land

MDPI

Article

What Are the Relationships between Public Transit and Gentrification Progress? An Empirical Study in the New York–Northern New Jersey–Long Island Areas

Yefu Chen [1], Hao Xi [2] and Junfeng Jiao [1,*]

[1] Urban Information Lab, The School of Architecture, The University of Texas at Austin, Austin, TX 78705, USA
[2] Center for Urban Future Research, Peking University, Beijing 100000, China
* Correspondence: jjiao@austin.utexas.edu

Abstract: Transit-oriented development has been a widely accepted tool among transportation planning practitioners; however, there are concerns about the risk of increasing residential property values leading to gentrification or displacements. Therefore, it is critical to provide precise investigations of the relationships between public transit and gentrification. Although numerous studies have explored this topic, few have discussed these relationships based on detailed measurements of gentrification from a regional perspective. This study aims to fill the research gap by measuring the gentrification subcategories through a hierarchical definition based on data in the New York–Northern New Jersey–Long Island areas and applying the transit desert concept as the measurement of transit services. Through multinomial logistic regression and machine-learning approaches, findings indicate that the rate of transit deserts in economically disadvantaged neighborhoods is higher than the others. In addition, the impacts of transit services are significant in gentrification but insignificant in super-gentrification. These findings can advance the knowledge of the role of the transit service in different gentrification progresses. Based on these findings, policymakers need to be careful when allocating public transit budgets and note the effects of these investments on neighborhoods with different socioeconomic statuses.

Keywords: gentrification; super-gentrification; public transport; transit desert; inequalities

Citation: Chen, Y.; Xi, H.; Jiao, J. What Are the Relationships between Public Transit and Gentrification Progress? An Empirical Study in the New York–Northern New Jersey–Long Island Areas. *Land* **2023**, *12*, 358. https://doi.org/10.3390/land12020358

Academic Editor: Muhammad Shafique

Received: 7 January 2023
Revised: 23 January 2023
Accepted: 27 January 2023
Published: 28 January 2023

1. Introduction

In recent decades, Transit-Oriented Development (TOD) has become one of the most popular policy concepts in the sustainable development toolkit. Based on this concept, authorities have invested in public transportation services to improve mobility and increase accessibility [1,2]. It can reduce greenhouse gas emissions and urban unaffordability issues by providing compact development, convenience, and affordable travel choices [3,4]. At the same time, this concept has received concerns about its disadvantages. The most significant concern is that it leads to gentrification/displacement. Researchers claimed that these developments could increase housing costs and appraisal values for property tax, which might cause gentrification and displacement [5,6]. For instance, public transit facilities can make a community attractive. Residential properties within these neighborhoods could be under high demand for purchases, causing increases in property values [7,8]. However, residents who cannot afford the rising living costs are forced to move out, which is regarded as displacement [9–12].

Previous studies have noted this debate, but there are two limitations worth noting from previous studies. First, the related study areas mostly were at the city level, and few were from the regional-level perspective [13,14], while, according to the gentrification researchers, it would be better to quantify gentrification progress based on region-level cases, such as metropolitan-level [15,16]. In addition, transportation planners have noticed that macro-level transportation planning and co-operation between local transportation agencies should be critical due to the increased commuting distance [17,18]. Second, previous studies mostly stopped at exploring the relationship between public transit and gentrification for low-income to median–high income populations, while few further dynamically explored the impacts of it on median-high income to high-income, which is called super-gentrification [19,20]. Hence, in terms of this dilemma and research gaps, we want to advance the knowledge on the relationship between public transit and gentrification in different subcategories in this study. There are two research questions:

1. What are the relationships between public transit and gentrification from a macro-level perspective?
2. How do these relationships change during the progress of gentrification?

To answer these questions, we introduced the transit desert concept and gentrification subcategories from previous studies as the macro-level measurements in transportation and gentrification based on the New York–Northern New Jersey–Long Island region [15,21]. The rest of the paper is formatted as follows: First, we reviewed the literature on the relationship between gentrification and public transit. Then, we introduced the data collection and methods in this study. Next, we presented the results of data descriptions, mapping, logistic models, and the gradient boosting decision tree algorithm to quantify the relationships between public transit factors and gentrification. In closing, we summarized the significant findings and policy implications, and this paper ends with a conclusion.

2. Literature Review

2.1. Gentrification and Measurements

Gentrification refers to changes in the sociodemographic and socioeconomic status of residents. It was first used to describe medium-income and educated residents who moved into neighborhoods where the low-income population lived in London [22]. Based on this concept, researchers defined gentrification as associated with advancing infrastructure and revitalizing residential areas, upgrading residential buildings, raising neighborhood consciousness, and upgrading the quality of residents' lives [23]. They also pointed out that gentrification specifically helped the middle class to improve their living conditions by moving to neighborhoods where costs were low. This process was concluded as an neighborhood economic change that can be a significant sign of gentrification [24]. Previous studies focused on the changes in economic characteristics of residents, while gentrification recently has been regarded as a complex form of neighborhood change that requires the analysis of social, political, and economic conditions. At the community level, it can be seen in the upgrading of community buildings as houses are renovated and new businesses are established [25]. Beyond these physical changes, the significant components of gentrification include the changes in residents' sociodemographic attributes and physical urban form redevelopments/improvements [6].

Recent studies highlighted the negative impacts of gentrification. During this process, the low-income population may suffer from displacements and relocations, and usually, they must move to neighborhoods with worse access to public services and assets than where they lived [26]. It can lead to myriad issues among the low-income population, including increasing costs of commuting and access to opportunities, which can result in a vicious cycle [27]. In addition, the emerging development type, public–private partnerships, can exacerbate the issues. Developers may highlight the advantages of redevelopments in urban form and life satisfaction while concealing the risks of displacements. Therefore, it is critical to study the factors of gentrification to avoid these negative impacts.

Some debates are worth noting in the previous measurements of gentrification. First, previous studies used binary categories to measure it, while recent studies prefer multilevel measurements [15]. The variations among measurements, especially binary and multilevel measurements, led to different findings and implications [28,29]. Second, some researchers only focused on changes within the central city [30], while some studies called for macro-level study due to the expansions in human activities [15,16]. Third, it is worth introducing other criteria in measuring gentrification instead of economic factors only [6].

2.2. Public Transit and Gentrification

The causes of gentrification are hotly argued. Studies have defined that infrastructure which benefits the quality of daily life can be regarded as the cause, such as recreational infrastructure, public transit facilities, postindustrial redevelopments, and green spaces [31–34]. Researchers claimed the mechanism of these causes is increasing property tax, forcing those who cannot afford it to move out. Although these facilities aim to improve life quality, they may serve the residents who can afford the increases instead of the existing residents, which can be a source of public service inequality if policymakers do not launch related anti-displacement plans. In addition, the demand for the analysis of these effects raises concerns at the post-pandemic stage because the inequity becomes more serious than before [35,36]. Hence, a clear view of the causes of gentrification should be critical for policymakers, and this study takes public transit as an example since it is a common method of policymakers allocating public services [37].

Public transit is potent for improving the residents' mobility, but its impacts are debated. Although the benefits of public transit are worth noting, there are arguments that public transit may not be equally beneficial to all population groups. Most studies involving TOD development and gentrification claimed that public transit could promote neighborhood renewal, and improvement in public transit facilities could lead to a rise in property values and potentially price out residents [7,38]. Under this situation, although mobility is increased, the most vulnerable groups might not benefit from these improvements since they were displaced. Moreover, the measurements of public transit in these studies are worth noting. Previous studies used accessibility to public transit services, such as the number of facilities and routes within catchments and the distance to the closest facilities [3,39]. This is criticized because a single measurement may not systematically present the transit services [21]. In addition, some studies highlighted the importance of region-level perspectives [40,41].

