Mapping Gentrification in London: A Comprehensive Analysis of Socio-economic and Physical Neighbourhood Traits.

Master Thesis

Carst de Weerd

MSc. Real Estate Studies

University of Groningen

Abstract

This study delves into the dynamics of gentrification in London, UK, aiming to find patterns and link socio-economic and physical neighbourhood characteristics to the likelihood of gentrification. This study uses a quantitative approach and uses secondary data from British governmental institutions and employs a timespan of 9 to 19 years. The findings shed light on positive associations between the likelihood of gentrification and the presence of historic buildings, the size of private gardens, limited improvements in public transport, a higher percentage of white residents and proximity of affluent areas. On the contrary, the increased distance to the city centre shows negative associations with the probability of gentrification. This study extends to the influence of recently gentrified areas on the likelihood of gentrification and shows a negative relationship, solely in some particular areas in the city. These nuanced insights contribute to the understanding the complexity of gentrification patterns in London.

Keywords: [Gentrification, Logistic Regression, GIS, Neighbourhood Characteristics]

Colofon

Title	Mapping Gentrification in London: A Comprehensive Analysis of Socio-economic and Physical Neighbourhood Traits.
Version	Final version
Author	C.E. de Weerd
Supervisor	Dr. N.M. Kuiper
Assessors	Dr. N.M. Kuiper, dr. M.N. Daams
E-mail	c.e.de.weerd.1@student.rug.nl
Date	19-02-2024
Word count	10,447

Disclaimer: "*Master Theses are preliminary materials to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the author and do not indicate concurrence by the supervisor or research staff.*"

Preface

This master this delves into the dynamics of gentrification in London, United Kingdom – a subject that first captured my attention during high school geography classes. Over the years, my interest in this process has grown, which eventually resulted as my main topic for my thesis.

This thesis focuses on the relationship between multiple neighbourhood characteristics and the likelihood of gentrification. Secondary data from several British governmental institutions made it possible to perform multiple statistical tests. Although the results provide some interesting insights, it crucial to acknowledge the limitations of the study.

I want to thank my supervisor, Niels Kuiper MSc, for his guidance and feedback on my drafts and I want to thank my family and close friends for their support during the process of this master thesis.

Carst de Weerd

Table of Contents

1. Introduction

"The East London neighbourhood is so gentrified, actors and coffee shops are 'pushing out' the working class" (Oluwalana, 2022). This phrase originates from a London news website and is more and more returning to the media. In recent decades, the urban landscape has witnessed profound transformations, reshaping the physical and social fabric of cities across the globe. Among these dynamics, gentrification is a concept which experienced more and more attention from researchers, policymakers and residents in recent decades (Slater, 2011). While some argue that gentrification improves the living standards in the gentrified area, it is often also associated with the displacement of the original residents and this triggers the discussion. Residents from London's borough of Brixton protested against the rising rents, forcing people to move out of the neighbourhood where they lived their whole lives (BBC, 2015). Following public discourse, academic research has also often focused on the negative effects of gentrification. In contrast, relatively few studies examine why and how neighbourhood gentrification emerges (Behrens et al., 2024). This thesis adds to this literature by examining which physical and socioeconomic neighbourhood characteristics are related to a higher likelihood of gentrification.

In the 1960s, Glass (1964) was one of the first researchers to describe the concept of gentrification. Glass began to describe a pattern where the working-class fabric of neighbourhoods in London, began to transform into wealthier, middle-income neighbourhoods, pushing out the original inhabitants. This initial conceptualization marked the start of increased attention on gentrification within spatial and urban studies. Gentrification marks the relatively fast change in socioeconomic status of a neighbourhood, where low-income residents are pushed out of their original neighbourhood by middle- and high-income classes (Smith, 1982; Guerrieri et al., 2013; Zuk et al., 2015). The factors that can induce gentrification are manifold, encompassing economic, cultural and social factors (Lang, 1982; Hammet, 2003; Atkinson, 2004; Zuk et al., 2015; Finio, 2022), while others argued that gentrification is the result of uneven urban development (Smith, 1979). Most of the time, gentrification in a neighbourhood is more of a plural process (Zukin et al., 1992). Gentrification is often characterised by the change in residents' characteristics. Walks and Maaranen (2008) state that gentrification leads to a decline in social mix. Next to that, gentrification leads to more ethnic homogeneity in the gentrified area. Gentrification is occasionally linked to urban revitalisation or urban redevelopment, but this claim is not well-grounded (Nwanna, 2012). Urban redevelopment is a term which describes large-scale urban renewal, for example, a new commercial centre which benefits all social classes. Although gentrification reduces the poverty rate and increases the GDP of the urban area, social displacement has to be taken seriously when arguing if this process is beneficial for all social classes (Lin et al., 2021). This master thesis will, based on the literature discussed previously, employ the following definition of gentrification:

‘*Gentrification is a process characterized by a rapid influx of affluent newcomers, leading to an escalation of living costs and eventually resulting in social displacement of original residents*’

This definition emphasises that a relatively fast increase in the average income of the neighbourhood, thus by the influx of affluent residents, eventually results in gentrification. Therefore, gentrification is operationalized by focusing on changes in the income composition of neighbourhoods for this master thesis.

Policies concerning urban equality adjust to the results of research on gentrification (Atkinson & Wulff, 2009). Strategies addressing gentrification have predominantly centred on mitigating the displacement of low-income residents from neighbourhoods undergoing gentrification (Thackway et al., 2023). Quantifying gentrification is recognised as a first step to addressing potential areas where displacement can occur.

Despite the amount of both qualitative and quantitative studies which focused on identifying and investigating the causes of urban gentrification, the results always tend to be limited by data. Both the availability of the data, combined with the inadequacy of the prediction models, has led to a major methodological border for creating accurate predictive gentrification models (Atkinson & Wulff, 2009). As this nature of research tends to be reactive, it allows policymakers to employ a ‘blame the market’ tactic, that can blame the neo-liberal nature of the housing market as the biggest cause of gentrification (Rigolan & Nemeth, 2019; Thackway et al., 2023). Therefore, measures which can prevent gentrification are seen as the most proactive intervention methods to prevent gentrification instead of measures which react to gentrification (Chapple & Zuk, 2016).

This master thesis will focus on describing patterns of gentrification. Next to that, this master thesis will examine which initial neighbourhood characteristics, both socioeconomic and physical, are associated with a higher likelihood of gentrification. Next to researching these characteristics, this study will contribute to existing research by examining if gentrification has spatiotemporal patterns. The relationship between the proximity of recently gentrified areas and the likelihood of gentrification will be investigated and if gentrification patterns differ between different levels of urbanisation within the city. The results of this master thesis can inform policymakers on factors that make certain neighbourhoods vulnerable to gentrification. The focus of this master’s thesis is the region of Greater London, UK. London is one of the biggest urban areas in Europe and is known for its big social and economic inequality (Higgins et al., 2014). This can potentially make the disadvantaged neighbourhoods of London vulnerable to gentrification and this makes it interesting for research about this phenomenon. This thesis will take relative income increase as a dominant factor in determining gentrification over a period stretching from 9 to 19 years.

To test whether neighbourhood characteristics increase the likelihood of gentrification, this study will try to answer the following research question:

To what extent are socio-economic and physical neighbourhood characteristics related to the probability of gentrification in London, United Kingdom?

This research question will be answered, by answering the following sub-research questions:

- *Which factors are theoretically associated with a higher likelihood of gentrification?*
- *Which socio-economic and physical neighbourhood characteristics increase the probability of gentrification?*
- *Are there any spatiotemporal patterns in gentrification?*

This master thesis is based on a quantitative approach. It will first use a literature review to seek factors which could have an impact on the likelihood of gentrification. This will be finished with the creation of a conceptual model. Next, the methodology will be discussed with a descriptive analysis. This will be followed up by the results, discussion and finished by the conclusion.

2. Literature Review

The next section will cover physical and socio-economic neighbourhood characteristics that, according to existing literature, are related to an increase in the probability of gentrification. Existing literature suggests a positive relationship between the likelihood of gentrification with historic buildings, green spaces and the quality of public transport. Next to that, the proximity of affluent neighbourhoods is considered as a positive indicator of gentrification. In the following subheadings, the different neighbourhood characteristics that are related to an increase in the likelihood of gentrification will be separately discussed below.

2.1 Old and historical Buildings

The state of the existing real estate affects gentrification. Helms (2003) researched gentrification and the age of the current buildings in Chicago, Illinois. The research conducted claims that older, low-density houses in older neighbourhoods have an increased chance of being renovated. As real estate is becoming older, people tend to appreciate the characteristics of the building more and therefore attract more investors. This claim is supported by Grevstad-Nordbrock and Vojnovic (2019). They researched gentrification and the effect of heritage sites in the area of Lincoln Park in Chicago, Illinois. This research concluded that historic neighbourhoods function as a locus of reinvestment. Historic sites attract investors and high-income class and this eventually leads to increasing housing prices. After reinvesting and redevelopment of these older, more aesthetic buildings, the new target group is often the higher middle class. Therefore, after renovation, housing prices increase and become unaffordable for lower-income residents. Listed buildings further act as an architectural form that is further vanishing in a globalising city (Chang, 2016). Been et al. (2016) focus on the effect of the designation of historic districts on local housing markets in New York City. This research states that designation influences decisions to construct in a specific area. Next to that, designation increases the prices of properties within the specific district and gives a boost to property values in direct surrounding areas. As designation reduces supply and increases demand, it increases property prices in areas that already have a high amenity value. As such, it might further strengthen gentrification by making housing unaffordable for low-income households. This suggests that the presence of listed buildings has a positive relationship towards house prices in surrounding areas. Rosenthal and Ross (2015) argue that protected historic city centres offer a distinctive urban amenity, drawing in families with higher incomes. This can potentially result in socio-economic shifts that may adversely affect the original residents.

2.2 Green Spaces

It is argued that green spaces tend to be a positive predictor of gentrification (Maantay & Maroko, 2018; Rigolon & Németh, 2020; Triguero-Mas et al., 2022). Maantay and Maroko (2018) researched community gardens and gentrification in Brooklyn, New York. Their analysis suggests that proximity to community gardens is associated with a significant increase in income per capita. This suggests that community gardens can function as a predictor of gentrification. The last part is contradicting the research of Hawes et al. (2022). This study researched urban agriculture and gardens in Detroit, Michigan, and did not find a positive relationship for gardens to function as a predictor of gentrification in the area. Green spaces are viewed as a positive component of urban liveability and provide health benefits, recreational opportunities and ecosystem services (Kim & Wu, 2022). These positive functions make it more attractive to live in an area with more green spaces and therefore might make areas with a larger number of green spaces more vulnerable to gentrification.

