Michael Chrzan
Mike O’Key
Laisha Ozuna

Wrong Place, Wrong Time?
Studying the Relationship Between Gentrification and Educational Outcomes

EDUC 423A/SOC 302A Final Paper
December 12, 2023

I. INTRODUCTION

The urban landscape of American cities is continually reshaped by a myriad of socio-economic and demographic forces, among which gentrification is a prominent agent of change. While the socio-economic effects of gentrification have been widely studied, less attention has been paid to its impact on educational outcomes. This research aims to fill this gap by focusing exclusively on Detroit, a city emblematic of industrial prominence and urban decline. Historically a cornerstone of American manufacturing, Detroit has undergone sizable demographic shifts and economic transformations in recent decades, as Smith (2021) noted. This study seeks to analyze the correlation between gentrification and academic performance in Detroit's public schools, providing a nuanced understanding of how urban economic changes relate to educational dynamics.

II. RESEARCH QUESTION

This study poses the question: What relationship, if any, exists between gentrification, as operationalized by the Pearman and Greene (2022) model, and educational outcomes in Detroit's public schools between 2010 and 2020? This period, marked by pronounced urban change, offers a critical lens to examine the interconnections between gentrification processes and the educational achievements within this historically rich and complex urban setting.

For our analysis, we modify the Pearman and Greene model to consider gentrification at the city level as opposed to neighborhood level.

Pearman and Greene Operationalization	Our Adjusted Operationalization
A ***neighborhood*** is becoming gentrified if it: 1. had a median income below the 50th percentile of its respective city 2. has a share of recently constructed housing 3. has an increase in housing prices 4. has an increase in college-educated households that is greater than the increase in college-educated households in its respective city	A ***city*** is becoming gentrified if… 1. housing construction has increased recently 2. there's an increase in housing prices 3. it has an increase in college-educated households

We do this for two reasons. First, the data we found available to make the connections between gentrification metrics and educational outcomes was most readily available at the city level, which became a limiting factor to our analysis. However, a second reason, pertinent to our analysis, is because we wanted to ensure that the results we found captured the displacement caused by neighborhood gentrification. By considering gentrification at the city level, we feel safe in our assumption that whatever neighborhood a family may have moved to (which may have caused a change in their child's school) will still be captured in the educational outcome data due to the broad reach of our operationalization.

III. DATASETS

The National Center for Education Statistics (NCES) data plays a pivotal role, providing comprehensive insights into various educational metrics. The NCES database includes critical variables such as student growth, SAT scores, and overall student performance, with these metrics further disaggregated by sex and racial demographics. This granularity enables a detailed analysis of educational outcomes in the context of Detroit's public schools. The inclusion of sex and racial demographics is particularly crucial, as it allows for an exploration of how gentrification might differentially impact students across these groups. Such an analysis is essential for understanding the nuanced effects of socioeconomic changes on diverse student populations, offering a clearer picture of the educational landscape in a gentrifying urban setting. By leveraging these variables, the study aims to uncover patterns and correlations that can inform policymakers, educators, and community stakeholders about the multifaceted impacts of urban economic shifts on educational equity and quality.

We also pulled two data profiles from the American Community Survey (ACS). Specifically, we selected the data profiles Selected Social Characteristics and Selected Housing Characteristics. With these profiles we were able to cover most of the data needed for our gentrification model, which was educational attainment and housing cost estimates for Detroit, Michigan. The ACS datasets offer a multifaceted lens to study trends, economic patterns, housing developments, and educational outcomes across the city of Detroit. For our particular research project, the ACS data works great because it displays detailed geographic information and provides data on a year-by-year basis unlike the decennial census. In order to explore gentrification and its outcomes on educational outcomes, this granularity in a dataset is important for our research question(s).