Regarding the above reviews, we find that ideal measurements in both gentrification and public transit should contain various criteria and be based on the macro-level/regional perspective. Hence, we applied an approach developed by a research team of New York University [15]. This approach has been validated through community engagement. In addition, we measured public transit based on the transit desert concept. It is based on the supply–demand concept, measuring the transit supply and demand. Then, we calculated the gap between supply and demand as the index measuring whether the local transit service can meet demands [21]. This approach has shown potential in analyzing public transit services across cases and from different scales [42,43]. We referenced the structures of both approaches and made modifications based on the final data collection.

3. Data and Method

3.1. Data

In this study, we chose the New York–Northern New Jersey–Long Island areas of the NY–NJ–CT–PA combined metropolitan statistical area as the study area [44]. This combined metropolitan statistical area includes twelve New Jersey counties, eleven New York counties, and one Pennsylvania county. Overall, there were more than eighteen million people (5.59% of the U.S.) living in this region, with higher median household income ($83,160) than the nationwide status ($68,703 in the U.S. as a whole).

We build the final data collection based on four data sources: the American Community Survey (ACS), the General Transit Feed Specification (GTFS), OpenStreetMap, and the 2014 U.S. Geological Survey. We collected related data at the census tract level and dropped those tracts whose population is less than five hundred. The final data collection contains 4057 census tracts. Moreover, we acknowledged one limitation in the final data collection. Since we captured some data from the open-source dataset, such as GTFS and OpenStreetMap, instead of local authorities directly, the reliability and accuracy could be worth noting. Given these datasets commonly used in previous studies, this concern can be minor.

Table 1 describes the factors to measure gentrification as the dependent variable. These factors are from two domains, socioeconomic and demographics, and housing, including population, education, race, income, in-migrants, employment status, housing prices, rent prices, and density factors [15]. We captured all these variables from ACS in 2010 and 2018.

Table 1. Variables for gentrification and transit desert measurements.

	Variables by Census Tract	Years
Gentrification variables		
Socioeconomic and demographics	Total population Adults (25 years old or older) Adults with a college degree Non-white population Median household income Low-income in-migration Employment density	2010, 2018
Housing	Housing units in pre-1950 buildings Median rent price Median home value	
Transit desert variables		
Transit supply	Number of transit facilities (e.g., subway, commuter train, bus/BRT, and ferry routes) Number of transit routes Number of intersections Length of low-speed road Length of bike lanes and pedestrian	2020
Transit demand	Population age 16 and over Population living in group quarters Vehicle ownership Nationwide carpooling ratio	2020

In addition, Table 1 presents the five criteria for measuring transit supply and four for transit demand. In this study, the independent variable was the transit desert index which is the difference between transit supply and demand [21]. Transit desert indices with a negative value are generally categorized as transit deserts, indicating areas where the existing supply cannot meet public transit demand. Conversely, positive transit desert index values indicate “transit oases” where the existing public transit facilities supply can meet the area’s demand. In this study, we captured the data involving the transit desert on 15 July 2020, from ACS, GTFS, and OpenStreetMap. For the transit supply, we defined the public transit system in the study area, including subway, commuter train, bus/BRT, and ferry routes from GTFS [45]. Built environment factors were captured from OpenStreetMap [46]. The intersections counted each intersection that involved primary roads, secondary roads, tertiary roads, and motorways. The low-speed roads measured in this analysis were limited to 35 miles per hour. The length of bike/pedestrian-friendly roads included tertiary roads, residential roads, bike lanes, and pedestrian facilities. We focused on transit-dependent adults based on four criteria density (i.e., divided by area of census tract) for transit demand. In the following section, we introduced in detail the process to define gentrification typologies, and transit supply and demand.

3.2. Methods

3.2.1. Gentrification Measurement

Gentrification is a complex process composed of sub-processes, but factors such as income, renting rate, race, public transit, and housing value have been considered sufficient criteria to measure gentrification progress [6]. Previous studies have noted the complexity in measuring gentrification progress, and a study has developed a method to quantify gentrification progress in New York metropolitan area [15]. They defined eight categories including not losing low income (low-income), at risk (low-income), ongoing displacement (low-income), ongoing gentrification (low-income), advanced gentrification (moderate to high income), stable exclusion (moderate to high income), ongoing exclusion (moderate to high income), super-gentrification (moderate to high income). Referencing their measurements, we considered eight subcategories of gentrification progress within three income categories: one subcategory of the very low-income group (VLI), three subcategories of the low-income group (LI), and four subcategories of the moderate–high income group (MHI) (Table 2). The reasons for modifications are as follows: First, studies claims that the poor and renters are more sensitive to changes in housing cost than other residents [47]. Therefore, it is worth adding a new category-specific neighborhood with a high rate of renters and low-income residents. Second, the difference between ongoing displacement (low-income) and ongoing gentrification (low-income) is not clear. We decided to merge them.

Table 2. Categories and subcategories of gentrification.

Category	**Subcategory**	**Definitions**
VLI (Very-low-income group)	Highly Vulnerable	Rate of renters > 50% and rate of poverty > 25%
	Stable	Low-Income Tract (<regional median household income) and not VLI in 2018
LI (Low-income group)		Low-Income Tract and not VLI in 2018
		2 out of the 3 of the following are true in 2018:
	At Risk	Has subway/rail station in tract
		% of units in pre-1950 buildings > regional median
		Employment density > regional median
		Low-Income Tract and not VLI in 2018
	Ongoing	Population stable or growing 2010–2018
		Loss of LI households 2010–2018 (absolute loss)
		Moderate-to-High-Income (~regional median household income) Tract in 2018
		Demographic change between 2010 and 2018 (at least 2 of 3 occurring):
	Advanced	Growth in % college educated > regional median
		Growth in real median household income (percent change) > regional median
MHI (Moderate–high income group)		Lost low-income households
	Stable	Moderate-to-High-Income (~regional median household income) Tract in 2018
		Moderate-to-High-Income (~regional median household income) Tract in 2018
		2 out of the 3 of the following are true in 2018:
	At Risk	Has subway/rail station in tract
		% of units in pre-1950 buildings > regional median
		Employment density > regional media
		Low-Income Tract and not VLI in 2018
	Exclusion	Population stable or growing 2010–2018
		Loss of LI households 2010–2018

Note. This approach referenced the work [15].

In detail, the first category is the very low-income group (VLI) which refers to the census tracts with more than 50% of residents as renters and 25% below the poverty rate. For neighborhoods in this type of census tract, a minor economic disruption (e.g., new facilities or investment) can significantly increase the property tax, leading to gentrification.

For the second category, census tracts fall under the low-income group (LI) census tracts if the share of low-income households (below 80% of the county's Annual Median Household Income) out of all households is larger than the regional median share of low-income households. The three groups, LI—Stable, Not Losing Low Income Households, LI—At Risk of Gentrification and Displacement, and LI—Ongoing Gentrification and Displacement, are used to define the gentrification progress of the tracts.

The transformation point, where a tract once considered an LI tract becomes an MHI tract, is represented by the subcategory MHI—Advanced Gentrification. Under this category, new residents tend to be highly educated. Low-income households are being forced out and leaving the neighborhood. The next phase of gentrification is a more pronounced version of that seen in the LI subcategories. Middle- and higher-income people continue to move in, while low-income households continue to be displaced. As more investment and infrastructure are added to the neighborhood, the degree of gentrification continues to increase. The last typology, Advanced Exclusion, describes a situation where low-income households cannot afford to move in due to high housing prices, and the gentrification process is completed. In this study, we considered the progress from LI—Stable to MHI—At Risk as gentrification, while the progress from MHI—At Risk to MHI—Exclusion is considered as super-gentrification.

