Gentrification and Local Employment Outcomes: Evidence from New York City

Liam Purkey*

Project Advisor: Sarah West, Department of Economics

28 April, 2021

Abstract

In this paper, I propose two methodological innovations to the gentrification literature. I first propose I data-driven procedure for defining gentrification. I then propose the Synthetic Difference in Differences Estimator presented in Arkhangelsky et al. (2020) as a way to identify the economic effects of gentrification without quasi-experimental variation in the incidence of gentrification. I use these methodological innovations to study the employment effects of gentrification in New York City from 2010 to 2018. I find that gentrification increases the number of neighborhood jobs in Accommodation and Food Services while reducing the number of neighborhood mid-wage and service jobs that are held by residents of those neighborhoods.

*lpurkey@macalester.edu. I am immensely indebted to my advisor Sarah West for her insights and insistence that I continuously revise and improve my work. I am also quite indebted to my committee members Gabriel Lade and Leslie Myint for their suggestions. I thank Amy Damon, Daniel Hartley, Pascal Noel, and Peter Ganong for providing suggestions at various stages of this project. Finally, I thank the member of my Honors Cohort for inspiring me to work harder through their own brilliant work.

Introduction

Although definitions vary, most media discussions and empirical analyses consider gentrification to be an increase in household income, education, and/or housing costs in previously low-income, central city neighborhoods (Ellen and Ding, 2016). No matter the definition, gentrification provokes more emotion than almost any other issue in urban policy. For example, Brooklyn Borough President and popular New York City mayoral candidate Eric L. Adams recently stated that new arrivals to the city should ”go back to Ohio” because ”New York City belongs to the people that was here and made New York City what it is” (Fitzsimmons, 2020).

Detractors of gentrification argue that gentrification raises rents, displaces long-term residents, and damages neighborhood character. Proponents argue that gentrification increases neighborhood economic activity in a way that benefits newcomers and long-term residents alike.

A small body of research attempts to evaluate these claims. The majority of papers in this literature study the relationship between gentrification and displacement, while only a handful of papers study how gentrification changes local economic activity or whether residents of gentrifying neighborhoods stand to benefit from those changes.[1]

No matter the outcome of interest, previous research struggles to address two important methodological problems: the lack of a consistent procedure for defining gentrification and the lack of quasi-experimental variation in the incidence of gentrification.

In this paper, I propose a set of solutions to these problems. To define gentrification in a data-driven manner, I use univariate k-means clustering to first determine which areas are eligible to gentrify and then to determine which neighborhoods ultimately gentrify. To identify the economic effects of gentrification without quasi-experimental variation, I use the synthetic difference in differences estimator presented in Arkhangelsky et al. (2020), which estimate counterfactual outcomes using a data-driven combination of outcomes in non-gentrifying units.

With these methodological innovations in hand, I evaluate the claim that gentrification increases neighborhood economic activity in a way that benefits newcomers

[1] While there is some disagreement, the empirical literature provides evidence that gentrification results in at most a small increase in displacement (McKinnish et al., 2010; Brummet and Reed, 2019; Dragan et al., 2019).

and long-term residents alike using a panel of New York City census tract-level employment data from 2002 to 2018. In particular, I first study how gentrification affects the number of jobs located in gentrifying census tracts by NAICS industry and wage group. I then study whether gentrification affects the number of neighborhood jobs that go to residents of gentrifying neighborhoods.

I find that over the 2011-2018 period, gentrification results in faster employment growth in accommodation and food services but otherwise does not meaningfully alter the industry or wage composition of employment in gentrifying census tracts. In my preferred specification, I find that on average gentrification increases the number of tract-level accommodation and food services jobs by 148 jobs per square mile. Relative to the average number of accommodation and food services jobs in gentrifying census tracts in 2010, this implies a 56% increase.

Turning to the relationship between gentrification and neighborhood residents' employment in gentrifying neighborhoods, I find that gentrification decreases both the number of service and mid-wage jobs that go to neighborhood residents. In my preferred specification, I find that on average gentrification decreases the number of neighborhood service jobs that go to neighborhood residents by 43 jobs per square mile while gentrification decreases the number of neighborhood mid-wage jobs that go to neighborhood residents by 54 jobs per square mile. Relative to average 2010 levels in gentrifying census tracts, these estimates imply a 10% and 17% decrease respectively. I provide suggestive evidence that these job losses are due to changes in the racial composition of gentrifying census tracts and the employment opportunities therein.

While gentrification may benefit neighborhood residents in other ways,[2] my results provide no evidence that gentrification benefits neighborhood residents by changing neighborhood economic activity. In fact, to the degree that residents of gentrifying neighborhoods derive utility from living and working in the same neighborhood, my results suggest that gentrification may harm neighborhood residents.

The rest of the paper proceeds as follows. In Section 1, I review the quantitative and qualitative literature that links gentrification to neighborhood employment outcomes. In Section 2, I discuss my data sources and propose the use of k-means clustering as a way to define gentrification. In Section 3, I verify that my definition

[2]For example, Autor et al. (2017) show that gentrification reduces neighborhood crime rates.

of gentrification is consistent with common perceptions of how gentrification changes neighborhoods and present descriptive results about the characteristics of gentrifying and non-gentrifying census tracts in 2010. In Section 4, I present the synthetic difference in differences estimator and argue that it improves on previous empirical strategies used in the gentrification literature. In Section 5, I present my results. In Section 6, I conclude with a discussion of this paper's limitations and directions for future research.

1 Literature Review

In this literature review, I discuss why gentrification might affect neighborhood employment outcomes and how gentrification correlates with neighborhood employment outcomes in previous empirical analyses. I then highlight a set of qualitative and quantitative papers which suggest that even if gentrification increases neighborhood economic activity overall, long-term residents may lose neighborhood jobs.

1.1 How Gentrification Affects the Types of Jobs in Gentrifying Neighborhoods

The recent gentrification of U.S. cities is driven by young, college-educated individuals who value urban service amenities like dining and nightlife and almost surely increase neighborhood level demand for such services (Couture and Handbury, 2017; Baum-Snow and Hartley, 2020). In this section, I review a set of empirical analyses showing that gentrification is in fact associated with both faster service employment growth and faster *luxury* service employment growth at the expense of manufacturing employment. I then argue that these occupational changes could lead to neighborhood-level wage polarization as in Autor and Dorn (2013).

Lester and Hartley (2014) examine the effects of gentrification on tract-level industrial composition using data for 20 large U.S. cities during the 1990-2008 period. They find that gentrification is associated with faster growth in service employment that is offset by faster contraction in manufacturing employment. Whereas Lester and Hartley (2014) study whether gentrification affects the transition from manufacturing to service employment, Glaeser et al. (2020) investigate the possibility that gentrification affects the transition from low-price services to luxury services.[3] Using Yelp data from New York City, Glaeser et al. (2020) find that gentrification correlates with both growth in retail establishments overall and faster transitions to higher-price point retail establishments.

By changing the industrial composition of employment in gentrifying neighborhoods, gentrification may also change the wage structure of neighborhood employment. Autor and Dorn (2013) document this phenomenon at the city level; they provide evidence that the shift in employment for less-educated individuals from clerical and goods-producing occupations to service occupations resulted in a secular decrease in the number of mid-wage jobs and a secular increase in the number of low-wage jobs going to less-educated individuals.

1.2 How Gentrification Affects whether Neighborhood Residents get Neighborhood Jobs

While there is a clear link between gentrification and the types of jobs located in gentrifying neighborhoods, the relationship between gentrification and whether neighborhood residents get neighborhood jobs is less clear. In this section, I draw on both quantitative and qualitative research to argue that even though gentrification may increase neighborhood employment opportunities overall, neighborhood residents may lose neighborhood jobs due to the racial dimension of gentrification.

Because the college-educated individuals who drive gentrification are overwhelmingly white, new businesses in gentrifying neighborhoods may cater to new white residents rather than long-term minority residents (Baum-Snow and Hartley, 2020). Additionally, these new businesses may replace businesses that cater to minority residents. Results from Agan and Starr (2020) suggest that this racial dimension of gentrification may make it harder for minority residents of gentrifying neighborhoods to get neighborhood jobs. Using data from a large field experiment in New York City during 2015 and 2016, Agan and Starr (2020) show that employers in whiter and less Black neighborhoods discriminate much more heavily against low-skill black applicants in favor of low-skill white applicants.

While Agan and Starr (2020) study employment discrimination in the same place and time period as this paper, they do not explicitly study the racial dimension of gentrification. A set of qualitative papers study this racial dimension and show

[3]For example, this transition may be from cheap eateries to upscale microbreweries and restaurants.

that long-term residents of gentrifying neighborhoods do negatively perceive the effect of gentrification on neighborhood economic activity. Goetz et al. (2019) investigate gentrification in the context of Minneapolis an St. Paul. They find that most residents of gentrifying neighborhoods are concerned that new economic activity does not correspond well to their needs. Sullivan and Shaw (2011) interview long-time Black residents of a gentrifying neighborhood in Portland, Oregon. They find that many Black residents have negative feelings about the retail changes associated with gentrification, and use racial language to express how they feel culturally excluded from those retail changes.

Meltzer and Ghorbani (2017) provide some evidence that neighborhood residents may lose neighborhood jobs due to gentrification. Using annual tract-level commuting data from the New York City Metropolitan Statistical Area during the 2002-2011 period, they find that gentrification is associated with low-wage job losses in residents' home census tracts while it is associated with low-wage job gains within a one-mile radius of residents' home census tract.

2 Data

To study the relation between gentrification and neighborhood employment outcomes, I get annual tract-level employment data from the Longitudinal Employer-Household Dynamics Origin-Destination Statistics (LODES) for the 2002-2018 period and tract-level demographic data from the American Community Survey for the 2006-2010 and 2014-2018 periods. I discuss these data sources in Section 2.1. In Section 2.2, I present a data-driven procedure for defining gentrification and argue that it improves on the ad-hoc definitions used in previous papers.

2.1 Data Sources

Using LODES data, I construct a panel of annual tract-level job counts from 2002-2018 for the 2,167 census tracts in New York City. I include counts of the number of jobs *located* in a census tract by industry, wage, and race. In particular, I include counts for NAICS sectors 72 (Accommodation and Food Services), 31-33 (Manufacturing), and 44-45 (Retail Trade).[4] I also include counts for low-wage jobs

[4]On average, three sectors alone account for almost 30% of 2010 employment in gentrifying census tracts.

(monthly earnings less than \$1,250), mid-wage jobs (monthly earnings between \$1,251 and \$3,333) and high-wage jobs (monthly earnings greater than \$3,333), as well as counts by race (white, Black, Asian) and hispanicity.

To study how gentrification affects the number of neighborhood residents who get neighborhood jobs, I include origin-destination statistics from the LODES data. Each observation in the origin-destination statistics dataset contains counts of individuals who live in tract r and work in tract w in year t. I aggregate these counts so that each observation contains counts of individuals who live in tract r and work in any tract w located in the same neighborhood as tract r in year t. I define neighborhoods using New York City's Neighborhood Tabulation Areas (NTAs), which are administrative boundaries that approximately correspond to historic New York City neighborhoods. For each observation, I include job counts for goods-producing industries and service-producing industries, as well as low-, mid-, and high-wage jobs.[5]

To describe tract-level demographic and economic characteristics, I pull in demographic and economic data from the American Community Survey (ACS) five-year estimates for the 2006-2010 and 2014-2018 periods.[6] I associate the 2006-2010 period estimates with the year 2010 in the LODES data, and the 2014-2018 period estimate with the year 2018 in the LODES data. Because I am interested in matching on pre-gentrification characteristics, I also pull in tract-level data on the share of residents with a college degree in 2000 to track pre-gentrification demographic changes over the 2000-2010 period.[7]

[5] It is important to note that there is no one-to-one relationship between the goods-producing and services-producing job counts in the origin-destination statistics and sector-specific location job counts. Instead, the goods-producing job counts contain information for all goods-producing NAICS sectors including manufacturing while the services-producing job counts contain information on all service-producing NAICS sectors including accommodation and food services and retail trade. Therefore, the services-producing jobs counts also include neighborhood residents who work in neighborhood information, financial activities and business services jobs. See Appendix Section A.2 for a full description of the NAICS sectors included in these job counts.

[6] ACS five-year estimates are generated using 60 months of collected data that are then weighted to produce period estimates. These data are surely generated with error. Unfortunately, it is difficult to assess the extent of this measurement error and its effect on point estimates using available data. Because more precise data are not available at the tract level, it is necessary to tolerate such error (Glaeser et al., 2018).

[7] Because some tract boundaries change between 2000 and 2010, data from the 2000 census must be harmonized to 2010 geographies. This procedure often produces substantial error. In this paper, I include data from the Longitudinal Tract Database-Differential Privacy estimates, which aggregates individual responses to the 2000 Census to 2010 Census geographies.

2.2 Encoding Gentrification

Across previous papers, definitions of gentrification vary widely with respect to the variables and cutoffs they use to define gentrification.[8] In my view, this disorder has two negative consequences for empirical research on the economic effects of gentrification. First, researchers have broad discretion in defining gentrification and consequently have broad discretion in choosing their sample. In the worst-case scenario, researchers may reach their definition of gentrification by iteratively redefining gentrification and assessing the statistical significance of the ensuing results. Second, it is nearly impossible to compare results across papers because each paper studies a different phenomenon.[9]

To remedy this situation, I propose the following solution. For a given unit of analysis:[10]

1. Split the sample into low-, medium-, and high-income groups using k-means clustering of unit median income with $k = 3$. The resulting sample of low-income units are *eligible to gentrify*.

2. Split the sample of low-income units into groups that experience growth in the share of residents with a college degree below the $50th$ percentile for all low-income units and units that experience growth in the share of residents with a college degree above the $50th$ percentile for all low-income units. Denote these groups of low-income units as below-median and above-median units respectively.

3. Split the sample of above-median units into large-growth units and not-large-growth units using k-means clustering of unit growth in the share of residents with a college degree with $k = 2$. Encode units in the large-growth group as gentrifying. Encode units in the below-median group as not gentrifying. Drop units in the not-large-growth group from the sample.

