

THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF FINANCE

Gentrification in New York City: The Effect on the Sales Volume of Small Businesses

ELIZABETH WALSH
SPRING 2022

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Finance
with honors in Finance

Reviewed and approved* by the following:

Brian Davis
Associate Clinical Professor of Finance
Thesis Advisor and Supervisor

Eric Silver
Professor of Sociology and Criminology
Thesis Reader

* Electronic approvals are on file.

ABSTRACT

This research investigates the effect of gentrification on the sales volume of small New York City businesses. I define gentrified or non-gentrified neighborhoods by collecting demographic data and then linking it to historical business information. I find that non-gentrified areas have higher average sales volumes than gentrified areas in each year from 2015 – 2020. In 2020, non-gentrified neighborhoods have an average sales volume 1,204 units higher than gentrified neighborhoods ($p < .01$). As well, New York City year, neighborhood, and zip code information all have a statistically significant relationship with the occurrence of gentrification and the average sales volume of small businesses.

TABLE OF CONTENTS

LIST OF FIGURES	iii
LIST OF TABLES	iv
ACKNOWLEDGEMENTS	v
Chapter 1 Introduction	1
Research Question & Hypothesis	2
Chapter 2 Literature Review	4
A Creative Approach in Analyzing Shifting Neighborhoods	4
Bias in Judging Neighborhoods	5
Business Retention and Displacement Rates in Gentrified Areas	6
Policy Action to Address Gentrification	9
Chapter 3 Data & Methodology	10
Defining a Gentrification Metric	11
Non-Gentrified Neighborhoods	13
Business Sales Volume	14
Data Transformation	15
Chapter 4 Results	19
Analysis of Variance (ANOVA).....	22
Tukey HSD	23
Analysis.....	26
Chapter 5 Conclusion.....	28
Further Research	29
Appendix A Gentrified New York City Zip Codes	30
Appendix B Non-Gentrified New York City Zip Codes	32

LIST OF FIGURES

Figure 1: Initial Residuals vs. Fitted Data Test.....	16
Figure 2: Initial Normal Q-Q Data Test.....	17
Figure 3: Transformed Residuals vs. Fitted Data Test	18
Figure 4: Transformed Normal Q-Q Data Test.....	18
Figure 5: Average Sales Volume Across All Zip Codes	19
Figure 6: Average Sales Volume – Case Neighborhood Specific	21

LIST OF TABLES

Table 1: Gentrified Zip Codes by Area.....	13
Table 2: ANOVA Test Results	22
Table 3: Tukey Interactions by Classification	24
Table 4: Tukey Interactions by Neighborhood	25

ACKNOWLEDGEMENTS

Dr. Brian Davis for his positivity and willingness to help at every step of the way. I met Dr. Davis through virtual class and connecting in person to collaborate on my thesis was very rewarding.

Dr. Eric Silver for being my go-to professor throughout my college career. It was an honor to work together for many semesters. As a professor of sociology, I appreciate his viewpoint on my subject matter.

My friend Kyle Wong who helped me navigate my statistical analysis over many cups of coffee.

My sister for being there from day one of my thesis development and inspiring the idea to study gentrification. I admire her accomplishments, and having her input is always invaluable.

My family, friends, and boyfriend for happily listening to my discussions on my thesis and offering their unwavering support of me.

Chapter 1

Introduction

Gentrification is the term used to describe changing demographics in neighborhoods around the United States: specifically, in common reference to when white, wealthy, or highly educated people move into neighborhoods that have been historically inhabited by minority populations with low-income. When gentrification occurs, existing residents are forced to adapt to rising rent prices, an increase in the general cost of living, and cultural shifts. If residents fail to match the rising expectations of the neighborhood, they are forced into the unpleasant alternative of moving to a different, and often less advantageous, neighborhood.

As there has been a natural progression in the quality of living over time, cities around the United States have generally experienced a rise in average income and prosperity. However, gentrification is an unnatural occurrence in which residing residents lack a fair chance to compete for their livelihood. So, while natural changes in city demographics are inevitable, gentrification addresses patterns of inequity.

According to “Advancing Our Understanding of Gentrification,” individuals may disagree on the various effects of gentrification, but research shows resounding clarity that gentrification is in fact occurring at a rapid pace around the country (Ellen and Ding). In each decade from 1980 to 2010, low-income city tracts in metropolitan areas have experienced an increase in each of the following gentrification related categories: income relative to the area, college educated individuals, percentage of white population, and rent prices. These trends show

irrefutable proof of the phenomenon (Ellen and Ding). Therefore, the informed debate does not question the occurrence of gentrification but rather its consequences.

In the field of gentrification studies, it appears unanimous that research on gentrification lacks complete development and leaves many questions unanswered. Most studies focus on residential displacement, analyzing if individuals can remain in their neighborhood as its demographics shift. It has been found that most residents are able to stay in their neighborhood, but the intersection of gentrification and economic outcomes is especially difficult to analyze (Dastrup and Ellen).

This study expands upon the limited existing research addressing economic outcomes in gentrified neighborhoods. Instead of analyzing individual economic outcomes, I look to see if small businesses in gentrified neighborhoods in New York City experience a positive or negative impact on their profitability, which is broadly interpreted by sales volume data, in comparison to non-gentrified neighborhoods. This research helps advance the understanding on how businesses react to gentrifying neighborhood environments, and since it focuses on small businesses, it sheds light on what small business owners experience in terms of their business prospects.

Research Question & Hypothesis

Research Question: How does gentrification change the sales volume of small businesses in New York City?

Hypothesis: When compared to small businesses in non-gentrified New York City neighborhoods, small businesses in gentrified neighborhoods will show an increase in average sales volume, as they experience an influx of wealthy residents who boost consumption rates.

I investigate past studies related to my research study in Chapter 2. I examine creative techniques and data collection methods that have attempted to research gentrification's impacts in real time. I also outline a study focused on small business retention or displacement rates in New York City that guided most of my own research.

To test my hypothesis, I define gentrified New York City zip codes and collect corresponding small business sales volume data as explained in Chapter 3. I utilize SimplyAnalytics to access Census Data and classify gentrified versus non-gentrified neighborhoods, and I use Data Axle to pull historical business information.

In my results section, Chapter 4, I show that there is a statistically significant difference in small business sales volumes for the following independent variables set forth in the study: occurrence of gentrification, year, neighborhood, and zip code. More specifically, I find that in 2020, small businesses in non-gentrified neighborhoods in New York City have a higher average sales volume than businesses in gentrified neighborhoods ($p < .01$).

Chapter 2

Literature Review

Current literature on gentrification focuses on a range of analysis including data on shifting neighborhood demographics, rare findings on financial impacts of gentrification, and potential policy to combat gentrification.

A Creative Approach in Analyzing Shifting Neighborhoods

Technology has enabled recent studies to observe changing neighborhood patterns in a more timely and efficient manner. For example, a study conducted by Harvard professors utilized Yelp data to analyze business trends, especially the movement of establishments into and out of the area (Glaeser et al.). Before applying the Yelp data, the study defined gentrified neighborhoods by using several publicly available metrics.