3.2.2. Transit Service Measurement

We used the transit desert concept to measure the public transit service. The transit desert concept is based on the supply and demand concept. The transit desert index is the gap between transit supply and demand [21]. We noted the various measurements in transit supply and demand [42,43,48]. In this study, we measured the transit desert as follows: First, we defined transit demand as residents who needed public transit to travel (Equation (1)), and the transit demand is the z-scored transit-dependent adults (Equation (2)). Additionally, we defined transit supply as the ability to move riders using the five criteria (Table 1). Then, the z-scored sum of z-score criteria density represents the comprehensive measurement of transit demand (Equation (3)). The transit desert index is defined as the gap between transit supply and demand (Equation (4)). A larger transit desert index means better public transit services since there is more surplus in transit supply than demand.

$$\text{Transit-dependent adults} = (\text{population age 16 and over}) - (\text{population living in group quarters}) - (\text{vehicles available}) * (\text{national-level carpooling ratio}) \tag{1}$$

$$\text{Transit demand} = \text{z-scored}(\text{Transit dependent adults}) \tag{2}$$

$$\text{Transit supply} = \text{z-scored}\left(\sum \text{z-scored criteria density}\right) \tag{3}$$

$$\text{Transit desert index} = \text{transit supply} - \text{transit demand} \tag{4}$$

3.2.3. Analysis Framework

This study aims to investigate the relationships between public transit services and gentrification progress, and we applied three analyses to achieve the research goal, including statistical/spatial description, regression models, and machine-learning algorithms. Through these analyses, the results can present a comprehensive and robust view of these relationships. First, we observed the rate of census tracts defined as transit deserts. Based on the results, the rates of census tracts defined as transit deserts at categories of gentrification can be calculated. In addition, the results can present a general idea about the locations of transit deserts, locations of gentrification categories, and the relationships between them spatially.

Second, regression models are applied to further identify the relationships. We used RStudio and multinomial logistic regression ("multinom" function of the "nnet" package) to explore the relationships [49]. It is a widely accepted method in modeling categorically dependent variables at multiple levels [50]. Results can contribute to the statistical understanding impacts of public transit on gentrification progress.

In addition, machine-learning algorithms raised research interest in modeling relationships without linear assumptions [51]. Moreover, studies highlighted robustness of these methods compared to regression models [52,53]. In terms of the advantages of machine-learning algorithms and comparing the results between machine-learning algorithms and regression models, we chose X-Gradient Boosting Decision Tree classifier (XGBclassifier). This branching structure allows regression trees to naturally learn the relationships without strict pre-assumption.

To achieve the XGBclassifier, we developed functions on Jupiter Notebook in Python 3.7 through the "statsmodels" and "xgboost" package. We first identified the hyperparameters through cross-validation grid searching. Then, we adopted the "SHAP" method to interpret the results [54]. Finally, it utilizes values from game theory to locally explain the contribution of each factor to the output of a predictive model (Equation (5)):

$$\mathrm{SHAP}_{\mathrm{feature}}(\mathrm{x}) = \sum_{\mathrm{set:feature} \epsilon \mathrm{set}} \frac{|\mathrm{A}|!(\mathrm{p} - |\mathrm{A}| - 1)!}{\mathrm{p}!}(\mathrm{val}(\mathrm{A} \cup \{\mathrm{x}\}) - \mathrm{val}(\mathrm{A}))) \quad (5)$$

where p is the number of features. A represents the subset of the feature. x is the vector of feature values of an instance to be explained. $\mathrm{val}(\mathrm{A})$ is the prediction for feature values in the set A.

4. Results

4.1. Distribution of Census Tracts in Different Gentrification Subcategories

We first examined the distribution of census tracts in different gentrification subcategories in total and in the transit desert only. Table 3 presents the statistical description of the gentrification status of all census tracts and those defined as transit deserts. In general, there were 3.7% of census tracts defined as VLI neighborhoods, which indicates that these areas were at extreme risk of being gentrified. In addition, over one-third of census tracts in the study area was defined as low-income neighborhoods (LI), and 19.6% were at risk of being gentrified and ongoing gentrification (LI—At Risk and LI—Ongoing). Moreover, focusing on the MHI neighborhoods, we found that 12.2% were at risk of ongoing exclusion, representing the risk of ongoing super-gentrification [19]. Moreover, the distribution of subcategories in LI and MHI was similar. In this study area, around 60% of neighborhoods were defined as stable, and more than 18% were at risk of gentrification (LI—At Risk and MHI—At Risk).

Table 3. Comparison between gentrification categories.

Gentrification Typology	All		Transit Desert	
	Number	Percentage	Number	Percentage
VLI—Highly Vulnerable	151	3.7%	26	17.9%
LI—Stable	1233	30.4%	24	16.6%
LI—At Risk	358	8.8%	35	24.1%
LI—Ongoing	440	10.8%	25	17.2%
MHI—Advanced	195	4.8%	7	4.8%
MHI—Stable	1186	29.2%	26	17.9%
MHI—At Risk	323	8.0%	1	0.7%
MHI—Exclusion	171	4.2%	1	0.7%
Sum	4057	100.0%	145	100.0%

There are various impacts of public transit on gentrification. Comparing the differences between the distribution of census tracts and transit deserts, we found that there was a higher rate of LI at-risk neighborhoods and a lower rate of LI stable neighborhoods in transit deserts than in the general situation. It indicates that low-income neighborhoods defined as transit deserts should be paid attention to when allocating budgets for improving local public transit services, which is consistent with previous studies [55]. Moreover, focusing on the MHI category, the rate of transit deserts is the same in the MHI—At Risk and MHI—Exclusion subcategories. It indicates that improving the transit service may not be the driver of super-gentrification.

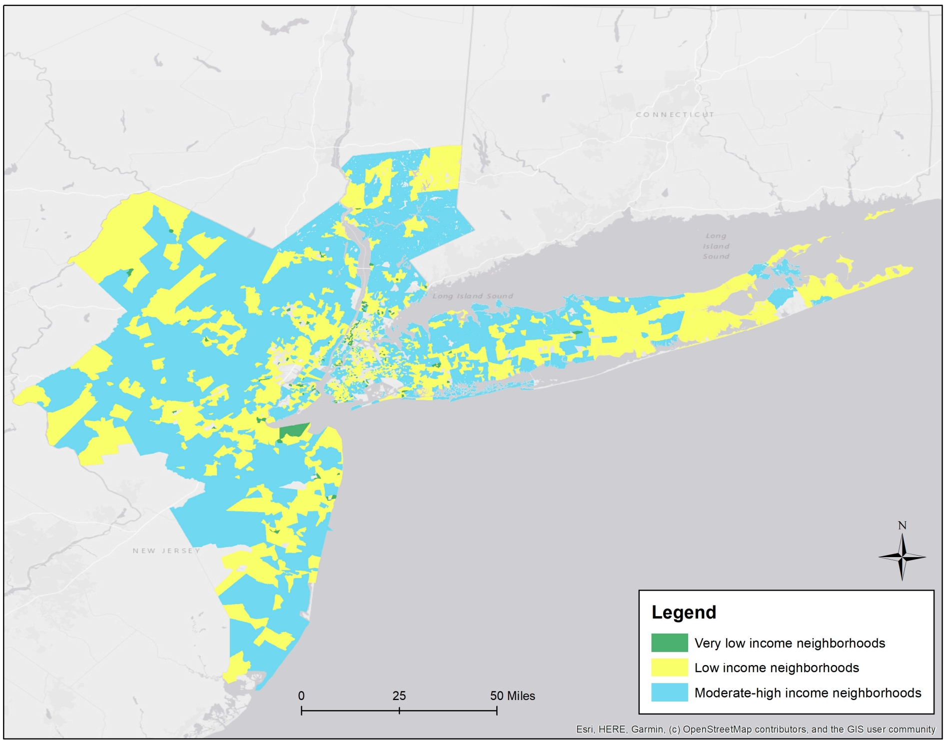
4.2. Spatial Distribution of Census Tracts in Different Gentrification Subcategories

