graduation, since data are only available for elementary and lower secondary schooling. State tests of high school students are critical to understand not only the quality of high school but the quality of all schooling leading up to high school.

The imputation model implemented in this study design makes a fair number of tacit assumptions that deserve to be investigated further. I have brought up how covariates and education outcomes are correlated, but underlying theory should be developed to clearly highlight how individual-level factors, combined with structural forces, influence educational outcomes. All of the models for students of color use White students' performance and retention as covariates, making these models inherently normative. While there is indeed an association when viewing data in aggregate, it is unclear the extent to which this association actually is manifested spatially and when accounting for other covariates. While the imputation algorithm utilized here proves to be useful at predicting these subgroups, performance is noticeably lessened as compared to its ability to predict White students' characteristics. As is the case with many non-parametric models, associations are not explicitly returned and cannot be investigated in depth.

The high school stock model also makes some assumptions that limit the validity of these results. High school stock data are measured at varying points within the school year, and so it is unclear to what extent dropout occurs between the start of 12th grade and graduation. Additionally, more students are technically graduating earlier than ever, with cross-listed courses at community colleges and the like, and it is unclear how counts of these individuals are tallied in the framework of the data provided to us. Ultimately, these drawbacks impel us to develop systems for tracking and linking administrative data across time and across districts at the individual level.

Future Studies

LAYS were developed as a component part of human capital, and so these estimates can be combined with other metrics such as functional health status to produce the first geospatial and race/ethnicity stratified metrics of US human capital. Likewise, learning and schooling outcomes could be used in a variety of methods to examine how labor stocks respond to changes in the demographic makeup of the labor pool. Ultimately, these estimates fill an important niche that the original SEDA estimates did not fill—a complete and comprehensive dataset of learning and of educational attainment for all school districts, and therefore the entire population, in the United States.

Aggregations at the county, congressional district, and region level could also be useful in investigating policy implications. While precision descriptive analytics are most useful for policy, aggregation is essential to dissemination and storytelling and for comparisons to other aggregated indicators not available at the school district level.

Promising results are seen in many states, with LAYS increasing in more places than they are decreasing. However, progress is not universal within states, and different race/ethnicities seen gains in different amounts. Targeted approaches are necessary to ensure equitable gains in educational outcomes over the coming years, and with geospatial and race/ethnicity disaggregation, such implementation practices are now feasible.

References

1 Mapping disparities in education across low- and middle-income countries. *Nature* 2020; **577**: 235–8.

2 Friedman J, York H, Graetz N, *et al.* Measuring and forecasting progress towards the education-related SDG targets. *Nature* 2020; **580**: 636–9.

3 Barro RJ, Lee JW. A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics* 2013; **104**: 184–98.

4 Card D, Krueger AB. Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States. *Journal of Political Economy* 1992; **100**: 1–40.

5 UN. Transforming our world: the 2030 Agenda for Sustainable Development. .

6 Lim SS, Updike RL, Kaldjian AS, *et al.* Measuring human capital: a systematic analysis of 195 countries and territories, 1990–2016. *The Lancet* 2018; **392**: 1217–34.

7 Burton NW, Ramist L. Predicting Success in College: SAT® Studies of Classes Graduating since 1980. Research Report No. 2001-2. College Entrance Examination Board, 2001 https://eric.ed.gov/?id=ED562836 (accessed June 14, 2020).

8 Clark M, Rothstein J, Schanzenbach DW. Selection bias in college admissions test scores. *Economics of Education Review* 2009; **28**: 295–307.

9 Walpole M, McDonough PM, Bauer CJ, Gibson C, Kanyi K, Toliver R. This Test is Unfair: Urban African American and Latino High School Students' Perceptions of Standardized College Admission Tests. *Urban Education* 2005; **40**: 321–49.

10 NAEP Nations Report Card - National Assessment of Educational Progress - NAEP. https://nces.ed.gov/nationsreportcard/ (accessed June 5, 2020).

11 Jaeger DA. Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers. *Journal of Business & Economic Statistics* 1997; **15**: 300–9.

12 The graduation rates from every school district in one map. The Hechinger Report. 2015; published online July 2. https://hechingerreport.org/the-gradation-rates-from-every-school-district-in-one-map/ (accessed May 14, 2020).

13 Odden A, Clune WH. School Finance Systems: Aging Structures in Need of Renovation. *Educational Evaluation and Policy Analysis* 1998; **20**: 157–77.

14 Digest of Education Statistics, 2012. https://nces.ed.gov/programs/digest/d12/tables/dt12_098.asp (accessed June 14, 2020).

15 Reardon SF. The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither opportunity* 2011; **1**: 91–116.

16 Roscigno VJ, Tomaskovic-Devey D, Crowley M. Education and the inequalities of place. *Social Forces* 2006; **84**: 2121–45.

17 Reardon SF, Ho AD, Shear BR, Fahle EM, Kalogrides D, DiSalvo R. Stanford education data archive (Version 2.1). *URL http://purl stanford edu/db586ns4974* 2017.

18 U.S. Dept. of Education. EDFacts Data Files | U.S. Department of Education. 2020; published online March 13. https://www2.ed.gov/about/inits/ed/edfacts/data-files/index.html (accessed March 17, 2020).

