Thesis

Katrina Wheelan

April 15, 2021

Abstract

Acknowledgements

1. Introduction

Social scientists have done extensive research on the effects of wealth and racial disparities between districts on lifetime outcomes, but there has been comparatively little research on how intra-school disparities affect educational outcomes. This thesis focuses on precisely these distributional effects. In this project, I investigate the extent to which the income and racial make-ups of schools and the neighborhoods they serve affect the distribution of outcomes over time. In essence, I ask whether traditional factors of positive educational outcomes – like an influx of wealth to a neighborhood – mask adverse distributional effects. I chose to focus on Chicago Public Schools because it comprises a single school district and because there are readily available and extensive data. Chicago schools are also notorious for their racial homogeneity; in more than half of Chicago public elementary schools, at least 90% of the students identify as the same race.[1]

To answer these questions, I use several high dimensional data sources. I rely heavily on the Illinois Department of Education data for each public elementary school in Chicago. These data include detailed demographic data, standardized testing scores, and other indicators at the school-level. I also use American Community Survey data, which provides detailed metrics about the characteristics of each Census tract. I employ spatio-temporal methods and traditional linear methods to analyze the effects of racial and income disparities on the distribution of testing outcomes across space (school zones) and time (school years).

2. Literature Review

3. Data Discussion

I combined and aggregated several sources to generate a high-dimensional dataframe containing the relevant demographic and student performance data. My final dataframe contains 96 variables across 398 schools in 2009 and 356 schools in 2019. This represents 165,079 students in 2009 and 192,617 students in 2019, as well as 2,715,471 residents in 2009 and 2,714,595 residents in 2019.

3.1 Demographic Data

For this analysis, I focus on data from a 30 year period, comparing results from 1999, 2009, and 2019. I rely on several public data sources. The Illinois Board of Education provides public data on annual school "report cards," including information about school demographics and aggregate test scores. The test score data includes the number of students in each grade who test at each proficiency level. After the 2013-2014

---

[1]Calculated from Illinois Board of Education Data

academic year, Illinois transitioned from the ISAT (Illinois Standardized Assessment Test) to the PARCC (Partnership for Assessment of Readiness for College and Career), which was in turn replaced by the IAR (Illinois Assessment of Readiness) in 2019. All three exams have multiple proficiency levels, but the ISAT has only four levels, while the PARCC and IAR have five levels (did not meet, partially met, approached, met, and exceeded expectation). Although these test score data are at the grade-level, the demographic data is exclusively at the school level. I supplement this state-provided data with public shape files from the Chicago Public Schools (CPS), which contain the geographic boundaries for public elementary school neighborhoods.

Since the school demographics often differ from the neighborhood demographics, I also use the decennial U.S. Census and the American Community Survey (ACS) for neighborhood demographics. The ACS is a longer form version of the Census questionaire, and the U.S. Census Bureau administers it to a random sample of American households every year. These annual results are aggregated over five year periods. Unlike the U.S. Census, the ACS collects information about individual and household income.

I use block-level data from the decennial Census, and I use block-group and tract level data from the ACS. A Census block is a highly granular geographic division (the United States is split into more than seven million Census blocks). A block-group typically consists of 39 Census blocks and between 600 and 3,000 residents. Census tracts are slightly larger and contain one or several block groups with an optimal population of 4,000 residents. Table 1 details the specific data sources and their granularity.

Variable	Source	Time Scale	Space Scale
Population	Decennial Census	10 years	block
Race/ethnicity of residents	Decennial Census	10 years	block
Median rent	ACS	5 years	block group
Median household income	ACS	5 years	block group
Personal income category	ACS	5 years	tract
Proportion of home owners	ACS	5 years	tract
School race/ethnicity proportions	ISBE	1 year	school boundary
School low income/English learner proportions	ISBE	1 year	school boundary
School test scores by grade and subject	ISBE	1 year	school boundary
Elementary school boundaries	CPS	1 year	school boundary

Table 1: Source Data Details

Although the Census classifies "Hispanic" as an ethnic identity, the Illinois Board of Education classifies "Hispanic" as a racial category. To reconcile this, I classify any individual who identifies as Hispanic in the Census as Hispanic, while White, Black, Asian, and other racial categories include individuals who identify with those racial categories and self-describe as "non-Hispanic" on the Census.

3.2 School Data

Most Chicago elementary schools limit enrollment to their elementary school boundaries, but some schools are open enrollment or magnet schools. Magnet schools do not have attendance boundaries, so their student body may not reflect the demographics of the neighborhood as a whole.

3.3 Response Rates

The annual response rate for the ACS is consistently above 92% with the exception of 2019, when the government shutdown paused surveytaking. The response rate in 2019 was 86%.[2] Non-response to the decennial Census is penalized by law, so its (pre-COVID) reponse rates have been consistently above 99%.[3] Standardized test participation is mandated for public school students by the state of Illinois, with few exceptions (such as significant cognitive disabilities). As a result, the student participation rate is close to comprehensive; in 2019, the mean participation rate across CPS elementary schools was 97.19%.[4]

4. Methodology

4.1 Weighting Demographic Data

I investigate the role of location, racial homogeneity, income inequality, and other demographic factors in determining the distribution of student outcomes. In order to use demographic variables, I had to reconcile the geographic boundaries of the Census block, block group, and tract level data with the elementary school boundaries. In Cook County, Illinois, there are nearly 100,000 blocks, distributed among 356 elementary school boundaries. Specifically, for my geographic area of interest:

$$m = \text{number of blocks} = 99,042$$
$$n = \text{number of block groups} = 3,993$$
$$t = \text{number of tracts} = 1,319$$
$$s = \text{number of schools} = 356$$

Each Census block is fully contained in a single school boundary, so I aggregated the block-level demographics to the elementary school boundary level. However, most Census block groups and tracts are not fully contained in a single school boundary. To get the appropriate school boundary level data, I used block populations to weight the block group and tract level data.

To transform the block group tract level data to elementary school boundary level data, I contructed matrices $\mathbf{POP}(\mathbf{BG})$ and $\mathbf{POP}(\mathbf{tract})$. To contruct these matrices, I use matrix $\mathbf{Z}$ to hold information on the blocks contained in each boundary. Row k of the $\mathbf{Z}$ matrix will contain a 1 in column i when boundary k contains block i. Similarly, row i of the $\mathbf{X}$ matrix will contain a 1 in column j when block group j contains block i, and row i of the $\mathbf{Y}$ matrix will contain a 1 in column p when tract p contains block i. The center matrix in both products below are diagonal matrices containing the population of block i in row and column i.

Multiplying as below yields $s \times n$ and $s \times t$ matrices, which we will use to weight block group and tract level data, respectively. Entry $\mathbf{POP}(\mathbf{BG})_{\mathbf{kj}}$ gives the population of the overlapping area between block group j and school boundary k. Similarly, $\mathbf{POP}(\mathbf{tract})_{\mathbf{kp}}$ gives the population of the overlapping area between tract p and school boundary k.

$$\mathbf{POP}(\text{BG})_{\mathbf{s \times n}} = \begin{pmatrix} z_{11} & \dots & z_{1m} \\ \vdots & & \vdots \\ z_{s1} & \dots & z_{sm} \end{pmatrix} \begin{pmatrix} pop_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & pop_m \end{pmatrix} \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \dots & x_{mn} \end{pmatrix} \tag{1}$$

$$\mathbf{POP}(\text{tract})_{\mathbf{s \times t}} = \begin{pmatrix} z_{11} & \dots & z_{1m} \\ \vdots & & \vdots \\ z_{s1} & \dots & z_{sm} \end{pmatrix} \begin{pmatrix} pop_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & pop_m \end{pmatrix} \begin{pmatrix} y_{11} & \dots & y_{1t} \\ \vdots & & \vdots \\ y_{m1} & \dots & y_{mt} \end{pmatrix} \tag{2}$$

---

[2]https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/response-rates/

[3]source

[4]Calculated from ISBE data

Where:

$$pop_i = \text{Population of block } i$$

$$z_{ki} = \begin{cases} 0 & \text{if block } i \text{ not in school boundary } k \\ 1 & \text{if block } i \text{ in school boundary } k \end{cases}$$

$$x_{ij} = \begin{cases} 0 & \text{if block } i \text{ not in block group } j \\ 1 & \text{if block } i \text{ in block group } j \end{cases}$$

$$y_{ip} = \begin{cases} 0 & \text{if block } i \text{ not in tract } p \\ 1 & \text{if block } i \text{ in tract } p \end{cases}$$