[8] For example, Meltzer and Ghorbani (2017) say a census tract is gentrifying if it experienced an increase in average income, Lester and Hartley (2014) say a census tract is gentrifying if it experienced a percentage increase in educational attainment greater than the percentage increase in educational attainment in the tract's Metropolitan Statistical Area, while Glaeser et al. (2020) say a zip code is gentrifying if its increase in median rent was greater than the median increase in median rent.

[9] In my data, the correlation between tract-level changes in median income and tract-level changes in the share of residents with a college degree is 0.23.

[10] The unit of analysis may be census tracts, zip codes, or neighborhood tabulation areas among others.

In what follows, I refer to this algorithm as the gentrification definition algorithm. Relative to previous ad-hoc definitions, this algorithm has two benefits. First, the use of changes in the share of residents with a college degree is consistent with the literature documenting that college-educated residents are responsible for the recent gentrification of U.S. cities (Edlund et al., 2015; Couture and Handbury, 2017; Baum-Snow and Hartley, 2020).[11] Second, the use of k-means clustering limits the researcher's discretion in choosing their sample in order to obtain statistically significant results.

In this study, I apply the gentrification definition algorithm at the neighborhood tabulation area level. As I discuss in Section 2.1, NTAs are administrative boundaries that approximately correspond to historic New York City neighborhoods and contain 14 census tracts on average. I choose to define gentrification at the NTA level for two reasons. First, census tracts in New York City are quite small.[12] Therefore, observed changes in educational attainment at the census tract level may be idiosyncratic rather than reflective of neighborhood-wide gentrification. Additionally, it is likely that labor market effects are difficult to detect in areas of less than a tenth of a square mile. Second, NTA data are produced using aggregated census tract data, which reduces the dependence of my gentrification measure on the sampling error that is present in tract-level data. I then say that a census tract is gentrifying if it is located in an NTA that is gentrifying and is not gentrifying if it is located in an NTA that is not gentrifying.

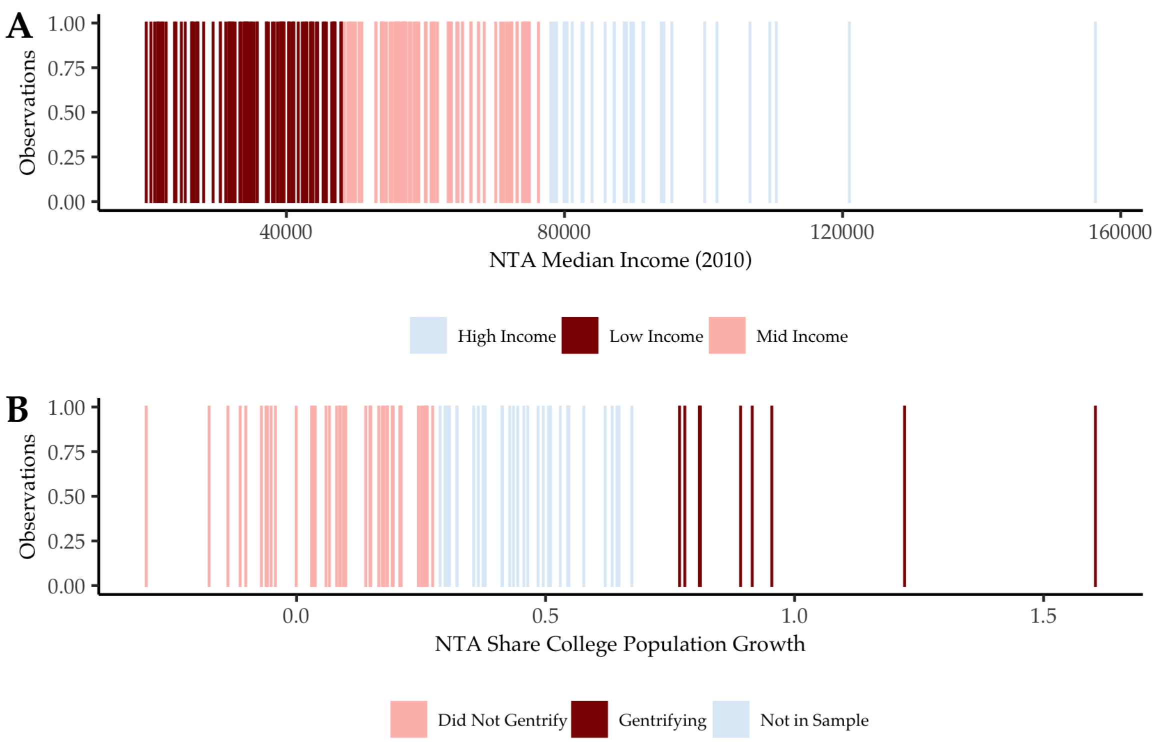
In Panel A of Figure 1, I first show how the gentrification definition algorithm classifies the neighborhood tabulation areas in New York City into low-, medium-, and high-income groups based on 2010 median income. This step of the algorithm produces 82 low-income NTAs from 188 NTAs in New York City. In Panel B of Figure 1, I show how the gentrification definition algorithm classifies the 82 low-income neighborhood tabulation areas into gentrifying and non-gentrifying groups based on 2010-2018 growth in the share of residents with a college degree. These steps of the algorithm produce 50 NTAs, of which 9 are gentrifying and 41 are not.

As I show in Table 15 of the Appendix, gentrification from 2010 to 2018 was limited to neighborhood tabulation areas in Brooklyn and The Bronx. Therefore, I

[11]See Section A.1 of the Appendix for a detailed discussion of the limitations of using alternative variables to define gentrification.

[12]The average size of the census tracts I use for estimation in Section 5 is slightly less than a tenth of a mile while the average size of the NTAs I use is approximately one mile

Figure 1: Visual Implementation of the Gentrification Definition Algorithm



Notes: In Panel A of Figure 1, I show how the gentrification definition algorithm classifies the 188 neighborhood tabulation areas in New York City into low-, medium-, and high-income groups based on 2010 median income. In Panel B of Figure 1, I show how the gentrification definition algorithm classifies the 82 low-income neighborhood tabulation areas into gentrifying and non-gentrifying groups based on 2010-2018 growth in the share of residents with a college degree.

drop the 12 non-gentrifying NTAs located in Manhattan, Staten Island and Queens so that comparisons are not made between census tracts in boroughs were gentrification did occur and boroughs where gentrification did not occur.[13] After dropping these 12 non-gentrifying NTAs, my sample includes 9 NTAs that gentrified from 2010 to 2018 and 29 NTAs that did not gentrify. At the census tract level, my sample includes 112 gentrifying census tracts and 331 non-gentrifying census tracts.

3 Descriptive Results

In Section 3, I first verify that the definition of gentrification in Section 2.2 captures the demographic changes commonly associated with gentrification. I then show that gentrifying and non-gentrifying census tracts differ with respect to numerous 2010 demographic and economic characteristics. I collapse these differences to a single measure using the propensity score estimation procedure in Imbens (2014) and show that tracts that gentrified from 2010 to 2018 were much more likely to gentrify given 2010 characteristics. I use these results to motivate the empirical strategy I present in Section 4.1.

3.1 Validating the Gentrification Measure

In Table 1, I show that the definition of gentrification I propose in Section 2.2 does well at capturing the tract-level demographic and economic changes that are commonly associated with gentrification. From 2010 to 2018, gentrifying census tracts had faster population growth overall (10% vs. 6%), as well as faster growth in the number and share of residents with a college degree (141% vs. 34% and 120% vs. 28% respectively). Gentrifying census tracts also experienced faster growth in the share of white residents (114% vs. 26%), which corresponds well with the literature documenting the racial dimension of gentrification (Sullivan and Shaw, 2011; Baum-Snow and Hartley, 2020). As detractors claim, median rent also grew at a faster rate in gentrifying census tracts (19% vs. 12%).

[13] See Figure 8 in Appendix Section A.7 for a map of the spatial distribution of gentrification in New York City.

Table 1: Average Demographic and Economic Changes from 2010 to 2018 in Gentrifying and Non-Gentrifying Tracts

	Gentrifying	Not Gentrifying	Difference
Population	10%	6%	4**
	(21)	(19)	(2.05)
College Population	141%	34%	107***
	(113)	(111)	(8.55)
Percent College	120%	28%	92***
	(97)	(100)	(8.51)
Percent Asian	143%	37%	106**
	(428)	(223)	(1.99)
Percent Black	15%	10%	0.05
	(92)	(83)	(0.53)
Percent Hispanic	11%	32%	21*
	(67)	(154)	(-1.9)
Percent White	114%	26%	88***
	(229)	(114)	(3.87)
Median Income	24%	3%	21***
	(36)	(27)	(5.71)
Median Rent	19%	12%	7***
	(18)	(14)	(4.02)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: Sample size of 112 gentrifying tracts and 331 non-gentrifying tracts. For each outcome, I report the average percentage change from 2010 to 2018 for gentrifying and non-gentrifying tracts and the difference in average percentage changes between gentrifying and non-gentrifying tracts. I report standard deviations and t-statistics in parentheses. For population variables I report percentage changes in raw population counts. For percent variables I report percentage changes in percents.

3.2 Assessing Pre-Gentrification Differences between Gentrifying and Non-Gentrifying Tracts

In addition to experiencing substantially different demographic and economic changes from 2010 to 2018, gentrifying and non-gentrifying tracts were quite different in 2010. In Table 2, I show that gentrifying tracts were less white (21% vs. 34%) and more Black (50% vs. 30%). Gentrifying tracts also had lower median incomes ($32,139 vs. $35,367). Consistent with results from Couture and Handbury (2017) and Baum-Snow and Hartley (2020), gentrifying tracts were closer to Midtown Manhattan (New York City's central business district) and to high-income neighborhoods. In terms of 2010 employment characteristics, gentrifying tracts had less retail (754 vs. 1,009) and low-wage employment (1,505 vs. 2,128) per square mile and more manufacturing (388 vs.195) employment per square mile. Together, these differences suggest that the young, college-educated individuals who drive gentrification were willing to live in poorer neighborhoods with less retail consumption opportunities in exchange for proximity both to Manhattan and high-income neighborhoods.

In Figure 2, I map these differences to a single measure using the propensity score estimation algorithm in Imbens (2014).[14] I show that census tracts that gentrified from 2010 to 2018 were much more likely to gentrify given 2010 characteristics. These large differences in the likelihood of gentrification raise the concern that observed and unobserved variables driving selection into gentrification also drive 2011-2018 employment outcomes.

[14] See Appendix Section A.3 for a more detailed discussion.

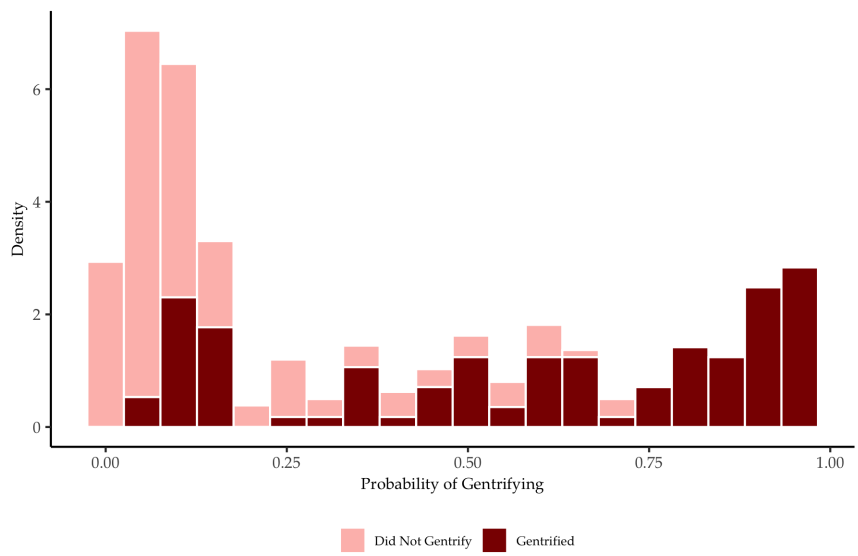
Table 2: Neighborhood and Employment Characteristics of Gentrifying and Non-Gentrifying Tracts in 2010

	Gentrifying	Not Gentrifying	Difference
Neighborhood Characteristics			
Percent College	0.09	0.10	0.01
	(0.06)	(0.06)	(-0.54)
Percent Asian	0.02	0.08	-0.06^{***}
	(0.03)	(0.14)	(-6.89)
Percent Black	0.50	0.30	0.20^{***}
	(0.30)	(0.27)	(6.13)
Percent Hispanic	0.41	0.40	0.01
	(0.30)	(0.26)	(0.24)
Percent White	0.21	0.34	-0.13^{***}
	(0.19)	(0.31)	(-5.13)
Median Income	32,139	35,367	$-3,228^{***}$
	(10,119)	(12,708)	(-2.7)
Median Rent	941	970	-29.0
	(205.75)	(217.43)	(-1.27)
Distance to CBD	5.66	9.24	-3.58^{***}
	(2.42)	(2.86)	(-12.77)
Distance to Nearest	2.57	4.26	-1.69^{***}
High Income Tract	(1.33)	(1.54)	(-11.04)
Employment Characteristics			
Food Service Employment	265.05	310.29	-45.25
	(327.37)	(420.90)	(-1.16)
Retail Employment	754.02	1009.34	-255.32^{*}
	(1,134.81)	(1,313.98)	(-1.96)
Manufacturing Employment	388.25	195.22	193.03^{*}
	(1,040.78)	(511.55)	(1.88)
Low wage Employment	1,505.32	2,128.54	-623.22^{**}
	(1,737.37)	(3,547.50)	(-2.40)
Mid wage Employment	2,808.69	3,048.04	-239.35
	(3,415.73)	(4,478.16)	(-0.58)
High wage Employment	1,523.08	2,187.41	-664.33
	(2,360.59)	(6,742.65)	(-1.50)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: Sample size of 112 gentrifying tracts and 331 non-gentrifying tracts. For each outcome, I report the 2010 mean for gentrifying and non-gentrifying tracts and the difference in means between gentrifying and non-gentrifying tracts. I report standard deviations and t-statistics in parentheses.

