Distance and Feeder Public Middle Schools for Specialized Public High Schools in New York City

December 12, 2023

Abstract

Specialized High Schools in New York City are considered elite public schools and require an admission exam. The objective of this paper is to identify "feeder" public middle schools for these high schools and their distances to lower income individuals. Even if a student meets the criteria to attend a feeder school, if they are unable to get to the school they will not be able to attend. The assumption is that Low Income students live further from feeder public middle schools than other students. The findings indicate that the distinction is more between Upper Class students and all others. The results also indicate that while there is a relationship between distance to a feeder public middle school and the number of individuals within an income bracket that live in a census tract, further analysis is needed.

1 Introduction

Applying to high schools in New York City is akin to applying to colleges. Middle school students are provided with a booklet with over 500 options, each with their own unique entrance requirements. Of these high schools, eight are considered "specialized" and require a rigorous exam called the Specialized High School Admissions Test (SHSAT), which is analogous to the SAT for college. These specialized high schools are feeder schools to top universities in the United States and thus are highly selective. From NYC Department of Education, "Offers to the eight testing Specialized High Schools are based on each student's SHSAT score, school preference order, and the number of seats at each school. These offers are made to each school in descending SHSAT score order until all seats at each school are filled."

In 2018 The New York Times published an article entitled, "See Where New York City's Elite High Schools Get Their Students", which focused on the public middle schools that "fed" the "feeder" schools. In their reporting, the Times discovered that, "Out of about 600 public middle schools, just 10 account for more than 1,200, or 25 percent, of the offers to attend one of the elite schools, according to Education Department data."

According to the New York Times, former Mayor Bill de Blasio proposed the "7 Percent Policy" in 2018, which would stop the SHSAT exam and "admit the top 7 percent of eighth graders from every middle school, based on class rank and state test scores". This proposal was meant to increase access to these specialized high schools for students who are Black and Hispanic. However, this initiative was met with a strong backlash and has since been tabled. Even if the former mayor's proposal to change the admissions process passed, students from lower performing middle schools will face significant challenges in high school when faced with students from higher performing middle schools.

In the absence of a new solution one potential option is to look at the current locations of these feeder public middle schools and their distance from lower income individuals. According to Burgess, S., & Briggs, A. (2010), "the single most important factor lying behind this differential chance is location, which is obviously priced. There is a clear link between poverty, location and school quality."

Therefore the hypothesis for this analysis is that Low Income students live further from feeder public middle schools than other students. For the purposes of this paper, a feeder middle school will be identified based on the percentage of students who take the high school admittance test who then are offered admissions to a specialized high school.

A second hypothesis is that distance from a feeder middle school is a predictor in the number of individuals within an income bracket that live in a census tract. This is in part to see if distance from a feeder public middle school has a relationship with where people of a certain income choose to live.

2 Data and Methods

2.1 Data Sources

Information related to public middle schools, school locations, and New York City boundaries is taken from the publicly available NYC Open Data. Household income information by census tract comes from Simply Analytics, which aggregates data from the US Census Bureau American Community Survey Data.

2.2 Data Manipulation

2.2.1 Defining Feeder Middle Schools

The first step in the analysis was defining the "feeder" middle schools. I downloaded the "2019-2020 SHSAT Admissions Test Offers By Sending School" from NYC Open Data into R. The dataset provides

information on offers, which is the number of offers to the specialized high schools. Any school which had an offer rate of "0-5" was removed from the dataset (which is congruent to the New York Times Analysis).

I then used a separate dataset from NYC Open Data called "2017-18 - 2021-22 Demographic Snapshot" for the total number of 8th graders for each middle school. I filtered the dataset for the 2019-2020 year and then merged this information with the specialized high school admissions data. I calculated the offers per total enrollment of 8th graders for each middle school (mean = 13.82%, median 8.89%) and plotted the distribution of the results. The right skew of the histogram indicates that the offer rate to specialized high schools per middle school is generally below 40% with a few outliers on the right tail. The schools on the right tail of the distribution (greater than or equal to 40% offer rate) will be considered "feeder schools".