The researchers first collected data from the Federal Housing Finance Agency (FHFA) to assess shifting real estate prices. They then cross-referenced American Community Survey (ACS) data and Census Zip Code Tabulation Areas (ZCTA) to analyze each of the locations for well-established markers of gentrification, such as percentages of people with a college degree, aged 25-34, and of white race. The study then compares Yelp data consisting of business listings, owners' business reports, and customers' online reviews to the collected housing prices from the FHFA and demographics from the ACS (Glaeser et al.).

The research hypothesized that Yelp data would be able to predict neighborhoods experiencing gentrification and would show shifts in consumer demands, as well as more densely packed businesses and storefronts with larger price tags. It even asserted that a Starbucks

appearing in a neighborhood on the Yelp App is a possible indicator of a wealthy demographic change. Starbucks is a valid representation for the effects of big corporations moving into various neighborhoods and is exceptionally accessible compared to other business data (Glaeser et al.).

The research concluded that for each new Starbucks in a given area each year, there is a .5 percent increase in housing prices. Past that, research showed that increases in college educated, young, and or white people correlated to positive increases in businesses such as groceries, laundromats, cafes, and expensive restaurants. Therefore, the study highlighted how gentrified areas focus on developing more service-based industries. The metrics used in the study are informative for feasible data collection strategies. Yelp data can be used to analyze the movements of businesses and then be compared to other publicly available databases that reflect gentrification patterns, such as the Census Bureau, FHFA, ZCTA, and ACS.

Bias in Judging Neighborhoods

Another study conducted by Naik et al. implements a similarly unique approach in identifying gentrification and investigates how a city's landscape appearance and social and economic development impact each other. The study inquires if social and economic development occurs before, during, or after changes in appearances in neighborhoods. The study utilizes Street Score, an algorithm which gathers Google Street View images to calculate the perception of safety for a given area (Naik et al.). The research focuses on cities known to experience gentrification, such as Baltimore, Boston, Detroit, New York, and Washington D.C. The researchers combine the Street Score, shown to be correlated to the wealth of the

neighborhood, with the previously mentioned metrics of the U.S. Census Data and American Community Survey Data (ACS).

Researchers found that higher levels of education and a higher population density correlated to positive changes in the perceived social safety of a neighborhood. When analyzing the reverse, they found that population densities and incomes were also higher in places after they had previously been perceived with more safety. The study's utilization of Street Score and perceived safety illustrates the social issues, stigma, and perpetual problems behind gentrification (Naik et al.). The research highlights how the effects of gentrification are amplified in poor communities as attitudes prevent community advancement and business investment.

Business Retention and Displacement Rates in Gentrified Areas

A study conducted on the effects of gentrification on small business closely links to my proposed area of research. The goal of the study, titled "Gentrification and Small Business: Threat or Opportunity?", is to analyze the retention and displacement rates of small businesses in New York City (Meltzer).

The study collected data from a private business and leveraged information from a database called the National Establishment Time-Series (NETS). The NETS has the special ability to track the movement of businesses in and out of exact locations over time. NETS data proves valuable since it is usually difficult to obtain general business information down to exact addresses. The NETS also tags businesses with a unique identification code so that they can be tracked if they migrate to a new neighborhood. In the NETS system, the North American

Industry Classification System (NAICS) code of the businesses is also reported, which is a 6-digit number that both broadly and specially classifies the industry of a given business (Meltzer).

To define a gentrified area, the study selects all New York City U.S. Census tracts with populations greater than 200 people. With that guideline in place, Meltzer pulls 1,990 Census tracts, which compose 95% of the Census tracts in New York City. Meltzer attempts to identify gentrified tracts by calculating if the neighborhoods have improved their economic position in the time period of the study. The researcher justifies this decision since there have already been proven causal links between an increase in the financial situations of a neighborhood, corresponding changes in consumer demand, and the given area's market potential for businesses (Meltzer).

In order to analyze changes in the economic positions of neighborhoods, Meltzer starts by identifying neighborhoods as low income if they are in the lowest 40% of income distribution at the start of her given time period. Once she establishes these neighborhoods as defined by lower income, Meltzer investigates if those neighborhoods experienced an increase in household income higher than the average for the city at the end of the time period in question. If the poor neighborhoods have increased their income over time, they fit gentrification criteria since economic gentrification is defined by the economic improvement of low-income neighborhoods throughout time. Of the 1,990 original Census tracts in question, Meltzer defines over 900 as experiencing gentrification.

With gentrified tracts established, the study then looks to analyze business retention and displacement in those areas. If the business is displaced, Meltzer further classifies if the building space is left empty or is occupied by an immediate replacement. Since she is looking at the

effects of retention and displacement for small businesses, she only analyzes businesses with less than 100 employees (Meltzer).

Meltzer also includes analysis on selected case neighborhoods in New York City, including East Harlem, Astoria, and Sunset Park. The study comments that case studies are beneficial since specific neighborhoods can more clearly reflect true gentrification patterns while a large-scale study struggles with refinement (Meltzer).

In its results, the study found that businesses do not have significant differences in retention versus displacement in gentrified versus non-gentrified neighborhoods, further meaning that it found no evidence that gentrification links to pushing small businesses out of their established neighborhoods. However, the research did conclude that when businesses are vacated, it takes longer to replace the establishment in gentrified areas compared to non-gentrified areas. The observation that buildings remain without tenants for longer in gentrified areas can connect back to Naik et al.'s study of perceived neighborhood safety — more closed businesses would not reflect positively on neighborhoods. The lack of perceived safety then contributes to gentrification's effects.

An overarching conclusion of Meltzer's study states that gentrification patterns are variable and case-specific, adding even more evidence that analysis of gentrification is nuanced. Meltzer argues that even without proven business displacement, gentrification still has harmful effects when storefronts are left vacant, and the services offered in their place disrupt the cultural basis of neighborhoods (Meltzer).

Policy Action to Address Gentrification

Various research suggests potential policy to address gentrification. One of the most common suggestions to address gentrification is to set policies that allow residents to maintain their previously affordable housing. Also important is promoting people of different socioeconomic statuses to interact once neighborhoods begin to shift. This could be encouraged through government funding of activities in the community to bridge a possible divide between social classes (Hyra).

One study sets forth extensive policy on neighborhoods with increasing rents and gentrification threats (Lubell). While many properties may be subsidized by the government, problems arise when those subsidies expire, properties require expensive upkeep projects, or properties are not actively operated by a responsive owner. To address these problems, owners can be surveyed on the intention for their properties, communities could facilitate transfers between owners, and owners could be offered tax abatements.

Protection techniques also suggest rent stabilization, which would guarantee limits on rent increases for tenants. As well, inclusionary policy seeks to include options for affordable housing right from the beginning of projects. For example, communities could require new developments to have a portion of affordable living space in their new housing complex. Charitable organizations, such as with The New York acquisition fund, set similar goals by acquiring land with the specific intention to build affordable housing (Lubell). Overall, advocacy from local governments and communities is vital to the success of passing legislation on gentrification and halting the detrimental impacts of shifting neighborhood demographics and exploited minority populations.

