CS 590.7 Computational Microeconomics Project Proposal

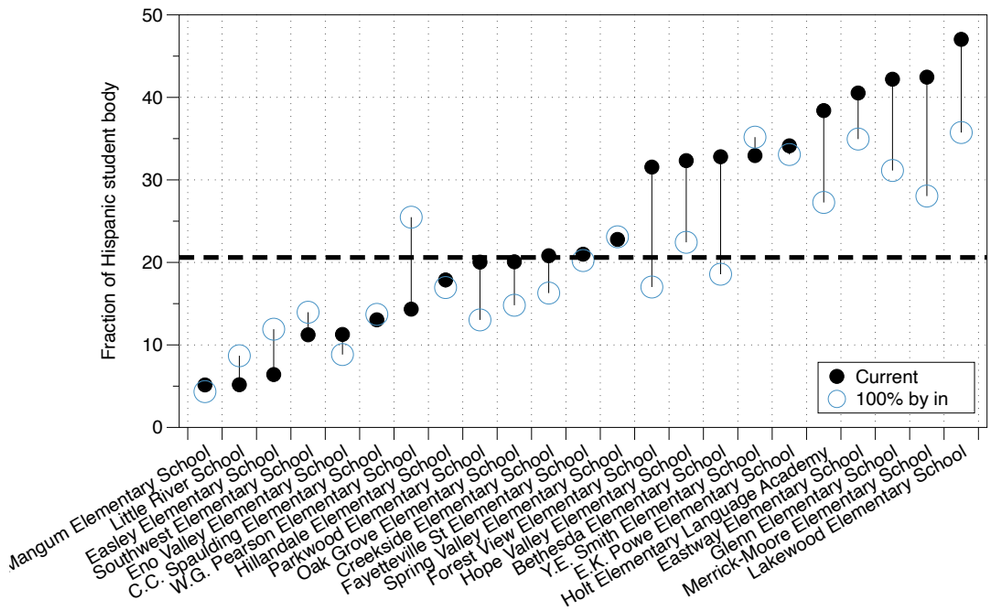
Xinchun(Atsushi) Hu Yuncong (Joe) Zuo

Oct 15 2020

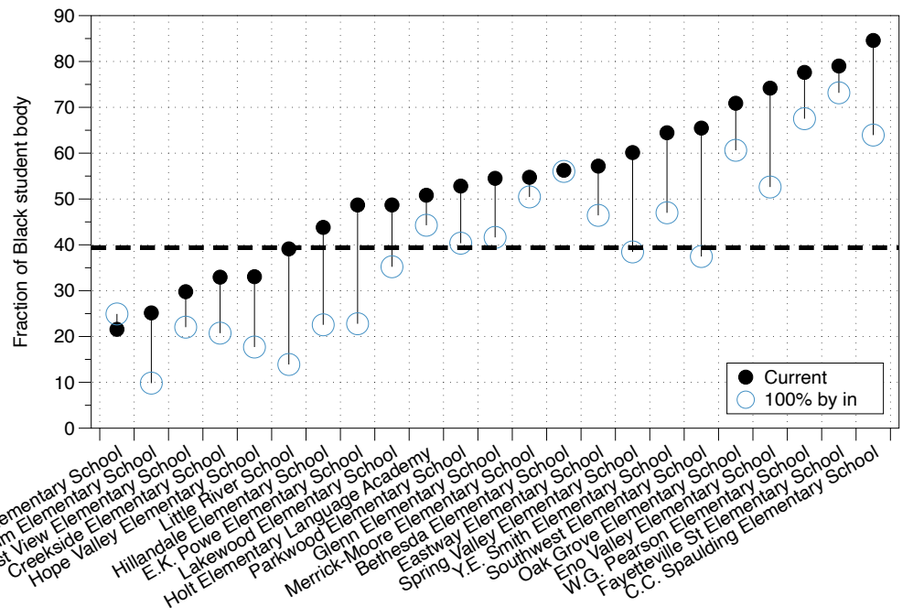
1 Problem Context

While the famous Brown vs Board of Education case embarked the official desegregation in schools across the United Stats, de facto segregation still happens in many places including Durham. Based on the most recent report, 48.4% of population in Durham are White, 39.3 % Black and 5.33% are Asian (the rest include other minority groups such as Hispanic and Pacific Islanders).[1] Nonetheless, the racial representation at public elementary schools (K-5 schools) can be very different from the above mentioned numbers, since private schools tend to attract students from more well-off families, and many of these families are White due to many historical and social reasons. For instance, Black students consist of over 80% of the student population of C.C Spaulding Elementary School (a public elementary school in Durham) based on data from North Carolina Education Research Data Center (NCERDC).