However, it is essential to acknowledge that ACS data, while informative, involves some projections and estimates rather than direct observations. For example, while the ACS provides annual estimates, there is a time lag between data collection and release. Furthermore, data for a specific year is calculated based on data from previous years. These projections can cause wide margins for sampling errors and contain collection and other data biases. Due to the size and expansiveness of the ACS data, there are also issues with variable consistency over time and across data profiles. While this does not entirely affect the integrity and interpretation of the data, this is another one of the weaknesses we discovered particularly when cleaning and manipulating the data further.

Despite the robustness of the ACS data on housing variables, one variable that it failed to include was information on new housing construction. While our group first tried to mutate the ACS data to deduce this variable, we found that to be unreliable and - due to the lack of estimates prior to 2010 - would have meant that we lost the ability to evaluate if a year counted as gentrifying for multiple years in the decade. Thus, to more accurately measure the changes in housing constructed, we pulled parcel data directly from the city of Detroit's website.

IV. DATA PROCESSING

As has been mentioned, educational attainment estimates for the city of Detroit were pulled straight from the ACS Selected Social Characteristics data profile. This data profile contains various population characteristics apart from educational attainment like marital status, disability status, language spoken at home, computer access, and much more. Due to the vast number of variables in this data, the first step was to filter out most variables in order to only include estimates and percentages for variables that corresponded with educational attainment. Since there is a different dataset for each year, this had to be done to 10 different tibbles. Once filtering to only include educational variables for all tibbles, the next step was more minimal restructuring of the data by shifting the rows and cleaning the names of each variable with clean_names() in order to make the data tidy.

Once each tibble had its unique variable name and observation for each row, the next step was to tweak and convert the data to be ready for visualization. The first step was to convert the educational estimate and percentage variables from characters to doubles. Since there were a fair amount of different tibbles, we sought to find a function that would help accelerate the process of converting the variable types for each variable in each tibble. For this we found a function that would convert a string to double for multiple datasets on Stack Overflow and modified it to work for our particular tibbles. Since we wanted to bind rows to combine all tibbles, the next step was to add a year to each tibble with mutate in order to distinguish which row came from what year.

There was also a discrepancy for the educational attainment data from 2018 as it only included estimates but not percentages. Thus, using the estimates, we were able to add a percentage for each variable by using mutate and dividing each estimate by total population to then be multiplied by 100. After this was done and all tibbles had matching columns, we were able to bind them to result in a larger dataset that had educational attainment measures for Detroit, MI from 2010 to 2020.

Perhaps unsurprisingly, we found similar issues within the housing estimates in the ACS data profiles. Once the data sets were imported and tidy, we selected variables that were directly related to our model - namely housing values and rents. These variables were only somewhat consistently named throughout each of the data sets we pulled, with a notable difference being in the naming convention of including (or not including) the broader category in the name of the variable. To assist with this tedious process of removing these variables for each year, we first wrote the code to remove the variables from the tibble for one year and then used Microsoft's Copilot AI to generate a loop that would iterate that code over each tibble. Once we had all the variables we needed from each year binded into one tibble, we then rename and unite the variables to remove this barrier to analysis.

Another notable difference within this data was the change in the demarcations for the number of units paying different levels of rent. In the earlier half of the decade, the ACS had a highest category of "occupied units paying $1500 or more in rent" but by the later half, the value for this bracket doubled to "... $3000 or more in rent". Other demarcations below that value also increased, perhaps pointing to the remarkable increase in the cost of living in the city and leading us early on to some assurance of gentrification in the city. To handle the discrepancy of the highest rents in the data, we evaluated the number of units listed in each category and decided that the most accurate way to combine these was to join all the values into the lowest limiting value, which would be $1500+ in rent across the decade. For the other demarcations, the ending values of the brackets increased and so they were able to neatly fit into new categories. For example, the categories from earlier in the decade - "rent less than $200", "rent between $200 and $299", and "rent between $300 and $499" - would all neatly fit into the category "rent less than $500" in the later half of the decade.