Public parks also seem to have a positive relationship with gentrification (Rigolon & Németh, 2020). In their quantitative research about public parks and gentrification in cities in the United States, the research concludes that park function and location are strong predictors of gentrification, while size is not. Public parks located close to downtown trigger gentrification more than public parks in the outskirts of the city, showing that location is of big importance. Next to that public parks with good access to public transport are seen as good predictors of gentrification. Triguero-Mas et al. (2022) conducted similar research about 28 cities across the US, Canada and Europe and confirmed the same results.

2.3 Proximity of Wealthy Residents

Another strong predictor of gentrification is the relative distance of already wealthy areas (Guerrieri et al., 2013; Wilhelmsson et al., 2021). Guerrieri et al. (2013) focused on gentrification and location and concluded that the proximity of a wealthier neighbourhood is of big importance for an area to gentrify. Wealthy residents prefer to live near even wealthier residents, therefore a neighbourhood which borders a wealthier neighbourhood has a higher chance of being gentrified in the future. Guerrieri et al. (2013) refer to this process as endogenous gentrification. As wealthier neighbourhoods have positive spillover effects, such as more and better amenities and lower crime rates, people tend to prefer to live close to these neighbourhoods. Wilhelmsson et al. (2021) add that gentrified neighbourhoods have positive spillover effects of housing prices on neighbouring areas, based on their case study of Stockholm, Sweden. Contradicting to results of Wilhelmsson et al. (2021) the research of Christafore and Leguizamon (2019) shows that the proximity of recently gentrified areas tends to negatively affect the likelihood of gentrification. As in a specific area gentrification occurs, original residents experience social displacement and seek new living areas close to their former place.

2.4 Public Transport Accessibility

Research by Lian and Yang (2019) focused on the relationship between metro stations and commercial gentrification in Taipei, Taiwan. The researchers conclude, using logit models, that as travel distance to metro stations decreased, the probability of commercial gentrification increased. These findings are being supported by Ostrensky et al. (2021). This research focused on metro stations in Sao Paulo, Brazil and also concludes that newly built metro stations have a positive effect on gentrification. In an area with a newly built metro station, income increased. Next to metro stations, the presence of train stations adds to the value of a neighbourhood (Debrezion et al., 2011). They conclude that a train station will add a significant impulse to house prices in the Netherlands. Although house prices do not directly insinuate an increase in household income and therefore gentrification, it can be a factor in determining the process of gentrification. This suggests that efficient and good working public transport has a positive effect on property values. Research from Bardaka et al. (2018) researched transit-induced gentrification in Denver, CO. Their analysis shows that the installation of a light rail station significantly stimulates household income and housing values in directly surrounding areas up to one mile from the station.

2.5 Conceptual Model

Based on the literature review above, a conceptual model is constructed, which is shown in Figure 1. This conceptual model is based on neighbourhood characteristics that can be split into physical and socio-economic characteristics, which together increase the likelihood of gentrification. Physical neighbourhood characteristics are divided into three different aspects. The presence of historic buildings increases the chance of gentrification because these often imbue a locality with a sense of cultural richness (Chang 2016; Been et al., 2016). Historic structures have drawn the attention of individuals seeking unique and character-rich living environments. This subsequently led to a demographic shift toward higher-income residents, which contributed to gentrification. The presence of green spaces and gardens has a positive influence on gentrification, as this can act as a trigger of gentrification. Lastly, the quality and accessibility of public transport can be argued as a positive trigger towards gentrification.

Socioeconomic neighbourhood characteristics are associated with the racial distribution of neighbourhoods and the proximity of wealthy areas. Guerrieri et al. (2013) state that the proximity of wealthy areas, known as endogenous gentrification, adds to the likelihood of gentrification and Walks and Maaranen (2008) argue that gentrification leads to a more homogenous distribution. Therefore, a neighbourhood with an already high percentage of white residents can act as a trigger towards the likelihood of gentrification.

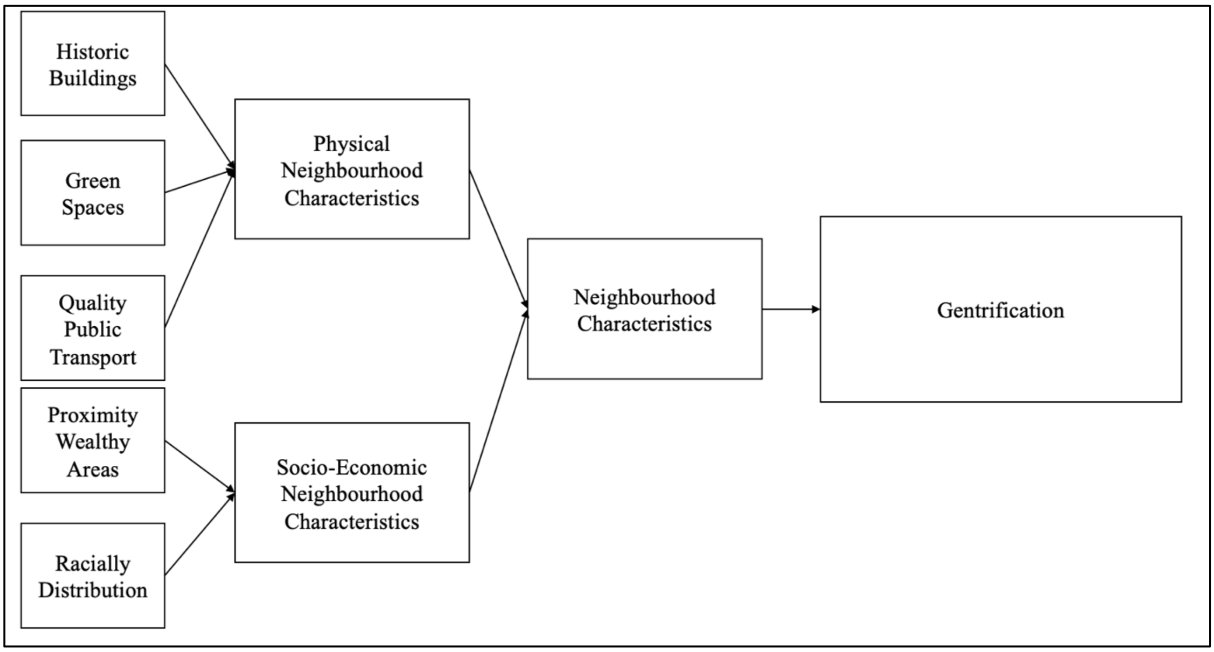

Figure 1, Conceptual model of gentrification, based on neighbourhood characteristics

By the creation of this conceptual model, the first sub-question, "*Which factors are theoretically associated with a higher likelihood of gentrification?"* is answered. This conceptual model forms the base for the construction of the model which tests to what extent neighbourhood characteristics increase the likelihood of gentrification, which will be further discussed in the next sections.

3. Data and Methodology

This section will cover the operationalisation of gentrification and the variables of interest in this study. First, the research area and period will be covered, followed by the data collection methods and the descriptive analysis. At the end of this section, the research design to answer the research question will be discussed.

3.1 Research Area and Period

This master thesis research area consists of the Greater London Region, UK. This consists of the City of London and the 32 boroughs of London. The map in Figure 2 shows the research area. The map shows the region of Greater London. The prominent black lines mark the administrative borders of the 33 boroughs. Within each of these boroughs, the smaller grey areas are the output areas, commonly known as the Middle layer Super Output Areas (MSOAs). The thesis unit of analysis is the MSOA.

MSOA is a geographical scale, used by the British government to gather specific area-bonded data. MSOAs comprise between 2.000 and 6.000 households, which usually have a resident population between 5.000 and 15.000 persons (ONS, 2024). Greater London in total, consists of 984 MSOAs.

To observe whether these MSOAs witnessed gentrification, data on the income composition of the MSOAs are collected in three moments: 2001, 2011 and 2020. These 19 years seem to be an adequate period to measure gentrification, as multiple kinds of research have periods from 6-15 years (Hedin et al., 2012; Wilhelmsson et al., 2021) up to 20 years (Rouwendal et al., 2018).

As discussed in the introduction, the area of Greater London was chosen, because of its long history with gentrification and therefore it provides interesting insights towards gentrification. London's image is characterised as a global city in an affluent region (Higgins et al., 2014). On the contrary, London has a big spatial inequality between 'inner' and 'outer' London. This inequality makes London an interesting research area, compared to other British cities. Next to that is the availability of secondary data affecting the choice of London. The British government and the London Datastore, a repository for statistical data, metadata and visualisations of the city region, gather and collect data on a very detailed level for different sectors and purposes. Most of these datasets are free and accessible to use for research. This facilitates the process of data gathering, compared to many other European countries where data gathered by governmental institutions are only available on request.

Figure 2, Research Area on borough and MSOA level.

3.2 Data Collection

This master thesis will make use of a quantitative method approach. Quantitative analysis tends to be the better option for this thesis, as it focuses on the process of gentrification therefore quantitative is more suitable, compared to qualitative. A qualitative approach is more able to highlight 'soft' findings, like the sense of place and residents' understanding. Nevertheless, this master thesis focuses on the process of gentrification and therefore quantitative analysis is a more appropriate approach.

The data used for the analysis is originating from multiple secondary data sources. These secondary data sources are all coming from different British governmental institutions, which are shortly described in Table 1. Table 1 shows that the collected data for the variables are stable over the period, which is primarily due to limited access to dynamic data over extended periods. While this limited temporal variation diminishes the robustness of the analysis, it still yields valuable insights for research purposes. The static nature of the dataset offers a unique snapshot and enables an exploration of the socio-economic and physical characteristics within the confined period. All of the data is collected on the MSOA geographical scale except for the PTAL scores, which are gathered at borough level.

GIS was used to create maps and merge different sorts of geographical data. The layers which are used for the gathering of the data originate from InFuse. InFuse is a free service providing layers of all different sorts of geographical scales, directly suitable for GIS. (UK Data Service, 2024). A layer with all the MSOAs of Greater London is used, combined with secondary data, which will be described shortly further in this section.

Table 1, Secondary data sources and year of collection

Used Secondary Data	Source	Year of collection
Income	National Statistics & Family Resource Survey	2001,2011,2020
Building Age	Valuation Office Agency	2014
Listed Buildings	National Heritage List for England	2023
Green Spaces	Office for National Statistics	2020
Public Transport	Transport for London	2016
Ethnicity	UK Census	2011

For the analysis, multiple independent variables will be used to test the model. All the variables which will be used are based on existing literature. The first variable focuses on the number of listed buildings per MSOA. Listed buildings are objects which have a special architectural and historic interest (Historic England, 2024). To make the location of listed buildings useable for analysis, 'listed building' point data, originating from the Historic England database, are merged with the polygons layer of all the MSOAs of Greater London. This created a new continuous variable, indicating the number of listed buildings of each MSOA.

The next variable covers the percentage of buildings which were constructed before 1940. The year 1940 was chosen for this analysis, because during 1940-1945, many bombardments took place in London and many buildings were destroyed, leading to a pronounced demarcation between pre- and post-war architectural landscapes. In the analysis, values will be used between 0 and 1, 0 implicating 0% and 1 implicating 100%.