Chapter 3

Data & Methodology

There is a lack of substantial research on the financial consequences for businesses in gentrified neighborhoods, due to previously mentioned challenges of ambiguous gentrification metrics, tracking businesses across time, and more focus spent on human rights considerations rather than business impacts. My goal is to calculate if businesses existing in neighborhoods prior to gentrification economically benefit or suffer after the effects of gentrification take place. More specifically, I seek to calculate the change in business profitability after a shift to a more gentrified neighborhood occurs. However, as profitability data for individual businesses is not readily accessible, I instead use businesses' individual sales volume data. Since sales volume represents the success of businesses in selling their goods and does not embed the price of the goods into the equation, it can adequately reflect changes in the general profitability of businesses when compared across years.

I target my data collection for 2015 through 2020. These years are also most feasible since data is available for these years on the two databases I utilize to analyze an effect of gentrification and business sales volumes: SimplyAnalytics and Data Axle.

My research is most significantly based on Meltzer's study in which she analyzes gentrification's effect on the retention or displacement of small businesses in New York City. Since Meltzer defines gentrification by the economic metric of household income, looks at its impact on small businesses, and uses case studies of neighborhoods in New York, the research closely aligns to my objectives.

While I use certain standard gentrification assumptions set by Meltzer, there are also key differences. Meltzer analyzes small business retention or displacement, while I analyze sales

volume to reflect general profitability of the business. As well, Meltzer uses a private NETS Database to collect precise business information, while I combine various sources that will be more thoroughly mentioned in my data collection (Meltzer).

Defining a Gentrification Metric

I leverage SimplyAnalytics data to define which New York City neighborhoods have experienced gentrification from 2015 to 2020 (SimplyAnalytics 2022). SimplyAnalytics pulls data directly from the U.S. Census and various private resources in order to create a more user-friendly platform for data customization and visualization. Since I later need to match SimplyAnalytics Data with Data Axle's Historical Business Data, I focus on aligning the data pulls. Data Axle only has business data linked to zip codes; therefore, I pull zip codes in place of the Census tracts analyzed in Meltzer's and other studies (Data Axle 2022).

SimplyAnalytics has median household income data for 182 zip codes in New York City. For each of the 182 zip codes, I select median household income data for 2015 through 2020. An increase in median household income over the five-year time period will serve as my main indicator for gentrification, since the phenomenon is defined by increasing neighborhood wealth. I also pull other data considered to be defining of gentrification: percentage of white residents, percentage of educational attainment – bachelor's degree and some college, housing tenure by white race, and housing values in the neighborhood. While these are not my main variables, I pull the data to verify the occurrence of gentrification after neighborhoods have indicated gentrification by median household income (SimplyAnalytics 2022).

As Meltzer did in her study, I define the lowest income neighborhoods in 2015 by the bottom two quintiles of median household income. Hence, the lowest 40% of 182 New York City zip codes results in 73 potential low-income zip codes to analyze for an increase in wealth. For these neighborhoods, an increase in wealth from 2015 to 2020 must be above the average increase in median household income for New York City from 2015 to 2020 which rounds closely to 20% when calculated. Only 38 of the 73 low-income zip codes in New York City experience an increase in median household income compared to the average, which means that 38 zip codes qualify to meet the set forth gentrification assumptions.¹

I check to verify that these neighborhoods experience gentrification outside of an increase in income. Out of the 38 neighborhoods classified as gentrified, 63% of neighborhoods experience an increase in white residents, and 89% of neighborhoods experience an increase in percentage of college educated population with a bachelor's degree from 2015 - 2020. Since both gentrification indicators are present in the majority of the 38 neighborhoods that experienced the household income change requirement, the metric proves sound to define gentrified neighborhoods.

As shown in Table 1, the gentrified zip codes are predominantly located in three New York City areas: Manhattan, Brooklyn, and the Bronx. My study first and foremost seeks to analyze overall trends in New York City, so my initial significance tests will include all gentrified zip codes. However, Manhattan, Brooklyn, and the Bronx will serve as potential areas for individual investigation, since Meltzer acknowledged that a more focused view at neighborhoods might reflect more clear results (Meltzer). Other zip codes include Staten Island, Long Island City, Astoria, Rego Park, and Jamaica.

¹ The complete list of gentrified New York City zip codes can be found in Appendix A.

Table 1: Gentrified Zip Codes by Area

Neighborhood	Number of Gentrified Zip Codes
Brooklyn, NY	16
New York, NY	9
Bronx, NY	7
Other	6

Non-Gentrified Neighborhoods

I require the same data on non-gentrified neighborhoods as gentrified neighborhoods so that I can set a comparison group and test if there is a significant difference in the change of sales volume from 2015 to 2020 between groups. I implement similar techniques as I did when defining gentrified zip codes and then apply opposite logic to determine non-gentrified zip codes.

To start, I use the same 182 New York City zip codes available from SimplyAnalytics. However, while I defined gentrified zip code areas by selecting the lowest two quintiles of median household income, I now select the highest singular quintile for median household income for non-gentrified neighborhoods. Thus, I look at the 36 zip codes with the high median income in 2015. I slightly alter this process because, by selecting neighborhoods in the top 20% of household income in New York City, I hope to select the best sample of non-gentrified neighborhoods. It is best for the comparison group to represent the absolute wealthiest sample from New York City.

Within the 36 zip codes, I select the five zip codes with the least amount of change in median household income between 2015 and 2020 when compared to the average increase in

New York City median household income. This logic applies since non-gentrified areas are defined by not experiencing extreme shifts in income brackets. The five zip codes will serve as my comparison group.²

The five selected zip codes that experience minimal change in median household income are located in wealthy Manhattan neighborhoods including Greenwich Village, the Upper East Side, and NoHo. Four of the five neighborhoods experienced a decrease in white residents and education at a bachelor's degree level from 2015 to 2020. With extremely high wealth levels but minimum change in economic status, race, and education, these neighborhoods fit the criteria for non-gentrified areas very well. In the results section of this study, tests for significance on classification refer to the gentrified versus non-gentrified groups from the sample.

Business Sales Volume

With neighborhoods classified as gentrified versus non-gentrified, I look to match the zip codes to New York City historical business information. Through UPenn's Wharton Research Data Services (WRDS) platform, I access Data Axle's application. Data Axle has the capability to list every business that operates in a given zip code each year along with its NAICS code, employee size by location, and sales volume by location (Data Axle 2022).

As part of the U.S. Census, The North American Industry Classification System (NAICS) classifies each business in the United States with an industry code ("North American Industry Classification System – NAICS).” Since I am studying gentrification's impact on businesses, and gentrification is linked to the movement of people through a neighborhood, it is logical to focus

² The complete list of non-gentrified New York City zip codes can be found in Appendix B.

on consumers' interactions with local business. Retail trade reflects consumerism well since it includes both necessities and goods that sell more with increased disposable income such as food and beverages, gasoline, clothing, electronics, tobacco, and home appliances. So, to focus my research on retail trade, I only select businesses for the study starting with the NAICS codes for retail: 44 and 45.

I also look to analyze gentrification's impacts on small business in New York City, given that chain businesses are usually the result of gentrification and are not as personally affected by initial demographic changes (Glaeser et al.). I define a small business in alignment with Meltzer, meaning I only include analysis for businesses with equal to or less than 100 employees.