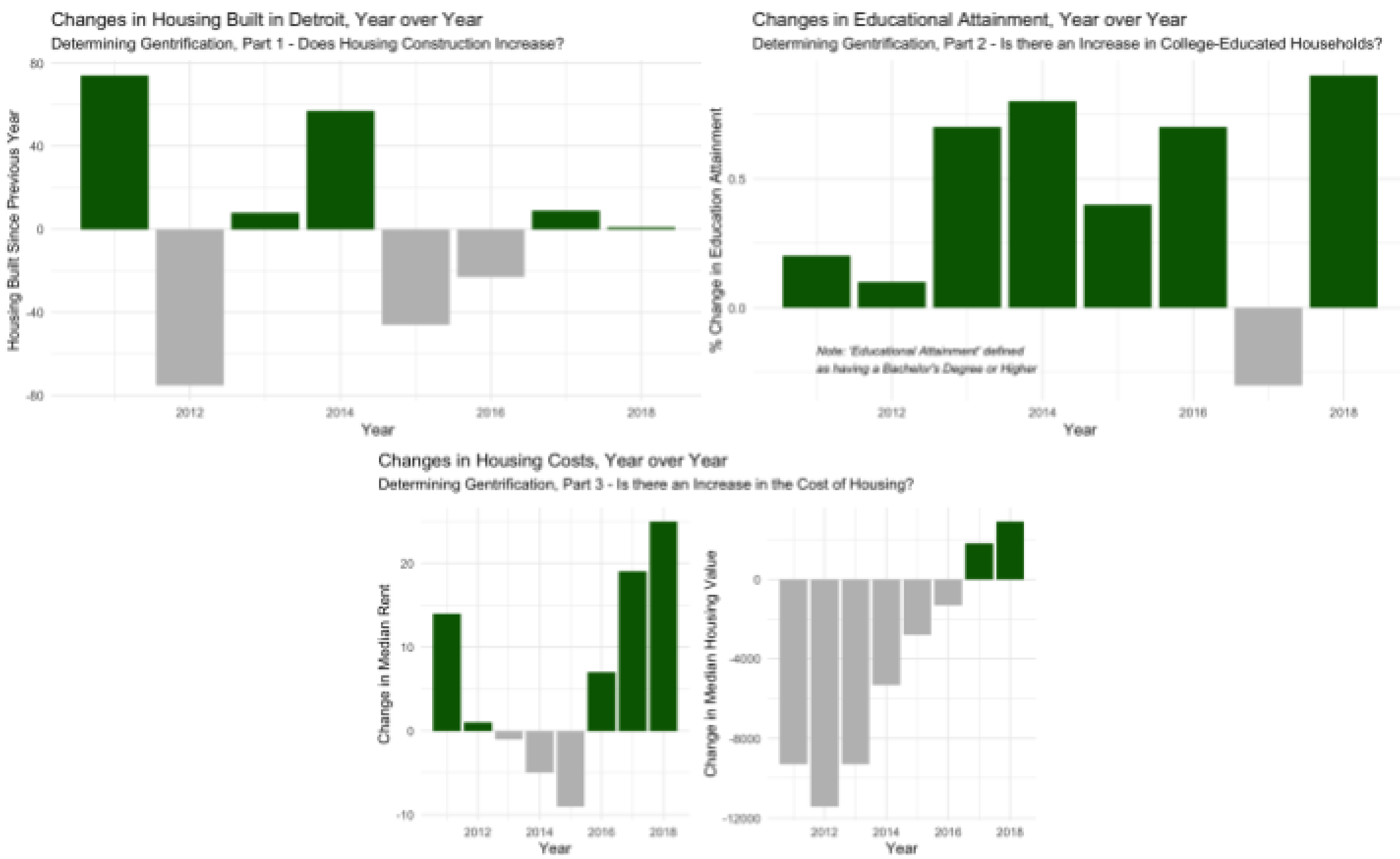
For our final data, the NCES data, we employed a similar process. We began with writing code to adjust the datatype of the variables and filter down to just the data for the Detroit Public Schools Community District within one tibble and employ the help of the Copilot AI to adjust this into a loop to avoid repeated code. Once each tibble is set, we bind them together to have a single tibble for all the data from each year combined. Due to the expansive nature of the NCES data, we decided to keep nearly all the variables so we could have a robust analysis of different types of educational outcomes to study against gentrifying years.