Figure 2: Probability of Gentrifying given 2010 Characteristics



Notes: In Figure 2 I plot the distribution of propensity scores calculated with 2010 characteristics for gentrifying and non-gentrifying tracts. I calculate the propensity score using the variable selection algorithm in Imbens (2014). See Table 9 in the Appendix for the model used to estimate the propensity score.

4 Empirical Strategy

The primary challenge to identifying the employment effects of gentrification is the possibility that potential employment outcomes are different across gentrifying and non-gentrifying census tracts. For example, college-educated individuals may choose to live in poorer neighborhoods because they anticipate that those neighborhoods will have faster growth in urban amenities such as food services at the same time that jobs in those neighborhoods increasingly go to workers that resemble them. A secondary challenge to identifying the employment effects of gentrification is appropriately adjusting for the large covariate differences between gentrifying and non-gentrifying tracts presented in Section 3.2.

To address these challenges, previous papers in the gentrification literature use a combination of ordinary least squares (OLS) and instrumental variables (IV) methods to estimate how employment outcomes would have evolved in gentrifying neighborhoods *absent gentrification* (Lester and Hartley, 2014; Meltzer and Ghorbani, 2017; Glaeser et al., 2020). However, these methods are unlikely to identify treatment effects because there is little to no quasi-experimental variation in the incidence of gentrification within cities.[15] In the following discussion I present the synthetic difference in differences (SDID) estimator, which to my knowledge has not been used in the gentrification literature, and argue that the estimator improves on previous methods (Arkhangelsky et al., 2020).

4.1 The Synthetic Difference in Differences Estimator: A Novel Approach

In this section, I briefly describe the synthetic difference in differences estimator presented in Arkhangelsky et al. (2020). I then argue that the synthetic difference in differences estimator improves on previous methods by transparently estimating counterfactual employment outcomes without quasi-experimental variation while minimizing the extent of extrapolation bias.

At a high level, the synthetic difference in differences estimator uses a combination of synthetic control weights that balance pre-treatment outcomes across gentrifying and non-gentrifying tracts and a two-way fixed effects model to model

[15] See Appendix Section A.4 for a detailed discussion of the limitations of OLS and IV methods in the context of gentrification.

counterfactual employment outcomes (Arkhangelsky et al., 2020). Because these weights are non-negative and sum to one, the SDID is both transparent about the source of the estimated counterfactual and is not subject to extrapolation bias (Abadie et al., 2015). Together with the double robustness properties I discuss in Appendix Section A.5, this makes the SDID an attractive estimator in the context of gentrification.

To discuss the synthetic difference in differences estimator in more detail, I introduce the following notation. I consider N tracts, of which the first $i = 1, ..., N_{co}$ do not gentrify and the last $i = N_{co} + 1, ..., N$ gentrify starting in period $T_{pre} + 1$. For each tract-year pair I observe an employment outcome of interest Y_{it} and a set of $c = 1, ..., C$ time-invariant covariates.

As in Arkhangelsky et al. (2020), I consider the following factor model for potential outcomes

$$\begin{aligned} Y_{it}(0) &= L_{it} + \varepsilon_{it} \\ Y_{it}(1) &= Y_{it}(0) + \tau W_{it} \end{aligned} \tag{1}$$

where L_{it} is the i,tth entry of the factor matrix $\mathbf{L} = \mathbf{\Gamma}\mathbf{\Upsilon}^\top$, W_{it} is a binary variable equal to one if tract i gentrifies in period t, and ε_{it} is a tract-year individual error term. In this framework, the $i \times t$ factor matrix $\mathbf{L}$ represents the systematic component of the $i \times t$ potential outcome matrix $\mathbf{Y}(0)$.

The factor matrix $\mathbf{L} = \mathbf{\Gamma}\mathbf{\Upsilon}^\top$ is quite general. For example, in the case where Y_{it} denotes mid-wage employment in tract i and time t, $\mathbf{\Gamma}$ is an $i \times j$ matrix whose ith row vector γ_i contains information on j time-invariant tract-level determinants of mid-wage employment. These determinants could include the share of employment in service occupations, population density, and distance to public transit. $\mathbf{\Upsilon}^\top$ is a $j \times t$ matrix whose tth column vector υ_t contains information on the year-specific importance of each determinant j, so that $L_{it} = \gamma_i \upsilon_t$ captures the tract-specific response to a set of common shocks in year t.

In this setting, the treatment assignment matrix $\mathbf{W}$ correlates with the potential outcome matrix $\mathbf{Y}(0)$ whenever gentrification correlates with the systematic component of potential outcomes, $\mathbf{L}$. Intuitively, this correlation will exist whenever the tract-level systematic determinants of employment also affect the college-educated demand for housing in that tract. Given such correlation, a naive comparison of gentrifying and non-gentrifying tracts will produce a biased estimate of the effect of gentrification on employment outcomes.[16]

In order to minimize this bias,[17] the SDID estimator choose tract weights ω to balance observed outcomes in the pre-gentrification period between gentrifying and non-gentrifying tracts. Because the motivating factor model requires tract-specific characteristics to be constant over time, the SDID estimator also chooses year weights λ to balance pre- and post-gentrification observed outcomes in non-gentrifying tracts. In applications, these time weights have the effect of placing most of the weight on the final pre-gentrification periods so that the time-invariance assumption is more likely to hold.

Formally, the SDID estimator chooses ω to solve the constrained optimization problem

$$(\hat{\omega}_0, \hat{\omega}^{sdid}) = \underset{\omega_0 \in \mathbb{R}, \omega \in \Omega}{\arg\min}\ l_{unit}(\omega_0, \omega) \tag{2}$$

where

$$\begin{aligned} l_{unit}(\omega_0, \omega) = &\sum_{t=1}^{T_{pre}} \nu_t \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 \\ &+ \sum_{c=1}^{C} \nu_c \left(\sum_{i=1}^{N_{co}} \omega_i C_{ic} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} C_{ic} \right)^2 + \zeta^2 T_{pre} \|\omega\|^2. \end{aligned} \tag{3}$$

and

$$\Omega = \left\{ \omega \in \mathbb{R}_+^{N_{co}} : \sum_{i=1}^{N_{co}} \omega_i = 1 \right\} \tag{4}$$

To guarantee that each outcome and covariate receives equal weight in the minimization problem, I include weights ν_t and ν_c that are equal to the inverse variance of the outcome in year t and the cth covariate respectively. As Arkhangelsky et al. (2020) suggest, I include an L^2 penalty $\zeta^2 T_{pre} \|\omega\|^2$ to ensure that the solution is unique given the large number of control. This regularization penalty increases the dispersion of the unit mass across non-gentrifying tracts, which means that the estimator is less subject to idiosyncratic tract-level employment shocks than if the unit mass concentrated on a small subset of non-gentrifying tracts.

[16] In most cases, a simple two-way fixed effects estimator overstates the absolute effect of gentrification relative to the synthetic difference in differences estimator. This suggests that college-educated residents knowingly or unknowingly choose to move to neighborhoods that would experience larger 2011-2018 employment changes even absent gentrification. These results are available on request from the author.

[17] See Appendix Section A.5 for a discussion of the conditions under which the synthetic difference in differences estimator successfully removes bias due to selection on **L**.

The SDID estimator chooses the pre-treatment year weights λ to solve a similar constrained optimization problem

$$(\hat{\lambda}_0, \hat{\lambda}^{sdid}) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda}{\arg\min}\, l_{time}(\lambda_0, \lambda) \tag{5}$$

where

$$l_{time}(\lambda_0, \lambda) = \sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_i Y_{it} - \frac{1}{T_{post}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2 \tag{6}$$

and

$$\Lambda = \left\{ \lambda \in \mathbb{R}_+^{T_{pre}} : \sum_{t=1}^{T_{pre}} \lambda_i = 1 \right\}. \tag{7}$$

With these sample weights $\hat{\omega}^{sdid}, \hat{\lambda}^{sdid}$ in hand, the SDID estimator is computed via the weighted DID regression

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - \tau W_{it})^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right) \tag{8}$$

where α_i is a tract fixed effect, and β_t is a year fixed effect.

I estimate Equation 8 using two sets of weights $\hat{\omega}$ that correspond to different vectors of covariates. In my baseline specification, I calculate $\hat{\omega}$ using pre-treatment outcomes as well as the share of residents with a college degree in 2000 and the share of residents with a college degree in 2010. I include these shares so that on average gentrifying and non-gentrifying tracts experienced similar pre-2011 levels of gentrification. Second, I calculate $\hat{\omega}$ using the covariates from the baseline specification as well as 2010 covariates that significantly predict variation in the average post-2010 level of the outcome of interest.[18] I include these covariates to address the possibility that the nine years of pre-gentrification data are not sufficient to balance the factor matrix **L** between gentrifying and non-gentrifying tracts given the large covariate differences presented in Section 3.[19]

To conduct inference on $\hat{\tau}$, I follow Arkhangelsky et al. (2020) and use a jackknife

[18] I choose covariates using the robust lasso estimator presented in Belloni et al. (2014).

[19] Unfortunately, nine years is a somewhat small number of pre-treatment periods relative to other applications of the synthetic control method.

estimator of the sampling variance V_τ:

$$\hat{V}_\tau = \frac{N-1}{N} \sum_{i=1}^{N} (\hat{\tau}^{-i} - \hat{\tau})^2 \tag{9}$$

where $\hat{\tau}^{-i}$ denotes the estimate of τ after dropping the ith observation. Arkhangelsky et al. (2020) show that in general this estimator of the variance is conservative. I note that this variance estimator assumes independently distributed errors between census tracts. While this assumption may not hold in the data, there is no asymptotic theory for inference on SDID estimates when errors are correlated between units. Because the variance estimator in Equation 9 does not account for correlated errors, I cannot exclude the possibility that results are driven by neighborhood-specific shocks.

4.1.1 Assessing the Robustness of Synthetic Difference in Differences Estimates

I assess the robustness of SDID estimates using two strategies that are common in the synthetic control literature and a separate strategy that is unique to this paper's context. First, I backdate the onset of gentrification in the data to 2006 to test if gentrifying and non-gentrifying tracts experienced divergent employment outcomes prior to the onset of gentrification in 2011. Second, I implement a ”leave-two-out” procedure similar to the leave-one-out procedure in Abadie et al. (2010) in which I iteratively drop all census tracts in every pair of high-weight[20] neighborhoods and then re-estimate the effect of interest.[21] This procedure assesses whether SDID estimates are dependent on idiosyncratic shocks to important comparison neighborhoods.

In the context of this analysis, an important concern is that estimated effects may be driven by distinct employment trends in census tracts that were extremely likely to gentrify and census tracts that were extremely unlikely to gentrify. To address this concern, I re-estimate significant effects using a trimmed sample that

[20]See Appendix Section A.6.3 for a discussion of how I determine which neighborhood tabulation areas are high-weight.

[21]In contrast, the leave-one-out iteratively drops each comparison unit and re-estimates the effect of interest. This is worthwhile in standard synthetic control settings because the unit weights cluster on a small subset of comparison units. Because the synthetic difference in differences unit weights are dispersed by design, the leave-one-out procedure is not applicable.

excludes census tracts that were extremely likely to gentrify and census tracts that were extremely unlikely to gentrify. In particular, I follow the suggestion of Imbens (2014) and drop all census tracts with propensity scores greater than 0.9 or less than 0.1.

5 Results

In Section 5, I present results from the estimation procedure in Section 4.1. I show that over the 2011-2018 period, gentrification resulted in faster employment growth in accommodation and food services but did not otherwise affect the types of neighborhood employment opportunities. In contrast, I show that gentrification did meaningfully decrease the number of neighborhood mid-wage and service jobs that went to neighborhood residents. I show that these results are robust to estimation using alternative samples in Appendix Section A.6.

In addition to reporting treatment effect estimates, I also report a set of statistics describing the quality of the outcome balances obtained with the synthetic difference in differences weights. For the unit weights $\hat{\omega}$ I report the normalized root mean square error between the synthetic control and the target path of pre-gentrification outcomes:

$$\text{NRMSE}_{\hat{\omega}} = \sqrt{\sum_{t=1}^{T_{pre}} \left(\hat{\omega}_0 + \sum_{i=1}^{N_{co}} \hat{\omega}_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2} \left(\frac{1}{N_{tr} T_{pre}} \sum_{t=1}^{T_{pre}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^{-1}. \tag{10}$$

For the time weights $\hat{\lambda}$ I report the normalized root mean square error between the synthetic control and the target vector of post-gentrification outcomes in non-gentrifying tracts:

$$\text{NRMSE}_{\hat{\lambda}} = \sqrt{\sum_{i=1}^{N_{co}} \left(\hat{\lambda}_0 + \sum_{t=1}^{T_{pre}} \hat{\lambda}_i Y_{it} - \frac{1}{T_{post}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2} \left(\frac{1}{N_{co} T_{post}} \sum_{t=T_{pre}+1}^{T_{post}} \sum_{i=1}^{N_{co}} Y_{it} \right)^{-1}. \tag{11}$$

For each covariate in $c = 1, ..., C$ I report the percent difference in means between gentrifying and non-gentrifying tracts:

$$\Delta\% = \frac{\sum_{i=1}^{N_{co}} \hat{\omega}_i C_i - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} C_i}{\sum_{i=N_{co}+1}^{N} C_i} \tag{12}$$

where $\sum_{i=1}^{N_{co}} \hat{\omega}_i C_i$ is the mean value of the covariate C in the re-weighted sample of comparison census tracts.