19 Reardon SF, Kalogrides D, Shores K. Stanford Education Data Archive (SEDA): Summary of data cleaning, estimation, and scaling procedures. *Stanford Center for Education Policy Analysis https://cepa stanford edu/sites/default/files/SEDA% 20data% 20construction% 20documentation% 20v01 docx* 2016.

20 National Center for Education Statistics. Common core of data. US Department of Education Washington, DC, 2012.

21 US Census Bureau. American Community Survey. US Census Bureau Washington, DC, 2006.

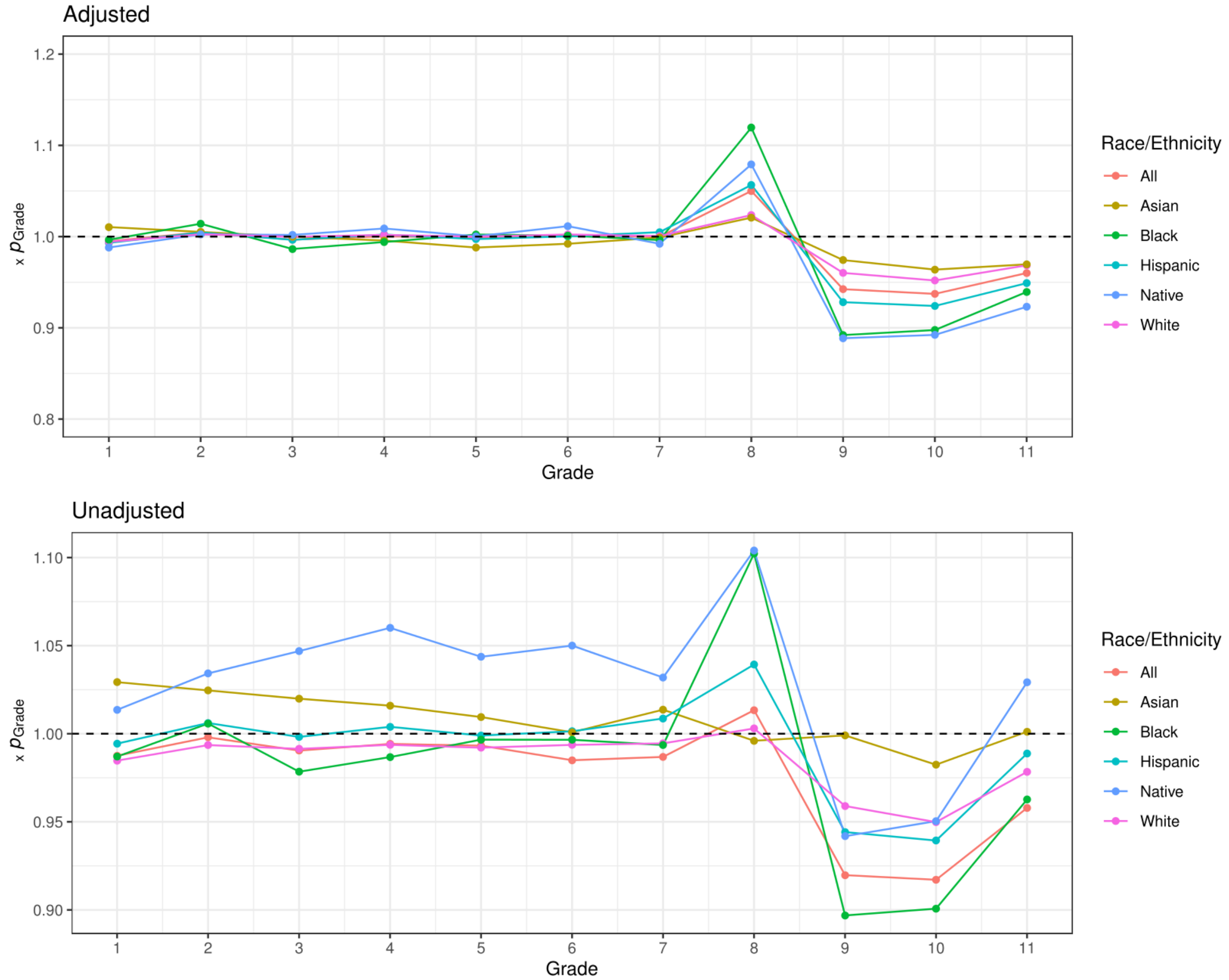
22 Walker K, Eberwein K, Herman M. Tidycensus: Load us census boundary and attribute data as' tidyverse'and'sf'-ready data frames. *URL https://CRAN R-project org/package= tidycensus R package version 04* 2018; **1**.

23 Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012; **28**: 112–8.

24 Podgursky M, Springer M. Teacher compensation systems in the United States K-12 public school system. *National Tax Journal* 2011; **64**: 165.

25 Tickamyer AR. Space Matters! Spatial Inequality in Future Sociology. *Contemporary Sociology* 2000; **29**: 805–13.

Supplement



Supplementary Figure 1. Average probabilities of advancement from one grade to the next, weighted by population, stratified by race/ethnicities.

All code used in performing these analyses is available at the following site: https://github.com/hunterwyork/geospatial_final_project