The only exceptions to this were the grade 11 reading and science "ready" scores. We removed these variables due to a lack of consistency across the decade in the reporting of the grade 11 standardized test scores. We noticed that in 2016 the readiness data started showing up in a separate file that we needed to join with the data prior. In investigating this difference, we inferred that this change would have been due to a change in the standardized test used that year, when the state of Michigan changed from using the ACT to the SAT (Higgins 2015). Due to how the SAT reports scores, there was no data available for science to compare to between the two tests. For the reading scores, we decided to use only the ACT section that is most aligned - in form and function - with the SAT's "Evidence-Based Reading and Writing" section, which was the "English" section of the ACT, and thus the reading scores were dropped from our analysis. Despite the change in tests, we believe that the use of the determination of "ready" benchmark from standardized data would be consistent enough to allow comparison across the two tests over the decade.

And lastly we discuss processing the city of Detroit's parcel data. This data contained information on every parcel of land in the city, including things like the owner information of the taxpayer, how much the property was sold for, what type of property it was (residential or commercial) and numerous other variables. For our analysis, we were only interested in residential properties (to analyze housing new residents might be moving into) so we filtered the dataset to only include residential properties and then group them by the year they were constructed so that we get a clear metric of how many residential structures were built in each year of the decade we are analyzing.

V. RESULTS

As we have mentioned, our gentrification model depends on the year-to-year changes for the variables that detail housing construction, housing costs, and educational attainment. We calculate and visualize those changes in *Figures 1-3* below.



Figures 1-3

In the figures above, we have visualized year-to-year changes with all three of the gentrification criteria. As we can observe, changes for each gentrification criteria not only deviate across years but also per each gentrification variable. While median housing value and educational attainment remains steadily increasing in Detroit (until 2017 for the later) , housing construction and median rents fluctuate with minimal patterns.

However, the years that are considered to be gentrifying for the city of Detroit are those that have a positive difference between years in all three gentrification variables. That signifies that, in Detroit, a year is only gentrifying if housing costs have increased, there is an increase in housing developments, and there is a greater percentage of people in Detroit who hold a bachelor's degree or higher than the year before.

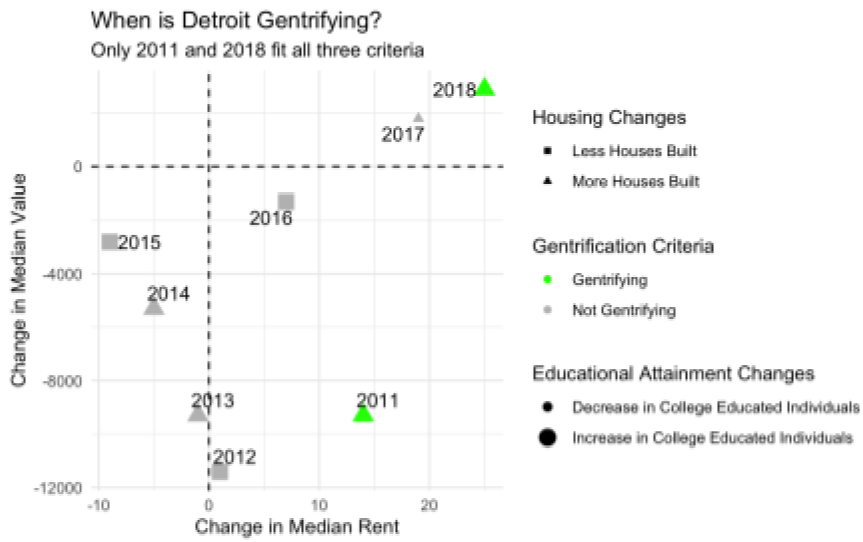


Figure 4

After calculating the combination of those differences, we can observe in *Figure 4* that under our gentrification model Detroit is only considered to be gentrifying during 2011 and 2018. Now that we have established when gentrification is occurring, we can consider the relationship between gentrifying years versus non-gentrifying years with educational outcomes.