In my final analysis, I remove businesses that reported zero units in sales volume or did not provide sales volume data in all years. They are removed from the calculation of the average, so they do not factor into data analysis conclusions. It is likely that some businesses failed to report sales volume data due to the impact of COVID-19 in 2020. Business might have also shut down all together in 2020, so reporting zero sales volume may have been accurate.

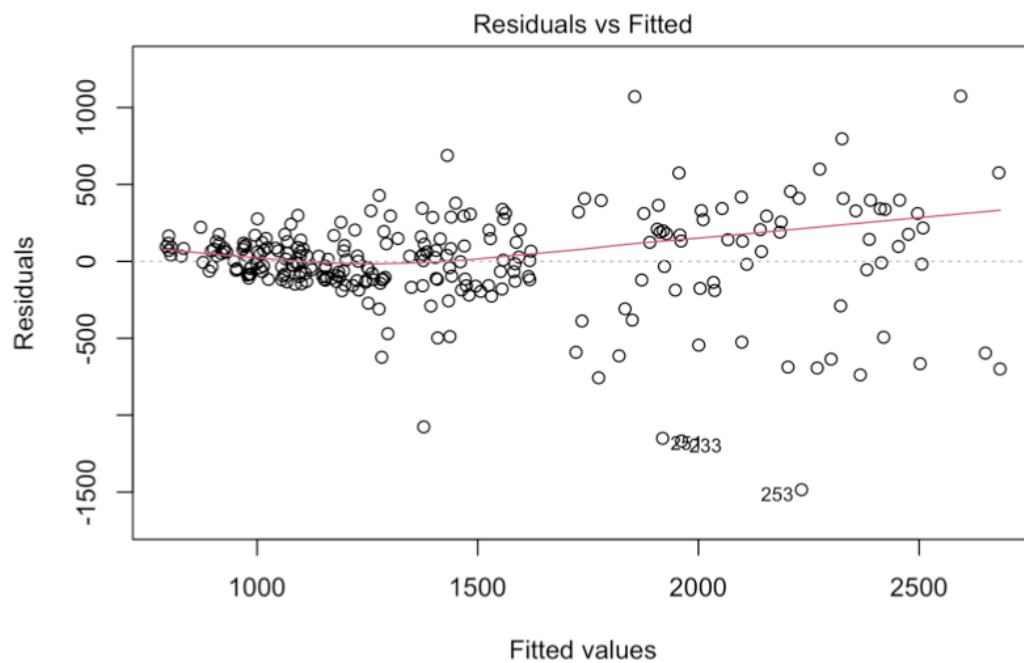
Data Transformation

In my results section, I use ANOVA to test if the difference in sales volume between non-gentrified and gentrified neighborhoods is statistically significant. I will also include significance tests for year, neighborhood, and zip code. After ANOVA results, I use Tukey HSD to analyze more specific relationships between variables and the directionality of their significant relationship. To use an ANOVA model and subsequent post-hoc tests, I must first test for data

normality. Using R Studio, I run both a Residuals vs. Fitted and Normal Q-Q test to check for data linearity (R Core Team).

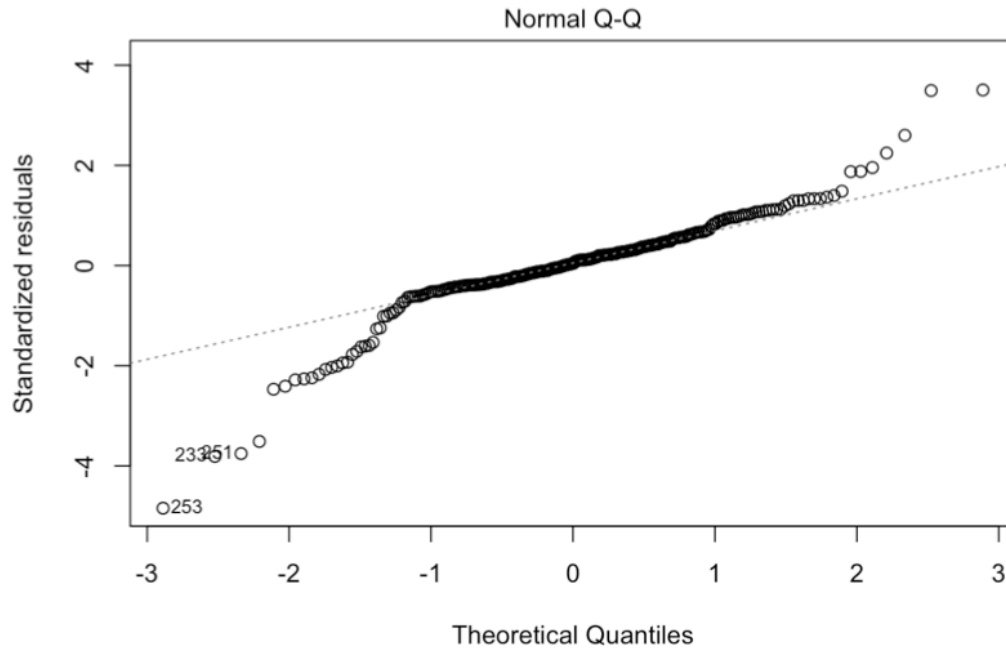
Residual values should not have high levels of deviation from the trend line of the data. Upon running a Residuals vs. Fitted test, it is clear there is a discrepancy between residual values and actual values – see Figure 1. The residual values have a conical shape, indicating a pattern of unequal variance since fitted values become larger on the right side of the graph. Therefore, this does not meet the assumption of equal variance and normality.