Two variables are based on green space presence and average green space size. The first represents the percentage of addresses with a (private) garden. The second variable is continuous, indicating the average size per (private) garden, measured in squared meters.

This analysis uses two similar variables, which indicate the proximity of wealthy neighbourhoods. The variable is created by measuring the numbers of shared borders with MSOAs from the affluent neighbourhoods in respectively 2001 and 2011. These variables will be separately used to test if the presence of wealthy residents can trigger gentrification.

One variable indicates the level of accessibility and quality of public transport. Public Transport Accessibility Levels (PTAL) is a measurement created by the 'Transport for London' and this measurement is used for the variable (Transport for London, 2016). PTAL measures the connectivity of any specific place in Greater London to public transport. This is calculated based on:

- Walking distance to the nearest station or stop
- Number of services and lines at the nearest station or stop
- Waiting times at the nearest station or stop

The PTAL values range from 0 to 6, whereby 0 represents the worst possible connectivity to public transport and 6 represents the highest connectivity. The values of 1 and 6 are split into 2 values (1a and 1b & 6a and 6b). The PTAL data was gathered at the borough level. This data is transformed to MSOA scale, by using the overarching borough value to the specific MSOA within these boroughs.

To test whether the distance from a specific neighbourhood to the city centre influenced the probability of gentrification, a continuous variable was created. This variable represents the distance from the centroid of each MSOA's polygon to the 'City of London', calculated in meters. 'City of London' in this case, is chosen as the city centre, because of the size of Greater London, the urban area accommodates multiple 'downtowns' where multiple sorts of activity take place. Due to the historic function and the central location of the 'City of London', this MSOA functions as the city centre in this study.

Although the absolute distance to the city centre can be an important variable related to the likelihood of gentrification, it can also be interesting to highlight if relative distance matters for the likelihood of gentrification. Therefore, an additional layer of analysis is introduced to explore the relative distance to the city centre. Consequently, the samples are divided into four distinct groups. The MSOAs located closest to the borough 'City of London' are categorised as 'City Centre/Downtown'. The second ring consists of MSOAs located at a relatively greater distance from the city centre and is called 'Urban'. The third ring is labelled as 'Sub-Urban' and the outermost ring is designated as the

‘Periphery’, encompassing MSOAs positioned at the outermost reaches of the urban area. The new classification is visualised in Figure 3. The new classification system allows one to investigate if there are differences between different relative distances to the city centre and gentrification processes within the designated regions.

One variable was included to test the effect of ethnicity on the probability of gentrification of an area. The variable indicates the percentage of white people, according to the UK Census, per MSOA.

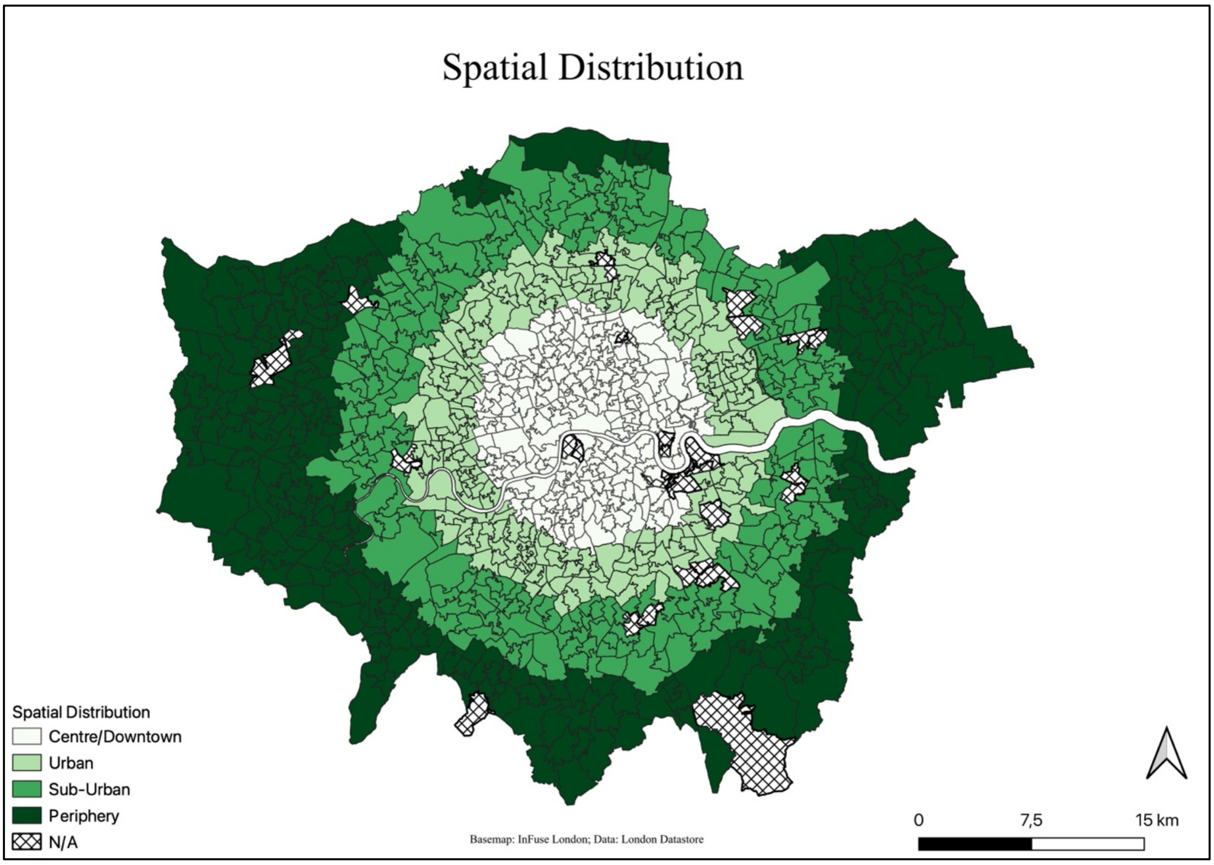

Figure 3, Spatial Distribution of London, divided by four subgroups.

3.3 The Gentrification Variable

There are many different ways to measure gentrification. Therefore, by taking a different way to measure gentrification, different outcomes will occur (Finio, 2022). Lyons (1995) saw displacement as the most suitable benchmark to measure gentrification, next to relative income change. The most used method to measure gentrification is by making use of income change, where education data can also be suitable for gentrification (Hedin et al., 2012). Wilhelmsson et al. (2021) determined gentrification by using the income per area, compared to the median income. Their research made use of a Getis-Ord statistic to locate gentrification hotspots on income and population changes. Gentrification occurs when in the first period, the income of an area was below the median of the city but eventually was above the median in the second period. Research by Behrens et al. (2024) states that income change and education are the best ways to measure gentrification. This research states that a neighbourhood gentrifies, if an area starts from a low-income level and experiences and substantial income increase.

The dependent variable for this research is a binary variable that indicates whether gentrification occurred in an MSOA. This variable is constructed out of the income data from 2001, 2011 and 2020. As described before, gentrification can be measured in many different ways and therefore resulting in very different results, compared to each other. For this master thesis, the measurement of the research of Wilhelmsson et al. (2021) is leading.

The analytical framework is centred on income quartiles, providing a dynamic lens to explore relative changes in income within the study population. The first step involved stratifying the entire dataset into quartiles based on income distribution for each respective year (2001, 2011 and 2020)[1]. These quartiles were constructed as follows: Group 1 represented the lowest 25% of the income of the observations, Group 2 encompassed observations in the 25-50% range, and so forth. This stratification facilitated the creation of a relative income variable, summarising the positioning of each observation within its annual cohort. A map with an overview of all the MSOAs divided by income quartiles is shown in Figure 4. As the map shows, there is a clear line from north to south where most of the wealthier neighbourhoods are located shown with the darker green colour.

The crux of the gentrification variable lies in discerning transitions across these income quartiles over time. Specifically, an observation was deemed to have undergone gentrification if it was in either Group 1 or 2 during the years of 2001 or 2011 but transitioned to Group 3 or 4 in a subsequent period. This temporal criterion describes the essence of gentrification, reflecting an upward trajectory in relative income. Observations which meet the requirements, are assigned with a 1 in the newly created variable. Observations which did not meet the requirements, are granted a 0. This created the new binary dependent variable. This binary variable is named 'gen_total'.

In the refinement of this created gentrification variable, two additional nuanced indicators are introduced. These variables, each distinct in their delineation of gentrification, afford a more comprehensive understanding of the income changes over time. These two extra binary variables are created to ensure a more comprehensive assessment of gentrification and cover any potential edge cases. Both will now be separately discussed. The initial binary gentrification variable exclusively considered whether the observation's income was below the median in the first period and above the median in the subsequent period. The introduction of the new variable incorporates an additional criterion by considering a minimum relative increase of 5 percentiles. Consequently, if an observation's income percentile is, for instance, at the 49th percentile in the initial period, it must exhibit a relative increase such that it attains at least the 54th percentile in the subsequent period. So, if an observation acquires both requirements, gentrification is experienced. This second binary dependent variable is called

[1] Income is measured differently in 2020 when compared to 2001 and 2011. However, as 2011 datasets contains both the 2001 and 2020 ways to measure income, 2011 could be used to see whether using either income variable results in different outcomes. The outcomes did not change.

‘gen_perc_5’. The second supplementary gentrification variable uses the same criterion as the first extra added dependent variable, but this added variable takes into account a 10% relative margin growth and is named ‘gen_perc_10’.