The spatial distribution of census tracts in different gentrification subcategories is worth noting. Figure 1 presents the distributions of the economic/income situations (Figure 1a) and gentrification statuses (Figure 1b). There are four findings worth noting. First, the distribution of economic/income situations resembled concentric circles, with Manhattan as the core gathering of MHI groups. Staten Island, The Bronx, Brooklyn, Queens, Newark, and other boroughs were the second ring where LI groups gathered. The third ring was where MHI groups gathered, including Nassau County, Suffolk County, Fairfield County, Sussex County, and other counties. There were some VLI groups located in Middlesex County. Second, given the distribution of gentrification status, we found that most of the tracts at risk of gentrification or ongoing gentrification were concentrated in New York City (NYC), and some surrounding areas were undergoing gentrification processes, which might be a "ripple effect" [56]. Third, although there can be a ripple effect in general, situations in different boundaries are worth noting. We found that most tracts in the northern part of NYC were experiencing ongoing gentrification. Most tracts in Queens and the southern part of Brooklyn were at risk of gentrification. In addition, gentrification progress from LI to MHI seemed nearly complete, as reflected by the category MHI—At Risk of the tracts in Brooklyn. Finally, focusing on the situation inside NYC, we found that most transit deserts are located in the Upper West Side (UWS) and Harlem neighborhoods. This may be caused by the high transit demand and low accessibility of transit services (e.g., lacking pedestrian and bike lanes).

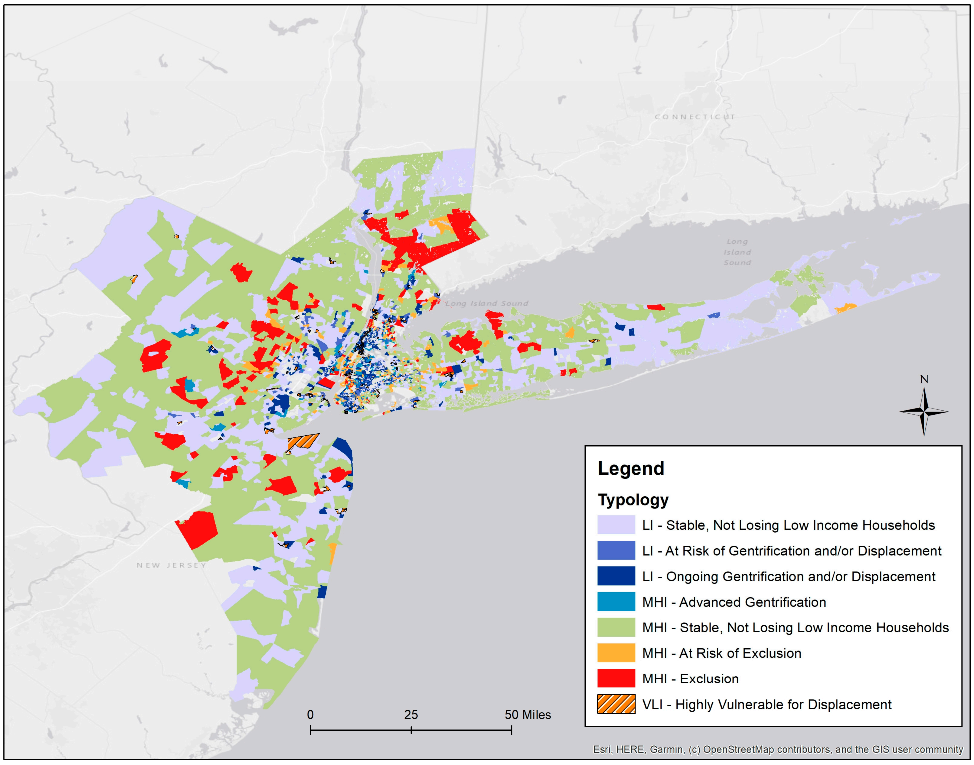
4.3. Multinomial Logit Regression Results

4.3.1. Relationship between Public Transit and Gentrification Categories

Table 4 presents the results of the multinominal models. Before modeling, we tested for the proportional odds assumption, and all dependent variables passed the test. In the first model, investigating the impacts of public transit factors on general economic groups, the coefficient of transit supply is significantly positive when comparing the VLI and MHI neighborhoods. The results indicate that a one-unit increase in transit supply could be related to an extra 50% possibility of neighborhoods defined as MHI than VLI, while the difference in transit supply between VLI and LI is not significant. At the same time, the impacts of transit demand are significantly negative, indicating that the demand for public transit service in VLI neighborhoods is likely much higher than in LI and MHI neighborhoods. A one-unit increase in public transit demand can be related to a lower possibility of being defined as LI (44%) and MHI (83%) compared to VLI.

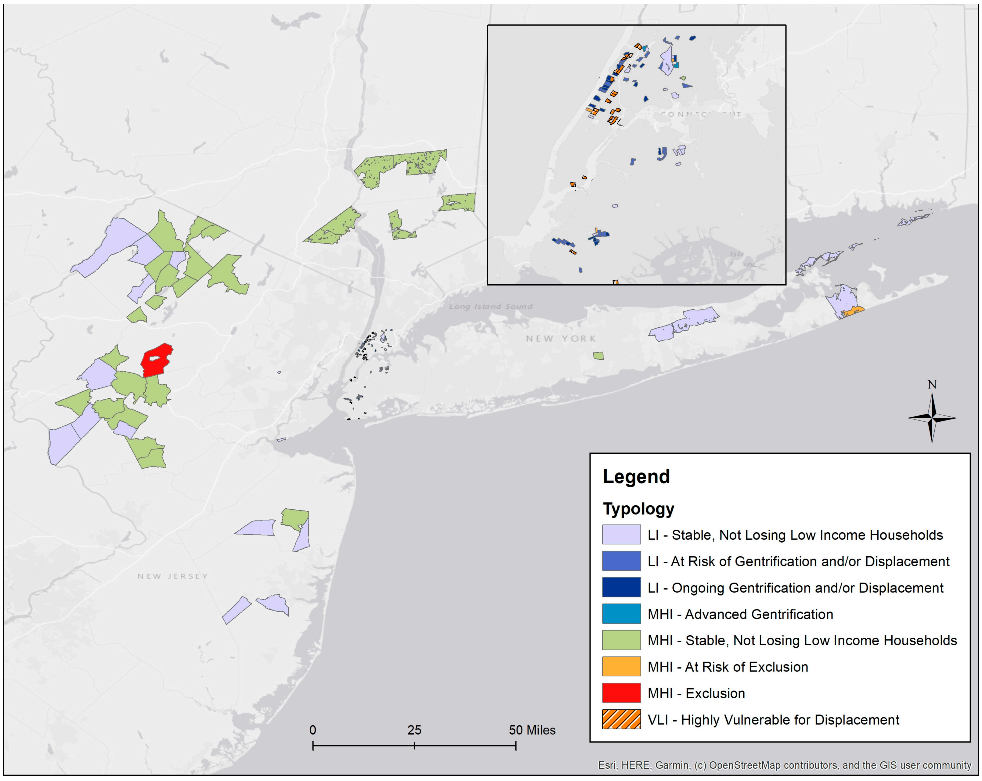


(**a**)



(**b**)

Figure 1. *Cont.*



(**c**)

Figure 1. Results of income categories, gentrification subcategories, and gentrification subcategories located in Transit Deserts ((**a**) census tracts by income categories; (**b**) census tracts by gentrification progress; (**c**) census tracts defined as transit deserts by gentrification progress).