I use the **POP(BG)** matrix to find the weighted average of the block group variables median household income and median rent. I multiply the population matrix with the vector of block level data (income and rent), and then I weight the product by the inverse of the population of the area contained in the school boundary. This yields the weighted mean of these variables for each school boundary.

$$\mathbf{INCOME_{s \times 1}} = \mathbf{POP(BG)} \begin{pmatrix} income_1 \\ \vdots \\ income_n \end{pmatrix} \begin{pmatrix} \frac{1}{\sum_{i \in school_1} pop_i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sum_{i \in school_s} pop_i} \end{pmatrix}$$

$$\mathbf{RENT_{s \times 1}} = \mathbf{POP(BG)} \begin{pmatrix} rent_1 \\ \vdots \\ rent_n \end{pmatrix} \begin{pmatrix} \frac{1}{\sum_{i \in school_1} pop_i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sum_{i \in school_s} pop_i} \end{pmatrix}$$

The tract-level variables are all counts (number of residents at each income level and number of residents who own homes). As a result, we need to contruct a weighting matrix slightly differently. We will scale **POP(tract)** with the population of each tract to get the proportion of residents living in each block of the tract. We'll call this new weight matrix **W**.

$$\mathbf{W_{s \times t}} = \mathbf{POP(tract)} \begin{pmatrix} f(1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(t) \end{pmatrix}$$

$$f(p) = \begin{cases} \frac{1}{\sum_{i \in \{1,...n\}} \mathbf{POP(tract)_{ip}}} & \text{if } \sum_{i \in \{1,...t\}} \mathbf{POP(tract)_{ip}} \neq \mathbf{0} \\ 0 & \text{if } \sum_{i \in \{1,...t\}} \mathbf{POP(tract)_{ip}} = \mathbf{0} \end{cases}$$

Then, we can weight the tract-level demographics, like the number of homeowners, as below:

$$\mathbf{HOMEOWNERS} = \mathbf{W} \begin{pmatrix} homeowners_1 \\ \vdots \\ homeowners_t \end{pmatrix}$$

This yields a an $s \times 1$ matrix with an estimate for the number of homeowners in each elementary school boundary. We can divide this result by a vector of total residents in each boundary (which we calculate similarly).

4.2 Measuring Racial Homogeneity

I considered two metrics to quantify racial homogeneity. First, I considered a Chi-squared statistic to compare a school's racial split to the Chicago Public School overall racial split. This metric essentially quantifies a school's deviation from the null distribution, but such a deviation does not necessarily mean greater racial homogeneity. I do not ultimately use this metric in my final analysis.

I also consider a second metric: the proportion of students who belong to the largest racial group in a school. This metric, lg, is rigorously defined below for school s across all racial classifications R:

$$lg_s = argmax_{i \in R}\{race.prop_{si}\} \tag{3}$$

4.3 Measuring Income Inequality

Creating a metric for income inequality in a school boundary created two significant challenges: approximating a continuous distribution from discrete, binned data; and defining a measure of disparity within this approximated distribution. The table below details the binned categorical data, already transformed from the tract to the school boundary level.

Income Range	Mean Proportion across School Boundaries
$0 - $10k	0.163
$10 - $15k	0.083
$15k - $25k	0.144
$25k - $35k	0.107
$30k - $50k	0.102
$50k - $65k	0.075
$65k - $75k	0.030
$75k+	0.123

Table 2: Personal Income

In order to meaure income inequality, we need a continuous series of income data. In order to transform the binned data into a continuous distribution, I used the `fitdistrplus` R package, which allows a user to fit a univariate distribution to "censored" data. The package's relevant function uses the Nelder-Mead method to optimize the distributions' parameters for maximum log-likelihood given the binned data.[5] Since income levels tend to follow a log-normal distribution, I fit a log-normal distribution. I also fit a gamma distribution as a sensitivity test on the inequality metric. Once I had the continuous distribution, I applied two metrics for disparity: standard deviation and the Gini coefficient.

The Gini coefficient measures half the relative mean absolute difference between every two values in a dataset, which is the average pairwise distance between sets of two points, divided by the overall mean. The denominator for the mean absolute difference is n^2 because there are n^2 pairs of values with resampling. Thus:

$$Gini = \frac{1}{2} \cdot \frac{\text{mean difference}}{\text{arithmetic mean}} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\bar{x}} \tag{4}$$

Equation 4 also has a graphical equivalent in the Lorenz Curve. The Lorenz Curve shows cumulative income share of total income on the y-axis that is held by the poorest x percent of the population. A perfectly equal

---

[5]https://www.rdocumentation.org/packages/fitdistrplus/versions/1.1-3/topics/fitdistcens

society would follow the line $y = x$, where, for instance, the poorest 80% of the population holds 80% of the total income. The Gini coefficient equals the area between a dataset's Lorenz curve and the line of equality (in blue below) divided by the total area under the line of equality (the blue and green areas together). A perfectly equal society has a Gini ceofficient of 0, and a perfectly unequal society (where one individual holds all the income) has a Gini coefficient of 1.

Lorenz Curve and Gini Coefficient
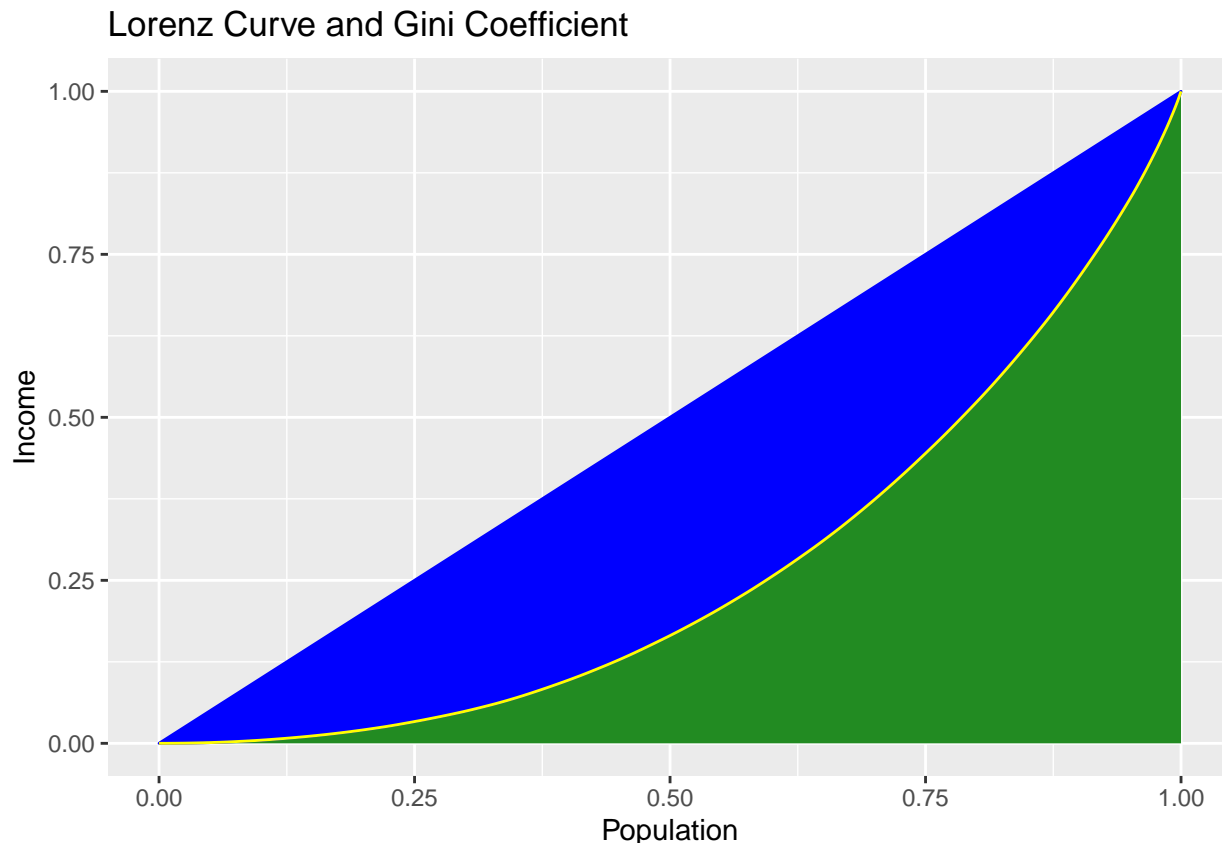
4.4 Measuring Disparities in Student Outcomes

Unfortunately, the most granular standardized testing data I could analyze was also binned data. I use math and English language arts (ELA) test scores from third and fifth grade students. The scores are binned by proficiency level. There were four levels before 2014, and there have been five levels since. The levels do not have equal bin size. See the cutoffs below for the 3rd grade ELA exam in 2019:

Proficiency Level	Lower Boundary	Upper Boundary
Did not meet expectations	650	700
Partially met expectation	700	725
Approached expectations	725	750
Met expectations	750	786
Exceeded expectations	786	850

Table 3: 2019 IAR Proficiency Levels

I considered several metrics to quantify the spread or disparity in test scores:

6

1. I assigned each proficiency level a value between 1 and 5 (and 1 to 4 for pre-2014 data), representing did not meet expectations to exceeds expectations. Then, I directly calculated the standard deviation of this set of integers. Since these values are really ordinal data and since the bins are not equal sizes for these levels, this is a fairly crude method.