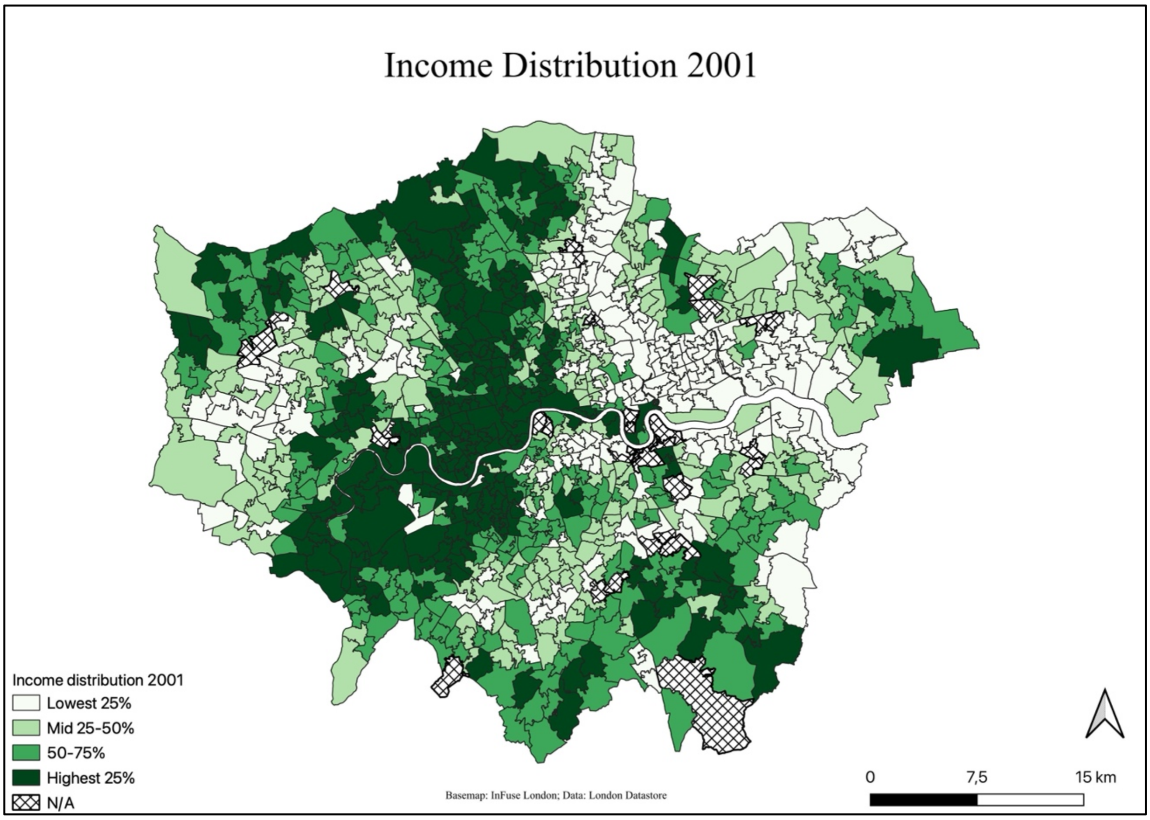

Figure 4, Income Distribution 2001, divided in four subgroups, based on relative income quartiles.

3.4 The Observation Pool for the Regression

This master thesis incorporates a dataset of 951 observations. The selection criteria for inclusion in the regression analysis are contingent upon the relative income status of the MSOAs. Specifically, only MSOAs falling below the median income in the first year of the period will be included in the model. This deliberate choice is predicated on the understanding that gentrification can only occur in economically disadvantaged neighbourhoods, as it is inherently incompatible with affluent areas (Behrens et al., 2024). Consequently, the analysis will exclusively focus on the observations which are situated below the median income threshold of 2001 or 2011. The total amount of MSOAs used in the analysis is therefore 477 in 2001 and 2011 (see Table 2). Furthermore, Table 2 visualises the differences between the three binary variables discussed above. As one can notice, the differences between the three binary variables are rather small, which shows the robustness of the calculation of the first dependent variable ‘gen_total’. Table 3 shows the amount of gentrified MSOAs, per borough of London. As the table shows, gentrification occurred in 20 of the in total 33 boroughs, with Brent, Hackney, Lambeth and Lewisham being the most represented boroughs. Figure 5 visualises the gentrified MSOAs, based on the variable ‘gen_total’.

Table 2, Observation Pool, overview of all gentrified MSOAs

All MSOAs	951
MSOAs with below median in either 2001 or 2011	477
Gentrification (gen_total)	
Yes	76
No	401
Gentrification (gen_perc_5)	
Yes	75
No	402
Gentrification (gen_perc_10)	
Yes	72
No	405

Table 3, Gentrification per Borough, based on assumptions of 'gen_total'.

Gentrification per Borough

Borough	Freq.	Percent	Cum.
Bexley	3	3.95	3.95
Brent	6	7.89	11.84
Bromley	3	3.95	15.79
Croydon	4	5.26	21.05
Ealing	3	3.95	25.00
Greenwich	2	2.63	27.63
Hackney	10	13.16	40.79
Haringey	5	6.58	47.37
Harrow	3	3.95	51.32
Hillingdon	2	2.63	53.95
Kingston upon Thames	1	1.32	55.26
Lambeth	6	7.89	63.16
Lewisham	7	9.21	72.37
Merton	1	1.32	73.68
Newham	3	3.95	77.63
Redbridge	2	2.63	80.26
Southwark	2	2.63	82.89
Sutton	4	5.26	88.16
Tower Hamlets	2	2.63	90.79
Waltham Forest	7	9.21	100.00
Total	76	100.00	

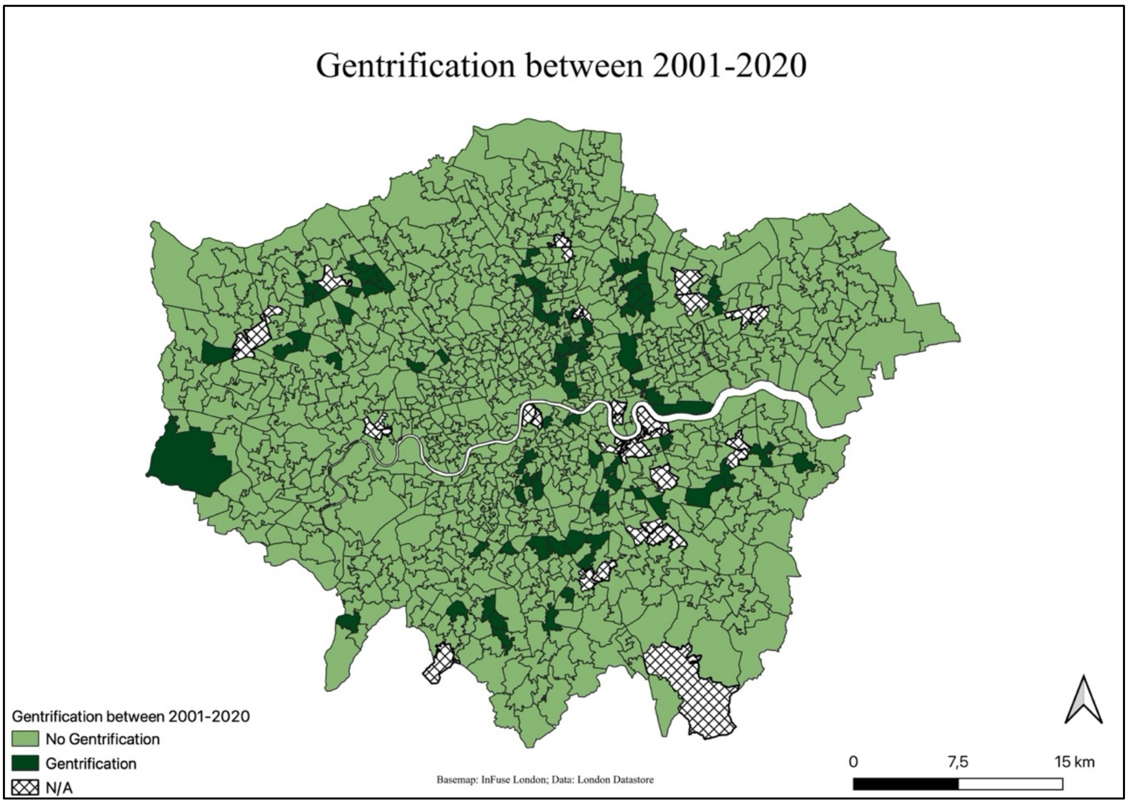

Figure 5, Gentrification between 2001-2020 of each MSOA in London, UK.

3.5 Descriptive Analysis

The following section will cover the descriptive analysis. In Table 4, descriptive statistics of all the observations in this dataset are shown, including the affluent neighbourhoods. Table 5 visualises the descriptive statistics for possible observations of 2001 or 2011. Overall, most of the variables show similar values for the variables between the two years. Looking at the descriptive statistics of the total observations and the observations where potential gentrification can occur, some interesting differences occur. First of all, the average amount of listed buildings for each MSOA is substantially higher in the total pool, with a mean of 20, compared to 8 of the possible observations of 2001. This initiates that wealthier neighbourhoods, overall accommodate more listed properties. Furthermore, there are substantial differences in the proximity of wealthy neighbourhoods. Potential MSOAs have a substantially lower average of 0.32 wealthy neighbouring areas, compared to 1.43 wealthy neighbouring areas in the whole observation pool. This can indicate that affluent neighbourhoods are already more clustered, which relates to the theory of 'endogenous gentrification' of Guerrieri et al. (2013).

Table 6 shows the descriptive statistics for the created spatial distribution, described before. Looking at the four distinct groups, some interesting differences appear. A noteworthy finding emerges as the average size of (private) gardens increases moving away further from the city centre. This can suggest that gentrification occurs in the MSOAs located in the outer rings, as existing literature suggests a positive relationship between green space and the likelihood of gentrification (Maantay & Maroko, 2018; Rigolon & Németh, 2020; Triguero-Mas et al., 2022). On the contrary, PTAL, representing the accessibility and quality of public transport, exhibits a declining trend as one moves away from the city centre. It suggests that the public transport in London is rather centre-focused. This finding can propound that gentrification will occur in the areas closer to the city centre, as the presence of public transport increases the likelihood of gentrification (Debrezion et al., 2011; Ostrensky et al., 2022). The proximity of wealthy neighbourhoods decreases as the distance to the city centre increases. This trend implies that areas further away may experience less gentrification, as gentrification often occurs in near-established affluent neighbourhoods (Guerrieri et al., 2013; Wilhelmsson et al., 2022)

Table 4, Descriptive statistics, total observation pool.

Total Observations (2001)				
Variable	Mean	Std. Dev.	Min	Max
Number of listed buildings	19.588	55.888	0	1024
% Buildings built before 1940	0.571	0.233	0	0.981
% addresses with (private) garden	0.817	0.139	0.072	0.996
Average size of (private) garden	175.854	114.004	27.3	1409.4
PTAL	3.694	1.452	2	8
Distance to city centre	11650.907	5717.858	0	26723
Spatial Distribution				
City Centre	4555	1811	0	7334
Urban	9169	1134	7336	11255
Sub-Urban	13556	1385	11282	16034
Periphery	19294	2395	16044	26723
Proximity of wealthy neighbourhoods in 2001	1.443	1.859	0	13
% of white people	0.607	0.193	0.061	0.962

Note: number of observations is 951. distance and size are in meters

Table 5, Potential Observation Pool, based on assumptions of 'gen_total'.

Possible Observations for Gentrification				
Variable	Mean	Std. Dev.	Min	Max
Number of listed buildings	8.036	13.028	0	139
% Buildings built before 1940	0.521	0.25	0	0.969
% addresses with (private) garden	0.818	0.13	0.208	0.996
Average size of (private) garden	148.849	62.041	52.2	663.7
PTAL	3.457	1.154	2	7
Proximity of wealthy neighbourhoods in 2001	0.323	0.805	0	5
Proximity of wealthy neighbourhoods in 2011	0.34	0.857	0	6
% of white people	0.504	0.189	0.061	0.935

Notes: *number of observations are 477; distance and size are in meters*

Table 6, Descriptive statistics, divided by spatial classification (2001); based on assumptions of 'gen_total'.