In fact, based on Atsushi's research over the summer[2] , if we assume all students at proper age for primary school, the racial representation at public schools in Durham would look significantly different from the current situation (see the graphs below):



[1] https://worldpopulationreview.com/us-cities/durham-nc-population
[2] https://www.overleaf.com/read/ghhdytqrwxhw



Comprehensive researches have shown that racial and social economic diversity in the classroom will contribute positively to students' academic, cognitive and social skills.[3]

One of the interesting questions to ask is, if we are to allocate all students at school age in Durham to a public school with some constraints that are to be listed later, what is the optimal allocation that will make the racial representation at each school as close to the overall racial composition as possible.

2 Problem Model

2.1 Assumptions

As we mention, our goal is to find a way to allocate students to public elementary schools in Durham, we will be treating the problem as a linear programming problem. We are having the following assumptions which are some of the simplifications we feel necessary for the project. We have the following assumptions:

1: Current public school assignment is based on school districts. We will not be adopting the current public school districts for our problem.

2: Each school has a capacity on how many students they can take due to the limit in educational resources.

3: Our alternative allocation for school districts is mostly distance based: a student would not be assigned to a public school that is too far geographically.

4: All students at proper age for elementary schools will go to a public school and we assume there is no private school.

5: Due to the limitation of available data, we will assume all students at a census block go to the same school, since we do not have the exact addresses of students and block is a best possible approximation.

[3]https://tcf.org/content/facts/the-benefits-of-socioeconomically-and-racially-integrated-schools-and-classrooms/?session=1

2.2 Parameters and Variables

After we have listed out the assumptions, we are parameterizing the relevant data in the following way to set up the model:

Parameters:

0: I represents the set of all students and J is the set of all schools. We are using i to index each student , and use j to index each school. R represents the set of all races and we use k to index each race (for example, $R = \{1, 2, 3\}$ where 1 = white, 2 = Black and 3 = Hispanic).

1; The race of a student, $r_i \in R$. This is an integer paramater.

2: Distance between a student to a public elementary school d_{ij}. Based on our assumption, for all students in the same census block, d_{ij} has the same value for any fixed school j.

3: Maximum possible distance a student needs to commute to attend a school D, We assume this is a fixed constant for all students for now.

4: The limit on the number of students at a public elementary school j is L_j

5: Percentage of students at primary school ages of a particular race k among all students at primary school ages (benchmark racial composition): B_k. $B_k \in [0, 1]$ for all k.

Variables:

1: Which school should a student be allocated to:A_i ($A_i \in J$ and is an integer).

2: The percentage of each race at a public school: P_{jk} represents the percentage of students that are of race k in school j, $P_{jk} \in [0, 1]$.

3: Quantified difference between the racial composition at a school and the benchmark racial composition: F_j . To obtain this value, we may have sum over $|B_k - P_{jk}|$ or the square of the difference.

2.3 Goal and Objective

The objective of our linear program is to minimize the sum of all the quantified differences between racial composition at each school and the benchmark racial composition, subject to all the aforementioned constraints.

Ideally, we will have a new way of assigning school districts, and we want to visualize in a Durham map. We also have the racial composition at each school after the new assignment. We can compare the racial composition based on our proposed districts to the existing racial composition.

3 Data and Resources

As the project is closely related to a related previous research we have conducted, relevant data have been already obtained. This include data about each public elemen-

tary schools in Durham,[4] US census data on details about Durham population, and the districting plan of public schools in Durham.[5] All the relevant data has already been put up on our github repository for the project:

https://github.com/Atsushihxc/CS590project

We will be working with Geopandas packages on Python for visualization and data extraction purpose, and will be using linear programming packages for computing the parameters and solving the problem.

[4] https://www.dpsnc.net/

[5] https://www.lib.ncsu.edu/gis/search/search.php?q=%22School+Districts%22rk=1s=sce=