Currently, just under half of the public feeder middle schools granted admission for students outside their zoned school district.

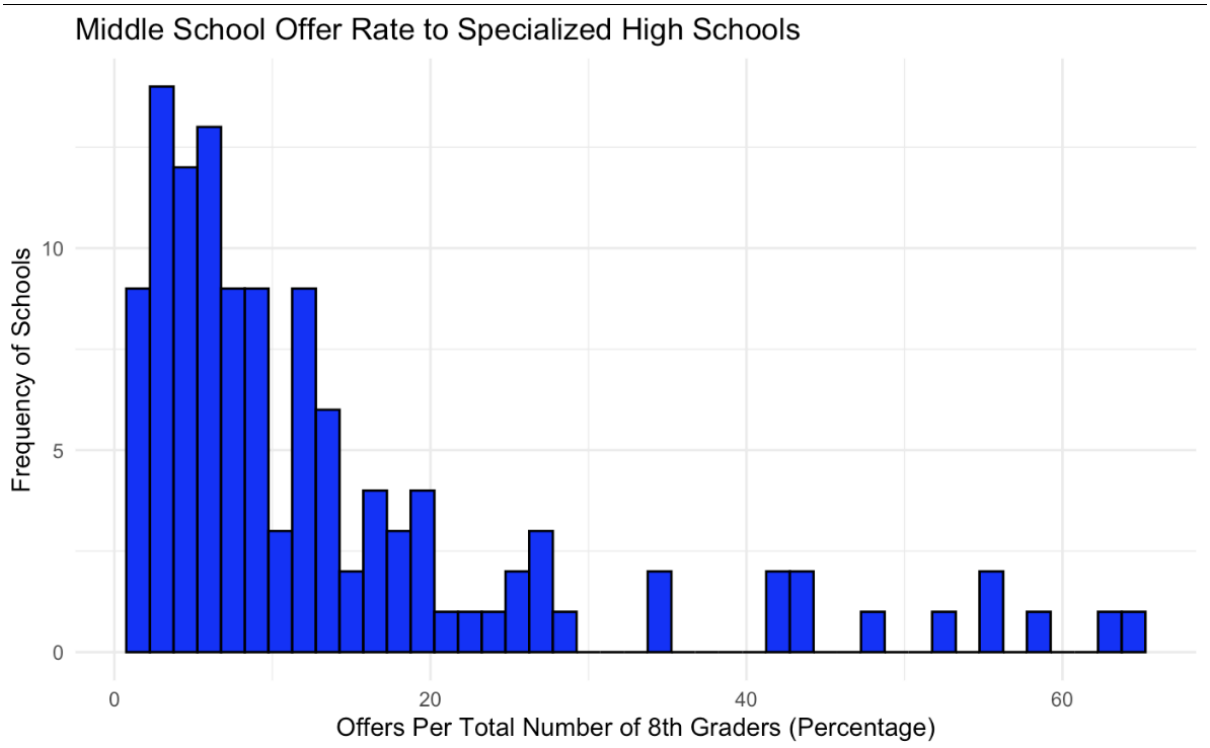

Figure 1: Histogram of Offer Rates per Middle Schools

Rank	Borough	School Name	Offer Per Total Number of 8th Graders	Admissions Criteria
1	Brooklyn	THE CHRISTA MCAULIFFE SCHOOL\I.S. 187 (20K187)	65.24%	Priority to District 20 students, test scores, and grades
2	Manhattan	NEW YORK CITY LAB MIDDLE SCHOOL FOR COLLABORATIVE STUDIES (02M312)	62.43%	Priority for siblings then to District 2 students, randomly selected
3	Manhattan	NEW EXPLORATIONS INTO SCIENCE, TECHNOLOGY AND MATH HIGH SCHOOL (01M539)	58.99%	Priority to continuing students then to all NYC students, based on grades
4	Manhattan	THE ANDERSON SCHOOL (03M334)	55.93%	Priority to continuing students then to all NYC students, based on grades
5	Queens	THE 30TH AVENUE SCHOOL (G&T CITYWIDE) (30Q300)	55.56%	Priority to continuing students then to all NYC students, based on grades
6	Queens	P.S. 122 MAMIE FAY (30Q122)	53.06%	District 30 students, based on grades
7	Manhattan	M.S. 255 SALK SCHOOL OF SCIENCE (02M255)	47.66%	Priority for siblings then to District 2 students, based on grades
8	Queens	BACCALAUREATE SCHOOL FOR GLOBAL EDUCATION (30Q580)	44.07%	All NYC students, assessment test

Figure 2: Feeder Middle Schools Ranked by Offer Rates to Specialized High Schools

2.2.2 Calculating Distance

In QGIS, I first loaded the shapefile of boundaries of New York City's five boroughs and then reprojected the layer to NAD83 / New York Long Island (ftUS). This projection (in feet) will serve as the projection for the entire analysis. The base layer was set to OSM's map of New York City to provide additional context.

For the locations of the middle schools, I used "School Point Locations" from NYC Open Data, which provided coordinates. I loaded the file into QGIS with the NAD83 projection.