Prior to discussing specific results, I highlight three properties of the unit weights $\hat{\omega}$ and the time weights $\hat{\lambda}$ that are common across all outcomes and specifications. First, the unit weights $\hat{\omega}$ largely succeed at balancing both pre-treatment outcomes and 2010 covariates across gentrifying and non-gentrifying tracts. Second, even after re-weighting, non-gentrifying tracts had a larger share of residents with a college degree in 2000 and a smaller share in 2010. This suggests that there was a degree of gentrification in gentrifying census tracts prior to 2010. However, I show that these 2000-2010 dynamics do not result in significantly different pre-gentrification employment outcomes across gentrifying and non-gentrifying tracts in Appendix Section A.6. Third, the time weights $\hat{\lambda}$ largely fail to balance pre and post-gentrification outcomes in non-gentrifying tracts. However, this does not invalidate estimates because the SDID estimator only requires that one of the unit weights or time weights produce an approximate match.[22]

5.1 How Does Gentrification Affect the Types of Jobs in Gentrifying Neighborhoods?

In Table 3 I report the average effects of gentrification on the industrial composition of gentrifying census tracts over the 2011-2018 period/ After re-weighting to adjust for different pre-trends and covariate values, I find that gentrification has large significant effects on the number of jobs in accommodation and food services but otherwise does not affect the industrial composition of gentrifying census tracts. Relative to the average number of accommodation and food service jobs in gentrifying census tracts in 2010, I find that on average gentrification increases employment in accommodation and food service by 56% (148/265).

I present this result graphically in Figure 3. In Panel B of Figure 3, I show that even after re-weighting non-gentrifying census with unit weights $\hat{\omega}$, there was much faster post-gentrification employment growth in gentrifying census tracts.[23]

In Table 4, I show that gentrification did not meaningfully alter the wage composition of employment in gentrifying tracts. Across specifications, gentrification has a slightly negative effect on low- and mid-wage employment while gentrification has a slightly positive effect on high-wage employment. However, none of these

[22] See Appendix Section A.5.

[23] For all insignificant results, I report figures in Section A.7 of the Appendix.

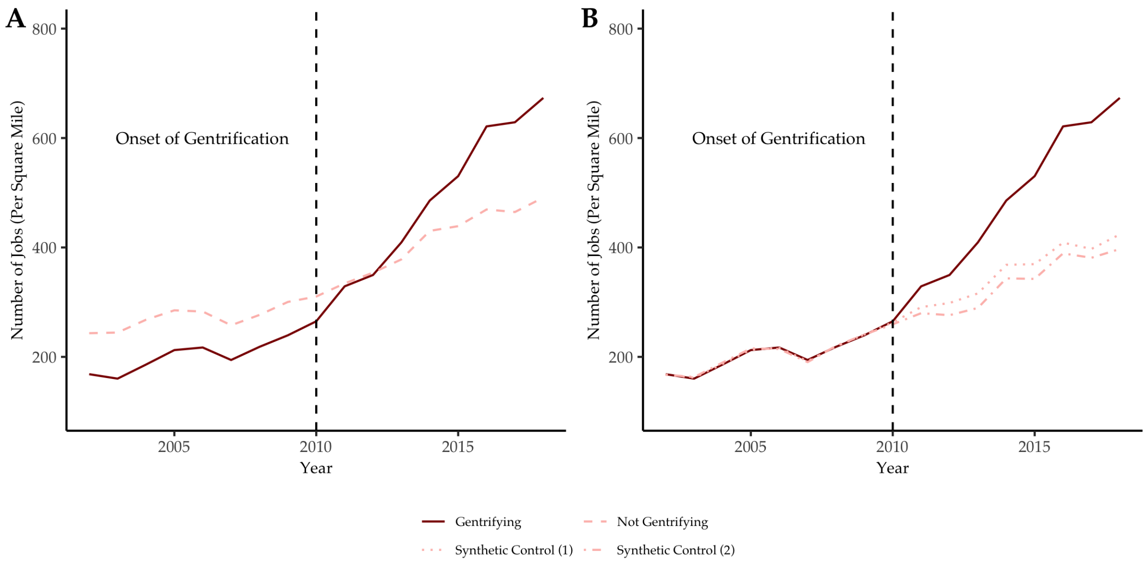
Table 3: Synthetic Difference in Differences Estimates for The Number of Jobs by Industry

	Food (1)	Food (2)	Retail (1)	Retail (2)	Manufacturing (1)	Manufacturing (2)
Gentrifying	139.30***	148.04***	18.25	32.08	−27.83	−25.25
	(3.80)	(4.04)	(0.42)	(0.74)	(-0.51)	(-0.47)
NRMSE						
$\hat{\omega}$	0.007	0.015	0.027	0.030	0.031	0.030
$\hat{\lambda}$	0.559	0.559	0.350	0.350	2.129	2.129
Percent Difference						
Percent College Graduates (2000)	0.080	0.104	0.120	0.079	0.200	0.078
Percent College Graduates (2010)	-0.032	-0.032	-0.013	-0.036	0.030	-0.034
Percent Asian		0.303				
Percent Hispanic		0.008				
Percent White				0.022		
Median Rent		0		0.0007		0.0013
Percent College Graduates (NTA)		-0.015				
Tracts	423	423	423	423	423	423
Observations	7191	7191	7191	7191	7191	7191

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: For each outcome, I report synthetic difference in differences estimates for my baseline specification (1) and a specification with covariates chosen by lasso (2). I report estimates in the number of jobs per square mile. I report t-statistics in parentheses.

Figure 3: Effect of Gentrification on Employment in Accommodation and Food Services



Notes: In Panel A of Figure 3, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 3, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 3.

effects are statistically significant. Overall, these null effects are consistent with the theory gentrification changes the wage composition of neighborhood employment by accelerating the transition from manufacturing to service employment.

5.2 How Does Gentrification Affect who Gets Jobs in Gentrifying Neighborhoods?

In Section 5.1, I show that gentrification results in faster employment growth in accommodation and food services but otherwise does not meaningfully change the number of employment opportunities by industry and wage in gentrifying census tracts. In this section, I show that even though gentrification does not change the number of neighborhood employment opportunities, gentrification does change the live-work patterns of residents of gentrifying neighborhoods.

In Table 5, I report the effect of gentrification on the number of neighborhood jobs that go to neighborhood jobs by industry. I find that gentrification significantly reduces the number of neighborhood service jobs that go to neighborhood residents without affecting the number of goods-producing jobs that go to neighborhood

Table 4: Synthetic Difference in Differences Estimates for The Number of Jobs by Wage

	Low-Wage (1)	Low-Wage (2)	Mid-Wage (1)	Mid-Wage (2)	High-Wage (1)	High-Wage (2)
Gentrifying	−65.54	−12.33	−323.64	−251.35	111.09	116.57
	(-0.52)	(-0.10)	(-1.37)	(-1.06)	(0.69)	(0.71)
NRMSE						
$\hat{\omega}$	0.021	0.025	0.020	0.027	0.034	0.072
$\hat{\lambda}$	0.632	0.632	0.737	0.737	0.651	0.651
Percent Difference						
Percent College Graduates (2000)	0.082	0.079	0.081	0.079	0.079	0.056
Percent College Graduates (2010)	-0.033	-0.039	-0.033	-0.038	-0.032	-0.047
Percent Asian		0.252		0.281		
Percent Hispanic						-0.047
Percent White		0.019		0.021		0.162
Median Rent		0.0003		0.0002		0.0007
Distance to CBD						0.087
Tracts	423	423	423	423	423	423
Observations	7191	7191	7191	7191	7191	7191

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: For each outcome, I report synthetic difference in differences estimates for my baseline specification (1) and a specification with covariates chosen by lasso (2). I report estimates in the number of jobs per square mile. I report t-statistics in parentheses.

residents. In my preferred specification, gentrification reduces the number of neighborhood service jobs per square mile that go to neighborhood residents by 10% (43/418) relative to the average 2010 level of employment in gentrifying tracts.

In Table 6, I show that gentrification decreases the number of neighborhood mid-wage jobs that go to neighborhood residents by a similar amount. In my primary specification, gentrification reduces the number of mid-wage jobs that go to neighborhood residents by 17% (54/311) relative to the average 2010 level in gentrifying tracts.

Table 5: Synthetic Difference in Differences Estimates for The Number of Jobs Going to Neighborhood Residents by Industry

	Goods (1)	Goods (2)	Services (1)	Services (2)
Gentrifying	-1.87	-1.02	-54.28^{**}	-43.23^{*}
	(-0.34)	(-0.19)	(-2.41)	(-1.94)
NRMSE				
$\hat{\omega}$	0.022	0.028	0.016	0.015
$\hat{\lambda}$	0.657	0.657	0.397	0.397
Percent Difference				
Percent College Graduates (2000)	0.085	0.086	0.091	0.105
Percent College Graduates (2010)	-0.034	-0.037	-0.029	-0.025
Percent Asian		0.257		0.325
Percent Hispanic		0.012		0.02
Percent White		0.024		0.056
Tracts	423	423	423	423
Observations	7191	7191	7191	7191

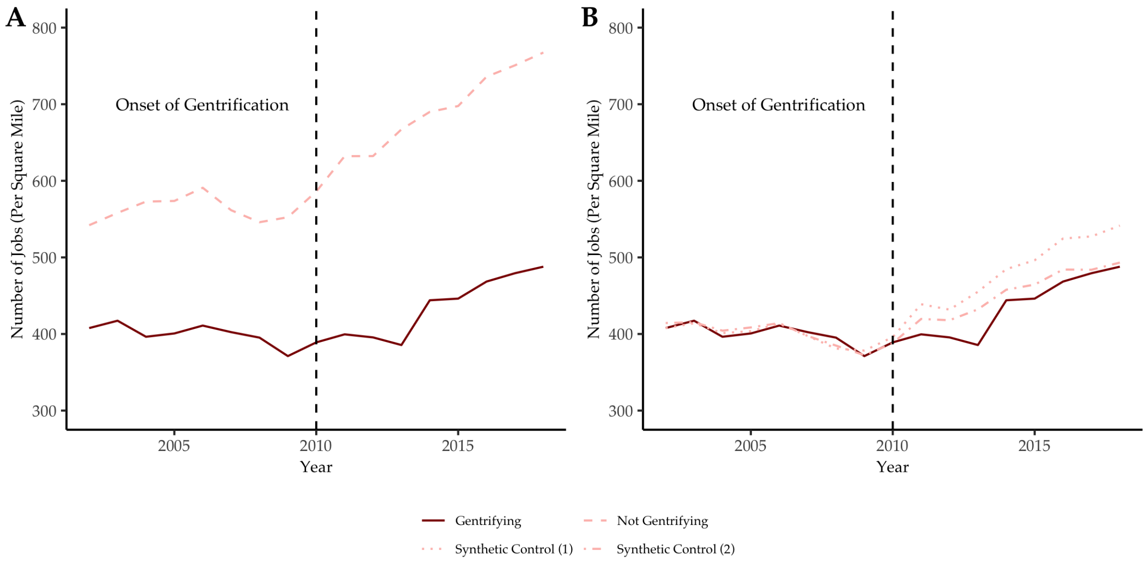
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: For each outcome, I report synthetic difference in differences estimates for my baseline specification (1) and a specification with covariates chosen by lasso (2). I report estimates in the number of jobs per square mile. I report t-statistics in parentheses.

5.2.1 Interpreting Live-Work Treatment Effects

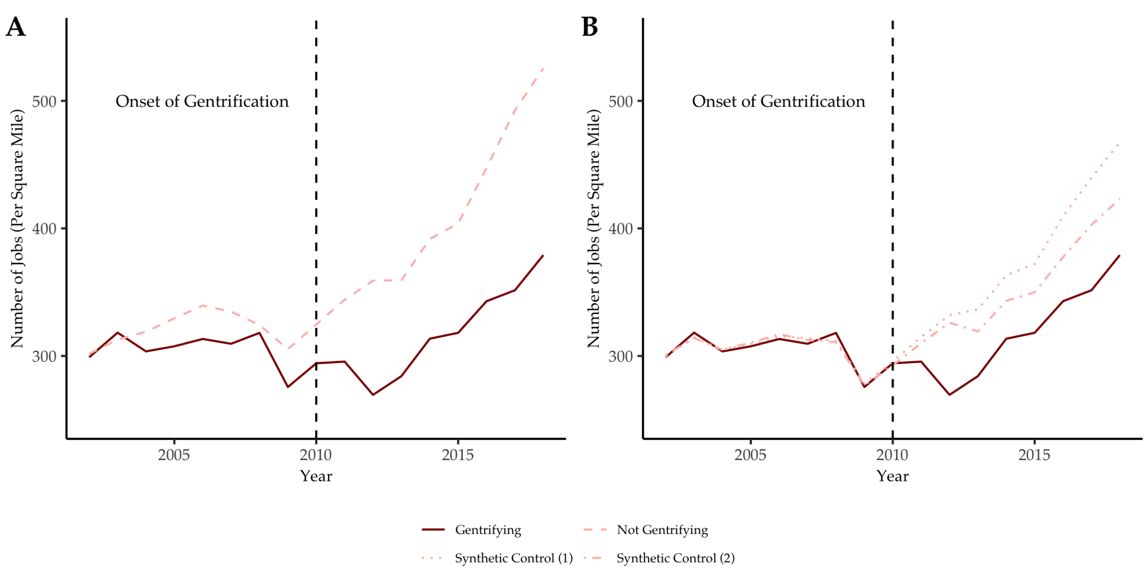
In Tables 5 and 6 I show that over the 2011-2018 period, gentrification results in a decrease in the number of residents who live in gentrifying census tracts and hold both service and mid-wage jobs in their census tract's neighborhood. However, the meaning of these results is not immediately clear because changes in aggregate data

Figure 4: Effect of Gentrification on Neighborhood Employment in Service-Producing Industries



Notes: In Panel A of Figure 4, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 4, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 5.

Figure 5: Effect of Gentrification on Employment in Mid-Wage Jobs



Notes: In Panel A of Figure 5, I plot the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel B of Figure 5, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 6.