Table 4. Results of logit models by categories.

	Dependent Variable: Baseline = VLI			
	LI	**MHI**	**LI**	**MHI**
	(1)	(2)	(3)	(4)
Transit supply	0.075 (0.068)	0.407 *** (0.073)		
Transit demand	−0.590 *** (0.078)	−1.823 *** (0.106)		
Transit desert index			0.250 *** (0.089)	0.403 *** (0.090)
Constant	2.832 *** (0.134)	2.913 *** (0.141)	2.325 (0.121)	2.033 *** (0.122)
Observations	4057		4057	
Akaike Inf. Crit.	5979.442		6094.968	

Note: *** $p < 0.01$.

In addition, the relationship between gentrification categories and the transit desert index is worth noting. Based on the results, areas with good public transit can be more likely defined as relatively high-income neighborhoods. In detail, compared to VLI, areas with one-unit increases in the transit desert index can be related to an extra possibility of being defined as LI (28%) and MHI (49%).

In summary, both results indicate that neighborhoods with good public transit can be highly likely defined as neighborhoods with the advanced progress of gentrification. These results are consistent with previous studies claiming that public transit in neighborhoods with higher incomes is better than those with lower incomes in general [57,58]. In addition, the results further identified that these mismatches are not only caused by high supply but also significant demand.

4.3.2. Relationship between Public Transit and Gentrification Subcategories

To compare transit service between different gentrification subcategories, we applied another multinomial logistic regression model choosing VLI as the reference (Table 5). First, the results indicate that the transit supply in VLI is less than most subcategories except the LI—Stable, but transit supply is not monotonically non-decreasing when the subcategories keep progressing. It is worth noting that the coefficients increase from LI—At Risk to MHI—Advanced Gentrification while they decrease from MHI—Advanced Gentrification to MHI—Stable. It can support the claim that the impacts of public transit on gentrification progress should be nonlinear.

A similar distribution happened to transit demand. In general, the results indicate that transit demand in VLI is greater than any other subcategories of gentrification progess, and the average coefficient in MHI groups is larger than in LI groups, indicating that MHI transit demand could be smaller than both VLI and LI groups. In addition, it is worth noting that the coefficients are not monotonically non-decreasing with increasing gentrification progress. For instance, both public transit demands in LI—At Risk and MHI—Stable are much smaller than VLI, but the increases in MHI—Stable are more significant than LI—At Risk.

Third, the transit desert index results are similar to the above two results. In general, public transit in most LI group and all MHI group neighborhoods is better than VLI. In addition, it is worth noting that the impacts of the transit desert index should be nonlinear. The coefficient reaches the highest in the MHI—Stable group, followed by the MHI—Advanced Gentrification group, and the difference in averages of the coefficients between LI and MHI groups is insignificant (0.46 vs. 0.54).

To sum up, the subcategories' results are consistent with the categories' results, highlighting that public transit in neighborhoods defined as Advanced Gentrification could be better than those defined as VLI. In addition, results mention that the impacts of public transit on gentrification progress should be nonlinear.

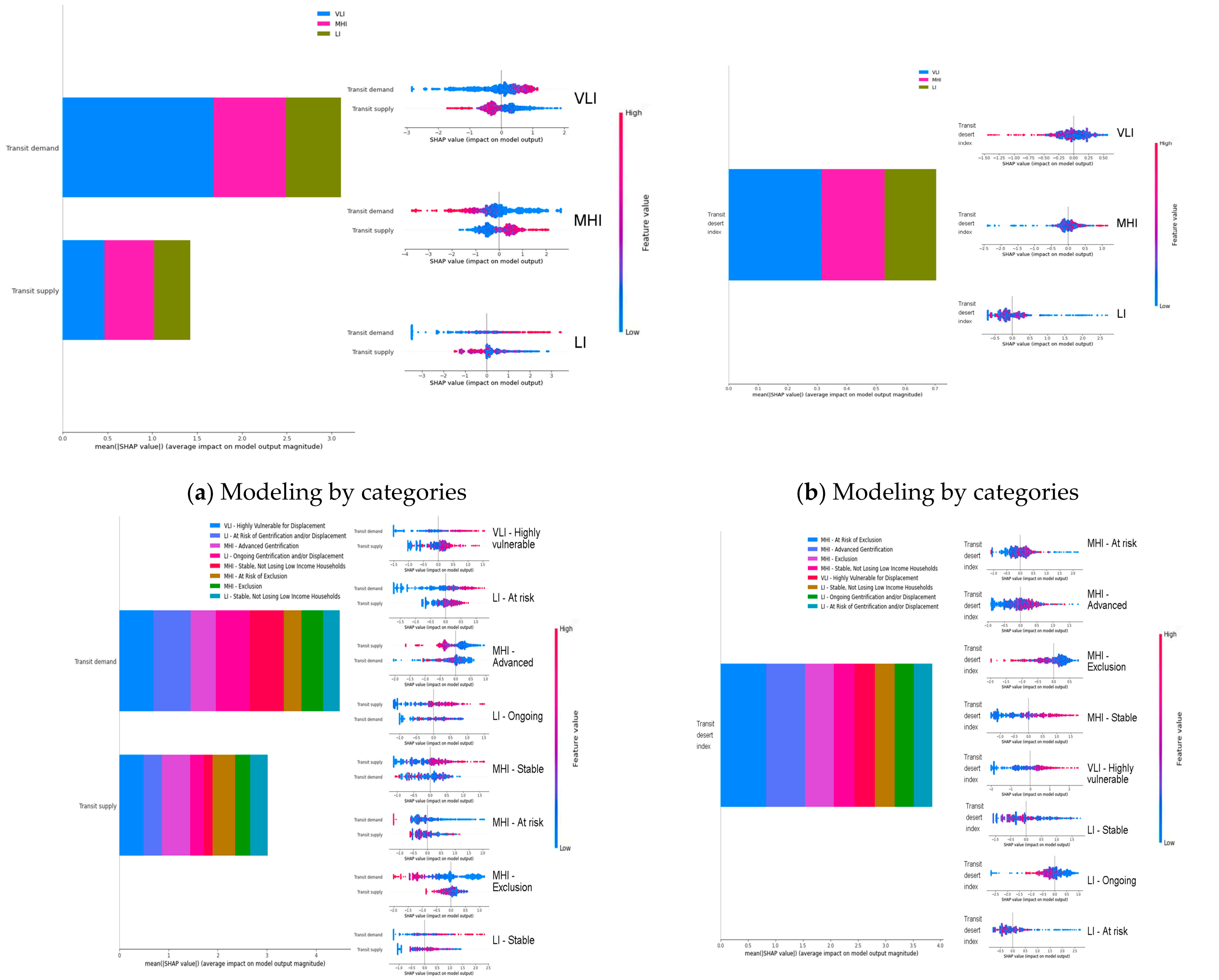
Table 5. Results of logit models by subcategories.