2. I calculate a Chi-squared statistic for the distribution of proficiency levels, using the overall system-wide proficiency level proportions as the null distributions for each grade, year, and test subject. This metric essentially measures the deviation from the average distribution of proficiency levels.

3. I use a similar method as with the income level data to generate a maximum likelihood continuous gamma distribution from the discrete, binned data for each school, subject, grade, and year. I then calculate the standard deviation of the continuous distribution.

4.5 Measuring Spatial Auto-correlation

This project deals with spatio-temporal data. A key question for this type of data is the extent of spatial auto-correlation: Do schools that are geographically close perform similarly? Does this pattern remain when we control for neighborhood charactersitcs such as race and income?

In order to answer these questions, I use Moran's I as a metric for the level of spatial auto-correlation. Moran's I quantifies the relationship between a single variable, x, with n values, and pairwise geographic distances between locations $\{l_1, \ldots, l_n\}$, which are incorporated in a distance weight matrix $\mathbf{W}$. Each entry w_{ij} represents a spatial weight related to the distance between points i and j. The diagonals on $\mathbf{W}$ are 0. For spatial weights, I use the inverse planar distance between a pair of school boundaries' centroids.

$$\mathbf{W} = \begin{pmatrix} 0 & \ldots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \ldots & 0 \end{pmatrix}, \ \mathbf{w_{ij}} = \frac{\mathbf{1}}{\mathbf{dist(centroid_i, centroid_j)}} \tag{5}$$

Then, I use $\mathbf{W}$ to calculate Moran's I:

$$I = \left(\frac{n}{\sum_i \sum_j w_{ij}}\right) \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \tag{6}$$

5. Descriptive Statistics

5.1 Neighborhood Demographics

As mentioned in the introduction, Chicago has a legacy of dramatic residential segregation. Historically, the richest and Whitest neighborhoods have been concentrated in the northern part of the city. These residental patterns have not changed dramatically over time. The plots below illustrate the median income for each and predominant racial or ethnic group in each elementary school boundary.

5.2 School Demographics

As a result of the stark residential segregation by race (pictured in the previous section), each neighborhood is highly homogeneous. For instance, the typical predominantly Black neighborhood is 0.8945821% Black. This neighborhood homogeneity spills over into the schools. More than half of public elementary schools in Chicago have 90% or more of its students identify with the same racial/ethnic group. This effect is particularly dramatic for Black students. Although Black students only comprised about 34% of the CPS student body in 2019, more than three quarters of Black students attended a school that was at least 75% Black.
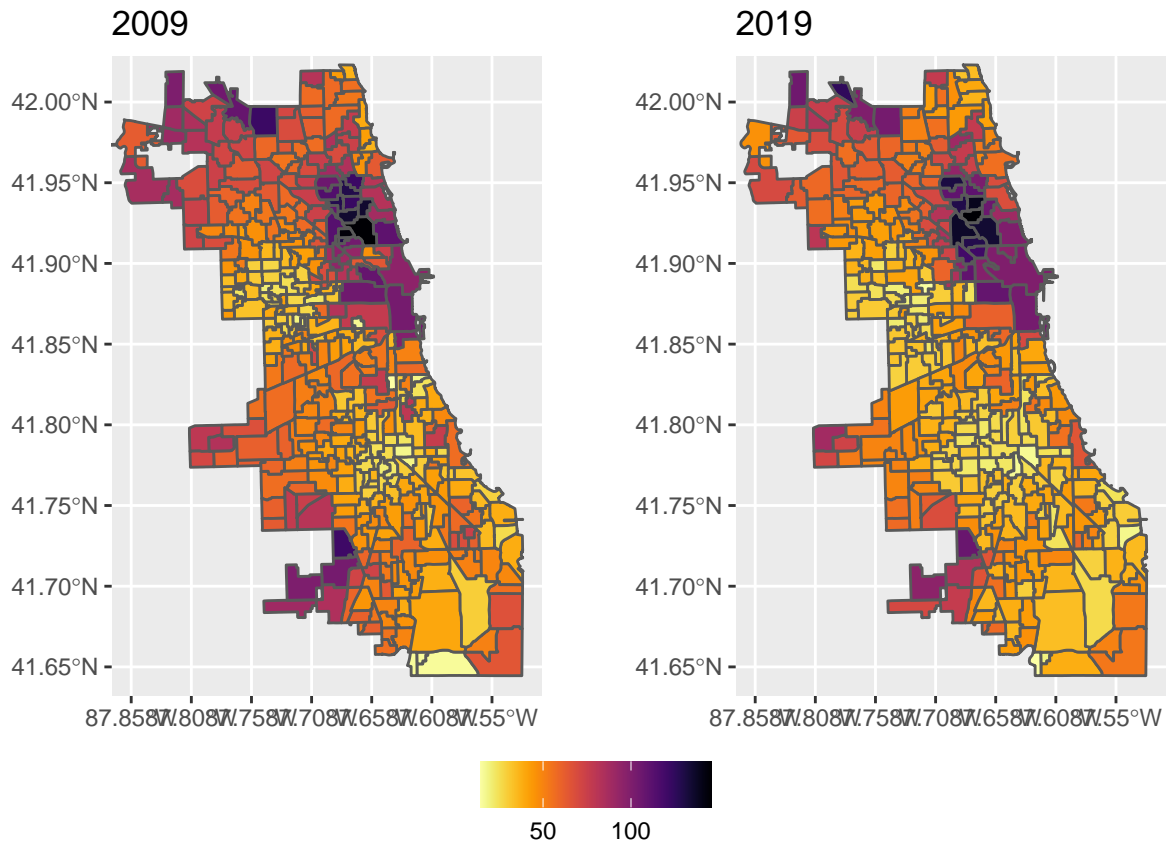
Figure 1: Median household income (in thousands of 2019 U.S. dollars) by elementary school boundary.
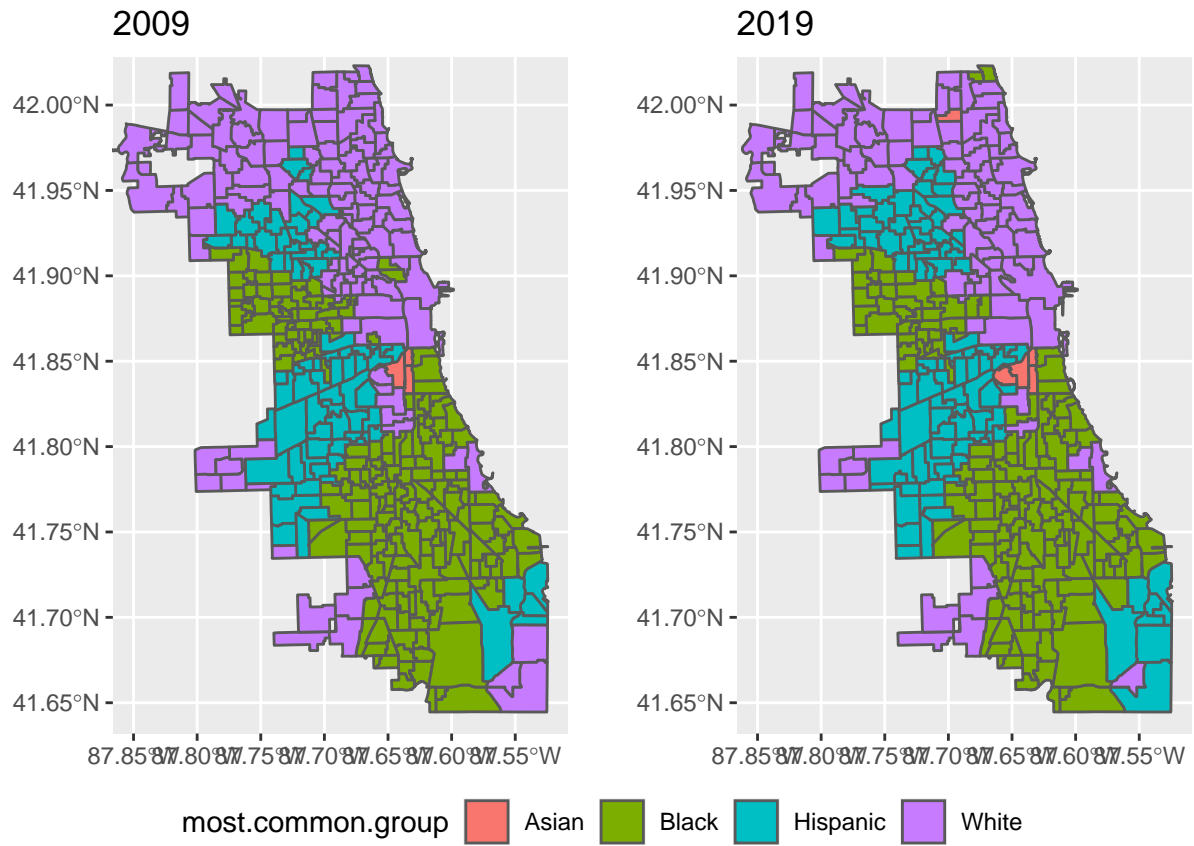
Figure 2: Largest racial/ethnic group by elementary school boundary. The largest group is defined as the most common racial/ethnic category for residents within a school boundary.