Possible Observations for Gentrification, divided by spatial distribution (2001)	*City Centre/Downtown*		*Urban*		*Sub-Urban*		*Periphery*	
Variable	Mean	*Std. Dev.*	Mean	*Std. Dev.*	Mean	*Std. Dev.*	Mean	*Std. Dev.*
Number of listed buildings	19.157	*19.895*	4.57	*6.534*	3.93	*6.577*	5.916	*9.324*
% Buildings built before 1940	0.378	*0.208*	0.637	*0.226*	0.596	*0.244*	0.45	*0.231*
% Addresses with (private) garden	0.705	*0.143*	0.85	*0.102*	0.843	*0.125*	0.861	*0.0845*
Average size of (private) garden	103.143	*30.921*	129.055	*53.345*	160.892	*49.042*	197.401	*65.629*
PTAL	5.065	*0.899*	3.57	*0.751*	2.86	*0.41*	2.253	*0.501*
Proximity of wealthy neighbourhoods in 2001	0.417	*0.908*	0.256	*0.69*	0.341	*0.906*	0.286	*0.691*
% of white people	0.477	*0.105*	0.447	*0.146*	0.486	*0.202*	0.605	*0.232*
Number of observations	*108*		*121*		*129*		*119*	

Note*: Size is in meters*

3.5 Logistic Regression

As this master thesis focuses on answering the research question: "To what extent are socio-economic and physical neighbourhood characteristics related to the probability of gentrification?", a logistic regression will be used. A logistic regression is a statistical model employed to test the relationship between one or more independent variables and a binary outcome (Peng et al., 2002). A logistic regression is especially useful when the dependent variable represents the probability of an event to occur, in this case, gentrification (1 =Yes, 0 = No). It estimates the log-odds of the event, based on the different coefficients of the independent variables. The analysis will consist of three different models. The first model will test to what extent socio-economic and physical neighbourhood characteristics increase the likelihood of gentrification. This model will make use of the following equation:

$$P(G) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_8 X_8)}} \tag{1}$$

In this equation, P stands for the probability of a neighbourhood to gentrify, G, which is the dependent binary variable that indicates whether gentrification occurred as specified in Section 3.3. For model 1, this is based on the first gentrification variable (gen_total, gen_perc_5 and gen_perc_10). β0 is the constant (the estimated mean of the dependent variable, if all independent variables are equal to 0. β1, β2, β3, β4, β5, β6, β7 and β8 are the logistic regression coefficients. This is the estimated change in the dependent variable for a one-unit change in each independent variable, only if all other variables remain constant. The independent variables are represented as X1, X2, X3, X4, X5, X6, X7 and X8. These independent variables are 'listed', 'pre1940', 'green', 'green_m2', 'prox_2001', 'ptal', 'dis_city_4' and 'white'.

Model 2 will make use of the same principles as Model 1, but the second model will test the four different spatial categories 'City centre', 'Urban', 'Sub-Urban' and 'Periphery'. This model will test if there are different predictors, based on differences in relative distance to the city centre. Therefore, the following equation will be used:

$$P(G) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_7 X_7)}} \tag{2}$$

This equation works the same as the one given for Model 1, the only difference is that it develops 4 different results, based on the level of urbanisation and excludes the categorical distance variable 'dis_city_4'.

Model 3 will test whether recently gentrified areas influence the likelihood of gentrification. Therefore, this model has a slightly different approach, compared to the previous two models. The time span of this model is 9 years, compared to the 19 years of Models 1 and 2. The dependent variable is based on whether an area experienced gentrification between 2011 and 2020 or not. This is calculated the same way as the first dependent variable 'gen_total', only with the first period of income of 2011 and the subsequent of 2020. This newly created binary variable is named 'gen2'. The added independent variable represents the amount of neighbouring MSOAs, which experienced gentrification in the period 2001-2011 (gen1_neigh). The first row will take the whole observation pool of 476 observations. Next, the model will test whether there are spatiotemporal differences between the four different levels of urbanisation. The equation of model 3 is shown below:

$$P(G) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_9 X_9)}} \quad (3)$$

P represents the probability of a neighbourhood to gentrify, the binary variable 'gen2'. β0 is the constant and β1, β2, β3, β4, β5, β6, β7 and β8 are the logistic regression coefficients. The independent variables are presented as X1, X2, X3, X4, X5, X6, X7, X8 and X9. These represent 'listed', 'pre1940', 'green', 'green_m2', 'prox_2011', 'ptal', 'dis_city_4', 'white' and 'gen1_neigh'.

3.6 Diagnostics Tests

The last part of the methodology section covers the diagnostics tests for the logistic regression model. Diagnostic tests are an important part of determining the model's performance and identifying potential issues. Firstly, a multicollinearity test is performed to test whether independent variables are highly correlated to each other. The correlation matrix is presented in Appendix 1.1. The matrix shows that none of the variables is highly correlated to each other (0.8 to 1 or -0.8 to -1), although the correlation between PTAL and the categorical distance to the city centre (dis_city_4) are strongly correlated to each other (-0.79). Despite the strong negative correlation, both variables remain suitable for analysis. Each variable may independently contribute valuable information about the process of gentrification.

A Hosmer-Lemeshow test was performed on model 1 to test whether the model provides a good representation between the predictor variables and the outcome variable. The outcome can be found in Appendix 1.2. With a p-value of 0.823, it suggests that there is no significant lack of fit and the model fits the data well.

A link test in Stata is performed to test whether there are specification errors among the variables. A table of the results of this test for Model 1 is located in Appendix 1.3. The test shows that the squared predicted values ('_hatsq') are not significant. For all the other models, the link test showed a non-significant squared predicted value too. This implies that the inclusion of squared terms in the models does not significantly increase the explanatory power of the model. These results show that the relationship between the predictors and log odds of the variable can be captured in linear terms. It also underscores the notion that the chosen independent variables can effectively capture the nuances of the observed dependent binary variable, gentrification.

4. Results

4.1 Logistic Regression Model

The next section of this master thesis focuses on the results of the multiple logistic regressions. Table 7 shows the results of model 1. In total, 474 observations are included in the analysis. three observations are lost in the analysis because the failure of these observations is perfectly predicting for PTAL score 6a. The table shows two columns for each model with the three different binary dependent variables: 'gen_total', 'gen_perc_5' and 'gen_perc_10'. The first column shows the log odds of the coefficient. For logistic regression, the odds thus, the likelihood of an event happening, in this case gentrification, is a binary outcome. Therefore, the outcome is a log-odds. To be able to interpret the coefficients more straightforwardly, an extra odds-ratio column is added. This is calculated by exponentiating the coefficients to get the new value. The formula used is shown below:

$$Odds\ Ratio = e^{\beta_x} \tag{4}$$

The Pseudo R2 of the first model is 22.65%. This implies that 22.65% of the variance in the dependent variable can be explained by the set of independent variables. As column 1 in Table 7 shows, the variable is containing information about the number of listed buildings significant, with a p-value of less than 0.01. This rejects the null hypothesis that there is no relationship between the number of listed buildings and the odds of gentrification, as measured by 'gen_total'. The log odd of this variable is 0,50291 and the odds ratio is 1.05157. This variable has a standard deviation of 13.028. This implies that keeping all other variables constant, the odds of an MSOA gentrifying will increase by 13.7% if one standard deviation is added to the number of listed buildings. The ratio of pre-1940 and post-1940 build structures show a significant relationship with the probability of gentrification ($P < 0.01$). This variable has a log odd of 5.0111 and an odd ratio of 150.077. While the odds ratio appears to be relatively elevated, it is imperative to recognize that the variable is expressed as transformed percentages, constrained within the range of 0 to 1. Notably, this transformation denotes values on a scale where 0 signifies 0%, and 1 denotes 100%. Therefore, the seemingly high odds ratio should be interpreted within the context of this percentage transformation, where the upper limit corresponds to the complete representation of the variable. A one standard deviation increase of this variable, which is 0.25 results in a 37.52% increase of the odds of gentrification. Although the percentage of addresses with a (private) garden tends to be not significant, the size of the (private) garden has a statistically significant relationship with the odds of gentrification ($P < 0.05$). Keeping all other variables constant, the odds of gentrification increase, by adding one standard deviation, with 62.51%. Model 1 shows no statistically significant relationship between the proximity of affluent neighbourhoods and the odds of gentrification. Of all the different levels of PTAL, only a PTAL score of 2 tends to be statistically significant, with a p-value of less than 0.05. For this analysis, the reference group for PTAL is 1b, as this is the lowest value in the used observation pool. Keeping a variable stable, moving from a PTAL score of 1b to 2, results in a 224% increase of the odds of gentrification. For the spatial distributions, the category ‘City centre/Downtown’ is the reference category. The choice of taking this category as a reference is based on the principle that this region is located the most central. While keeping all other factors constant, the log odds associated with gentrification exhibit a marked decline in the 'Urban,' 'Sub-Urban,' and 'Periphery' categories when contrasted with the baseline category ‘City Centre’. Specifically, within the 'Urban' category, there is a decrease of -1.14501, indicating an approximately 68.18% reduction in the odds of gentrification. The 'Sub-Urban' category experiences a more substantial decline of -1.86248, corresponding to an approximate 84.47% decrease in gentrification odds. Furthermore, the 'Periphery' category demonstrates the most pronounced decrease, with a decline of -2.18581, equating to an approximately 88.76% reduction in the odds of gentrification. The variable containing the ratio of white people in the MSOA shows a statistically significant relationship with the dependent variable ($P < 0.01$). A one standard deviation increase, keeping all other variables constant, results in an 8.5% increase in the odds of gentrification.

The second and third dependent variables, respectively ‘gen_perc_5’ and ‘gen_perc_10’ are added to the model to test its robustness. As described in the methodology, these two variables exclude so-called edge cases. Looking at the Pseudo R2 of the second variable, it slightly increases from 22.65% to 23.25%. This implies that the second model is slightly better at explaining the dependent variables by the set of independent variables. On the other hand, the third model is slightly less able to perform this, because of a lower Pseudo R2 (22.28%). The Bayesian Information Criterion (BIC) decreases in the second and third models, compared to the original model. A diminishing BIC suggests an improvement in the balance and complexity of the model. This trend indicates that the models provide a better trade-off between explanatory power and simplicity. The coefficients across the three models exhibit minimal variation, implying a robust model.

Appendix 2.1 examines whether there is heterogenicity in the factors that are associated with gentrification for MSOAs that gentrified between 2001-2011 and those that gentrified between 2011-2020. To do this, separate logistic regressions for the periods 2001-2011 and 2011-2020 are created. Both logistic regressions are based on the same principles as model 1. ‘Gen1’ and ‘gen2’ are calculated on the same principles as ‘gen_total’ and the robustness check is performed on the same principles too. In both periods 2001-2011 and 2011-2020, the variables representing the number of listed buildings and the ratio of pre-1940 and post-1940 build structures are both statistically significant and both exhibit minimal variation compared to model 1. Interesting to notice is that the proximity of wealthy neighbourhoods has a positive significant relationship with the likelihood of gentrification in the period 2001-2011.