Figure 1: Initial Residuals vs. Fitted Data Test



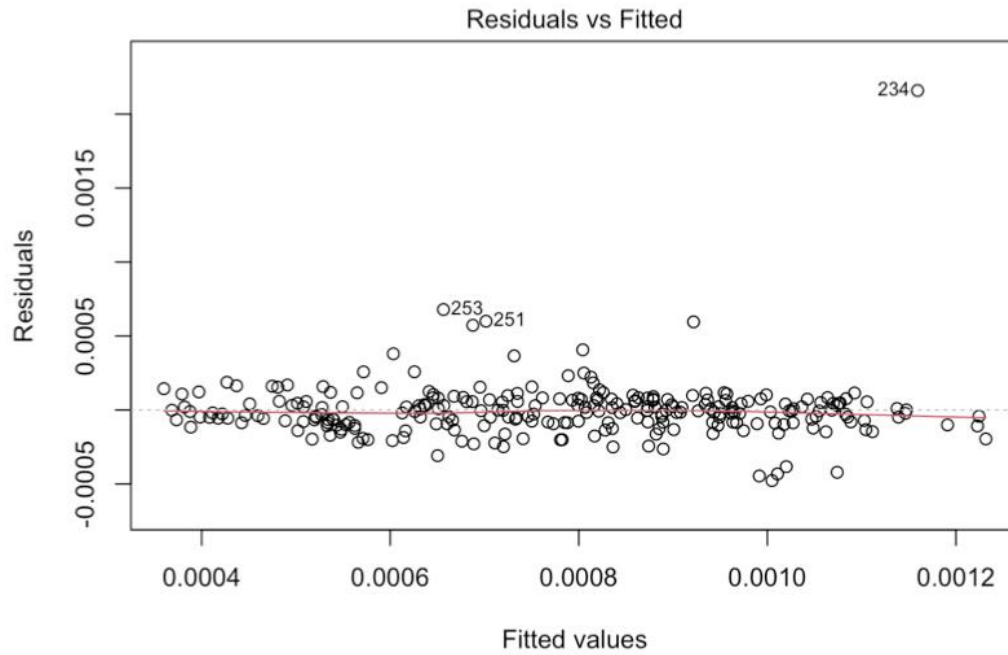
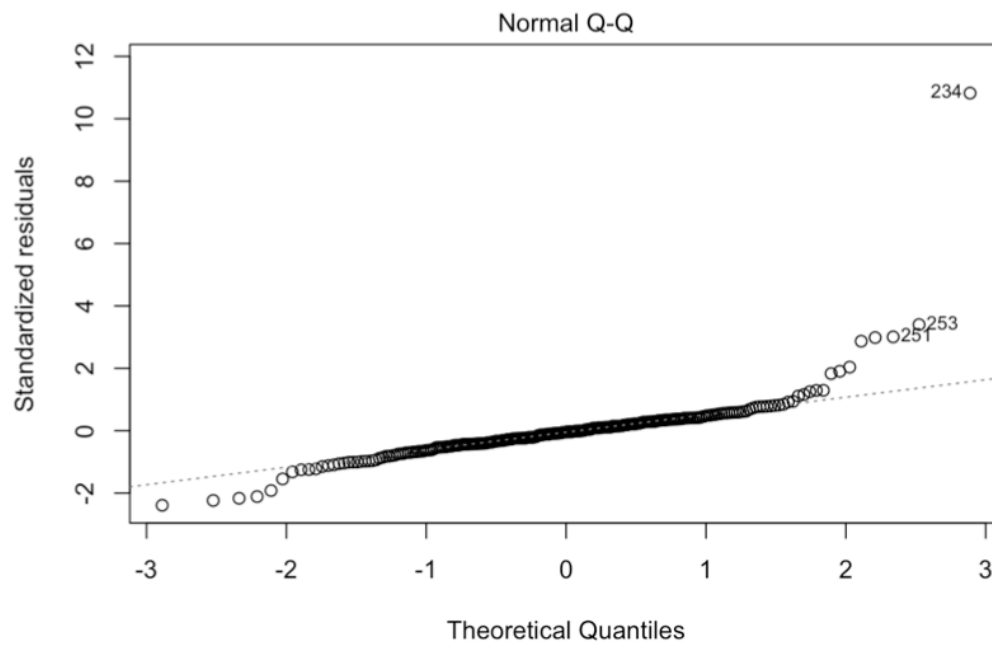
In Figure 2, the Normal Q-Q Data test also shows that the standardized residuals deviate from expected linear normality. There is over dispersed data at both tail ends of the graph, which does not satisfy normality assumptions.

Figure 2: Initial Normal Q-Q Data Test



To correct the lack of data normality, the data can be transformed with the inverse. Using the inverse of the data does not affect the calculation of statistical significance but does cause the sales volume number to be inaccurate. Therefore, after statistical significance is shown, the inverse can be disregarded, and the original data can be utilized in a Tukey HSD test to calculate differences in specific sales volume numbers.

Transforming the data with its inverse proves effective as shown by the improved linearity in Figure 3 and 4. The Residuals vs. Fitted shows more equal variance, satisfying normality's assumption of equal variance of errors. The Normal Q-Q also shows much less skewing at the tails of the graph, with only slight skewing which is largely attributable to outlier values. With this in mind, the transformed data is utilized in the ANOVA test in the Chapter 4 results section.

Figure 3: Transformed Residuals vs. Fitted Data Test**Figure 4: Transformed Normal Q-Q Data Test**

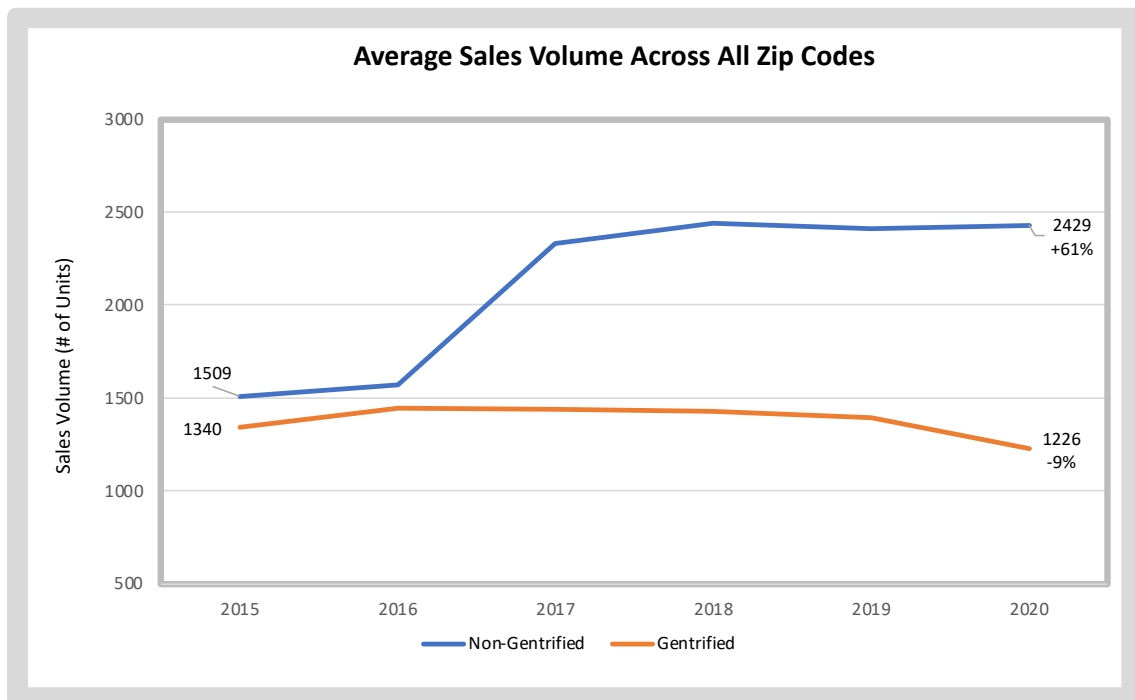
Chapter 4

Results

It is important to note that the following results are based on combined averages instead of a pure total of sales volume. For example, in Figure 5, the number of 1,509 sales volume units in non-gentrified neighborhoods in 2015 is the average of the 5 non-gentrified neighborhoods. Figures 5 and 6 serve to visualize the data, while Tables 2 - 4 include statistical analysis calculated with R Studio computer programming (R Core Team).

In Figure 5, the data shows that small businesses in non-gentrified neighborhoods started with a higher average sales volume than gentrified neighborhoods at the beginning of the study.