Table 6: Synthetic Difference in Differences Estimates for The Number of Jobs Going to Neighborhood Residents by Wage

	Low-Wage (1)	Low-Wage (2)	Mid-Wage (1)	Mid-Wage (2)	High-Wage (1)	High-Wage (2)
Gentrifying	-5.67	1.85	-63.34^{***}	-54.41^{***}	6.48	9.59
	(-0.31)	(0.10)	(-3.99)	(-3.46)	(0.84)	(1.28)
$\hat{\omega}$	0.016	0.021	0.010	0.012	0.039	0.047
$\hat{\lambda}$	0.377	0.377	0.455	0.455	0.513	0.513
NRMSE						
Percent College Graduates (2000)	0.080	0.084	0.091	0.083	0.081	0.076
Percent College Graduates (2000)	-0.033	-0.035	-0.029	-0.036	-0.033	-0.032
Percent Difference						
Percent Asian		0.236		0.243		
Percent Black						-0.018
Percent Hispanic		0.013		0.013		
Percent White		0.026		0.029		-0.009
Median Rent		-0.0003				-0.0003
Tracts	423	423	423	423	423	423
Observations	7191	7191	7191	7191	7191	7191

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: For each outcome, I report synthetic difference in differences estimates for my baseline specification (1) and a specification with covariates chosen by lasso (2). I report estimates in the number of jobs per square mile. I report t-statistics in parentheses.

do not reveal which types of neighborhood residents are losing jobs. In this section, I present a simple decomposition of the SDID estimator to clarify the meaning of results in Tables 5 and 6. I then argue that these results likely capture the fact that gentrification makes it harder for neighborhood residents to get neighborhood jobs by changing the racial composition of employment opportunities within industries and wage groups, although I cannot definitively show that this is the case with available data.

In a given year t, census tract i's population is composed of non-gentrifying residents L_{it} and gentrifiers G_{it}. Non-gentrifying residents may be individuals who lived in tract i prior to the onset of gentrification or residents who did not live in tract i prior to the onset of gentrification but are similar to them with respect to socio-economic status. In this context, the number of individuals who live in tract i and work in tract i's neighborhood (by wage or industry) can be decomposed as

$$Y_{it} = L_{it} + G_{it}. \tag{13}$$

Now the synthetic difference in differences estimator is simply a weighted two-way fixed effects estimator. Therefore, the Frisch-Waugh theorem implies that the SDID estimator can be written as

$$\hat{\tau}^{\text{sdid}} = \frac{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t Y_{it} \tilde{W}_{it}}{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t \tilde{W}_{it}^2} \tag{14}$$

where $\tilde{W}_{it}$ is the gentrification indicator with unit and time means removed. Substituting $L_{it} + G_{it}$ for Y_{it} in Equation 14 and taking expectations produces

$$\begin{aligned}
\mathbb{E}\left[\hat{\tau}^{\text{sdid}}\right] &= \mathbb{E}\left[\frac{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t Y_{it} \tilde{W}_{it}}{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t \tilde{W}_{it}^2}\right] \\
&= \mathbb{E}\left[\frac{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t (L_{it} + G_{it}) \tilde{W}_{it}}{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t \tilde{W}_{it}^2}\right] \\
&= \mathbb{E}\left[\frac{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t L_{it} \tilde{W}_{it}}{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t \tilde{W}_{it}^2} + \frac{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t G_{it} \tilde{W}_{it}}{\sum_i \sum_t \hat{\omega}_i \hat{\lambda}_t \tilde{W}_{it}^2}\right] \\
&= \mathbb{E}\left[\hat{\tau}_L^{\text{sdid}}\right] + \mathbb{E}\left[\hat{\tau}_G^{\text{sdid}}\right]
\end{aligned} \tag{15}$$

where $\hat{\tau}_L^{\text{sdid}}$ is the SDID estimator of the number of neighborhood jobs that go to non-gentrifying residents and $\hat{\tau}_G^{\text{sdid}}$ is the SDID estimator of the number of neighborhood jobs that go to gentrifying residents.

Because the number of gentrifiers who move into non-gentrifying neighborhoods is small by definition, it is likely that $\mathbf{E}\left[\hat{\tau}_G^{\text{sdid}}\right] > 0$ and $\mathbf{E}\left[\hat{\tau}^{\text{sdid}}\right] > \mathbf{E}\left[\hat{\tau}_L^{\text{sdid}}\right]$. It follows that on average the synthetic difference in difference estimator using aggregate data is an upper bound on the synthetic difference in difference estimator using disaggregated employment data for non-gentrifying residents. Therefore, the negative estimates in Tables 5 and 6 suggest job losses for non-gentrifiers in gentrifying neighborhoods.

In Section 1.2, I emphasize that non-gentrifying residents of gentrifying neighborhoods may lose neighborhood jobs due to racial dimension of gentrification: relative to non-gentrifying neighborhoods, gentrifying neighborhoods become significantly more white. In what follows, I provide suggestive evidence in support of this mechanism. I then discuss three alternative mechanisms and argue that they are not likely given the results in this paper and the gentrification literature.

In Table 7, I show that from 2010 to 2018, the share of jobs held by white workers increased significantly faster in gentrifying census tracts even though white residents were a minority in gentrifying census tracts in 2010. In contrast, I show that the share of jobs held by Black and Hispanic increased significantly slower in gentrifying census tracts despite the fact that Black and Hispanic residents accounted for a majority of the population in gentrifying census tracts in 2010.

Table 7: Average 2010-2018 Changes in the Racial Composition of Tract Employment

	Gentrifying	Not Gentrifying	Difference
Percent White	0.16	-0.03	0.19^{***}
	(0.39)	(0.22)	(4.88)
Percent Black	-0.01	0.17	-0.18^{***}
	(0.36)	(0.63)	(-3.63)
Percent Asian	0.30	0.27	0.04
	(0.8)	(0.65)	(0.45)
Percent Hispanic	0.00	0.09	-0.09^{**}
	(0.3)	(0.43)	(-2.51)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: Sample size of 112 gentrifying tracts and 331 non-gentrifying tracts. For each outcome, I report the average percentage change from 2010 to 2018 for gentrifying and non-gentrifying tracts and the difference in average percentage changes between gentrifying and non-gentrifying tracts. I report standard deviations and t-statistics in parentheses.

In Table 8, I assess whether the employment effects of gentrification vary by 2010 racial composition. To do so I modify Equation 8:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}, \hat{\rho}) = \underset{\tau,\mu,\alpha,\beta}{\arg\min} \left(\sum_{i=1}^{n} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - \tau W_{it} - \rho W_{it} * X_i)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right) \tag{16}$$

where $W_{it} * X_i$ is the interaction of the gentrification indicator and X_i is the racial composition of tract i's population. I find that in census tracts that were more white in 2010, service and mid-wage job losses for neighborhood residents were significantly less severe. In contrast, I find that in census tracts that were more Hispanic in 2010, job losses for neighborhood residents were significantly more severe.

Table 8: Differences in the Employment Effects of Gentrification by 2010 Racial Composition

	Percent White	Percent Black	Percent Hispanic	Percent Asian
Service Employment	257.73	5.96	-154.53^{***}	-168.87
	(1.16)	(0.08)	(-2.72)	(-0.36)
Midwage Employment	223.05^{*}	-5.94	-104.05^{***}	-318.42
	(1.71)	(-0.12)	(-2.62)	(-0.75)
Tracts	423	423	423	423
Observations	7191	7191	7191	7191

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: In Table 8 I report estimates of heterogeneity by the 2010 racial composition of gentrifying neighborhoods. The i, jth entry of the matrix reports the estimated coefficient on the interaction term $W_{it} * X_i$ for variable i and racial composition j.

While the results in Tables 7 and 8 do not decisively show that the racial dimension of gentrification causes job losses for neighborhood residents, they are largely consistent with this mechanism. However, other mechanisms are possible.

First, job losses in gentrifying neighborhoods may be due to the displacement of non-gentrifying residents. However, previous empirical research on the relationship between gentrification and displacement finds that in the short run,[24] gentrification results in at most a small increase in the displacement of long-term residents, while multiple studies find no short term relationship at all (McKinnish et al., 2010;

[24]Most papers study the relationship between gentrification and displacement over a 10 year period.

Brummet and Reed, 2019; Dragan et al., 2019). If this relationship holds in the data, then the negative relationship between gentrification and the number of residents who hold neighborhood jobs is unlikely to be driven by displacement.

Second, residents of gentrifying neighborhoods may have easier access to other attractive employment opportunities during the 2011-2018 period because gentrifying census tracts are closer to both Midtown Manhattan and high-income neighborhoods. This is unlikely because the effects of gentrification in Tables 5 and 6 are largely similar to the trimmed sample effects I present in Section A.6 of the Appendix, in which I trim my sample based on a propensity score calculated using distance to Midtown Manhattan and distance to high income neighborhoods as predictors.

Third, the job losses in service-producing industries may be due to changing commuting patterns of college-educated individuals who work in business, finance, or information sectors. This is unlikely because there is no significant relationship between gentrification and the number of neighborhood high-wage jobs that go to neighborhood residents, as I show in Table 6.

6 Conclusion

Across U.S. cities, gentrification is both increasingly common and increasingly contentious. Unfortunately, there is still a lack of empirical evidence documenting how gentrification affects gentrifying neighborhoods. This problem is exacerbated by the fact that researchers have not yet adequately addressed the methodological challenges of defining gentrification and identifying the economic effects of gentrification without quasi-experimental variation in the data.

In this paper, I first propose solutions to these methodological challenges and then use these methodological innovations to study how gentrification changes neighborhood economic activity. In contrast to the claims of proponents of gentrification, I find no evidence that gentrification changes neighborhood economic activity in ways that benefit long-term residents. In conjunction with a robust qualitative literature documenting that long-term residents of gentrifying neighborhoods often feel racially and culturally excluded from the economic changes associated with gentrification, my findings suggest that the economic changes associated with gentrification likely make long-term neighborhood residents worse-off.[25]

Despite the methodological innovations I propose in this paper, these results should be considered within the limitations of the study. First, it is not obvious that results from New York City should apply to other U.S. cities. Second, I am not able to conclusively establish the mechanism driving neighborhood residents' post-gentrification job losses. Third, while the synthetic difference in differences estimator is able to identify treatment effects without quasi-experimental variation in treatment assignment, it is only able to do so with sufficient pre-treatment information. Because I only have nine years of pre-gentrification data, it is possible that the SDID estimator does not actually remove bias due to selection on the factor matrix $\mathbf{L}$.

These limitations suggest three important directions for future research. First, future research should attempt to answer this paper's questions (or similar questions) in the context of other U.S. cities. Second, future research should attempt to exploit micro-data to better understand how gentrification affects the employment outcomes of neighborhood residents. Third, researchers should look to longer and longer panels of data as a solution to the challenging problem of drawing inferences about the way gentrification affects neighborhoods and the people living in them.

[25]This discussion abstracts from the physical and cultural displacement associated with gentrification, which likely also make long-term residents worse-off.

References

Abadie, A., Diamond, A., and Hainmuller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490).

Abadie, A., Diamond, A., and Hainmuller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Agan, A. and Starr, S. (2020). Employer neighborhoods and racial discrimination. *NBER Working Papers*.

Arkhangelsky, D., Athey, S., Hirshberg, D., Imbens, G., and Wager, S. (2020). Synthetic differences in differences. *NBER Working Papers*.

Autor, D. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–1597.

Autor, D., Palmer, C., and Pathak, P. (2017). Sgentrification and the amenity value of crime reductions: Evidence from rent deregulation. *NBER Working Papers*.

Baum-Snow, N. and Ferreira, F. (2014). Causal inference in urban and regional economics. *NBER Working Papers*.

Baum-Snow, N. and Hartley, D. (2020). Accounting for central neighborhood change, 1980–2010. *Journal of Urban Economics*, 117:103–228.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650.

Brummet, Q. and Reed, D. (2019). The effects of gentrification on the well-being and opportunity of original resident adults and children. Working paper, Federal Researve Bank of Philadelphia.

Couture, V. and Handbury, J. (2017). Urban revival in america, 2000-2010. Working paper, National Bureau of Economic Research.

Dragan, K., Ellen, I., and Glied, S. (2019). Does gentrification displace poor children and their families? new evidence from medicaid data in new york city. *Regional Science and Urban Economics*.

Edlund, L., Machado, C., and Sviatschi, M. M. (2015). Gentrification and the rising return to skill. *NBER Working Papers*.

Ellen, I. and Ding, L. (2016). Advancing our understanding of gentrification. *Cityscape: A Journal of Policy Development and Research*, 18(3).

Fitzsimmons, E. (2020). In a crowded new york city, should newcomers 'go back to iowa'? The New York Times.

Glaeser, E., Kim, H., and Luca, M. (2018). Nowcasting gentrification: Using yelp data to quantify neighborhood change. *AEA Papers and Proceedings*, 108:77–82.

Glaeser, E., Luca, M., and Moszkowski, E. (2020). Gentrification and neighborhood change: Evidence from yelp. *NBER Working Papers*.

Goetz, E., Lewis, B., Damiano, A., and Calhoun, M. (2019). The diversity of gentrification: Multiple forms of gentrification in minneapolis and st. paul. Technical report, University of Minnesota.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Imbens, G. (2014). Matching methods in practice: Three examples. *IZA Discussion Paper Series*.

Lester, T. W. and Hartley, D. (2014). The long term employment impacts of gentrification in the 1990s. *Regional Science and Urban Economics*, 45:80–89.

McKinnish, T., Walsh, R., and White, T. K. (2010). Who gentrifies low-income neighborhoods. *Journal of Urban Economics*, 67(2):180–193.

Meltzer, R. and Ghorbani, P. (2017). Does gentrification increase employment opportunities in low-income neighborhoods? *Regional Science and Urban Economics*, 66:52–73.

Sullivan, D. and Shaw, S. (2011). Retail gentrification and race: The case of alberta street in portland, oregon. *Urban Affairs Review*, 47(3):413–432.

A Appendix

A.1 Limitations of Alternative Methods of Defining Gentrification

The most common variables used to define gentrification are unit-level changes in the number of share of residents with a college degree, unit-level changes in median income (Meltzer and Ghorbani, 2017), and unit-level changes in median rent (Glaeser et al., 2020). In this section, I build on the discussion in Section 2.2 and argue that there is little justification for using changes in median income or changes in median rent to define gentrification.