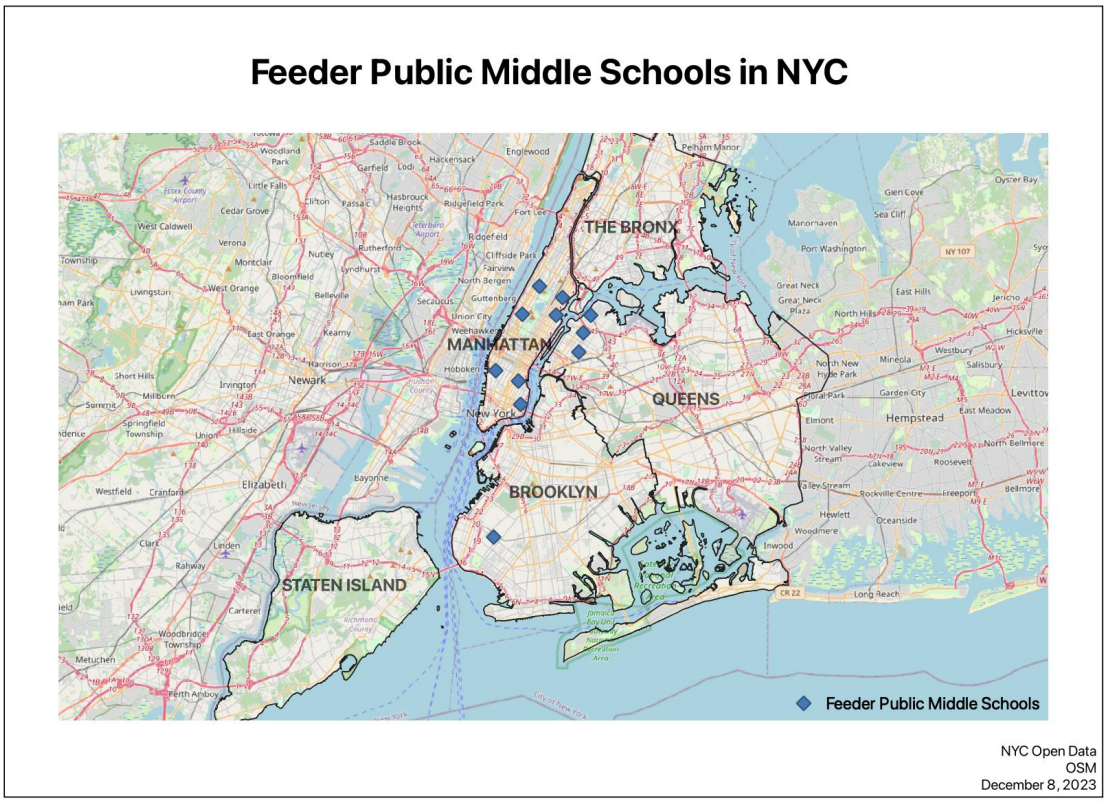

Figure 3: Feeder Public Middle Schools Locations

The next step was to incorporate income information across New York City. Income from the survey was defined as a household of four (2 adults and 2 children) and was collected at a census tract level. For this analysis I used four income brackets: "Low Income" (less than $40,000), "Working Class" ($40,000 - $74,999), "Middle Class" ($75,000 - $199,999), and "Upper Class" ($200,000 and above). The income brackets were based on the poverty index in New York City in 2019 was $36,262 and Pew Research's definition of middle class ($75,000 - $200,000). The Simply Analytics income data was provided in a shape file with census tract information which I imported into R.

To account for variations in income within the census tract I created a weighting system for a better representation of the data. For each income bracket I divided the count of each census tract by the total population of that bracket to create the weight. I then reprojected the file in R to NAD83, transformed the data frame into a shape file, and loaded the file into QGIS.

From there I clipped the file to match the borough boundaries and then created centroids for each census tract. With the feeder middle school layer, I used the GRASS v.distance tool to calculate the distance for each centroid to the closest feeder middle school. I converted the distance from feet to miles by dividing by 5,280 and created a new column to capture the miles.

To finalize the calculations, I downloaded the layer into R and multiplied the weights by the distance in miles. Finally, I took an average of all data points to determine the average weighted distance by income to feeder middle schools.

2.2.3 Regression Analysis

I conducted both an OLS regression for each income bracket to determine whether there is a relationship between household income and distance from a feeder middle school. I first regressed the total number of individuals in each tract on distance in miles from a feeder middle school. As there was

heteroskedasticity in the data I also log-transformed the distance to account for the unequal variances. I also conducted a robust standard errors test, which did not alter the results.

3 Results

The original hypothesis for this analysis was that low income students live further from feeder middle schools than other students. As shown in Figure 4, this is not actually the case. On average, most individuals outside of the upper class live approximately 4 miles away from feeder middle schools.