Using the data from the NCES, we were also able to match educational outcomes to our gentrification criteria by each year in Detroit. We first observed the relationship between gentrification and student enrollment.

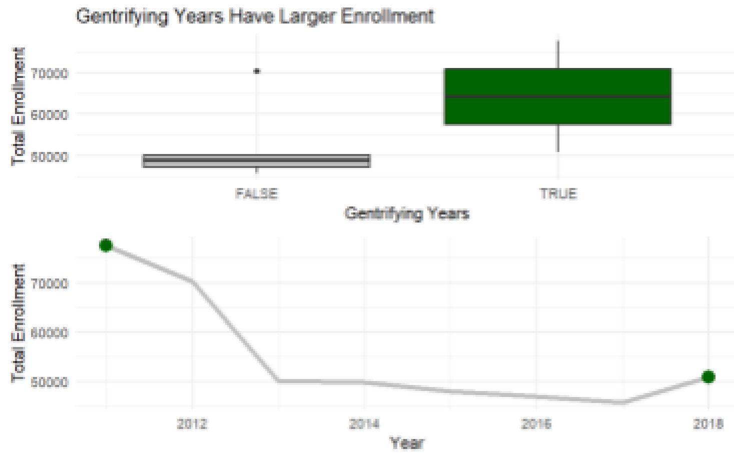


Figure 5

In the years that are considered gentrifying in Detroit by our operationalization, overall student enrollment increased. While the figure above demonstrates that student enrollment has greatly decreased overall from 2011, we can see after 2017 there is an increase in student enrollment. In fact, 2018 is the first year in the decade that reveals an increase in school enrollment from the previous year.