	Dependent Variable: Baseline = VLI													
	LI Groups	**MHI Groups**	**LI Groups**					**MHI Groups**						
	Stable	**At Risk**	**Ongoing**	**Advanced Gentrification**	**Stable**	**At Risk**	**Exclusion**	**Stable**	**At Risk**	**Ongoing**	**Advanced Gentrification**	**Stable**	**At Risk**	**Exclusion**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Transit supply	−0.129	0.175 **	0.288 ***	0.514 ***	0.308 ***	0.612 ***	0.430 ***							
	(0.084)	(0.081)	(0.080)	(0.086)	(0.087)	(0.083)	(0.103)							
Transit demand	−1.192 ***	−0.248 ***	−0.483 ***	−1.078 ***	−3.316 ***	−1.500 **	−2.684 ***							
	(0.106)	(0.085)	(0.089)	(0.125)	(0.177)	(0.125)	(0.261)							
Transit desert index								0.125	0.383 ***	0.551 ***	0.748 ***	0.192 **	0.814 ***	0.421 ***
								(0.097)	(0.105)	(0.101)	(0.103)	(0.098)	(0.100)	(0.114)
Constant	2.610 ***	0.760 ***	0.0908 ***	−0.119	1.377 ***	0.241	−0.322	1.954 ***	0.452 ***	0.42 1 ***	−0.739 ***	1.830 ***	−0.383 **	−0.395 **
	(0.158)	(0.170)	(0.166)	(0.185)	(0.180)	(0.174)	(0.226)	(0.127)	(0.146)	(0.142)	(0.163)	(0.128)	(0.151)	(0.170)
Observations	4057							4057						
Akaike Inf. Crit.	12,422.500							13,856.560						

Note: ** $p < 0.05$; *** $p < 0.01$.

4.4. Machine-Learning Classification Results

Figure 2 shows the results of the XGBclassifier. In general, these results indicate that the importance of transit demand in explaining gentrification in both categories and subcategories is higher than transit supply. In addition, a large transit demand has high and positive impacts on determining neighborhoods as VLI, while a large transit supply has high but negative impacts. Also, more transit demand can be related to a lower likelihood of being classified as LI, while less transit demand can be related to lower likelihood of being classified as MHI. At the same time, more transit supply can contribute to a higher possibility of being defined as MHI, but the trend in LI is unclear.



(**a**) Modeling by categories (**b**) Modeling by categories

(**c**) Modeling by subcategories (**d**) Modeling by subcategories

Figure 2. SHAP values of XGBDT results. ((**a**) impacts of transit demand and supply on income categories; (**b**) impacts of transit desert index on income categories; (**c**) impacts of transit demand and supply on gentrification subcategories; (**d**) impacts of transit desert index on gentrification subcategories).

Given Figure 2b, it is worth noting that the impacts of the transit desert index in all three categories are not as clear as the above. We claim that a large number of transit desert indices has high and positive impacts on determining neighborhoods as VLI but negative impacts on MHI neighborhoods; however, the impacts of a median number of indices can be various since the trends are unclear. For the LI neighborhoods, we claim that a small transit desert index can contribute to determining neighborhoods as LI.

Figure 2c,d present the results by subcategories. Similar to the above results, the importance of transit supply is globally lower than demand, but in some cases, the contribution of transit supply is larger than transit demand, including MH—Advanced, LI—Ongoing, and MHI—Stable. In addition, focusing on the impacts of transit demand, we claim that high transit demand could contribute to defining neighborhoods as VLI, LI—At Risk, and LI—Stable, and low transit demand could be related to MHI—Advanced, MHI—At Risk, and MHI—Exclusion. In addition, transit supply in LI—At Risk, LI—Ongoing, MHI—Stable, and LI—Stable neighborhoods could be higher than the others, while in MHI—Advanced neighborhoods, it could be significantly lower. Figure 2d presents the results of the transit desert index. It indicates that a high transit desert index could be related to neighborhoods defined as MHI—Stable and VLI, while a low transit desert index could happen in MHI—Exclusion and all LI group neighborhoods.

In short, the findings indicate that the relationships between public transit and gentrification progress could be myriad. These relationships and the importance of these factors could be various at the different points of gentrification progress. In addition, these findings indicate that the impacts of public transit should change even within the same subcategory in terms of knots in the beeswarm figures.

5. Discussion

Gentrification raises concerns among researchers. Previous studies explored the relationships between gentrification and public transit. At the same time, few considered the relationship between gentrification and transit from a regional perspective and applied a dynamic perspective on gentrification progress. In addition, previous studies focused on whether transportation investment led to gentrification, lacking views of the impacts on gentrification progress. By subdividing the categories of gentrification, we examined the impacts of transit services on gentrification. The results indicate that the rate of neighborhoods defined as very-low-income and low-income groups in transit deserts is higher than the average. In comparison, the rate of moderate–high-income groups is much lower in transit deserts. In addition, the level of public transit service is positively correlated with gentrification, while its impacts are insignificant in the super-gentrification progress. In addition, the leading cause of transit deserts in this study can be high demand besides lacking supply.

Based on these findings, there are policy implications that can be generated. First, local authorities must be aware that public transit could not meet the demands inside NYC and the periphery among low-income neighborhoods. This study identifies transit deserts, most of which are in NYC and the periphery. In addition, the rate of economically disadvantaged neighborhoods (i.e., VLI and LI categories) in transit deserts is higher than in others. These findings are consistent with previous reports [59,60]. Notably, these unmet demands might lead to a vicious circle of low-income households living in neighborhoods where they must spend extra time and money on transportation [61]. Hence, we call local authorities to invest in public transit in these low-income neighborhoods. Still, they should be careful in allocating these budgets to avoid displacement since this study proves transit's impacts on gentrification progress. To improve the transit service without displacement, we encourage authorities to reference other empirical cases, such as TOD in Pacoima [62]. In addition, it is crucial to co-operate with planning authorities to provide housing credit programs and ensure community engagement during the development [63].

Secondly, the issue of transit demand is more important than the issue of supply. In this study, results prove that the importance of transit demand is higher than transit supply in explaining gentrification progress, demonstrating that the gentrification status of neighborhoods correlated with the transit-dependent populations. Although there are minor differences between results of regression models and machine-learning approaches, it is worth noting that both highlight the general importance of transit demand on gentrification progress. Hence, it would be better for transportation authorities to co-operate with the land use department, using zonings to allocate affordable housing to avoid gathering. For neighborhoods with high transit-dependent populations, local authorities can allocate budgets to improve the service and accessibility of public transit facilities, such as bike lanes and pedestrian lanes.

Thirdly, the transit demand from moderate–high-income neighborhoods should not be ignored. Although this study argues that the transit demand in moderate–high-income neighborhoods is lower than in low-income neighborhoods in general, we identified that there are two subcategories of neighborhoods in the MHI category whose transit demand may be like the low-income one, including MHI—Advanced Gentrified and MHI—At Risk. We note that the transit supply in these areas is also high, but it does not mean that these services can be reduced. Therefore, transportation authorities do not need to improve the transit service in those areas but should keep the existing service to meet the high demand. We also call for future studies to examine the role of public transit in middle-class neighborhoods in attracting residents from driving alone to taking public transit.

In addition, we encourage future studies to reference the study framework applying different methods to explore the impacts of gentrification progress. In this study, the results highlight the complexity and nonlinearity of the relationships between gentrification progress and public transit. These not only happened to the marginal impacts of public transit during gentrification progress but also inside each progress according to the machine-learning results. Therefore, future studies should be aware of these and apply different models to quantify these nonlinearities.