The plot below illustrates the magnitude of this homogeneity. For each district, the plot shows the proportion of students who identify with the predominant racial or ethnic group in the school. For instance, a value of 0.90 for the largest group corresponds to a school where 90% of the students identify as the same race. This metric is exactly the "largest group" metric described in equation 3 in section 4.2. The most diverse school had a "largest group" metric of 27.7% in 2019, and the least diverse schools had a metric of 100%, with complete racially homogeneous. In 2019, 144 schools had a largest group of more than 95%.
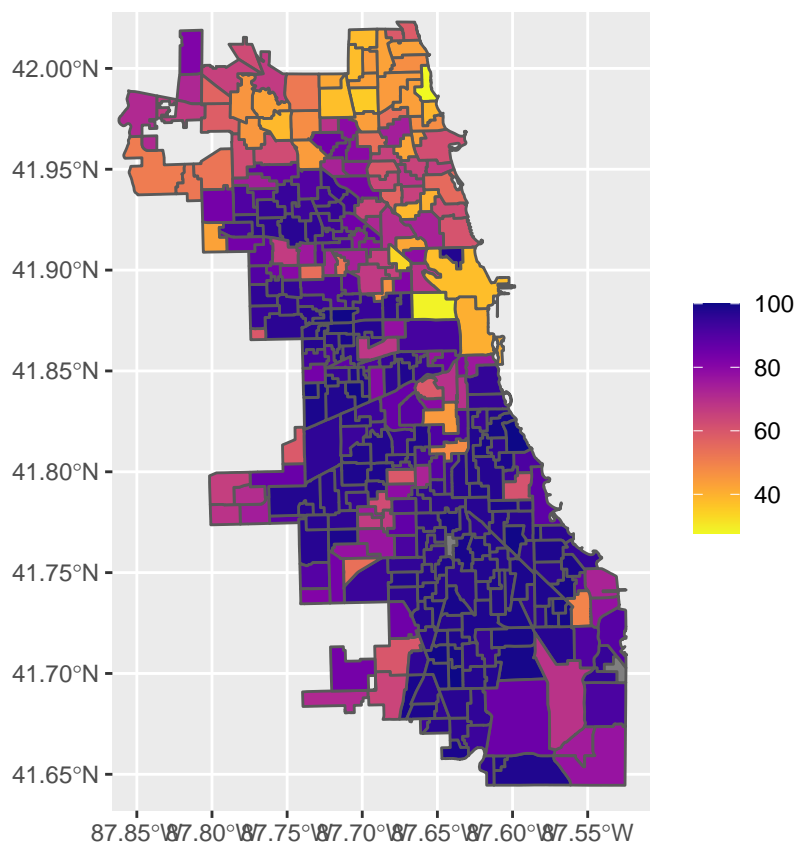

Figure 3: Proportion of students who identify with the predominant racial or ethnic group in the school. This is a measure of racial homogeneity that corresponds to equation 3 in section 4.2. Higher values indicate greater homogeneity.

School demographics generally closely follow the demographics of a boundary's district, with some exceptions. In the most recent data, 31.73% of Chicago residents were White, but only 12.5% of Chicago public elementary school students were White. This could reflects both age differences (if more White Chicagoans are not school age) and any families who opt out of the public school system (and into private schools or homeschooling).
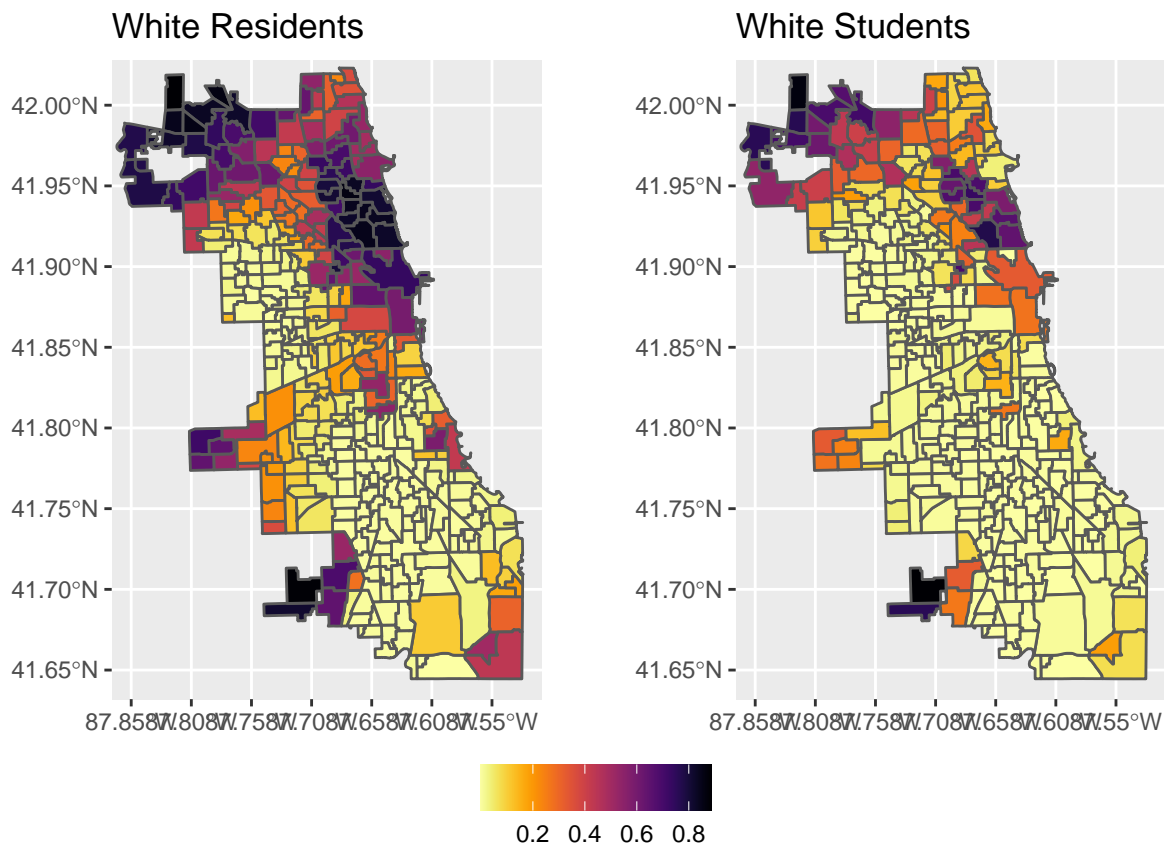

Figure 4: Proportion of White residents versus White students in 2019. In the most recent data, 31.73% of Chicago residents were White, but only 12.5% of Chicago public elementary school students were White.

5.3 School Outcomes

I measure school outcomes at the grade, subject, year, and school level. As mentioned earlier, this data is only available as the proportion of students at each the proficiency level. The data come from third and fifth grade standardized test scores, collected at 356 elementary schools serving a total of 192,617 students in 2019. In 2009, the data comes from 398, 97 schools serving 165,079 students.