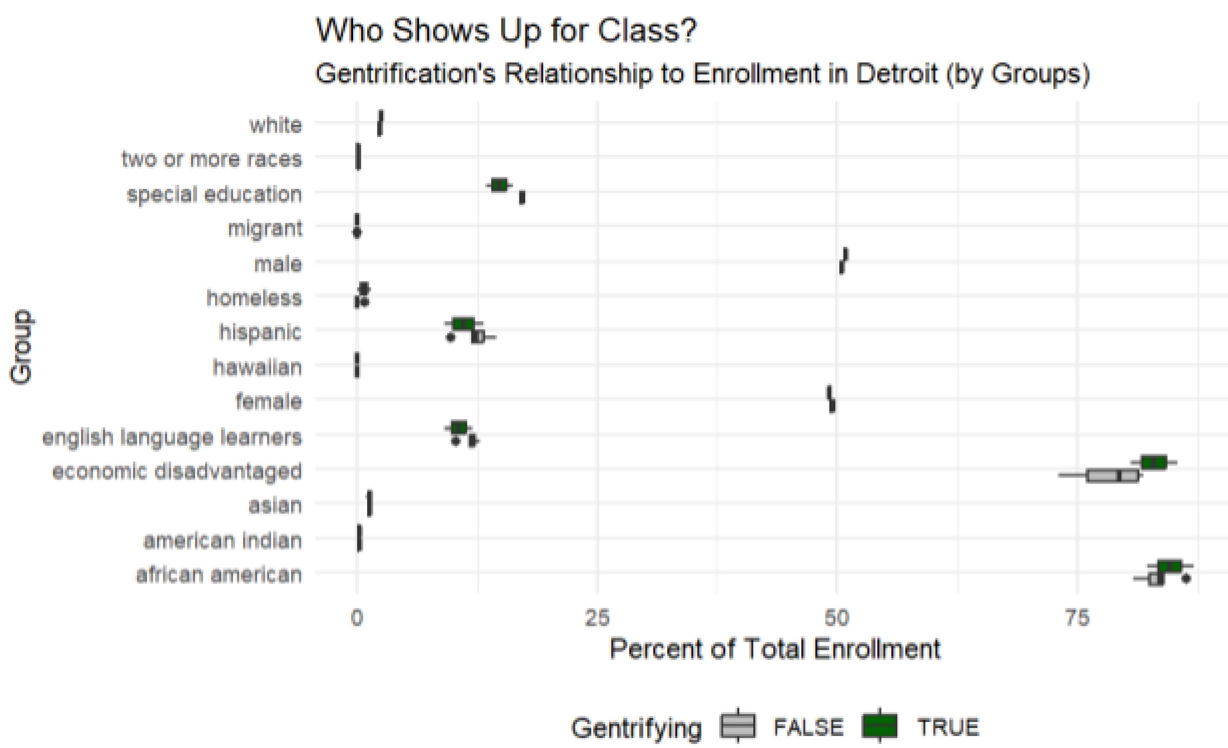
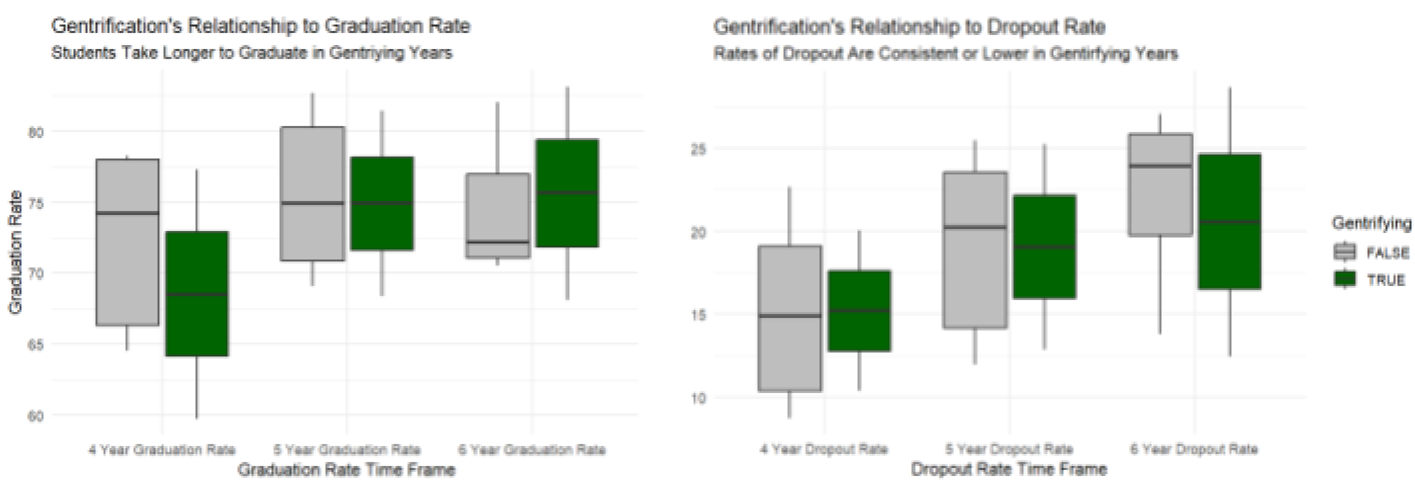


Figure 6

When distinguishing for student demographic groups (race, gender, socioeconomic status, etc.), we find that African-American, economically disadvantaged, and Homeless student enrollment goes up in gentrifying years while Special education, English Language Learners (ELL), and Hispanic enrollment tend to decrease in gentrifying years. This is detailed in *Figure 6*.

The NCES data also included data on graduation rates and dropout rates over time, including how long it took students to graduate. In examining the graduation rate, students seem to graduate faster under non-gentrifying years, as students in gentrifying years are more likely to graduate in a 6 year graduation time frame than a 4 and 5 year rate when compared to non-gentrifying years, where the medians appear to be nearly steady. We also find that gentrification has a relationship with dropout rates as students tend to consistently drop out less in gentrifying years. The dropout rate for a 4, 5, and 6 year time frame is lower for years distinguished with gentrification and the gap gets larger for students taking the longest to graduate.