4.2 Differences in Spatial Distribution

To test whether there is heterogeneity in the factors that are associated with gentrification for MSOAs that belong to different groups in terms of spatial distribution, a new model (Model 2) is created. This model is based on the dependent variable ‘gen_total’ and divides the data sample into four distinctive groups, based on the distance from the city centre which is discussed in Section 3.2. Table 8 shows the results of the logistic regression models. The number of observations differs between the distinctive degrees of urbanisation. This can be explained, by due that for all four groups, only observations which could experience gentrification (income below median in 2001 or 2011), are included. The first thing to notice is that the observations are the lowest in the city centre (99) and are almost equal in the other three geographically distinctive groups (121-124). In the city centre, the presence of listed buildings significantly contributes to the odds of gentrification (Odds ratio = 1.07632, $p < 0.01$). This underscores the role of historical architecture as a positive trigger of gentrification in city centres. Suburban areas and the periphery exhibit a dynamic where a higher percentage of buildings constructed before 1940 strongly correlates with elevated odds of gentrification (Odds ratio = 105.863, $p < 0.05$ & odds ratio = 2342.403, $p < 0.01$). In sub-urban MSOAs, the size of (private) gardens seems to have a positive influence on the likelihood of gentrification, with a statistically significant odds ratio of 1.01746. In the geographical regions 'City centre', 'Urban' and 'Periphery', the percentage of white people is associated with a statistically significant increase in the probability of gentrification. The 'City centre' stands out with the highest odd ratio. This demographic factor shows the importance of demographic distribution as a factor in the likelihood of gentrification. The model representing the city centre experiences the highest Pseudo R2, with a value of 39.29%. This suggests that this model demonstrates the highest explanatory power among the four.

4.3 Proximity of Recently Gentrified Neighbourhoods

This final section examines spatiotemporal patterns in gentrification. This is accomplished by introducing the effect of recently gentrified neighbourhoods on the likelihood of gentrification. this model, as described in the methodology (Section 3.5), takes into account the number of proximate areas which gentrified during the period 2001-2011. The model is visualised in Table 9. This is calculated on the three dependent variables 'gen1', 'gen_perc_5_2011' and 'gen_perc_10_2011'. The last two binary variables are calculated on the same principles as 'gen_perc_5' and gen_perc_10', only these two variables take into account the period 2011-2020. All the three regressions show no statistically significant relationship with the proximity of recently gentrified neighbourhoods and the likelihood of gentrification.

The next step in researching the effect of recently gentrified neighbourhoods on the probability of gentrification, is by calculating the coefficients for the four different degrees of urbanisation. This gives insights if areas with different levels of urbanisation potentially have other triggers for gentrification. Table 10 shows the regression model of the four different levels of urbanisation. In the 'Urban' areas, the proximity of both wealthy and recently gentrified neighbourhoods is statistically significant. The proximity of wealthy neighbourhoods has a positive relationship with the odds of gentrification.

Contrary, in the 'Urban' area, the proximity of recently gentrified neighbourhoods exhibits a negative relationship with the odds of gentrification. Specially, adding one extra neighbouring area, which experienced gentrification in the period 2001-2011, results in a decrease of 89% in of the odds of gentrification.

Table 7, Logistic Regression Model 1, based on the period 2001-2020

Variables	*gen_total*		*gen_perc_5*		*gen_perc_10*	
	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio
Number of Listed Buildings	0.50291***	1.05157***	0.0509***	1.05222***	0.04459***	1.0456***
Percentage of buildings built before 1940	5.011149***	150.077***	5.16118***	174.3696***	4.91002***	135.6424***
Percentage of addresses with (private) garden	-1.0753	0.3412	-1.44835	0.23496	-0.63005	0.52257
Average size of (private) garden, in meters	0.007535**	1.00756**	0.00721***	1.00723***	0.00576**	1.00576**
Proximity of wealthy neighbourhoods in 2001	0.05095	1.05227	-0.016985	0.98316	0.06808	1.07045
PTAL *reference 1b*						
2	1.17552**	3.23983**	1.18865**	3.28267**	1.02646*	2.79116*
3	0.9809	2.66686	1.01169	2.75026	0.84727	2.33326
4	0.5327	1.70353	0.51767	1.67811	0.47247	1.60296
5	0.68258	1.97901	0.69424	2.00218	0.6663	1.94702
Spatial distribution *reference City Centre*						
Urban	-1.14501**	0.31822**	-1.12042**	0.32614**	-1.14263**	0.31898**
Sub-Urban	-1.86248***	0.155287***	-1.92111***	0.14644***	-2.19264***	0.11162***
Periphery	-2.18581***	0.112387***	-2.14043***	0.1176***	-2.11519***	0.12061***
Percentage of white people	3.80657***	44.9956***	4.13045***	62.20607***	4.00472***	54.8565***
Constant	-6.95211***	0.00096***	-6.87873***	0.00103***	-6.95447***	0.00095***
Observations	474		474		474	
Pseudo R2	22.65%		23.25%		22.28%	
Bayesian crit. (BIC)	409.088		403.614		400.101	

Notes: *****p<,01, **p<,05, *p<,1*

.

Table 8, Model 2, Differences by spatial distribution, based on dependent variable 'gen_total'. .

Variables (*gen_total*)	*City centre*		*Urban*		*Sub-Urban*		*Periphery*	
	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio
Number of Listed Buildings	0.07366***	1.07632***	0.06333	106.537	0.078968	108.217	0.034	103.438
Percentage of buildings build before 1940	339.758	2.989.166	282.064	1.678.757	4.66215**	105.863**	7.75893***	2342.403***
Percentage of addresses with (private) garden	-0.14476	0.86523	417.375	6.495.854	-171.654	0.17969	-0.61524	0.54051
Average size of (private) garden. in meters	0.01534	101.546	-0.00013	0.99986	0.017305**	1.01746**	0.00626	100.628
Proximity of wealthy neighbourhoods in 2001	-0.09124	0.9128	0.42645	153.181	-0.24737	0.78085	0.48777	162.869
PTAL *reference 1b*								
2			0.29532	134.355	0.17806	11.949	184.228	631.091
3	148.357	440.865	-0.0705	0.93193				
4	0.02123	102.145						
5			133.255	379.069				
Percentage of white people	10.18118**	26401.6**	8.84759***	6957.602***	0.27731	131.956	4.2546*	70.42893*
Constant	-11.04535***	0.00002***	-11.91217***	0.00001***	-7.01296**	0.0009**	-11.7507**	-0.00001***
Observations		99		122		124		121
Pseudo R2	39.29%		23.19%		17.65%		27.22%	
Bayesian crit. (BIC)		111.791		136.486		120.188		88.593

*Notes: ***p<.01. **p<.05. *p<.1*

Table 9, Logistic Regression Model 3, effect of recently gentrified neighbourhoods

Variables	*gen2*		*gen_perc_5_2011*		*gen_perc_10_2011*	
	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio
Number of Listed Buildings	0.05844***	1.06018***	0.05454***	1.05606***	0.055216***	1.05677***
Percentage of buildings built before 1940	4.90346***	134.7557***	4.79805***	121.2732***	4.54799***	94.44222***
Percentage of addresses with (private) garden	-0.34442	0.70863	-0.17468	0.83972	-0.07705	0.92584
Average size of (private) garden. in meters	0.0046*	1.00461*	0.00421	100.422	0.00404	100.405
Proximity of wealthy neighbourhoods in 2011	0.09373	109.827	0.12979	113.859	0.14065	115.102
Proximity of recently gentrified neighbourhoods between 2001-2011	-0.29078	0.74768	-0.3971	0.67227	-0.38251	0.68215
PTAL *reference 1b*						
2	1.63917***	5.15092***	1.33422**	3.79703**	1.50152**	4.4885**
3	1.71211**	5.54067**	1.52578**	4.59874**	1.66838**	5.30354**
4	0.69751	200.874	0.72565	206.608	0.85245	234.538
5	119.584	330.634	0.859049	2.360.914	0.95468	259.783
Spatial distribution *reference city centre*						
Urban	-0.181	0.41941	-0.82672	0.43748	-0.78864	0.45446
Sub-Urban	-1.32166*	0.26669*	-1.31962*	0.26724*	-1.27312*	0.27996*
Periphery	-0.80261	0.44816	-0.70052	0.49633	-0.72045	0.48654
Percentage of white people	3.15638**	23.48539**	2.56444***	12.99335***	2.38153**	10.82148**
Constant	-7.72834***	0.00044***	-7.31205***	0.00067***	-7.30241***	0.00067***
Observations	472		472		472	
Pseudo R2	17.37%		15.67%		15.42%	
Bayesian crit. (BIC)	407.987		398.786		397.095	

*Notes: ***p<.01. **p<.05. *p<.1*

Table 10, Regression Model 3, divided by spatial distribution, including spatiotemporal patterns.

Variables (*gen2*)	*City centre*		*Urban*		*Sub-Urban*		*Periphery*	
	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio
Number of Listed Buildings	0.06946**	1.07193**	0.10478**	1.11047**	0.12782**	1.13635**	0.10974**	1.11599**
Percentage of buildings build before 1940	141.059	409.834	253.607	1.262.998	4.79979**	121.4843**	8.65248***	5724.4328***
Percentage of addresses with (private) garden	0.84052	231.756	642.766	6.187.223	-0.63344	0.53076	657.766	7.188.541
Average size of (private) garden. in meters	0.00908	100.912	0.00479	10.048	0.01355*	101.364	-0.0012	0.9988
Proximity of wealthy neighbourhoods in 2011	0.15058	116.251	0.70879*	2.03153*	-135.245	0.25861	0.34572	141.301
Proximity of recently gentrified neighbourhoods between 2001-2011	-0.23592	0.78984	-2.20616**	0.11012**	0.91643	250.034	136.649	392.157
PTAL *reference 1b*								
2			0.34531	141.243	266.371	527.897	2.28258**	9.80198**
3	1.94661*	7.0049*	0.31872	137.537				
4	0.091181	109.547						
5			272.303	1.522.645				
Percentage of white people	9.35346*	11538.63*	10.07127***	23653.54***	-0.16405	0.8487	4.94384**	140.308**
Constant	-10.2133***	0.00014***	15.13937***	2.66e-07***	-8.81331**	0.00015**	-17.85972***	1.75e-08***
Observations	83		120		129		132	
Pseudo R2	30.05%		27.7%		19.43%		33.72%	
Bayesian crit. (BIC)	96.877		130.848		121.668		108.573	

*Notes: ***p<.01. **p<.05. *p<.1*

5. Discussion

Section 4 covered the results of the multiple logistic regressions. The following section will link the results to the existing literature and explain the outcome. This section is structured as follows: first all physical and socio-economic neighbourhood characteristics which have influence on the likelihood of gentrification will be discussed. This will be followed by exploring any heterogeneity between the four levels of urbanisation and spatiotemporal patterns of gentrification.