It is well-documented in the literature that gentrification is driven by young, college-educated individuals (Edlund et al., 2015; Couture and Handbury, 2017; Baum-Snow and Hartley, 2020). Therefore, the only conditions under which it is appropriate to use a non-education-based measure of gentrification is if the alternative variable is a reliable proxy for tract-level changes in educational composition and is measured with less error than tract-level changes in educational-composition. It is unlikely that both of these conditions are met when using aggregate data from the American Community Survey or the U.S. Census.

Tract-level changes in median income are likely a good proxy for changes in the share of residents with a college degree, but tract-level changes in median income are likely measured with significantly more error than tract-level changes in the share of residents with a college degree. From year to year, an individual's income is subject to idiosyncratic shocks, whereas her educational attainment is not. As a consequence, observed aggregate changes in neighborhood income could be driven by correlated shocks to incumbent residents' income rather than the in-migration of wealthier individuals. In contrast, it is unlikely that large increases in the tract-level count of individuals with a college degree could be driven by changes in incumbent residents' educational attainment alone. Perhaps more importantly, previous papers in the urban change literature document that while higher-income college graduates are increasingly choosing to live in cities, low-income city residents are increasingly choosing to live in suburbs (Baum-Snow and Hartley, 2020). Therefore, changes in tract-level median income could be driven by the departure of low-income individuals rather than the in-migration of college graduates.

In contrast, tract-level changes in median rent may not be measured with significant error but do require strong assumptions about the relationship between gentrification and housing prices. In particular, the use of tract-level changes in median rent assumes that housing supply is relatively inelastic so that increased demand for neighborhood demand is reflected in rents. While this assumption likely holds in places like San Francisco and New York City, it is less likely to hold in the southern cities that college-educated individuals are increasingly choosing to move to.

A.2 Goods-Producing and Services-Producing Job Counts

In this section, I explain in further detail the goods-producing and services-producing job counts I describe in Section 2.1. From the LODES data, I collect counts of the number of individuals in a census tract who work in that census tract's neighborhoods. These counts are not available at the detailed NAICS sector level. Instead, they are available for goods-producing industries and services-producing industries. The goods-producing job counts report the number of residents of census tract r who work in neighborhood jobs in the following industries:

1. Natural Resources and Mining
2. Construction
3. Manufacturing

The services-producing job counts report the number of residents of census tract r who work in neighborhood jobs in the following industries:

1. Trade, Transportation, and Utilities
2. Information
3. Financial Activities
4. Professional and Business Services
5. Educational and Health Services
6. Leisure and Hospitality

A.3 Estimating the Propensity Score

In Section 3.2, I collapse the multidimensional covariate differences between gentrifying and non-gentrifying census tracts to a single measure using the propensity score estimation algorithm in Imbens (2014). The algorithm approximates the population propensity score function with a quadratic model where variables and interactions are chosen based on changes to the log likelihood. See Imbens (2014) for a detailed description of the algorithm. The algorithm produces the following model:

Table 9: Model of the Propensity Score for Gentrification

	Gentrifying
Distance to CBD (Miles)	-0.830^{***}
	(0.170)
Distance to Nearest High Income Tract	-0.934^{***}
High Income Tract	(0.354)
Percent White (2010)	-4.354^{***}
	(0.691)
Distance to CBD × Distance to Nearest High Income Tract	0.090^{***}
	(0.034)
Constant	6.925^{***}
	(1.119)
Observations	423
Log Likelihood	-150.500

A.4 Limitations of Previous Empirical Strategies

The Bartik instrument is commonly used to identify causal effects in the urban and regional economics literature (Baum-Snow and Ferreira, 2014). For example, in the gentrification literature Meltzer and Ghorbani (2017) instrument use a version of this instrument. In particular, Meltzer and Ghorbani (2017) instrument for an income-based[26] measure of gentrification by interacting tract-level industry shares for industries traditionally requiring a college degree with national industry earnings growth rates. They argue that the instrument relates to gentrification because census tracts with a higher concentration of the included industries experience larger income increases over the period of study.

[26]In Appendix Section A.1, I argue that income-based measures of gentrification are in almost all circumstances inferior to education-based measures of gentrification.

In my reading of the literature, the relationship the Bartik instrument in Meltzer and Ghorbani (2017) and gentrification is not conceptually clear. It is well-established that gentrification occurs when college-educated individuals move into low-income neighborhoods rather than when college-educated individuals in low-income neighborhoods experience an increase in income (Couture and Handbury, 2017; Baum-Snow and Hartley, 2020). Therefore, the Bartik instrument in Meltzer and Ghorbani (2017) is only related to gentrification if the share of neighborhood residents working in included industries enters into the residential choice of college-educated individuals. To my knowledge, no study finds conclusive evidence that this is the case. In the plausible case that initial residents' industry of employment does not matter to potential gentrifiers, the Bartik instrument in Meltzer and Ghorbani (2017) is unlikely to measure differential exposure to a city-level increase in demand for housing by college-educated individuals, which is the standard argument for using a Bartik instrument (Goldsmith-Pinkham et al., 2020).

Even if regression-based identification strategies as in Meltzer and Ghorbani (2017) randomize gentrification with respect to potential outcomes, the estimated treatment effects may suffer from extrapolation bias due to large differences in the covariate distributions between gentrifying and non-gentrifying tracts. To adjust for covariate imbalances between treated and comparison units, regression methods rely on the strong assumption that the population conditional expectation function of the outcome of interest $E[Y_i|W_i, X_i]$ is linear in the relevant covariates X_i (Imbens, 2014). If $E[Y_i|W_i, X_i]$ is non-linear in X_i and the distribution X_i differs substantially[27] across treated and control units, as I show is the case in Section 3.2, then OLS estimates of the treatment effect of interest will be biased (Imbens, 2014).

A.5 Properties of the Synthetic Difference in Differences Estimator

In Section 4.1, I state that the goal of the SDID estimator is to remove bias due to selection on the factor matrix **L** by re-weighting non-gentrifying tracts to have approximately equal pre-treatment outcomes as gentrifying tracts. Importantly, this bias-removal procedure works well under relatively mild assumptions even though

[27] Mathematically, this condition can be written as $E[X_i|W_i = 1] \neq E[X_i|W_i = 0]$.

there is a discrepancy between the observed pre-treatment outcome matrix **Y** and the underlying factor matrix **L** (Arkhangelsky et al., 2020). First, there must be little serial correlation within units so that trends in Y_{it} can be attributed to trends in L_{it} rather than trends in ε_{it}. Second, there must be sufficient pre-intervention periods to estimate L_{it}.

Beyond removing any bias induced by the factor matrix **L**, the SDID estimator has desirable robustness properties. First, the SDID estimator approximately removes bias from selection on **L** if $\hat{\omega}^{sdid}$ balances outcomes and covariates between gentrifying and non-gentrifying tracts before gentrification or if $\hat{\lambda}^{sdid}$ balances outcomes pre-gentrification outcomes for non-gentrifying units with post-gentrification outcomes for those units. Second, even if neither $\hat{\omega}^{sdid}$ or $\hat{\lambda}^{sdid}$ balance outcomes, the SDID estimator is unbiased if the true model for potential outcomes is not the general factor model in Equation 1 but a two-way fixed effects model.

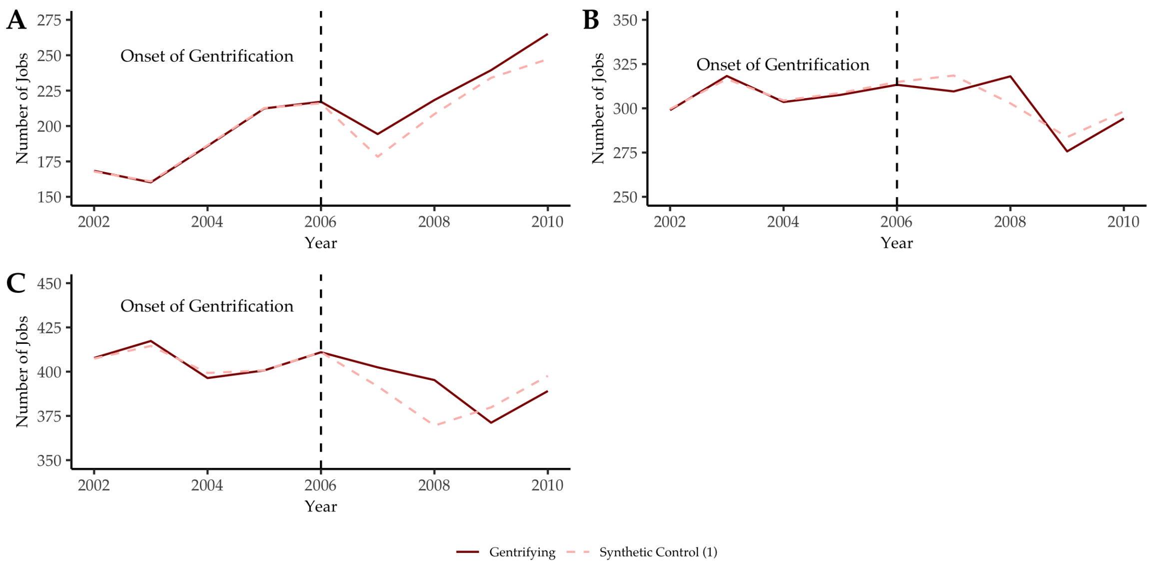
A.6 Robustness Checks

In this section, I assess the robustness of the estimates I report in Section 5 using the methods I discuss in Section 4.1.1. I first provide evidence that estimates are not driven by pre-gentrification trends. I then show that estimates are largely similar in magnitude and sign after trimming the sample based on propensity scores. Finally, I show that estimates are also robust to the leave-two-out procedure I describe in Section 4.1.1.

A.6.1 Backdated Estimates

In Table 10, I report estimates of the 2007-2010 effect of gentrification after artificially backdating the onset of gentrification to 2006. I find that in contrast to the results in Tables 3, 5, and 6, backdated gentrification has no significant effect on jobs in accommodation and food services, service jobs that go to neighborhood residents, or mid-wage jobs that go to neighborhood residents. In Figure 6, I present these results graphically. While the synthetic difference in differences estimator does not perfectly replicate the 2007-2010 outcomes in the average gentrifying tract, it does not systematically over- or under-estimate these outcomes, which lends credibility to interpretation that the large observed divergences in Figures 3, 4, and 5 are due to gentrification.

Figure 6: Backdated Effect of Gentrification



Notes: In Panel A of Figure 6, I plot the average annual number of jobs in accommodation and food services for gentrifying tracts from 2002 to 2010. I also plot the backdated synthetic control from Table 10. In Panel B of Figure 6, I plot the average annual number of mid-wage jobs held by neighborhood residents for gentrifying tracts from 2002 to 2010. I also plot the backdated synthetic control from Table 10. In Panel C of Figure 6, I plot the average annual number of service jobs held by neighborhood residents for gentrifying tracts from 2002 to 2010. I also plot the backdated synthetic control from Table 10.

Table 10: Synthetic Difference in Differences Estimates with Backdated Onset of Gentrification

	Food	Mid-Wage	Services
Gentrifying	13.74	−1.8	3.37
	(0.72)	(-0.09)	(0.20)
NRMSE			
$\hat{\omega}$	0.004	0.004	0.004
$\hat{\lambda}$	0.720	0.301	0.391
Percent Difference			
Percent College Graduates (2000)	0.050	0.059	0.058
Percent College Graduates (2010)	-0.020	-0.0170	-0.018
Tracts	423	423	423
Observations	3807	3807	3807

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Notes: For each outcome, I report synthetic difference in differences estimates for my baseline specification when the onset of gentrification is backdated in the data to 2006. I report estimates in the number of jobs per square mile. I report t-statistics in parentheses.

A.6.2 Trimmed Sample Estimates

As I show in Figure 2, a subset of tracts that gentrified from 2010 to 2018 were extremely likely to gentrify given their 2010 characteristics while a subset of tracts that gentrified from 2010 to 2018 were extremely unlikely to gentrify given their 2010 characteristics. To assess whether outcomes are driven by these outliers, I follow the recommendation of Imbens (2014) and restrict the sample to census tracts with a 2010 probability of gentrifying greater than 0.1 and less than 0.9. This produces a sample of 187 census tracts, of which 75 tracts were gentrifying and 112 tracts were not gentrifying. In Table 11, I show that estimates are largely similar in sign, magnitude, and significance in this trimmed sample. In Figure 7, I show that even with fewer comparison units, the synthetic difference in difference estimator is able to accurately approximate pre-gentrification outcomes in the average gentrifying tracts. I also show that the post-gentrification divergence in outcomes in the trimmed sample largely resembles the post-gentrification divergence in outcomes in the full sample. Therefore, it is unlikely that the observed effects are driven by distinct trends in census tracts that were extremely likely to gentrify and in census tracts that were extremely unlikely to gentrify.