Variable	Grade 3 ELA	Grade 3 Math	Grade 5 ELA	Grade 5 Math
2009				
Proportion at Level 1	0.1	0.1	0.01	0.01
Proportion at Level 2	0.35	0.21	0.43	0.33
Proportion at Level 3	0.42	0.46	0.41	0.58
Proportion at Level 4	0.13	0.22	0.13	0.06
2019				
Proportion at Level 1	0.26	0.17	0.21	0.19
Proportion at Level 2	0.16	0.25	0.28	0.34
Proportion at Level 3	0.19	0.24	0.26	0.24
Proportion at Level 4	0.33	0.27	0.21	0.17
Proportion at Level 5	0.06	0.06	0.01	0.03

Table 4: City-wide proficiency levels by year, grade, and subject.

These proficiency levels vary dramatically by school. The plots below show density plots for 3rd grade math and ELA scores in 2009 and 2019. Each colored line on the plot represents a proficiency level. The density is across all schools, so the curve represents the density (y-axis) of schools having a given proportion of students (x-axis) testing at the line's proficiency level. For instance, the green line on the top, right plot shows that most schools have around 50% of its students testing at level 3 proficiency for math. The highest proficiency levels (the blue level 4 line for the 2009, and the purple level 5 line for 2019) are heavily right-skewed, which shows that very few schools have many very high preforming students.
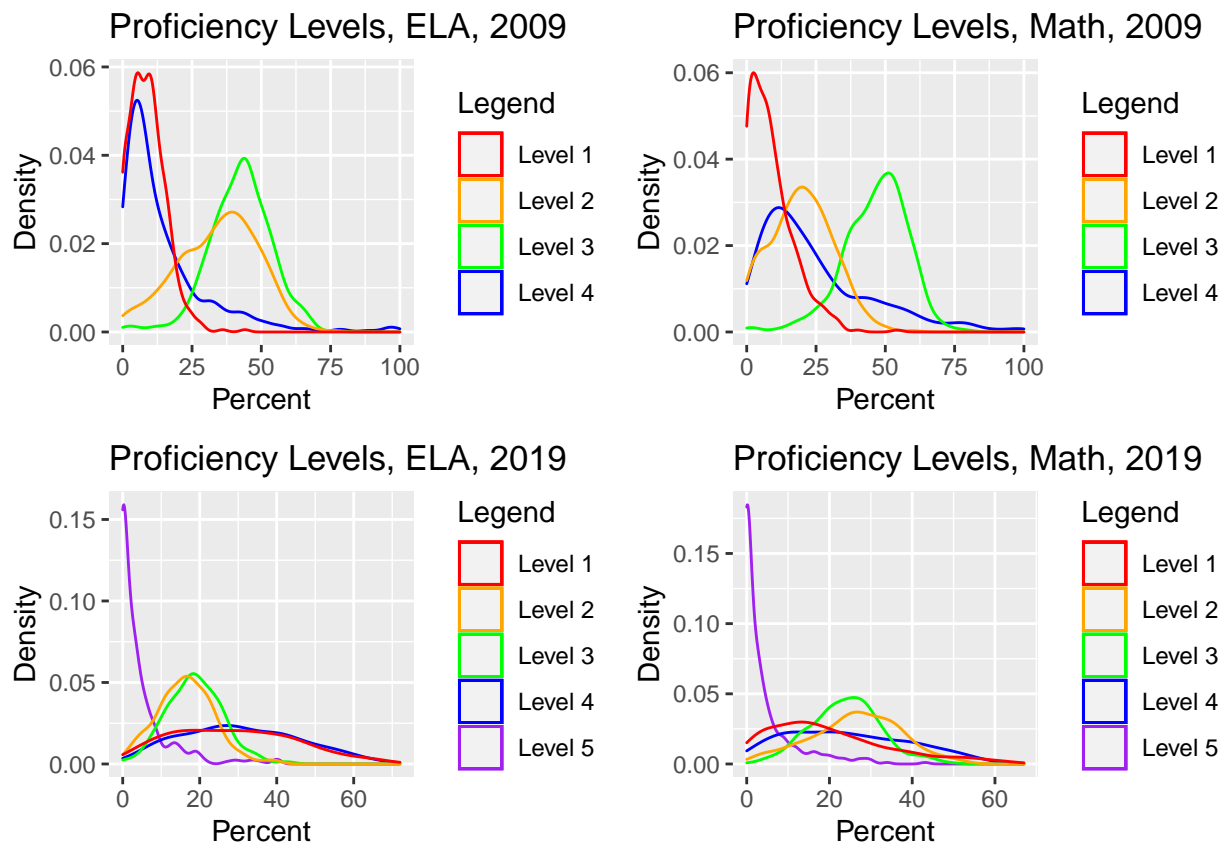
Figure 5: Density curves for proficiency levels. Each colored line on the plot represents a proficiency level. The density is across all schools, so the curve represents the density (y-axis) of schools having a given proportion of students (x-axis) testing at the line's proficiency level. or instance, the green line on the top, right plot shows that most schools have around 50% of its students testing at level 3 proficiency for math.

As mentioned in the methods section, I use three metrics for score disparities: standard deviation of proficiency levels, Chi-squared statistic, and standard deviation of an interpolated continuous distribution of scores. I ultimately chose to use the Chi-squared statistic as a metric for score disparities. Its distribution is approximately log-normal, and it is a good metric for a schools variation from the default proficiency level distribution. I discuss this decision in greater depth in the Analysis section.
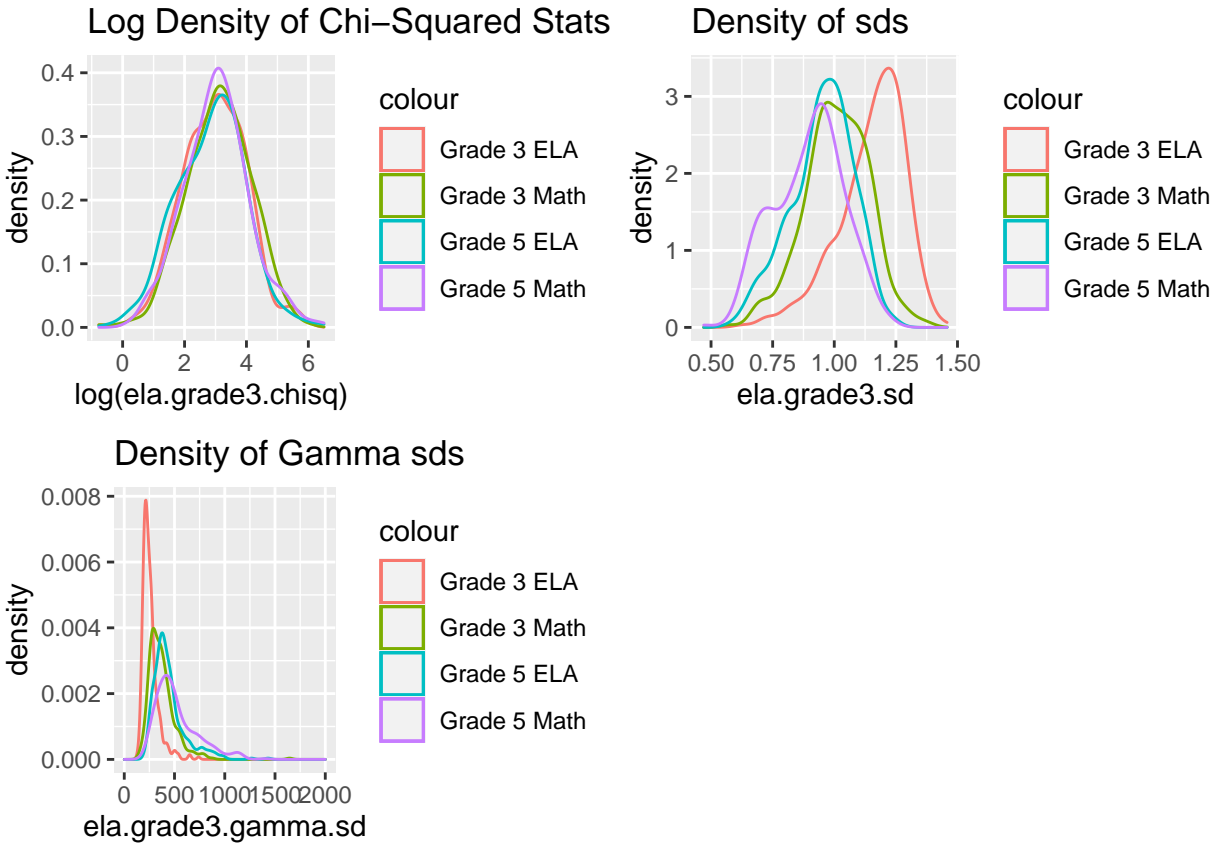

Figure 6: Density plots for response variables measuring spread of student scores

5.4 Spatial Autocorrelation

I use Moran's I (equation 6) to calculate the level of spatial auto-correlation for our reponse variables (Chi-squared statistics) of standardized test scores. Tests on both grade levels and subjects suggest a significant and positive spatial auto-correlation. This means that schools that are near each other tend to have similar proficiency level disparities.