Figure 5: Average Sales Volume Across All Zip Codes



However, counter to my hypothesis that gentrified areas would experience an increase in average sales volume over time, Figure 5 shows that gentrified areas experienced a 9% decrease in average sales volume over the 5-year test period. Although, this decrease in average sales

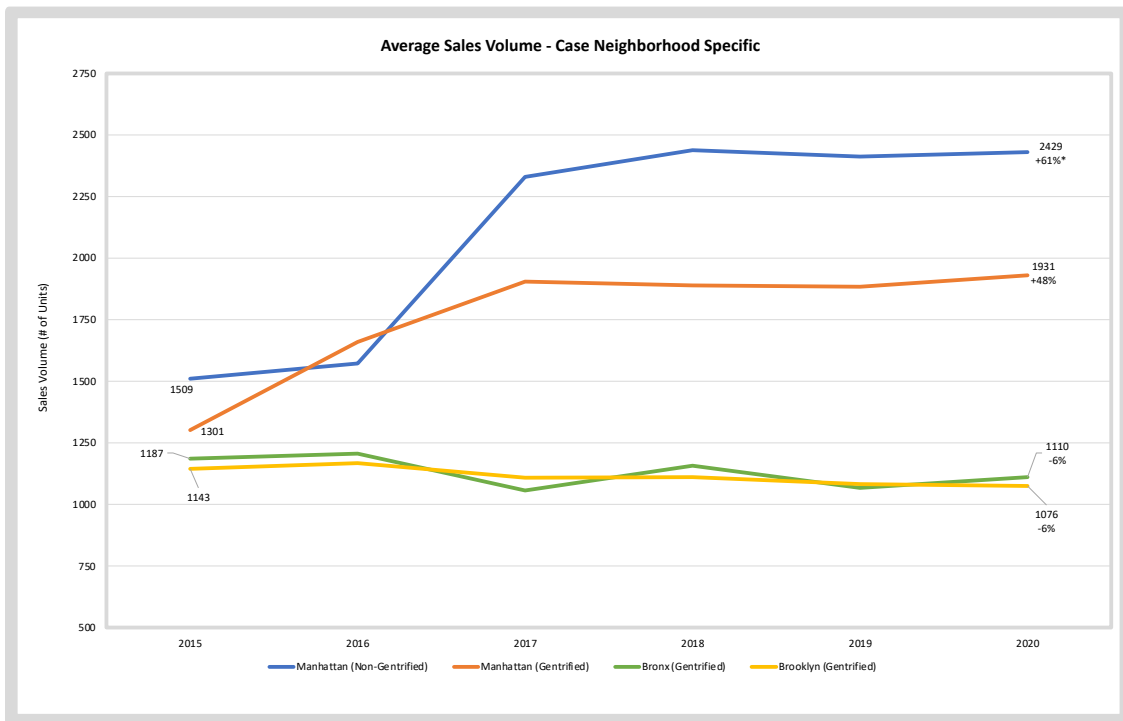
volume in gentrified neighborhoods occurs exclusively from 2019 - 2020, since gentrified neighborhoods experienced a 4% increase up until 2019. This may have been due to a possible pandemic confound from COVID-19, which is discussed further in the analysis section.

Meanwhile, non-gentrified zip codes enjoyed a 61% increase in their overall average sales volume over the course of the study. As is clear in Figure 5, non-gentrified areas experienced a strong increase in average sales volume from 2016 - 2017. This may be explained by the fact that GDP growth in the United States was down -1.36% from 2015 in 2016 but rose back up to an increase of +.62% in 2017 (“GDP Growth (Annual %) - United States”). Alongside this change, the state of New York had a wage growth and personal income growth of just over 2% and around 3.5%, respectively, in 2016. However, in 2017, the state of New York had a wage growth and personal income growth of around 5.25% and almost 7%, respectively (“Economic and Demographic Trends”). Therefore, the New York economy was experiencing a very positive upturn into 2017. The U.S. GDP and New York State growth are relevant to New York City as they effect the general economy and the tourism related to consumer based small businesses in NYC.

However, while the annual data explains why New York City businesses may have greatly benefitted in 2016 - 2017, it does not explain why overall economic conditions only benefitted non-gentrified neighborhoods. It may be possible that non-gentrified neighborhoods benefit more than gentrified neighborhoods from a general increase in economic prosperity. It is also possible that Manhattan benefits more than other neighborhoods from economic growth, especially since it is a neighborhood that experiences more tourism. Even with those considerations, Figure 6 still shows that there are differences in average sales volumes between non-gentrified and gentrified Manhattan neighborhoods.

In Figure 6, the non-gentrified average sales volume data is compared to three case areas of Manhattan, Brooklyn, and the Bronx. On the average, non-gentrified areas in Manhattan strongly outperform gentrified zip codes in Manhattan, Brooklyn, and the Bronx. It is interesting to note that even gentrified areas of Manhattan outperform gentrified areas of the Bronx and Brooklyn. While gentrified Manhattan areas experienced an overall increase of 48% over 5 years, the Bronx and Brooklyn each experience a 6% decrease in average sales volume. This raises the question if small businesses in Manhattan experiences less of an impact from gentrification than other neighborhoods, or perhaps if Manhattan experienced initial gentrification impacts before peripheral neighborhoods did and therefore appear as an outlier.

Figure 6: Average Sales Volume – Case Neighborhood Specific



Analysis of Variance (ANOVA)

I select an ANOVA analysis to test if there is significance between the differences in independent variables. In Table 2, classification refers to the two independent variables of gentrified and non-gentrified neighborhoods. Year includes 2015 - 2020. The ANOVA analysis includes the 8 different neighborhoods and the 43 zip codes determined for the study.

In relation to my main research question, my null hypothesis is that there is not a statistical difference in the average sales volume of small businesses in gentrified versus non-gentrified neighborhoods from 2015 - 2020, while my alternative hypothesis is that there is a difference. The ANOVA analysis proves that the difference in small business sales volumes between gentrified and non-gentrified neighborhoods from 2015 - 2020 is statistically significant ($p < .001$). Therefore, I can reject the null hypothesis and accept my alternative hypothesis.

Table 2: ANOVA Test Results

	<u>Df</u>	<u>Sum Sq.</u>	<u>Mean Sq.</u>	<u>F Value</u>	<u>P-value</u>
Classification	1	1 2.5472e-06	2.5472e-06	52.0798	9.567e-12 ***
Year	5	8.6730e-07	1.7347e-07	3.5467	0.004233 **
Neighborhood	7	4.7682e-06	6.8117e-07	13.9273	8.608e-15 ***
Zip Code	34	3.3247e-06	9.7780e-08	1.9993	0.001710 **
Residuals	210	1.0271e-05	4.8910e-08		
P value key	***[0,.001]	**(.001, .01]	*(.01, .05)		
<i>Computation by RStudio</i>					

Table 2 highlights that the year shows a statistical significance in sales volume ($p < .01$). As well, there is a statistical difference in sales volume for both the 8 New York City neighborhoods ($p < .001$) and the 43 zip codes ($p < .01$).

As briefly mentioned in Chapter 3, ANOVA does not show the direction of the significance in differences. To test the direction of the relationship between the average sales volume of non-gentrified or gentrified neighborhoods, I run a Tukey HSD.