Table 11: Synthetic Difference in Differences Estimates with Trimmed Sample

	Food	Mid-Wage	Services
Gentrifying	144.19^{***}	-61.55^{**}	-39.7
	(3.52)	(-2.41)	(-1.21)
NRMSE			
$\hat{\omega}$	0.037	0.015	0.028
$\hat{\lambda}$	0.562	0.461	0.316
Percent Difference			
Percent College Graduates (2000)	0.034	0.042	0.043
Percent College Graduates (2010)	-0.035	-0.036	-0.036
Percent White	-0.0009	0.0017	0.0006
Percent Hispanic	-0.023		
Percent Asian	1.002	0.848	0.794
Median Rent	-0.010		
Tracts	187	187	187
Observations	3179	3179	3179

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

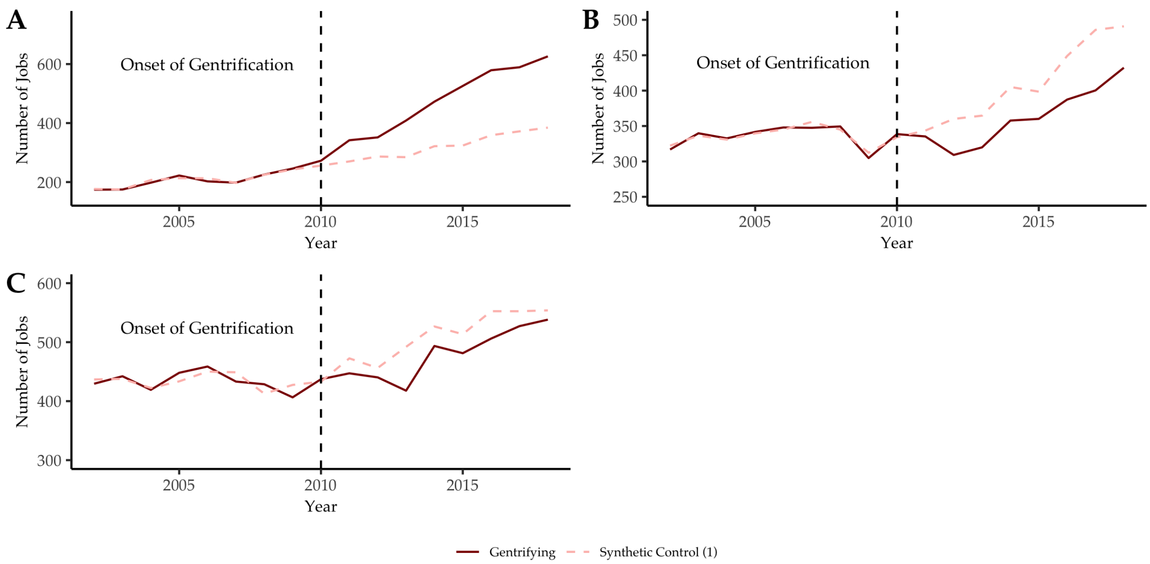
Notes: For each outcome, I report synthetic difference in differences estimates for the specification with covariates using the trimmed sample I describe in Section 4.1.1. I report estimates in the number of jobs per square mile. I report t-statistics in parentheses.

A.6.3 Leave Two Out Estimates

In Section 4.1.1, I discuss the leave-two-out procedure I use to assess the possibility that estimates are driven by correlated shocks to census tracts in high-weight neighborhood tabulation areas. In this discussion, I explain how I determine which neighborhood tabulation areas are high-weight and which neighborhoods tabulation areas are not and report results.

For a given outcome and specification, I determine which neighborhoods are high-weight as follows. I first sum the census tract unit weights by neighborhood to create neighborhood weights. I then say a neighborhood is high-weight if it had one of the 10 largest neighborhood weights (out of 29 neighborhoods). Implementing the leave-two-out procedure with these 10 neighborhoods results in 45 (10 choose 2) different estimates of the effect of interest.

Figure 7: Effect of Gentrification in Trimmed Sample



Notes: In Panel A of Figure 7, I plot the average annual number of jobs in accommodation and food services for gentrifying tracts from 2002 to 2010. I also plot the trimmed sample synthetic control from Table 11. In Panel B of Figure 7, I plot the average annual number of mid-wage jobs held by neighborhood residents for gentrifying tracts from 2002 to 2010. I also plot the trimmed sample synthetic control from Table 11. In Panel C of Figure 7, I plot the average annual number of service jobs held by neighborhood residents for gentrifying tracts from 2002 to 2010. I also plot the trimmed sample synthetic control from Table 11.

In Tables 12 - 14, I implement the leave-two-out procedure. For each outcome, I report the mean, minimum, $25th$ percentile, $50th$ percentile, $75th$ percentile, and maximum estimates from the 45 (10 choose 2) leave-two-out estimates. I show that even after dropping census tracts in pairs of high-weight neighborhoods, the synthetic difference in differences estimator is able to closely approximate the average gentrifying census tract while results are substantively unchanged.

Table 12: Leave-Two-Outs Estimates of Gentrification's Effect on Accommodation and Food Service Employment

	Mean	Minimum	$25th$ Percentile	$50th$ Percentile	$75th$ Percentile	Maximum
Gentrifying	145.20***	129.60***	141.37***	145.90***	149.53***	156.23***
	(3.92)	(3.46)	(3.81)	(3.93)	(4.06)	(4.26)
NRMSE						
$\hat{\omega}$	0.02	0.01	0.01	0.02	0.02	0.02
$\hat{\lambda}$	0.54	0.51	0.53	0.55	0.55	0.56
Percent Difference						
Percent College (2000)	0.11	0.10	0.11	0.11	0.12	0.13
Percent College (2010)	-0.03	-0.04	-0.04	-0.03	-0.03	-0.03
Percent Asian	0.35	0.26	0.30	0.33	0.39	0.53
Percent Hispanic	0.01	0.00	0.01	0.01	0.01	0.02
Percent College (NTA)	-0.02	-0.03	-0.02	-0.01	-0.01	-0.01
Median Rent	-0.00	-0.00	-0.00	0.00	0.00	0.00
Tracts	390.00	371.00	384.00	391.00	396.00	400.00
Observations	6630.00	6307.00	6528.00	6647.00	6732.00	6800.00

Table 13: Leave-Two-Outs Estimates of Gentrification's Effect on Service Employment

	Mean	Minimum	$25th$ Percentile	$50th$ Percentile	$75th$ Percentile	Maximum
Gentrifying	-46.91^{**}	-55.32^{**}	-49.45^{**}	-46.83^{**}	-44.44^{*}	-38.06^{*}
	(-2.05)	(-2.34)	(-2.15)	(-2.03)	(-1.93)	(-1.70)
NRMSE						
$\hat{\omega}$	0.02	0.01	0.01	0.02	0.02	0.03
$\hat{\lambda}$	0.39	0.38	0.39	0.39	0.39	0.40
Percent Difference						
Percent College (2000)	0.11	0.10	0.11	0.11	0.12	0.13
Percent College (2010)	-0.03	-0.05	-0.03	-0.03	-0.02	-0.02
Percent Asian	0.38	0.31	0.35	0.37	0.39	0.46
Percent Hispanic	0.02	0.02	0.02	0.03	0.03	0.05
Percent White	0.07	0.05	0.06	0.07	0.08	0.10
Tracts	395.00	379.00	391.00	398.00	400.00	401.00
Observations	6715.00	6443.00	6647.00	6766.00	6800.00	6817.00

Table 14: Leave-Two-Outs Estimates of Gentrification's Effect on Mid-Wage Employment

	Mean	Minimum	25*th* Percentile	50*th* Percentile	75*th* Percentile	Maximum
Gentrifying	-56.74^{***}	-65.86^{***}	-59.16^{***}	-56.39^{***}	-53.76^{***}	-49.83^{***}
	(-3.47)	(-3.92)	(-3.61)	(-3.43)	(-3.33)	(-3.11)
NRMSE						
$\hat{\omega}$	0.01	0.01	0.01	0.01	0.01	0.02
$\hat{\lambda}$	0.45	0.44	0.45	0.45	0.46	0.46
Percent Difference						
Percent College (2000)	0.09	0.09	0.09	0.09	0.10	0.12
Percent College (2010)	-0.04	-0.05	-0.04	-0.04	-0.04	-0.04
Percent Asian	0.29	0.25	0.27	0.28	0.30	0.39
Percent Hispanic	0.02	0.01	0.01	0.02	0.02	0.04
Percent White	0.04	0.03	0.03	0.04	0.05	0.09
Tracts	395.00	379.00	391.00	398.00	400.00	401.00
Observations	6715.00	6443.00	6647.00	6766.00	6800.00	6817.00

A.7 Plots

Figure 8: Spatial Distribution of Gentrification in New York City

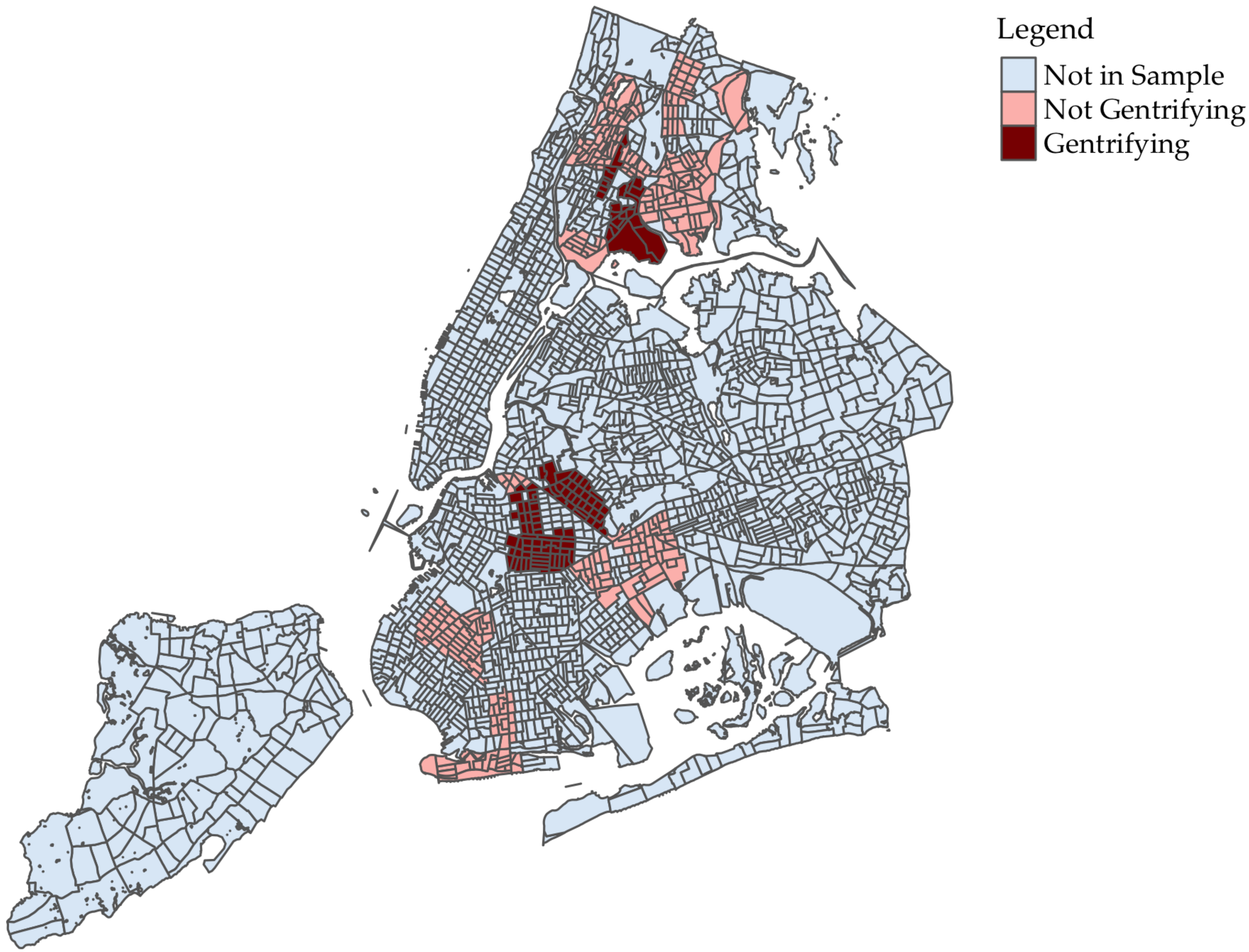
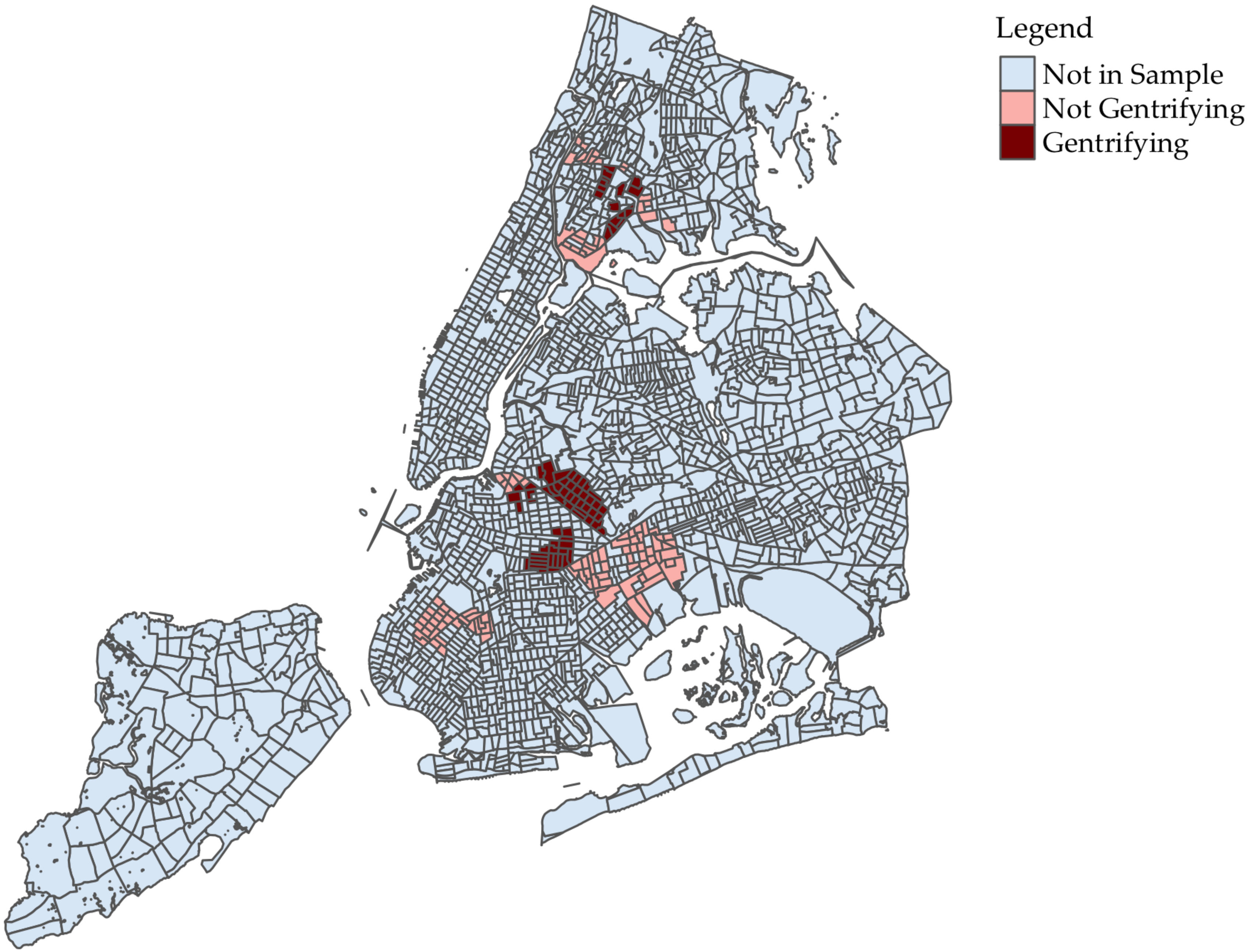
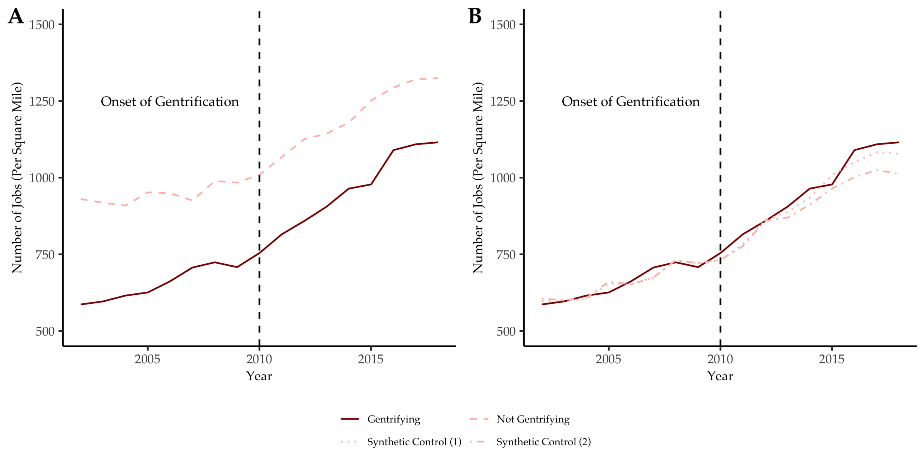