The results of the models overall suggest a positive relationship between the presence of historic buildings and the probability of gentrification. this result is supported by existing literature (e.g., Helms, 2013; Been et al., 2016). The presence of listed buildings results in a significant increase in the probability of gentrification. This last claim is supported by the research of Rosenthal and Ross (2015). London is a historical city, which therefore serves as an old and aesthetic location. Old historic city centres differ from modern city centres with the presence of older buildings with more heritage and aesthetic structures (Rosenthal and Ross, 2015). They provide a unique amenity, attracting households with elevated income levels. This appreciation of older and historic buildings has a higher chance of being renovated (Helms, 2003) and eventually results in higher house prices and living costs. This can eventually result in local displacement of the original residents. This last claim also is supported by the significant positive relationship of percentage buildings structured before 1940 and the likelihood of gentrification.

Model 1 suggests that the percentage of addresses with a garden tends to be not significantly related to the likelihood of gentrification. In the sub-sample of gentrification in 2011-2020 (shown in Table 10), a significant positive effect of the percentage of gardens is suggested. This connects to the results of Maantay & Maroko (2018). Their results suggest that the proximity of community gardens increases the income per capita, which can initiate gentrification. This implicates environmental justice, as existing lower-income residents have an increased chance of displacement in an area with a higher proximity to community gardens.

The size of gardens and the likelihood of gentrification have a positive relationship. According to the results of the regression. This contradicts existing research about the size of community gardens and the likelihood of gentrification (Hawes et al., 2022). One reason for the difference in results is the

function of community gardens and private gardens. Where the function of community gardens is mostly focused on multiple households and residents, private gardens mainly serve the owner of the property and therefore it is important to view the results with nuance. Access to bigger gardens is associated with better subjective well-being and higher self-rated health (Poortinga et al., 2021). Affluent people can afford more spending to increase their subjective well-being, although this claim is to be made with caution (Stevenson & Wolfers, 2013).

In Model 3, a higher PTAL score results in a significant positive effect on the likelihood of gentrification. By dividing the PTAL into categories with 1b as a reference, 2 out of the total 7 levels of PTAL show a significant positive effect on gentrification. As the quality accessibility of public transport improves, it also increases the probability of gentrification. The results of the logistic regression models support the claims of existing literature (Lian & Yang, 2019; Ostrensky et al., 2021). The results do not show a significant relationship between the higher levels of quality and accessibility of public transport and gentrification. A possible explanation is that a higher quality of public transport overall is more located in the already wealthy areas, and therefore gentrification cannot take place. Important to notice is the relatively strong negative correlation between distance to the city centre and quality and access to public transport. However, it is important to approach the findings with a nuanced interpretation. While the PTAL score provides an overall assessment of public transport accessibility and quality, it does not specifically show the impact of a new public transport hub, as is explained in the existing literature. Therefore, the effects of the higher accessibility and quality of public transport on the likelihood of gentrification should be interpreted with nuance.

Across all three models, there seems to be a positive relation between the likelihood of gentrification and racial composition. The size of the effect fluctuates between the models, the positive effect is consentient. As Waalks and Maaranen (2008) suggest, neighbourhoods with an already high percentage of white people are more vulnerable to gentrification.

The analysis of spatiotemporal patterns in gentrification reveals compelling insights into the dynamics of urban transformation. Notably, the logistic regression results of the period 2001-2011 (see Table 10) underscore the influence of the proximity of affluent neighbourhoods on gentrification, aligning with the concept of 'endogenous gentrification' of Guerrieri et al. (2013). This finding suggests a notable tendency for affluent residents to seek proximity to other affluent individuals, forming larger clusters within a city.

While Wilhelmsson et al. (2021) propose positive spillover effects from recently gentrified areas, the findings of the present study reveal a conflicting narrative, particularly in the 'Urban' areas, where a significant negative relationship appears between proximity of recently gentrified areas and the likelihood of gentrification. This unexpected observation prompts a closer examination of related outcomes by the research of Christafore and Leguizamon (2019), whose research also highlights a negative relationship between the proximity of recently neighbouring areas and the probability of gentrification. This research suggests that low-income households, priced out of gentrified areas, seek housing in neighbouring low-income areas.

6. Conclusions

6.1 Main Findings

This research examined gentrification patterns in London, UK, and discerned the physical and socio-economic neighbourhood characteristics contributing to the likelihood of this phenomenon. By adding sub-research questions guided by existing literature, this study explored the impact of distinct key physical features on gentrification including historic buildings, public transport quality, green spaces, and the percentage of white residents. Furthermore, shows existing literature a positive relationship between the likelihood of gentrification and the proximity of wealthy surrounding areas. By making use of a quantitative approach with British governmental data, gentrification was measured by a relative income increase over time spanning from 9 to 19 years.

Multiple logistic regressions were applied to systematically evaluate the impact of the identified socioeconomic and physical neighbourhood characteristics on the likelihood of gentrification. The analysis revealed positive associations between gentrification and the presence of historic buildings limited public transport improvements, average size of private gardens, percentage of white residents and the proximity of affluent areas. Contrarily, an increase in distance from the city is associated with a negative impact on the likelihood of gentrification. It is noteworthy that, in the 'Urban' areas, proximity to wealthy neighbourhoods increased the probability of gentrification, whereas no statistically significant effects were discerned in the remaining three spatial subgroups.

The third research question of this master thesis focussed on understanding the impact of the proximity of recently gentrified areas on the likelihood of further gentrification. Remarkably, a discernible negative relationship in this context was observed solely within the 'Urban' area, while the logistic regression models for the other three areas showed no statistically significant result. In conclusion, these findings furnish nuanced insights into the dynamics of gentrification in London. This master thesis underscores the intricate interplay of neighbourhood characteristics that impact gentrification and provides a foundation for future investigations aimed at providing more insights into the phenomenon of gentrification.

6.2 Limitations and Future Research Recommendations

This study's findings should be interpreted with regard to several limitations. The dataset used is characterized by a relatively modest number of variables, posing challenges for establishing intricate patterns and constraining the comprehensive considerations of all potential influencing factors. This study uses nine variables to determine neighbourhood characteristics, adding more variables increases the strength of the model is can possibly be able to capture more patterns which are related to gentrification.

Additionally, the static nature of the data, where most of the variables remain stable over time, affords a snapshot rather than a dynamic image of the evolving urban landscape, potentially limiting a total understanding of gentrification's temporal dynamics. Gentrification is a dynamic process which can change over time; therefore, the static data of this study may impose constraints on providing a dynamic portrayal of the urban environment. Concerns about the measurement and definition of variables, particularly the operationalisation of gentrification, introduce potential bias in the results. Taking another definition or measurement of gentrification could possibly result in different outcomes. This study focussed on a specific spatial scale, which may hinder the generalizability of the findings and the temporal scope, concentrated on specific intervals, may overlook crucial periods of change. To address these limitations, future research should focus on implementing more dynamic data sources.

Furthermore, should future research add a larger and wider set of variables, in terms of numbers and diversity. Additionally, future research should explore the execution of comparative spatial analyses to enhance a better understanding of spatial variations in gentrification. The integration of qualitative methodologies can provide valuable insights into the nuanced factors influencing gentrification. Further research into policy implications is crucial for developing strategies that can address and tackle potentially vulnerable areas.

7. References

Atkinson R and Wulff M (2009), *Gentrification and displacement: a review of approaches and findings in the literature*. AHURI Positioning Paper No. 115, Australian Housing and Urban Research Institute Limited, Melbourne

Atkinson, R. (2004), "The evidence on the impact of gentrification: new lessons for the urban renaissance?", European Journal of Housing Policy, Vol. 4 No. 1, pp. 107-131.

Bardaka, E., Delgado, M. S., & Florax, R. J. (2018). Causal identification of transit-induced gentrification and spatial spillover effects: The case of the Denver light rail. Journal of Transport Geography, 71, 15-31.

BBC. (2015). Police use CS spray on Brixton gentrification protesters. Retrieved on January 28, 2024 from: https://www.bbc.com/news/uk-england-london-32463425

Been, V., Ellen, I. G., Gedal, M., Glaeser, E., & McCabe, B. J. (2016). Preserving history or restricting development? The heterogeneous effects of historic districts on local housing markets in New York City. *Journal of Urban Economics*, *92*, 16-30.

Behrens, K., Boualam, B., Martin, J., & Mayneris, F. (2024). Gentrification and pioneer businesses. *Review of Economics and Statistics*, 1-14.

Bockarjova, M., Botzen, W. J. W., Van Schie, M. H., & Koetse, M. J. (2020). Property price effects of green interventions in cities: A meta-analysis and implications for gentrification. *Environmental science & policy*, *112*, 293-304.

Chapple, K., & Zuk, M. (2016). Forewarned: The use of neighborhood early warning systems for gentrification and displacement. *Cityscape*, *18*(3), 109-130.

Chapple, K. (2017). Income inequality and urban displacement: the new gentrification. In *New Labor Forum* (Vol. 26, No. 1, pp. 84-93). Sage CA: Los Angeles, CA: SAGE Publications.

Christafore, D., & Leguizamon, S. (2019). Neighbourhood inequality spillover effects of gentrification. *Papers in Regional Science*, *98*(3), 1469-1484.

Debrezion, G., Pels, E., & Rietveld, P. (2011). The impact of rail transport on real estate prices: an empirical analysis of the Dutch housing market. *Urban studies*, *48*(5), 997-1015.

Diamond, R. (2016). The determinants and welfare implications of US workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

Finio, N. (2022). Measurement and definition of gentrification in urban studies and planning. *Journal of Planning Literature*, *37*(2), 249-264.

Glass, R. (1964). *London: Aspects of change*. London, McGibben & Kee

Grevstad-Nordbrock, T., & Vojnovic, I. (2019). Heritage-fueled gentrification: A cautionary tale from Chicago. *Journal of Cultural Heritage*, *38*, 261-270.

Guerrieri, V., Hartley, D., & Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, *100*, 45-60.

Hamnett, C. (2003). Gentrification and the middle-class remaking of inner London, 1961-2001. *Urban studies*, *40*(12), 2401-2426.

Hawes, J. K., Gounaridis, D., & Newell, J. P. (2022). Does urban agriculture lead to gentrification?. *Landscape and Urban Planning*, *225*, 104447.

Hedin, K., Clark, E., Lundholm, E., & Malmberg, G. (2012). Neoliberalization of housing in Sweden: Gentrification, filtering, and social polarization. *Annals of the association of American geographers*, *102*(2), 443-463

Helms, A. C. (2003). Understanding gentrification: an empirical analysis of the determinants of urban housing renovation. *Journal of urban economics*, *54*(3), 474-498.