Figures 7 & 8

Furthermore, we also wanted to observe gentrification's relationship with academic achievement across different subjects, which is provided by the "percent ready" metric given in the NCES data, which measures 11th grade students considered "ready" for collegiate work based on their standardized test score.

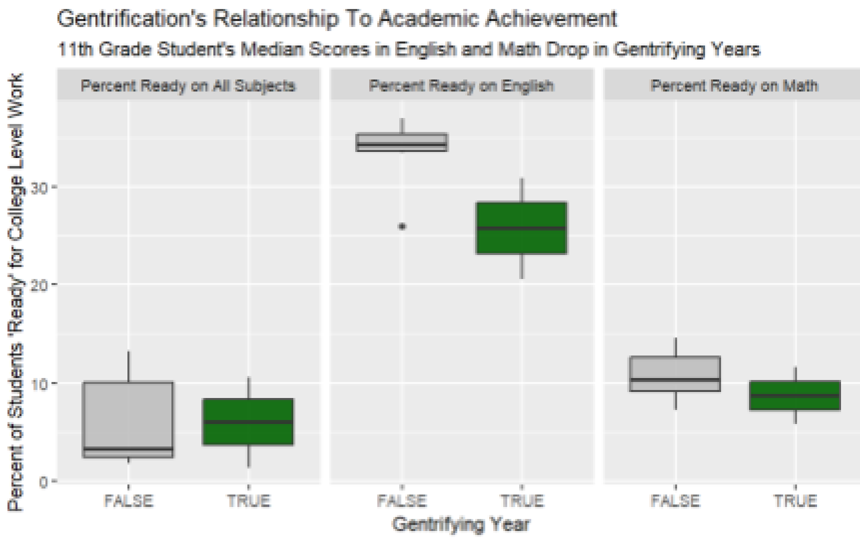


Figure 9

We found that, overall, gentrification seems to have a minimal or slightly negative relationship with the percentage of students "ready" in a specific subject. Moreover, the median percentage of students ready in Math and English is lower, while the median percentage of students ready for all subjects tends to be higher in gentrifying years.

However in gentrifying years, the actual scores are much lower relative to the surrounding years.