	Grade 3 ELA	Grade 3 Math	Grade 5 ELA	Grade 5 Math
Expected	-0.0028409	-0.0028409	-0.0028409	-0.0028409
Observed	0.0443965	0.0256845	0.0182377	0.0245635
Standard Deviation	0.0041149	0.0041233	0.0038737	0.0040405
P-value	0	$4.5770054 \times 10^{-12}$	5.2857567×10^{-8}	$1.1823653 \times 10^{-11}$

Table 5: Assessing Spatial Auto-Correlation Using Moran's I, 2019

This level of auto-correlation is unsurprising given the level of income and racial segregation in the city and the strong correlation of income and race with test results. Note that median household income and race both have strong spatial auto-correlations.

	Median Household Income	Proportion White Students	Proportion Black Students
Expected	-0.0028409	-0.0028409	-0.0028409
Observed	0.1928371	0.1852793	0.2683243
Standard Deviation	0.0042029	0.0041938	0.0042289
P-value	0	0	0

Table 6: Assessing Spatial Auto-Correlation Using Moran's I, Demographics, 2019

We test whether there is still spatial auto-correlation when we control for race and income by testing Moran's I for the residuals on a regression that includes race and income.

$$resid = \chi^2 - (b_0 + b_1(income) + b_2(propWhite) + b_3(propWhite \times income)) \tag{7}$$
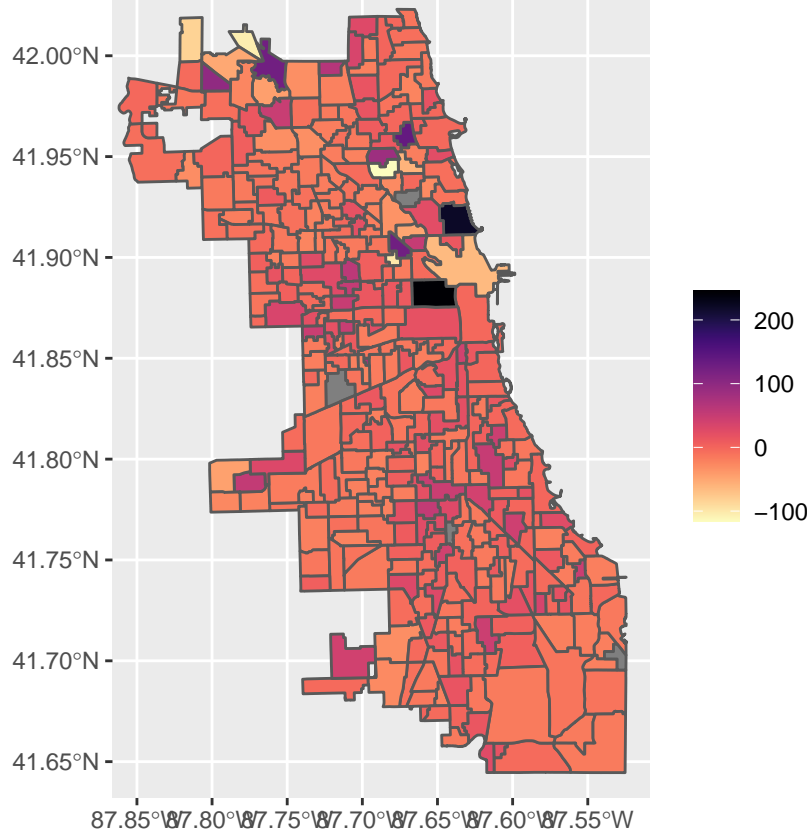
Figure 7: Plot of residuals for the Chi-squared statistic of 3rd grade ELA proficiency levels, accounting for race and income (equation 7)

Using the residuals from the regression equation above, we get a p-value of 0.7167723 from calculating Moran's I. This suggests that race and income account for most of the spatial auto-correlation. We will discuss the outliers in this plot in section 7.2.3.

5.5 Variable Relationships

In order to investigate the effect of racial and income homogeneity on elementary school performance, we must isolate the effects of these variables by controlling for potentially confounding variables or colinearity.

First, we consider the relationships between various explanatory variables. Many of these relationships are not surprising. For instance, the proportion of White residents in a neighborhood is highly correlated with the proportion of White students, with a correlation coefficient of 0.87. There are several other relationships of note. The proportion of White students and the level of median household income are both negatively correlated with the proportion of low income students in a school, while income is strongly and positively correlated with the proportion of White students in a school. This makes sense in a city where White residents have significantly higher income than residents of color. The correlation coefficient for the relationship between proportion of White residents and median household income is 0.7.

Also noticeably, the proportion of White residents has a strong negative association with the size of the largest same-race group in a school. This makes sense since only 25 of the schools have student bodies that are more than 50% White. In general, schools that have more White students are less homogeneous.

Because of the level of residential segregation by race, there is strong negative association between the proportion of Hispanic and Black students, with a correlation coefficient of -0.82. Schools with more Hispanic

16

students tend to have more English learners, so the association between proportion of Hispanic students and proportion of English learners is large and positive, while the association between proportion of Black students and English learners is large and negative.
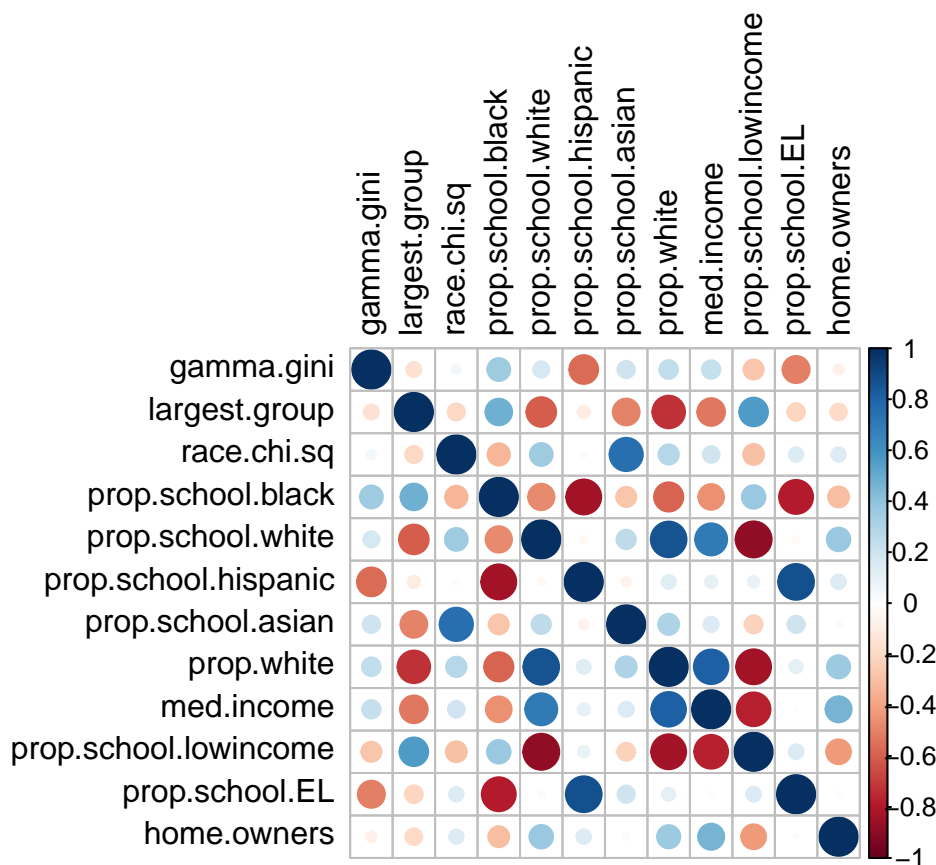

Figure 8: Investigating relationships between predictor variables

I also analyzed the relationships between potential response variables. The variables from the interpolated gamma distribution are highly correlated with each other and with the standard deviation calculated directly from the proficiency levels. This is because the mean and standard deviation of a gamma distribution are directly proportional.