Higgins, P., Campanera, J., & Nobajas, A. (2014). Quality of life and spatial inequality in London. *European Urban and Regional Studies*, *21*(1), 42-59.

Historic England, (2024) Listed Buildings. Retrieved on January 28, 2024 from: https://historicengland.org.uk/listing/what-is-designation/listed-buildings/

Kim, S. K., & Wu, L. (2022). Do the characteristics of new green space contribute to gentrification?. *Urban Studies*, *59*(2), 360-380.

Lang, M. H. (1982). Gentrification amid urban decline: Strategies for America's older cities. *(No Title)*.

Lin, J. J., & Yang, S. H. (2019). Proximity to metro stations and commercial gentrification. *Transport Policy*, *77*, 79-89.

Lin, L., Di, L., Zhang, C., Guo, L., & Di, Y. (2021). Remote Sensing of Urban Poverty and Gentrification. *Remote Sensing*, *13*(20), 4022.

Lyons, M. (1995). Inmigration, gentrification and neighbourhood change in London: can they inform theoretical debate. *Gentrification in Europe. London: Wiley*.

Maantay, J. A., & Maroko, A. R. (2018). Brownfields to greenfields: Environmental justice versus environmental gentrification. *International journal of environmental research and public health*, *15*(10), 2233.

Nwanna, C. (2012). Gentrification in Lagos state: Challenges and prospects.

Oluwalana, A. (2022). *The East London neighbourhood so gentrified actors and coffee shops are 'pushing out' the working class.* Retrieved on January 14, 2024 from https://www.mylondon.news/news/east-london-news/east-london-neighbourhood-gentrified-actors-25590727. My London News.

ONS. (2024). Census 2021 geographies. Retrieved on January 28, 2024 from:https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies

Ostrensky, V. P., Porsse, A. A., & da Frota, L. M. (2022). Public transport and gentrification. Evidence from São Paulo metro new stations. *Regional Science Policy & Practice*, *14*(6), 254-269.

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The journal of educational research*, *96*(1), 3-14

Poortinga, W., Bird, N., Hallingberg, B., Phillips, R., & Williams, D. (2021). The role of perceived public and private green space in subjective health and wellbeing during and after the first peak of the COVID-19 outbreak. *Landscape and Urban Planning*, *211*, 104092.

Rigolon A & Nemeth J (2019), *Towards a socioecological model of gentrification: How people, place and policy shape neighbourhood change*. Journal of Urban Affairs, 41(7), pp. 887-909,

Rigolon, A., & Németh, J. (2020). Green gentrification or 'just green enough': Do park location, size and function affect whether a place gentrifies or not?. *Urban Studies*, *57*(2), 402-420.

Rosenthal, S. S., & Ross, S. L. (2015). Change and persistence in the economic status of neighborhoods and cities. *Handbook of regional and urban economics*, *5*, 1047-1120.

Rouwendal, J., Keus, A., & Dekkers, J. (2018). Gentrification through the sale of rental housing? Evidence from Amsterdam. *Journal of Housing Economics*, *42*, 30-43.

Slater, T. (2011). Gentrification of the city. *The New Blackwell companion to the city*, 571-585.

Smith, N. (1979). Toward a theory of gentrification a back to the city movement by capital, not people. *Journal of the American planning association*, *45*(4), 538-548.

Smith, N. (1982). Gentrification and uneven development. *Economic geography*, *58*(2), 139-155.

Stevenson, B., & Wolfers, J. (2013). Subjective well-being and income: Is there any evidence of satiation?. *American Economic Review*, *103*(3), 598-604.

Thackway, W., Ng, M., Lee, C. L., & Pettit, C. (2023). Building a predictive machine learning model of gentrification in Sydney. *Cities*, *134*, 104192.

Transport for London. (2015). Assessing transport connectivity in London. *Mayor of London.*

Triguero-Mas, M., Anguelovski, I., Connolly, J. J., Martin, N., Matheney, A., Cole, H. V., ... & Baró, F. (2022). Exploring green gentrification in 28 global North cities: the role of urban parks and other types of greenspaces. *Environmental research letters*, *17*(10), 104035.

UK Data Service. (2024) InFuse. Retrieved on January 28, 2024 from: https://infuse.ukdataservice.ac.uk

Walks, R. A., & Maaranen, R. (2008). Gentrification, social mix, and social polarization: Testing the linkages in large Canadian cities. *Urban Geography*, *29*(4), 293-326.

Wilhelmsson, M., Ismail, M., & Warsame, A. (2022). Gentrification effects on housing prices in neighbouring areas. *International Journal of Housing Markets and Analysis*, *15*(4), 910-929.

Zuk, M., Bierbaum, A. H., Chapple, K., Gorska, K., Loukaitou-Sideris, A., Ong, P., & Thomas, T. (2015). Gentrification, displacement and the role of public investment: a literature review. In *Federal Reserve Bank of San Francisco*(Vol. 79).

Zukin, S., Lash, S., & Friedman, J. (1992). *Postmodern urban landscapes: mapping culture and power*.

Appendix

Appendix 1.1

Tabulation of gen_total gen_perc_5

	gen_perc_5		
gentrification between 2001-2020	0	1	Total
0	401	0	401
1	1	75	76
Total	402	75	477

Tabulation of gen_perc_5 gen_perc_10

	gen_perc_10		
gen_perc_5	0	1	Total
0	402	0	402
1	3	72	75
Total	405	72	477

Appendix 1.2

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) listed	1.000							
(2) pre1940	-0.285	1.000						
(3) green	-0.515	0.515	1.000					
(4) green_m2	-0.127	0.033	0.195	1.000				
(5) prox_2001	0.178	-0.171	-0.263	0.077	1.000			
(6) ptal	0.379	-0.204	-0.438	-0.471	0.090	1.000		
(7) dis_city	-0.333	0.066	0.414	0.561	-0.058	-0.789	1.000	
(8) White	0.072	-0.165	0.125	0.394	0.108	-0.266	0.263	1.000

Appendix 1.3

note: obs collapsed on 10 quantiles of estimated probabilities.
Goodness-of-fit test after logistic model
Variable: gen_total
Number of observations = 474
Number of groups = 10
Hosmer,ÄìLemeshow chi2(8) = 4.36
Prob > chi2 = 0.8236

Appendix 1.4

Iteration 0: log likelihood = -208.66864
Iteration 1: log likelihood = -165.57053
Iteration 2: log likelihood = -161.002
Iteration 3: log likelihood = -160.66607
Iteration 4: log likelihood = -160.66231
Iteration 5: log likelihood = -160.66231
Logistic regression

Number of obs. = 474
LR chi2(2) = 96.01
Prob > chi2 = 0.0000
Log likelihood = -160.66231 Pseudo R2 = 0.2301

gen_total	Coefficient	Std. err.	z	P>z	[95% conf. interval]
_hat	1.130	0.246	4.590	0.000	0.648	1.613
_hatsq	0.046	0.070	0.650	0.517	-0.092	0.184
_cons	0.033	0.225	0.150	0.882	-0.408	0.475

Appendix 2.1

Regression Model 1 (2001-2011)

Variables	*gen1*		*gen_perc_5_2001*		*gen_perc_10_2011*	
	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio
Number of Listed Buildings	0,03553**	1,03617**	0,03875***	1,03951***	0,04062*	1,04145*
Percentage of buildings build before 1940	4,27313***	71,74611***	3,59767**	36,51286***	1,61298	5,01774
Percentage of addresses with (private) garden	0,0759	1,07886	0,95895	2,60895	1,46433	4,32465
Average size of (private) garden, in meters	0,0154***	1,01552***	0,014444***	1,01455***	0,008151	1,00818
Proximity of wealthy neighbourhoods in 2001	0,72771***	2,07034***	0,65293***	1,92116***	0,55183*	1,73642*
PTAL *reference 1b*						
2	0,40007	1,49193	0,19607	1,21661	-0,77967	0,45856
3	0,84264	2,32249	0,81467	225.842	0,30681	1,35909
4	1,35655	3,88277	1,2713	3,5656	0,43815	1,54984
5	0,40067	1,49283	0,41364	1,51231	-0,87924	0,4151
Degree of urbanisation *reference city centre*						
Urban	-1,74591**	0,17445***	-2,60745*	0,2004*	-1,35034	0,25915
Sub-Urban	-3,98522***	0,018588***	-4,1007***	0,016561***	-4,44706***	0,011713***
Periphery	-4,63473***	0,00971***	-5,32955***	0,00485***		
Percentage of white people	7,58766***	1973,688***	9,11857***	9123,143***	11,33481***	83644,51***
Constant	-11,36926**	0,00001***	-12,36891**	4,25e-06***	-11,49291**	0,00001***
Observations	474		474		353	
Pseudo R2	39,97%		40,77%		40,11%	
Bayesian crit. (BIC)	223,742		218,736		181,249	

Notes: ****p<,01, **p<,05, *p<,1; in model 'gen_perc_10_2011' 121 observations are left out of the analysis. These are origniated from 'Degree of urbanisation; Periphery'. All the observations predict failure perfectly.*

Regression Model 1 (2011-2020)

Variables	gen2		gen_perc_5_2011		gen_perc_10_2011	
	Log odds	Odds ratio	Log odds	Odds ratio	Log odds	Odds ratio
Number of Listed Buildings	0,05743***	1,05911***	0,05965***	1,06147***	0,0553***	1,05686***
Percentage of buildings build before 1940	4,88461***	132,2388***	4,79895***	121,3834***	4,83705***	126,0968***
Percentage of addresses with (private) garden	-0,1957	0,82226	-0,10414	0,9011	-0,33323	0,71661
Average size of (private) garden, in meters	0,0043	1,00431	0,00466*	1,00467*	-0,00451	1,00452
Proximity of wealthy neighbourhoods in 2011	0,1117	1,11818	0,10894	1,11509	0,13189	1,14098
PTAL *reference 1b*						
2	1,6252***	5,07944***	1,47517***	4,37178***	1,36669**	3,92234**
3	1,75152**	5,55782**	1,67768**	5,35312**	1,55952**	4,75656**
4	0,54398	1,72284	0,4801	1,61624	0,39731	1,48783
5	1,27556	3,58069	1,17653	3,2431	1,09765	2,99711
Degree of urbanisation *reference city centre*						
Urban	-0,82745	0,43716	-0,86926	0,41926	-0,89412	0,40896
Sub-Urban	-1,20621	0,29932	-1,11794	0,32695	-1,21373	0,29709
Periphery	-0,65588	0,51898	-0,68025	0,50649	-0,68262	0,50478
Percentage of white people	3,01996***	20,49053***	2,64726***	14,11529***	2,42761**	11,33177**
Constant	-7,84857***	0,00039***	-7,6915***	0,00046***	-7,24952***	0,00071**
Observations	472		472		472	
Pseudo R2	17,18%		16,46%		15,87%	
Bayesian crit. (BIC)	402,582		399,169		398,269	

Notes: *****p<,01, **p<,05, *p<,1*