Moreover, future studies need to apply dynamic and regional perspectives on gentrification and transit. In the literature review, we note that the emerging gentrification studies focused on regional analysis, while few studies involving transportation and gentrification noticed it [3,4,15]. This can be caused by lacking efficient approaches and data analyzing regional transportation services. Our study applies dynamic gentrification typologies and transit desert concepts that can be applied in future studies and suggests that future studies consider gentrification and public transit from dynamic and regional perspectives. Since the different relationships occurred within economic groups, it is critical to introduce large-scale and multiple-group analysis.

Finally, we acknowledge the limitations. First, as we mentioned in the data section, there is a limitation in the data collection, and we call for future studies to work with regional databases to deal with this. Second, it would be better to introduce other confounding variables in the regression models [64]. For example, access to public green space and new investments in such amenities should be at play. Future studies should include factors such as these to control for their effects. Third, we applied two approaches to investigate the impacts of public transit on gentrification progress and developed functions to find the best-fit model of XGBclassifier but did not further compare the performance with other machine-learning algorithms. We encourage future studies to consider multiple machine-learning approaches, such as random forest, elastic net, and other approaches for finding the best model.

6. Conclusions

This paper explored the relationship between gentrification and public transit service from a dynamic and regional perspective based on data in the New York–Northern New Jersey–Long Island area. We presented the spatial relationship between gentrification and public transit service and applied regressions and a machine-learning classifier to explore the relationship between gentrification and transit. Results identify the gentrification subcategories of these transit deserts and prove the significance of transit in gentrification progress. Based on the results, we suggested that policymakers need to be careful in allocating public transit budgets. In addition, future studies should explore related topics from a dynamic and regional perspective and consider the nonlinearity in the relationships.

Author Contributions: Conceptualization, Y.C. and H.X.; methodology, Y.C.; software, Y.C. and H.X.; validation, Y.C.; formal analysis, Y.C. and H.X.; analysis, Y.C.; writing—original draft preparation, Y.C.; writing—review and editing, J.J. and H.X.; visualization, Y.C. and H.X.; supervision J.J.; project administration, J.J.; funding acquisition, J.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the NSF Grants (2043060, 2133302, 1952193, 2125858, 2236305) USDOT consortium of Cooperative Mobility for Competitive Megaregions, Good System at the University of Texas at Austin and The Mitre Corporation. The authors would like to acknowledge these supporters.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Niu, S.; Hu, A.; Shen, Z.; Lau, S.S.Y.; Gan, X. Study on land use characteristics of rail transit TOD sites in new towns—Taking Singapore as an example. *J. Asian Arch. Build. Eng.* **2019**, *18*, 16–27. [CrossRef]
2. Yi, C. Impact of Public Transit on Employment Status: Disaggregate Analysis of Houston, Texas. *Transp. Res. Rec.* **2006**, *1986*, 137–144. [CrossRef]
3. Dong, H. Rail-transit-induced gentrification and the affordability paradox of TOD. *J. Transp. Geogr.* **2017**, *63*, 1–10. [CrossRef]
4. Jamme, H.-T.; Rodriguez, J.; Bahl, D.; Banerjee, T. A Twenty-Five-Year Biography of the TOD Concept: From Design to Policy, Planning, and Implementation. *J. Plan. Educ. Res.* **2019**, *39*, 409–428. [CrossRef]
5. Rayle, L. Investigating the Connection Between Transit-Oriented Development and Displacement: Four Hypotheses. *Hous. Policy Debate* **2015**, *25*, 531–548. [CrossRef]
6. Zuk, M.; Bierbaum, A.H.; Chapple, K.; Gorska, K.; Loukaitou-Sideris, A. Gentrification, Displacement, and the Role of Public Investment. *J. Plan. Lit.* **2018**, *33*, 31–44. [CrossRef]
7. Padeiro, M.; Louro, A.; da Costa, N.M. Transit-oriented development and gentrification: A systematic review. *Transp. Rev.* **2019**, *39*, 733–754. [CrossRef]
8. Revington, N. Gentrification, Transit, and Land Use: Moving Beyond Neoclassical Theory. *Geogr. Compass* **2015**, *9*, 152–163. [CrossRef]
9. Dawkins, C.; Moeckel, R. Transit-Induced Gentrification: Who Will Stay, and Who Will Go? *Hous. Policy Debate* **2016**, *26*, 801–818. [CrossRef]
10. Kahn, M.E. Gentrification Trends in New Transit-Oriented Communities: Evidence from 14 Cities That Expanded and Built Rail Transit Systems. *Real Estate Econ.* **2007**, *35*, 155–182. [CrossRef]
11. Rosenthal, T. Op-Ed: Transit-Oriented Development? More like Transit Rider Displacement. Los Angeles Times. Available online: https://www.latimes.com/opinion/op-ed/la-oe-rosenthal-transit-gentrification-metro-ridership-20180220-story.html (accessed on 27 March 2021).
12. Turrentine, J. When Public Transportation Leads to Gentrification. NRDC. Available online: https://www.nrdc.org/onearth/when-public-transportation-leads-gentrification (accessed on 27 March 2021).
13. Baker, D.M.; Lee, B. How Does Light Rail Transit (LRT) Impact Gentrification? Evidence from Fourteen US Urbanized Areas. *J. Plan. Educ. Res.* **2017**, *39*, 35–49. [CrossRef]
14. Margono, R.B.; Zuraida, S.; Abadi, A.A. Transit-induced Gentrification in Bangkok, Thailand: A Review. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *532*, 012013. [CrossRef]
15. Chapple, K.; Kleiman, N.; Sobolevsky, S.; Reshef, D.; Xi, H.; Vazquez, G.; Hambardzumyan, R. *Map of Gentrification and Displacement for The Greater New York*; New York University: New York, NY, USA, 2018.
16. Atkinson, R.; Wulff, M. *Gentrification and Displacement: A Review of Approaches and Findings in the Literature*; Australian Housing and Urban Research Institute: Melbourne, Australia, 2009.
17. Handy, S. Regional transportation planning in the US: An examination of changes in technical aspects of the planning process in response to changing goals. *Transp. Policy* **2008**, *15*, 113–126. [CrossRef]
18. Zegras, C.; Sussman, J.; Conklin, C. Scenario Planning for Strategic Regional Transportation Planning. *J. Urban Plan. Dev.* **2004**, *130*, 2–13. [CrossRef]
19. Lees, L. Super-gentrification: The case of Brooklyn Heights, New York City. *Urban Stud.* **2003**, *40*, 2487–2509. [CrossRef]
20. Shi, J.; Duan, K.; Xu, Q.; Li, J. Effect analysis of the driving factors of super-gentrification using structural equation modeling. *PLoS ONE* **2021**, *16*, e0248265. [CrossRef]
21. Jiao, J.; Dillivan, M. Ball State University Transit Deserts: The Gap between Demand and Supply. *J. Public Transp.* **2013**, *16*, 23–39. [CrossRef]