Meanwhile, the direct mean from the levels is only very weakly ($cor = 0.15$) associated with the mean calculated from the gamma distribution. This deeply undermines my confidence in the interpolated distribution as an appropriate measure of spread. Meanwhile, the Chi-squared statistic is not heavily associated with the mean proficiency level. The plot below also demonstrates that, as we would expect, the Chi-squared statistic for different grades and subjects are positively associated with each other.

7. Results

7.1 Response Variable

I ultimately chose to use the Chi-squared statistic as a metric for score disparties. It measures the deviation from the city-wide distribution of test scores. However, one weakness of using Chi-squared as the response variable is that it cannot distinguish between distributions of proficiency levels that are equally different from the null distribution, but indicate different levels of score spread. For instance, consider the following
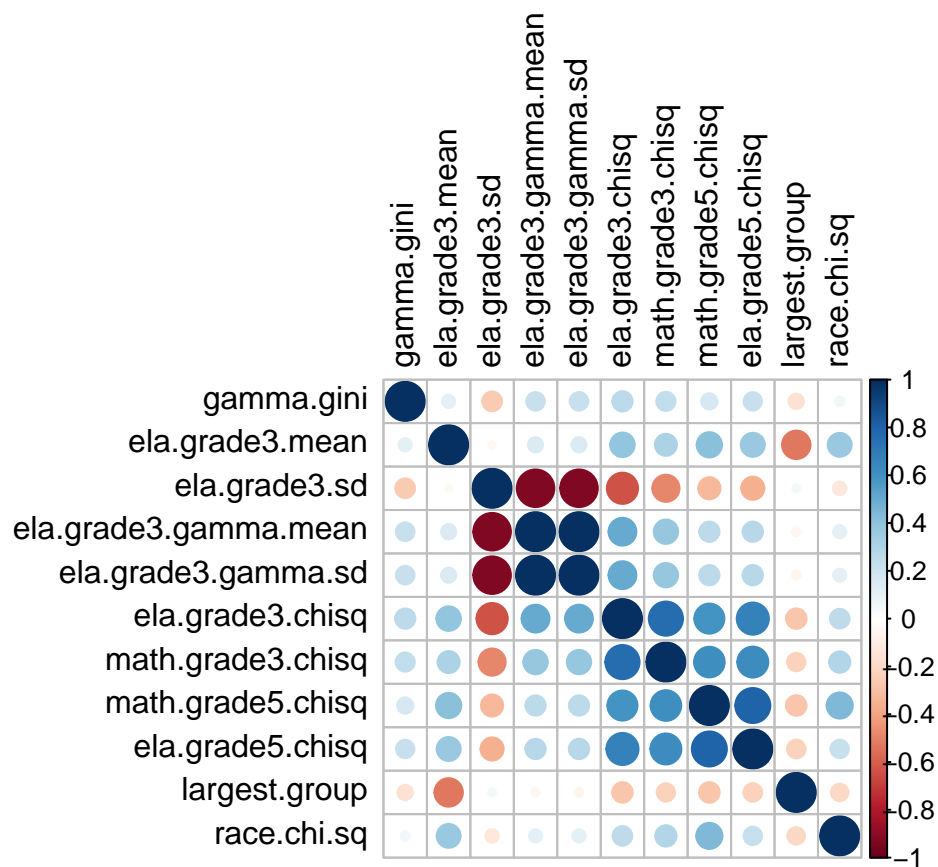
Figure 9: Investigating relationships between response variables

hypothetical sets of students in a district where the null score distribution is equal probabilities for each proficiency level:

	# at Level 1	# at Level 2	# at Level 3	# at Level 4	# at Level 5	χ^2	SD	Mean
School A	1	2	3	4	5	3.33	1.29	3.67
School B	1	4	5	3	2	3.33	1.16	3.07

Table 7: Chi-Squared Example

Although the total number of students are the same, and the Chi-squared statistic is the same, the distributions of proficiency levels vary significantly, and the standard deviations and means differ.

Despite these limitations, we will proceed cautiously with Chi-squared as the response variable. As shown in section 5.3, the Chi-squared statistic is not normally distributed. It is roughly log-normal, so I use the log of the Chi-squared statistics as a proxy for the spread of student outcomes.
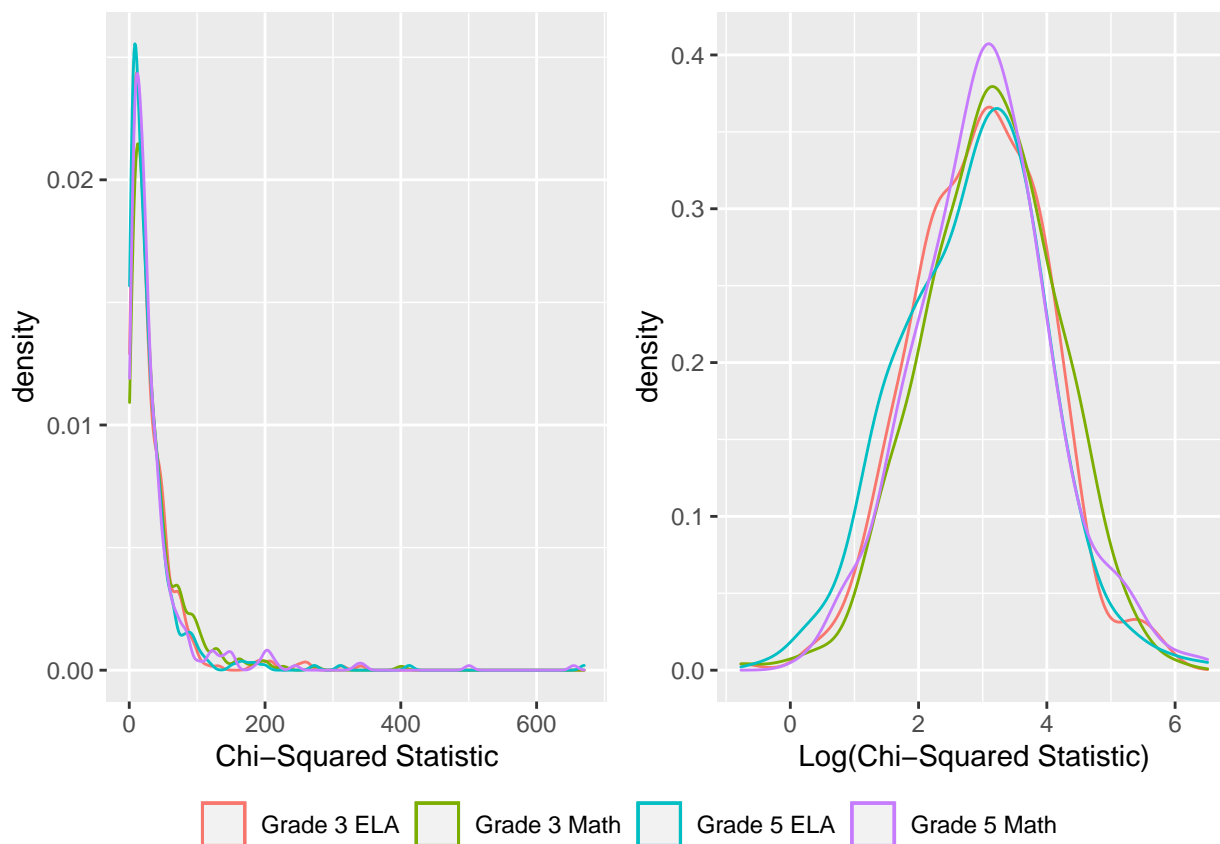

Figure 10: Chi-squared density plots. For all grades and subjects, Chi-squared statistics appear to be generally log-normal

7.2 Modeling Score Disparities

7.2.1 The Model Finally, we fit a linear regression to quantify the influence of income inequality and school homogeneity on disparities in test scores.

$$\hat{log}(\chi^2) = b_0 +$$
$$b_1 gini +$$
$$b_2 largest.group +$$
$$b_3 prop.black +$$
$$b_4 median.income + \quad (8)$$
$$b_5 prop.lowincome +$$
$$b_6 prop.black \times prop.lowincome +$$
$$b_7 largest.group \times median.income$$