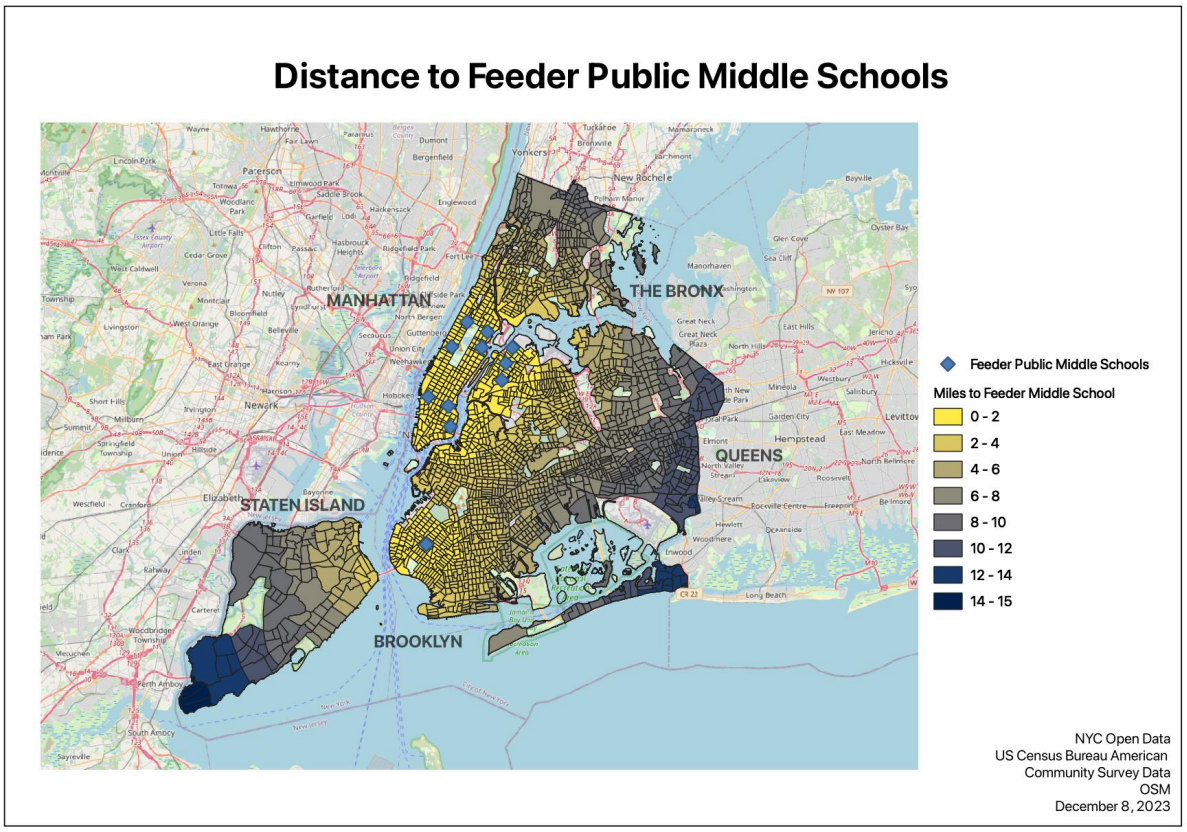

Figure 4: Distance to Feeder Public Middle Schools at the Census Tract Level

Diving deeper, I looked at whether the distance to a feeder middle school had any correlation with the number of individuals in a particular income bracket in a particular census tract.

For the Low Income bracket, an OLS regression showed that for every 10% increase in distance from a feeder middle school the number of people in the Low Income bracket who live that distance away decreases by -0.049 ($F(1, 2104) = 49.98$, $p < .001$, $R2 = 0.0232$).

```
Residuals:
   Min     1Q Median     3Q    Max
-529.9 -213.9  -91.5  121.0 3910.7

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   900.538     71.965   12.51  < 2e-16 ***
log_distance  -52.237      7.389   -7.07 2.11e-12 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 315.2 on 2104 degrees of freedom
Multiple R-squared:  0.0232,    Adjusted R-squared:  0.02274
F-statistic: 49.98 on 1 and 2104 DF,  p-value: 2.107e-12
```

Figure 5: Low Income OLS Results

For the Working Class bracket, an OLS regression showed that for every 10% increase in distance from a feeder middle school the number of people in the Working Class bracket who live that distance away decreases by -0.022 ($F(1, 2105) = 21.73$, $p < .001$, $R2 = 0.0102$)

```
Residuals:
   Min     1Q Median     3Q    Max
-335.4 -144.3  -49.2   89.9 3334.6

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)   545.895     49.926  10.934  < 2e-16 ***
log_distance  -23.894      5.125  -4.662 3.33e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 218.6 on 2105 degrees of freedom
Multiple R-squared:  0.01022,   Adjusted R-squared:  0.00975
F-statistic: 21.73 on 1 and 2105 DF,  p-value: 3.327e-06
```

Figure 6: Working Class OLS Results

For the Middle Class bracket, an OLS regression showed that for every 10% increase in distance from a feeder middle school the number of people in the Middle Class bracket who live that distance away decreases by -0.086 ($F(1, 2102) = 111.3$, $p < .001$, $R2 = 0.050$).

```
Residuals:
   Min     1Q Median     3Q    Max
-671.9 -217.9  -88.8  122.4 3896.8

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1375.341     83.573   16.46   <2e-16 ***
log_distance  -90.540      8.582  -10.55   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 366.3 on 2102 degrees of freedom
Multiple R-squared:  0.05029,   Adjusted R-squared:  0.04983
F-statistic: 111.3 on 1 and 2102 DF,  p-value: < 2.2e-16
```