Figure 9: Spatial Distribution of Trimmed Sample



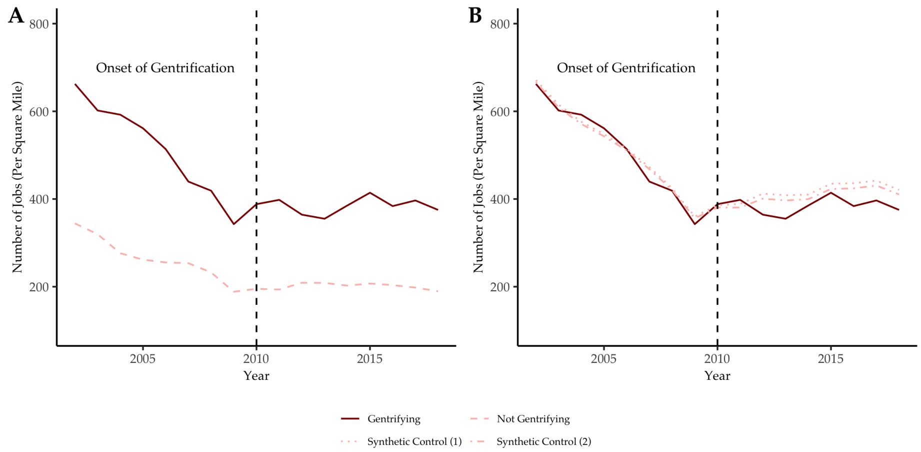
A.7.1 Results

Figure 10: Effect of Gentrification on Employment in Retail Trade



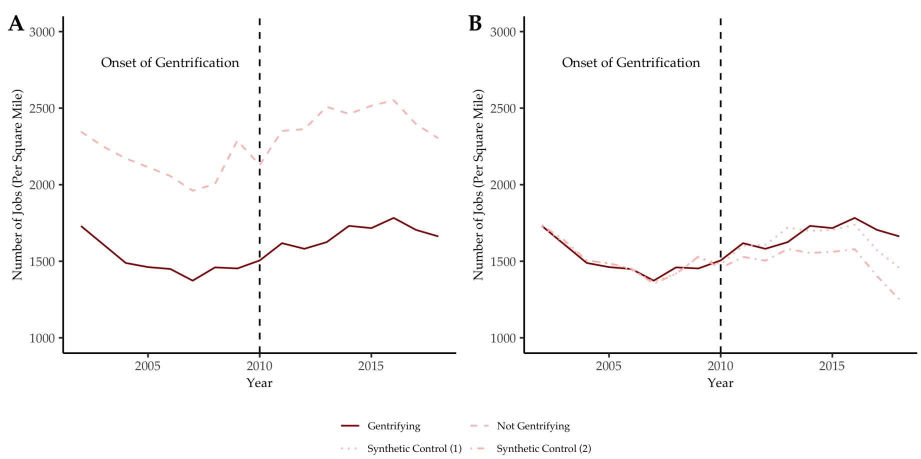
Notes: In Panel A of Figure 10, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 10, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 3.

Figure 11: Effect of Gentrification on Employment in Manufacturing



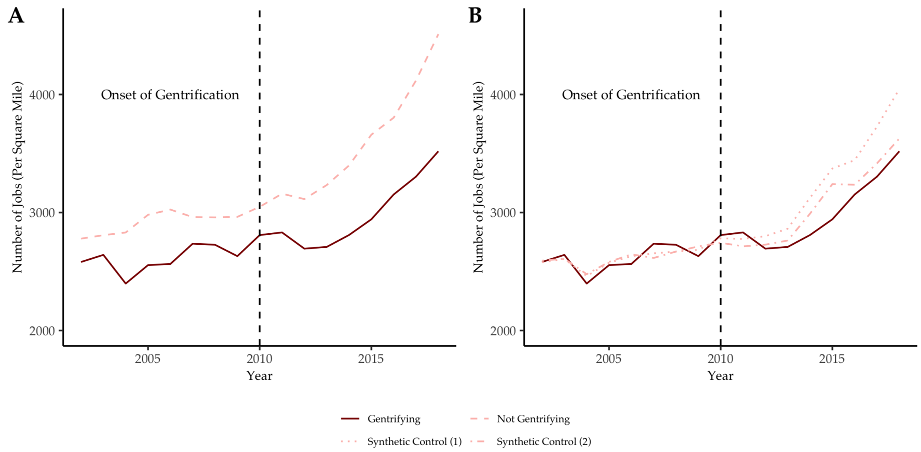
Notes: In Panel A of Figure 11, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 11, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 3.

Figure 12: Effect of Gentrification on Employment in Low-Wage Jobs



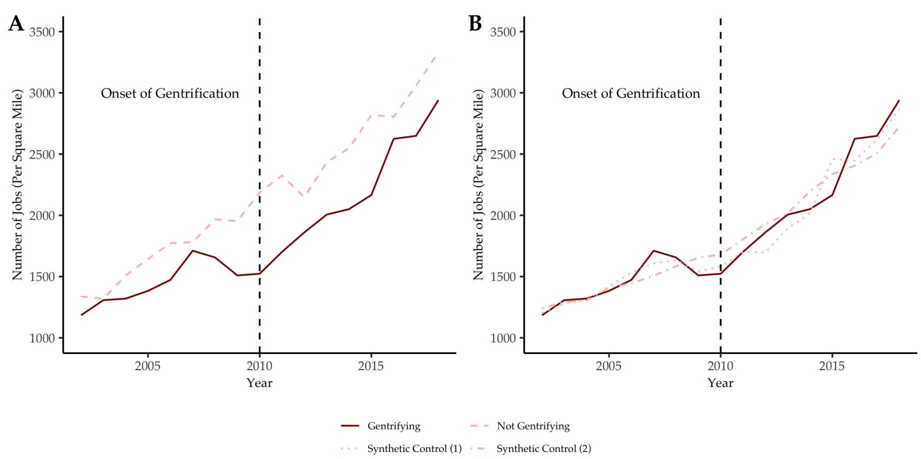
Notes: In Panel A of Figure 12, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 12, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 4.

Figure 13: Effect of Gentrification on Employment in Mid-Wage Jobs



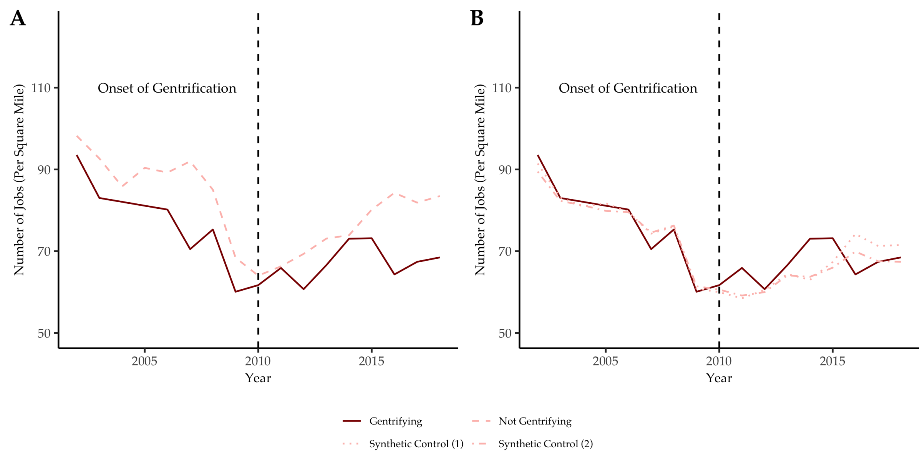
Notes: In Panel A of Figure 13, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 13, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 4.

Figure 14: Effect of Gentrification on Employment in High-Wage Jobs



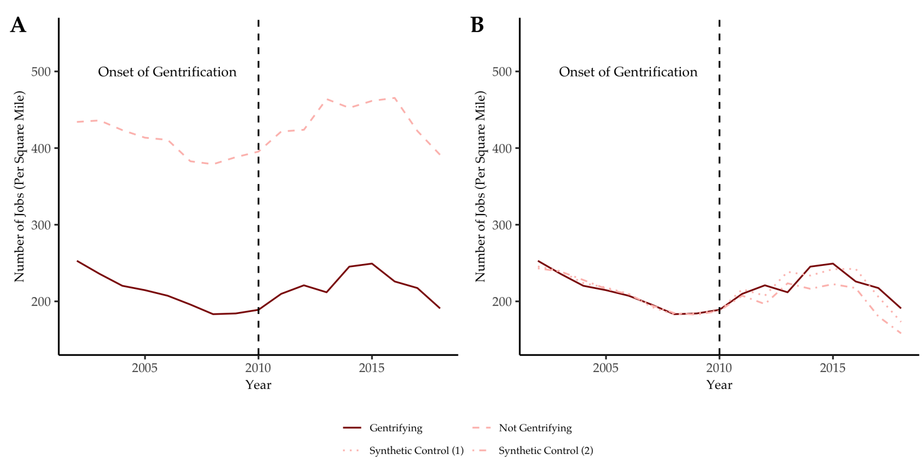
Notes: In Panel A of Figure 14, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 14, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 4.

Figure 15: Effect of Gentrification on Neighborhood Employment in Goods-Producing Industries



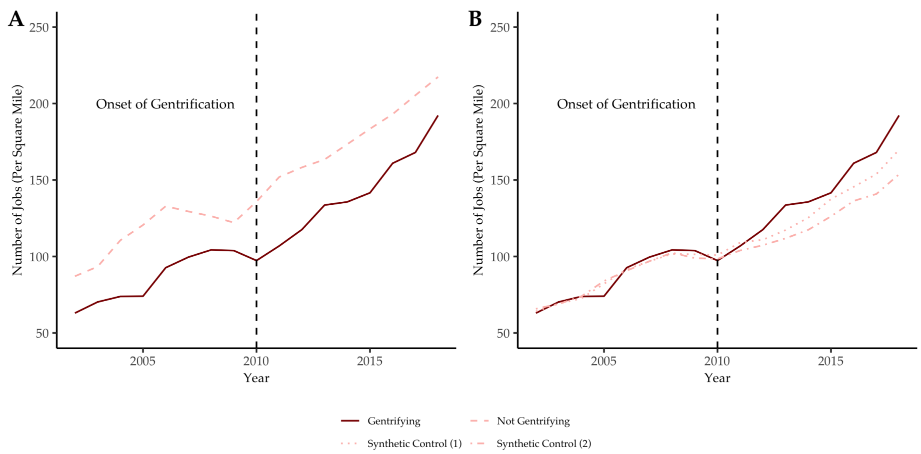
Notes: In Panel A of Figure 15, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 15, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 5.

Figure 16: Effect of Gentrification on Neighborhood Low-Wage Employment



Notes: In Panel A of Figure 16, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 16, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 6.

Figure 17: Effect of Gentrification on Neighborhood High-Wage Employment



Notes: In Panel A of Figure 17, I report the average annual number of jobs for gentrifying and non-gentrifying tracts from 2002 to 2018. In Panel A of Figure 17, I report the average annual number of jobs in accommodation and food service for gentrifying tracts and the synthetic controls from specifications (1) and (2) in in Table 6.

A.8 Tables

A.8.1 Summary Statistics

Table 15: Neighborhood Tabulations Areas that Gentrified from 2010 to 2018

Neighborhood	Borough	Share College Population Growth
Claremont-Bathgate	The Bronx	89%
Crotona Park East	The Bronx	95%
Hunts Point	The Bronx	78%
Longwood	The Bronx	91%
Bedford	Brooklyn	81%
Bushwick North	Brooklyn	160%
Bushwick South	Brooklyn	122%
Crown Heights North	Brooklyn	81%
Crown Heights South	Brooklyn	77%

Notes: In Table 15, I report the nine New York City neighborhood tabulation areas that gentrified from 2010 to 2018. I also report the neighborhood tabulation area's borough and the percentage growth in the share of residents with a college degree.