7.2.2 Diagnostics These models generally satisfies the assumptions for linear regression. See the diagnostic plots below for 3rd grade ELA scores. Although there is some horizontal clustering of the residuals, there is no obvious vertical pattern, and the QQ-plot suggests that the residuals are normally distributed. I did not display the diagnostic plots for other grades and subjects, but they similarly satisfy the conditions for linear regression. There are a few outliers, which we address next.
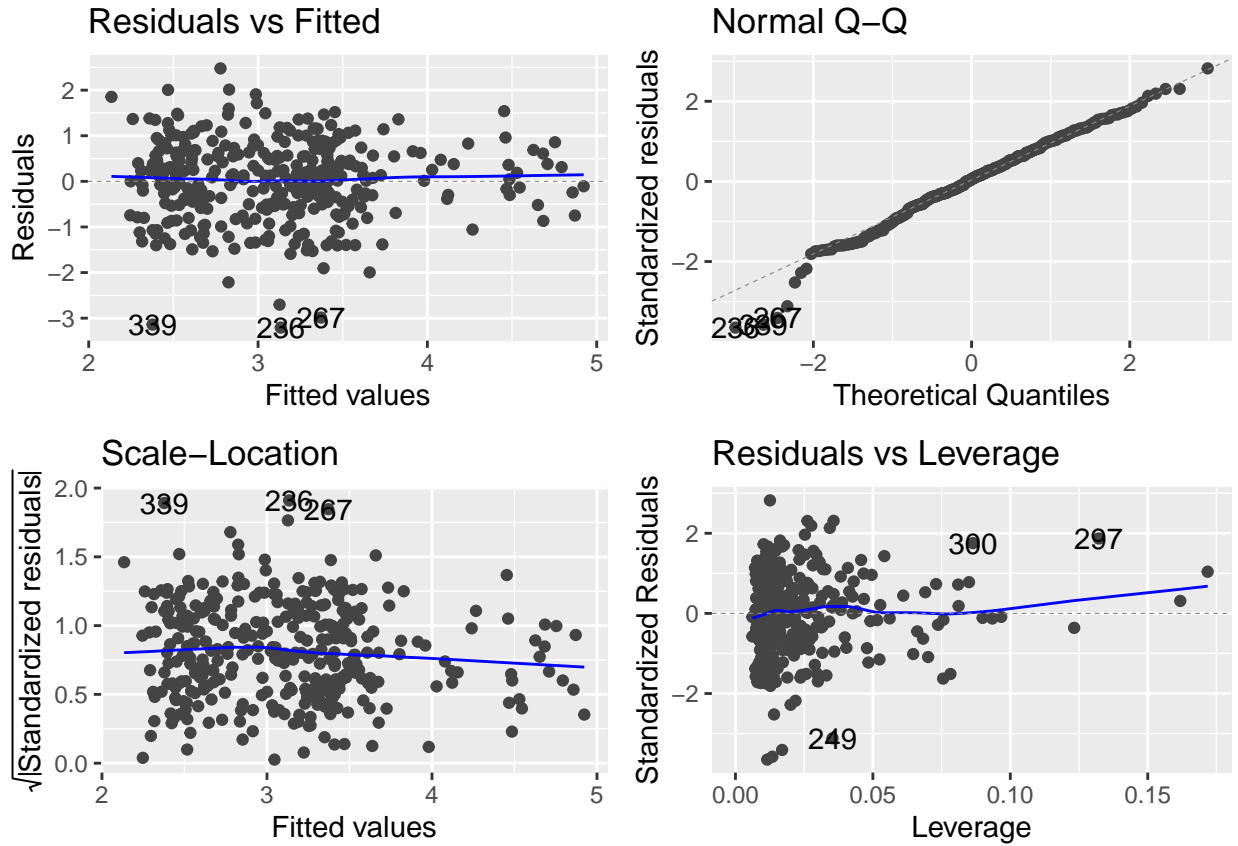

Figure 11: Diagnostic plots for a linear model of score disparities for 3rd grade ELA scores in 2019. The QQ-plot suggests normality for the residuals, and although there is some horizontal clustering of the residuals, there is no obvious vertical pattern.

7.2.3 Outliers

7.2.4 Results   Using equation 8 to model score disparities, we get the following results for the effect of various factors on the spread of scores. The table includes results for the Chi-squared statistic of proficiency levels for English language arts and math standardized test scores for third and fifth graders.

	Grade 3 ELA	Grade 3 Math	Grade 5 ELA	Grade 5 Math
Intercept	-0.508	-0.318	4.673*	5.361*
	(1.83)	(1.958)	(2.046)	(2.068)
Gini	4.719*	5.463*	0.074	-0.868
	(2.242)	(2.399)	(2.498)	(2.526)
Largest Group	2.242*	1.799*	0.47	0.406
	(0.681)	(0.729)	(0.765)	(0.773)
Proportion Black Students	-2.22*	-3.447*	-2.637*	-3.116*
	(0.699)	(0.749)	(0.775)	(0.783)
Median Income ($1,000s)	0.028*	0.017*	0.007	0.006
	(0.008)	(0.008)	(0.009)	(0.009)
Proportion of Low-Income Students	-2.538*	-2.622*	-3.282*	-2.858*
	(0.368)	(0.394)	(0.409)	(0.413)
Black Students × Low-Income	3.126*	4.401*	3.967*	4.097*
	(0.768)	(0.822)	(0.852)	(0.861)
Largest Group × Median Income	-0.034*	-0.023*	-0.01	-0.007
	(0.01)	(0.01)	(0.011)	(0.011)

Table 8: Regressing Score Spreads, 2019

* = Significant at the 0.05 level

	Grade 3 ELA	Grade 3 Math	Grade 5 ELA	Grade 5 Math
Intercept	-0.276	3.677	1.723	3.247
	(2.291)	(2.139)	(2.429)	(2.385)
Gini	2.654	-1.965	0.59	-2.857
	(2.961)	(2.766)	(3.137)	(3.079)
Largest Group	1.5	0.294	0.234	0.965
	(0.938)	(0.876)	(0.996)	(0.977)
Proportion of Black Students	0.004	0.336	0.228	0.382
	(0.237)	(0.221)	(0.251)	(0.246)
Median Income	0.041*	0.034*	0.031*	0.042*
	(0.012)	(0.011)	(0.013)	(0.013)
Proportion of Low Income Students	-0.85*	-0.672*	-0.966*	-0.309
	(0.218)	(0.204)	(0.23)	(0.226)
Black Students × Low Income	0.179	-0.311	0.912	0.428
	(0.676)	(0.631)	(0.715)	(0.701)
Largest Group × Median Income	-0.053*	-0.037*	-0.034*	-0.042*
	(0.016)	(0.015)	(0.017)	(0.017)

Table 9: Regressing Score Disparities, 2009

* = Significant at the 0.05 level

7.3 Modeling Mean Scores

To investigate the effect of the same factors on the magnitude of scores, we run a similar regression with mean proficiency level as the explanatory variable.

$$
\begin{aligned}
\hat{mean.level}_s = b_0 + \\
b_1 gini + \\
b_2 prop.black + \\
b_3 largest.group + \\
b_4 median.income + \\
b_5 home.owners + \\
b_6 prop.lowincome + \\
b_7 prop.English.learners + \\
b_8 prop.black \times largest.group + \\
b_9 prop.black \times prop.lowincome
\end{aligned}
\tag{9}
$$

This equation yields the following results:

	Grade 3 ELA	Grade 3 Math	Grade 5 ELA	Grade 5 Math
Intercept	4.51*	4.313*	3.418*	3.634*
	(0.862)	(0.859)	(0.69)	(0.778)
Gini	0.023	0.222	0.502	0.408
	(1.072)	(1.069)	(0.856)	(0.965)
Largest Group	-0.748*	-0.6*	-0.293	-0.447*
	(0.215)	(0.215)	(0.175)	(0.197)
Proportion of Black Students	-2.314*	-1.978*	-1.172*	-1.435*
	(0.569)	(0.567)	(0.456)	(0.514)
Median Income	-0.002	-0.003	-0.001	-0.002
	(0.002)	(0.002)	(0.001)	(0.001)
Home Owners	0.298*	0.199	0.297*	0.208
	(0.136)	(0.135)	(0.109)	(0.122)
Proportion of Low Income Students	-0.964*	-0.908*	-1.26*	-1.31*
	(0.2)	(0.199)	(0.159)	(0.179)
Proportion of English Learners	-0.009*	-0.007*	0.001	0.002
	(0.003)	(0.003)	(0.002)	(0.003)
Black Students × Largest Group	1.262*	0.874	0.765	1.135*
	(0.535)	(0.534)	(0.429)	(0.484)
Black Students × Median Income	0.011*	0.009*	0.003	0.001
	(0.003)	(0.003)	(0.002)	(0.003)

Table 10: Regressing Mean Scores, 2019

* = Significant at the 0.05 level

7.4 Sensitivity Analysis for Income Inequality Metric

8. Discussion

9. Conclusion

References

Appendices