Figure 7: Middle Class OLS Results

For the Upper Class bracket, an OLS regression showed that for every 10% increase in distance from a feeder middle school the number of people in the Upper Class bracket who live that distance away decreases by -0.125 ($F(1, 1987) = 390.1$, $p < .001$, $R2 = 0.164$).

```
Residuals:
    Min      1Q  Median      3Q     Max
-602.52 -126.51  -51.53   39.01 2151.12

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1449.042     64.576   22.44   <2e-16 ***
log_distance -131.125      6.639  -19.75   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 279.9 on 1987 degrees of freedom
Multiple R-squared:  0.1641,    Adjusted R-squared:  0.1637
F-statistic: 390.1 on 1 and 1987 DF,  p-value: < 2.2e-16
```

Figure 8: Upper Class OLS Results

4 Conclusion

Overall, the results of this analysis showed that the relationship between distance to feeder middle schools and income bracket is not clear. When looking at the average distance to schools between different income brackets the data shows a distinction between the Upper Class compared to the rest of the groups. This goes against the original hypothesis that solely Low Income households would be farther from feeder middle schools. The results could be explained, in part, by the concentration of feeder public middle schools either in Manhattan or within a close distance to Manhattan where higher incomes are more prevalent.

Although there are many factors that determine a child's academic success, the ability to physically access top schools is important. Currently, just under half of the public feeder middle schools granted admission for students outside their zoned school district and as stated by Burgess, S., & Briggs, A. (2010), "children from poor families face a reduced chance of being assigned to a good school in large part because of where they live." This means that most students in New York City can only attend feeder middle schools that reach for students beyond their school zone. While New York City should invest in upleveling schools in other parts of the city, in the current state this information reinforces the importance of getting students to and from feeder schools.

The regression analysis produced significant results however, the residuals were homoskedastic. This leads me to believe that a different type of analysis could provide better results and is an option for future research. The current results demonstrate that living farther from a feeder middle school regardless of income bracket has a negative relationship with the number of individuals who live in that census tract.

Further research could include a multi-modal analysis as many individuals in New York City use public transportation as well as cars, bikes, ferries, school buses and walking as methods of getting to school. Incorporation information on cost of transportation (e.g. a MetroCard) as well travel times and access to different modalities may shed additional light on this topic.

5 Bibliography

Bennett, J., Fry, R., & Kochhar, R. (2020, July 23). *Are you in the American middle class? Find out with our income calculator.* Pew Research Center. https://www.pewresearch.org/short-reads/2020/07/23/are-you-in-the-american-middle-class/

Burgess, S., & Briggs, A. (2010). School assignment, school choice and social mobility. *Economics of Education Review*, *29*(4), 639-649

The New York Times. (2018, June 29). *See Where New York City's Elite High Schools Get Their Students*. The New York Times. https://www.nytimes.com/interactive/2018/06/29/nyregion/nyc-high-schools-middle-schools-shsat-students.html

City of New York. (2022). *Data*. https://opendata.cityofnewyork.us/data/

NYC Department of Education. (n.d.). *Specialized High Schools.* https://www.schools.nyc.gov/enrollment/enroll-grade-by-grade/specialized-high-schools

NYC Mayor's Office for Economic Opportunity. (2021). *Poverty in New York City 2019: The New York City Government Poverty Measure, 2005-2019*. https://www.nyc.gov/assets/opportunity/pdf/21_poverty_measure_report.pdf

NYC MySchools. (n.d.). *Admissions.* https://www.myschools.nyc/en/