22. Freeman, L.M. Commentary: 21st Century Gentrification. *Cityscape* **2016**, *18*, 163–168.
23. Byrne, J.P. Two Cheers for Gentrification Urban Poverty & Gentrification: An Exchange. *Howard Law J.* **2002**, *46*, 405–432.
24. Lees, L.; Shin, H.B.; López-Morales, E. *Planetary Gentrification*; John Wiley & Sons: Hoboken, NJ, USA, 2016.
25. Kreager, D.A.; Lyons, C.J.; Hays, Z.R. Urban Revitalization and Seattle Crime, 1982–2000. *Soc. Probl.* **2011**, *58*, 615–639. [CrossRef]
26. Atkinson, R. *Alpha City: How London Was Captured by the Super-Rich*; Verso Books: New York, NY, USA, 2020.
27. Revington, N.; Zwick, A.; Hartt, M.; Schlosser, J. Universities and urban social structure: Gentrification, studentification, and youthification in five United States legacy cities. *Urban Geogr.* **2021**, 1–22. [CrossRef]
28. Brummet, Q.; Reed, D. *The Effects of Gentrification on the Well-Being and Opportunity of Original Resident Adults and Children*; Federal Reserve Bank of Philadelphia: Philadelphia, PA, USA, 2019. [CrossRef]
29. Drew, R. Does the Definition of Gentrification Matter? Enterprise Community Partners. Available online: https://www.enterprisecommunity.org/blog/2018/10/many-definitions-gentrification (accessed on 28 March 2021).
30. Grifith, J.M. Gentrification: Perspectives on the Return to the Central City. *J. Plan. Lit.* **1996**, *11*, 241–255. [CrossRef]
31. Camerin, F. From "Ribera Plan" to "Diagonal Mar", passing through 1992 "Vila Olímpica". How urban renewal took place as urban regeneration in Poblenou district (Barcelona). *Land Use Policy* **2019**, *89*, 104226. [CrossRef]
32. Hwang, J.; Li, J. What Have We Learned About the Causes of Recent Gentrification? *Cityscape* **2016**, *18*, 9–26.
33. Sklair, L. *The Icon Project: Architecture, Cities, and Capitalist Globalization*; Oxford University Press: New York, NY, USA, 2017.
34. Stein, S. *Capital City: Gentrification and the Real Estate State*; Verso Books: New York, NY, USA, 2019.
35. Honey-Rosés, J.; Anguelovski, I.; Chireh, V.K.; Daher, C.; van den Bosch, C.K.; Litt, J.S.; Mawani, V.; McCall, M.K.; Orellana, A.; Oscilowicz, E.; et al. The impact of COVID-19 on public space: An early review of the emerging questions–design, perceptions and inequities. *Cities Health* **2020**, *5*, S263–S279. [CrossRef]
36. Mueller, N.; Rojas-Rueda, D.; Khreis, H.; Cirach, M.; Andrés, D.; Ballester, J.; Bartoll, X.; Daher, C.; Deluca, A.; Echave, C.; et al. Changing the urban design of cities for health: The superblock model. *Environ. Int.* **2020**, *134*, 105132. [CrossRef]
37. Hörcher, D.; Tirachini, A. A review of public transport economics. *Econ. Transp.* **2021**, *25*, 100196. [CrossRef]
38. Chapple, K.; Loukaitou-Sideri, A. *Transit-Oriented Displacement or Community Dividends? Understanding the Effects of Smarter Growth on Communities*; MIT Press: Cambridge, MA, USA, 2019.
39. Grube-Cavers, A.; Patterson, Z. Urban rapid rail transit and gentrification in Canadian urban centres: A survival analysis approach. *Urban Stud.* **2015**, *52*, 178–194. [CrossRef]
40. Berube, A.; Holmes, N. *City and Metropolitan Inequality on the Rise, Driven by Declining Incomes*; Brookings Institution: Washington, DC, USA, 2016.
41. Paul, J.; Taylor, B.D. Who lives in transit-friendly neighborhoods? An analysis of California neighborhoods over time. *Transp. Res. Interdiscip. Perspect.* **2021**, *10*, 100341. [CrossRef]
42. Chen, Y.; Jiao, J. Are There Transit Deserts in Europe? A Study Focusing on Four European Cases through Publicly Available Data. *Sustainability* **2022**, *14*, 13182. [CrossRef]
43. Vanderschuren, M.; Cameron, R.; Newlands, A.; Schalekamp, H. Geographical Modelling of Transit Deserts in Cape Town. *Sustainability* **2021**, *13*, 997. [CrossRef]
44. NY Department of Labor. New York-Northern New Jersey-Long Island, NY-NJ-PA Economy at a Glance. Available online: https://www.bls.gov/eag/eag.ny_newyork_msa.htm (accessed on 28 March 2021).
45. OpenMobilityData. TransitFeeds Is Now OpenMobilityData—OpenMobilityData. Available online: https://transitfeeds.com/news/open-mobility-data (accessed on 12 April 2021).
46. Geofabrik GmbH Geofabrik Downloads—OpenStreetMap Data. Available online: http://download.geofabrik.de/ (accessed on 24 August 2016).
47. NCRC. *Bloomberg: Poverty Is a Bigger Problem Than Gentrification*; NCRC: Ann Arbor, MI, USA, 2019.
48. Aman, J.J.C.; Smith-Colin, J. Transit Deserts: Equity analysis of public transit accessibility. *J. Transp. Geogr.* **2020**, *89*, 102869. [CrossRef]
49. RDocument. Glm Function—RDocumentation. Available online: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/glm (accessed on 28 March 2021).
50. Fullerton, A.S. A Conceptual Framework for Ordered Logistic Regression Models. *Sociol. Methods Res.* **2009**, *38*, 306–347. [CrossRef]
51. Zeng, J.; Yue, Y.; Gao, Q.; Gu, Y.; Ma, C. Identifying localized amenities for gentrification using a machine learning-based framework. *Appl. Geogr.* **2022**, *145*, 102748. [CrossRef]
52. Afonso, B.; Melo, L.; Oliveira, W.; Sousa, S.; Berton, L. Housing Prices Prediction with a Deep Learning and Random Forest Ensemble. In *Anais Do XVI Encontro Nacional de Inteligência Artificial e Computacional*; SBC Sociedade Brasileira de Cardiologia: Rio de Janeiro, Brazil, 2019.
53. Revend, W. *Predicting House Prices on the Countryside Using Boosted Decision Trees*; KTH Royal Institute of Technology: Stockholm, Sweden, 2020; Available online: https://www.diva-portal.org/smash/get/diva2:1464970/FULLTEXT01.pdf (accessed on 7 January 2023).
54. Chakraborty, D.; Başağaoğlu, H.; Winterle, J. Interpretable vs. noninterpretable machine learning models for data-driven hydro-climatological process modeling. *Expert Syst. Appl.* **2021**, *170*, 114498. [CrossRef]

55. Tehrani, S.O.; Wu, S.J.; Roberts, J.D. The Color of Health: Residential Segregation, Light Rail Transit Developments, and Gentrification in the United States. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3683. [CrossRef]
56. Elvinvest. *What Is the Ripple Effect, and How Does It Affect Property Prices?* ELVInvest: Geneva, Switzerland, 2019.
57. Giuliano, G. Low Income, Public Transit, and Mobility. *Transp. Res. Rec.* **2005**, *1927*, 63–70. [CrossRef]
58. Zhao, P.; Li, P. Travel satisfaction inequality and the role of the urban metro system. *Transp. Policy* **2019**, *79*, 66–81. [CrossRef]
59. Chen, S. Who's Afraid of a Transit Desert? *The New York Times*, 11 October 2019.
60. Suga, M.; We, L. Mapping Where the New York City Subway Doesn't Go. Available online: Bloomberg.com (accessed on 1 November 2021).
61. Tan, S.; Fowers, A.; Keating, D.; Tierney, L. Amid the Pandemic, Public Transit Is Highlighting Inequalities in Cities. *Washington Post*, 15 May 2020.
62. Bhusal, S. *Transit-Oriented Development Without Displacement: Strategies to Help Pacoima Businesses Thrive*; UCLA Institute of Transportation Studies: Los Angeles, CA, USA, 2021.
63. Berlin, L. Development without Displacement: An Interview with Isela Gracian. Urban Institute Next 50. Available online: https://next50.urban.org/article/development-without-displacement-interview-isela-gracian (accessed on 20 October 2021).
64. Trinh, S. Keeping Gentrification from Following Green Space. *Shelterforce*, 8 September 2020.