Tukey HSD

A Tukey HSD is a post-hoc ANOVA analysis to determine the significance in the difference of means and values between more specific groups. Tukey HSD tables directly compare each of the independent variables and report the significance. For example, the average sales volume in gentrified neighborhoods in 2015 would be compared against gentrified and non-gentrified neighborhoods in every other year. If the Tukey HSD test shows significance between two independent variables, I remove the inverse transformation from the data, use the original data, and pull specific sales volume numbers. All of the Tukey tables in this section are comprised only of the significant independent variable relationships.

Table 3 shows that that only non-gentrified neighborhoods from 2017 to 2020 have significant differences in average sales volume as compared to gentrified neighborhoods in 2020. All other data comparisons did not show significance and were not included. The Tukey HSD also highlights that non-gentrified neighborhoods experienced a statistically significant greater sales volume starting from 2017 as compared to gentrified 2020 neighborhoods.

The most important information that Table 3 offers is the direct and significant relationship between non-gentrified and gentrified neighborhoods in 2020. Non-gentrified neighborhoods in 2020 had an average sales volume that is 1,204 units greater than gentrified neighborhoods ($p < .01$). Table 3 disputes my research hypothesis that gentrified neighborhoods would experience an increase in average sales volume over time when compared to non-gentrified neighborhoods.

Table 3: Tukey Interactions by Classification

<u>Tukey Interaction</u>	<u>Difference in Sales Volume</u>	<u>P-Value</u>
Non-Gentrified: 2017 - Gentrified: 2020	1104	0.0031657**
Non-Gentrified: 2018 - Gentrified: 2020	1212	0.0021366**
Non-Gentrified: 2019 - Gentrified: 2020	1187	0.0022157**
Non-Gentrified: 2020 - Gentrified: 2020	1204	0.0021217**
P value key ***[0,.001] **(.001, .01) *(.01, .05)		
<i>Computation by RStudio</i>		

On a neighborhood case basis, Table 4 shows that 12 comparisons between sample neighborhoods prove significant. Therefore, while gentrification occurs on a large scale across

New York City, certain gentrified neighborhoods maintain a much higher sales volume than others. The data shows that areas such as Jamaica, Long Island City, Manhattan, and Rego Park all have statistically significant higher average sales volumes than Astoria, Brooklyn, and the Bronx.

Table 4: Tukey Interactions by Neighborhood

<u>Tukey Interaction</u>	<u>Difference in Sales Volume</u>	<u>P-Value</u>
Jamaica - Astoria	457	0.0035316**
Jamaica - Bronx	803	0.0043028**
Jamaica - Brooklyn	820	0.0001959***
Long Island City - Brooklyn	946	0.0044298**
Long Island City - Bronx	929	0.02368*
Long Island City - Astoria	584	0.0067256**
Manhattan - Brooklyn	773	0.00***
Manhattan - Astoria	411	0.0060627**
Manhattan - Bronx	756	0.00013***
Rego Park - Brooklyn	903	0.0075464**
Rego Park - Astoria	541	0.0098139**
Rego Park - Bronx	886	0.0367426*
P value key ***[0,.001] **(.001, .01] *(.01, .05)		
<i>Computation by RStudio</i>		

Analysis

It is expected that neighborhoods not experiencing active gentrification, classified in the study as already highly educated and wealthy, would have initially higher average sales volumes than poor neighborhoods experiencing gentrification. However, it is surprising that, on average, gentrified neighborhoods experienced a 9% decrease in the average sales volumes over the 5-year period. Research would support that gentrified neighborhoods, with actively increasing education and wealth status, should be improving their economic condition. A possible explanation for this occurrence could be that the effects of gentrified neighborhoods have not had enough time to set in over a five-year time and future research is suggested.

As well, COVID-19 serves as a large externality for the study, since many businesses either temporarily closed their businesses for many months in 2020 or experienced an extreme dip in consumerism. However, the effect of COVID-19 is minimized since this study only studies businesses with less than 100 employees. Small businesses should have all experienced similar impacts from COVID-19; therefore, the study avoids the inclusion of large corporate businesses that would have systems in place to stay open in unexpected circumstances. To further this point — as seen in Figure 5 and 6, there were consistent trends in sales volume data as 2019 transitioned to 2020. Non-gentrified areas are continuously higher in its average sales volume than gentrified areas, and non-gentrified neighborhoods and gentrified Manhattan areas consistently outperformed the Bronx and Brooklyn. 2019 data would only cause the Bronx and Brooklyn's sales volume data to be slightly positive rather than slightly negative.

This study showed that sales volume is significantly impacted by the occurrence of gentrification, which relates to the overall profitability of businesses. If sales volume is accepted

as a representation of profitability, then the study highlights the likely possibility that non-gentrified areas are more profitable than gentrified areas, at least in the short term.

Overall, my study found that depending on the year, neighborhood, and specific zip code, the effects of gentrification differ. This supports prior research that gentrification is case specific, variable, and difficult to track. To properly combat gentrification, the government must spearhead standardized data collection efforts catered toward gentrification metrics, and sooner rather than later.

Chapter 5

Conclusion

This study intended to shed light on an issue lacking proper research: gentrification's financial impacts on small businesses. The goal of the paper was to examine whether non-gentrified and gentrified neighborhoods experience different outcomes in their average sales volume over time. Average sales volume analysis serves as the available placeholder metric for general business profitability. This research is valuable since it can contribute economic evidence to the argument over gentrification policy.

I defined gentrified areas in New York City by selecting zip codes in the lowest two quintiles of average household income that increased their economic conditions over the five-year period as compared to the average. Non-gentrified zip codes were selected from the highest quintile of average household income, and experienced close to stagnant change in average household income over the five-year period as compared to the average.

After defining non-gentrified and gentrified neighborhoods in New York City, an ANOVA analysis found that there is statistical significance between the differences of all variables in the study — gentrification, year, neighborhood, and zip code. A post-hoc Tukey HSD analysis showed that most statistical significance exists between specific variable relationships rather than the entire data set, which is logical given that gentrification has been proven to have more observable small-scale effects rather than broadly applicable effects. This research indicates that while gentrification affects residents' circumstances, the impact extends to

the neighborhood's consumer patterns, business prosperity, and outcomes for small business owners.

Further Research

Since the data for this research spanned in data collection from 2015 - 2020, it would be beneficial to determine whether a longer time frame would result in similar results. Given that the last year of the study, 2020, coincided with the effects of COVID-19, future years could offer more external validity.

To further study the difference in financial outcomes between specific New York City neighborhoods, it would be helpful for future research to define non-gentrified zip codes in neighborhoods other than Manhattan.

With the data resources available, it was only possible to use average sales volume to reflect more general patterns of business profitability. It would be useful, therefore, for future research to leverage actual, business specific profitability data to study the effects of gentrification.