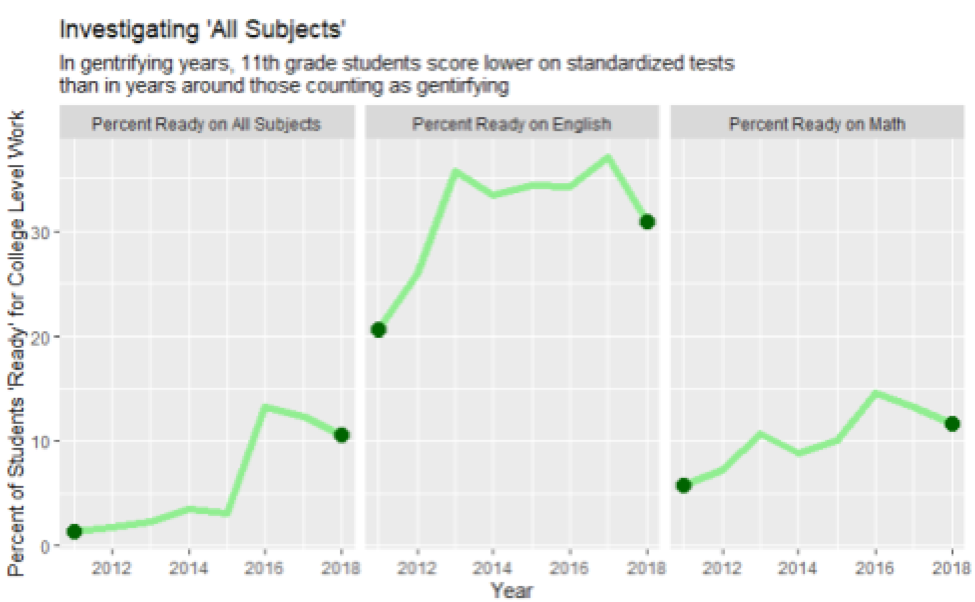


Figure 10

Figure 10 displays the marked drops in test scores in gentrifying years. The lowest scores, relative to the years near them, come from the year 2011 and in 2018 there is a percentage ready drop for every subject area, which includes English, Math, and All Subjects. Thus, the slight median increase in all subject readiness in *Figure 9* might be misleading since it is likely reflecting a change in standardized tests providers. As discussed in our data processing section, this is because the state of Michigan required that public schools change the type of standardized test offered from the American College Testing (ACT) to the Scholastic Aptitude Test (SAT) in 2016, aligning with the drop in readiness *post* 2016 in *Figure 10* (Higgins 2015).

VI. CONCLUSION

In this study, we examined the relationship between gentrification and educational outcomes in Detroit's Public Schools Community District, focusing on three criteria for gentrification: housing construction, housing costs, and educational attainment. Our key takeaways from this research are:

- During gentrifying years, overall student enrollment increased, with specific demographic groups showing varying trends.
- Gentrification appears to have a relationship to high school completion rates, with students appearing to take longer to graduate in gentrifying years, yet being less likely to drop out in those years.
- Gentrification appears to have a negative relationship with academic achievement, with test scores being the lowest or dropping in gentrifying years.

These conclusions add to our lacking understanding of the ways gentrification may impact educational outcomes via a case study of one quintessential city for America’s struggles with educating all students toward equitable outcomes. Future research should refine the current model to capture the complexities of gentrification more accurately. This includes integrating variables such as housing market dynamics, income levels, and local government policies, which might influence gentrification and educational outcomes.

The current study provides a city-wide perspective on the impacts of gentrification. Future research could also compare these findings with neighborhood-specific gentrification models. Such comparative studies could reveal whether gentrification's educational impacts are uniform across the city or vary significantly between neighborhoods and the significance of that difference.

Conducting longitudinal studies that extend beyond the current decade framework could also provide deeper insights into the long-term effects of gentrification on educational outcomes. This would allow researchers to observe trends and changes over an extended period, offering a more comprehensive understanding of the phenomena. Additionally, incorporating qualitative research methods, such as interviews and focus groups with students, parents, educators, and community members, could provide valuable context to the quantitative findings here. This approach would enable an exploration of the personal experiences and perceptions of those directly impacted by gentrification and educational changes.

Finally, expanding the research to include other cities with varying degrees and forms of gentrification could provide a broader perspective. This comparative approach would allow for a better understanding of whether and how the dynamics observed in Detroit are unique or part of a broader national or global trend. We would recommend that this research get data directly from the city of interest, if possible, when considering factors like housing constructed and housing costs, as these will potentially be more reliable than ACS estimates of those metrics.

By pursuing these research directions, scholars and policymakers can gain a more nuanced and comprehensive understanding of the interplay between gentrification and educational outcomes, ultimately contributing to more informed and effective urban planning and educational strategies.

REFERENCES

Pearman, F. A., & Marie Greene, D. (2022). School Closures and the Gentrification of the Black Metropolis. *Sociology of Education*, *95*(3), 233-253.

Higgins, L. (2015). Michigan to dump act for SAT as high school requirement. Detroit Free Press.

https://www.freep.com/story/news/local/michigan/2015/01/07/michigan-replaces-act-sat/21385299/

Smith, J. (2021). The Changing Demographics of American Cities: Insights from Detroit's Population Dynamics. Journal of Urban Studies, 59(2), 312-327.

APPENDIX

To see all code for this project, visit our repository at [mlchrzan/Booo-Gentrification (github.com)]