Appendix A

Gentrified New York City Zip Codes

Neighborhood	Zip Code
New York	10026
New York	10027
New York	10030
New York	10031
New York	10032
New York	10033
New York	10034
New York	10037
New York	10039
Staten Island	10303
Bronx	10452
Bronx	10453
Bronx	10455
Bronx	10458
Bronx	10459
Bronx	10460
Bronx	10466
Long Island City	11101
Astoria	11102

Brooklyn	11205
Brooklyn	11206
Brooklyn	11213
Brooklyn	11216
Brooklyn	11220
Brooklyn	11221
Brooklyn	11223
Brooklyn	11224
Brooklyn	11225
Brooklyn	11226
Brooklyn	11229
Brooklyn	11230
Brooklyn	11232
Brooklyn	11233
Brooklyn	11235
Brooklyn	11237
Rego Park	11374
Jamaica	11433
Jamaica	11435

Appendix B**Non-Gentrified New York City Zip Codes**

Neighborhood	Zip Code
New York	10003
New York	10014
New York	10016
New York	10023
New York	10128

BIBLIOGRAPHY

Dastrup, Samuel, and Ingrid Gould Ellen. "Linking Residents to Opportunity: Gentrification and Public Housing." *Cityscape*, vol. 18, no. 3, US Department of Housing and Urban Development, 2016, pp. 87–108, <http://www.jstor.org/stable/26328274>.

Data Axle (2022). *Historical Business Data, 2015-2020*. Wharton Research Data Services, 2020. Web. 4 February 2022. <https://wrds-www.wharton.upenn.edu>.

"Economic and Demographic Trends." *Office of the New York State Comptroller*, <https://www.osc.state.ny.us/reports/finance/2020-fcr/economic-and-demographic-trends>.

Ellen, Ingrid Gould, and Lei Ding. "Guest Editors' Introduction: Advancing Our Understanding of Gentrification." *Cityscape*, vol. 18, no. 3, US Department of Housing and Urban Development, 2016, pp. 3–8, <http://www.jstor.org/stable/26328270>.

"GDP Growth (Annual %) - United States." *The World Bank*, <https://data.worldbank.org/indicator/locations=US>.

Glaeser, Edward L., et al. "Nowcasting Gentrification: Using Yelp Data to Quantify Neighborhood Change." *AEA Papers and Proceedings*, vol. 108, American Economic Association, 2018, pp. 77–82.

Hyra, Derek. "Commentary: Causes and Consequences of Gentrification and the Future of Equitable Development Policy." *Cityscape*, vol. 18, no. 3, US Department of Housing and Urban Development, 2016, pp. 169–78, <http://www.jstor.org/stable/26328279>.

Lubell, Jeffrey. "Preserving and Expanding Affordability in Neighborhoods Experiencing Rising Rents and Property Values." *Cityscape*, vol. 18, no. 3, US Department of Housing and Urban Development, 2016, pp. 131–50, <http://www.jstor.org/stable/26328276>.

Meltzer, Rachel. “Gentrification and Small Business: Threat or Opportunity?” *Cityscape*, vol. 18, no. 3, US Department of Housing and Urban Development, 2016, pp. 57–86, <http://www.jstor.org/stable/26328273>.

Naik, Nikhil, et al. “Cities Are Physical Too: Using Computer Vision to Measure the Quality and Impact of Urban Appearance.” *The American Economic Review*, vol. 106, no. 5, American Economic Association, 2016, pp. 128–32, <http://www.jstor.org/stable/43861001>.

“North American Industry Classification System - NAICS.” *United States Census Bureau*, <https://www.census.gov/naics/2022>.

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from <https://www.R-project.org/>

SimplyAnalytics (2022). *Report with Zip Code Data, 2015-2020*. SimplyAnalytics Database, 2020. Web. 2 February 2022. <https://simplyanalytics.com>.

ACADEMIC VITA

Elizabeth Walsh

elizabeth.walsh450@gmail.com

Education:

The Pennsylvania State University Schreyer Honors College Smeal College of Business B.S. in Finance with minor in International Business College of the Liberal Arts Minor in Psychology	University Park, PA Class of 2022
Temple University Study Abroad Program	Rome, Italy/Virtual Jan 2020 - May 2020

Work Experience:

The Pennsylvania State University <i>Teaching Assistant</i> <ul style="list-style-type: none">Managerial TA for a psychology class with over 300 students, tasked with leading fellow TAs' roles and responsibilitiesTA for an honors management class with 30 students, responsible for interactive class discussions, grading, and review sessions	University Park, PA Jan 2021 - Present
PricewaterhouseCoopers <i>Management Consulting Intern - Deals Advisory</i> <ul style="list-style-type: none">Staffed on the finance workstream for a sell-side project worth billions of dollars in projected revenueCreated process flows, developed transition service agreements, and managed separation roadmaps across 6 business unitsIn charge of analyzing spend, reviewing contracts, and finalizing application inventories for 3 IT due diligence projects	New York, NY - Virtual Jun 2021 - Aug 2021
University Professional and Continuing Education Association <i>Research Intern</i> <ul style="list-style-type: none">Utilized SPSS Statistics platform daily to analyze survey data for thousands of university respondentsDeveloped consulting reports for various institutions nationwide, with a specific focus on graduate degree planning	Washington, DC - Virtual Jan 2021 - May 2021
Johnson & Johnson <i>Finance Intern</i> <ul style="list-style-type: none">Created a Power BI dashboard to increase efficiency in consolidating and visualizing quarterly gross profit data for the 2 franchises and 11 platforms of J&J's medical devices company, EthiconWorked with the supply chain team to complete frequent headcount reports and capital budgets	Somerville, NJ - Virtual Jun 2020 - Aug 2020

Leadership:

Phi Beta Lambda Business Fraternity <i>New Member Educator Executive Board Member Professional Development Chair Event Coordinator</i> <ul style="list-style-type: none">Elected to 4 positions in efforts to advance the professionalism of the more than 100 members of PBLNetworked with companies such as Pepsi Co. and J&J to organize exclusive student recruitment eventsServed on the executive board to promote organization wide initiatives and uphold standardsCurrently acting as New Member Educator to lead 30 new members through a semester-long process as they focus on building professional skills surrounding job hunting, resume crafting, and interview techniques	University Park, PA Sept 2019 - Present
Penn State Lion Scout <i>Admissions Tour Guide</i> <ul style="list-style-type: none">In charge of leading at least 3, 1.5 hour tours on a monthly basis for an average of 15 university visitorsResponsible for providing extensive information to students and their families about transitioning to college, enrolling in academic classes, and enjoying student life at the university	University Park, PA Oct 2018 - Present
Nittany Consulting Group <i>Senior Consultant Event Coordinator</i> <ul style="list-style-type: none">Graduated from a 10-week student-led training program that prepares students for a career in the consulting industry and successfully interviewed to join the organization's core team of consultantsConsulted on a professional business engagement with the multi-media company Campus Live and provided a final deliverable on consumer preferences, best practices, and business expansion options	University Park, PA Jan 2019 - May 2020
Penn State Alternative Breaks <i>Volunteer</i> <ul style="list-style-type: none">Aided a non-profit organization, New American Pathways, with their mission of refugee resettlement in AtlantaWorked with the NPO Helping Mamas by sorting donations and reorganizing their distribution center	Atlanta, GA Feb 2019 - Mar 2019

Honors: Dean's List (7/7), 1st Place in J&J's Intern Case Competition, 2nd Place in PwC's Case Competition, Valedictorian

Interests: Cooking Italian food (especially carbonara), alternative rock music, writing sonnets