Do segregated housing markets have a spillover effect on housing prices in nearby residential areas?

Mohammad Ismail, Abukar Warsame and Mats Wilhelmsson

Working Paper 2020:6

Division of Real Estate Economics and Finance
Division of Real Estate Business and Financial Systems
Department of Real Estate and Construction Management
School of Architecture and the Built Environment
KTH Royal Institute of Technology

Do segregated housing markets have a spillover effect on housing prices in nearby residential areas?

Royal Institute of Technology
Department of Real Estate and Construction Management
Division of Real Estate Economics and Finance
Stockholm, Sweden

Mohammad Ismail
Email:mohammad.mri82@gmail.com

Abukar Warsame
Email: abukar.warsame@abe.kth.se

Mats Wilhelmsson
Email: mats.wilhelmsson@abe.kth.se

Abstract:

From an international perspective, housing segregation in Sweden is considered relatively low. However, in recent years the issue has been raised and problematized. For example, some studies show that ethnic housing segregation increased in 199 of Sweden's 285 municipalities during the years 2005- 2015. The purpose of this project is to analyze the trends regarding housing segregation over the past 10-20 years, and whether housing segregation has a spillover effect on neighboring housing areas. Namely, does proximity to a specific type of segregated housing market has a negative impact, while another type of segregation has a positive impact, on nearby housing markets. The results indicate that segregation measured as income sorting has increased over time in some of the housing markets. Its effects on housing values in neighboring housing areas are significant and statistically significant.

Keywords: segregation, spillover effect, housing values

JEL Classification: R30, R20, R23

Acknowledgment: We thank the research project Housing 2.0 (Bostad 2.0) for financial support and Mäklarstatistik AB for transaction data. We also thank IVA and their Jobbsprånget initiative.

1 Introduction

The segregation refers to the residential separation of population groups based on the characteristics of population groups such as ethnic, racial, social, age, and income. Although all kinds of population groups can be the subject of segregation, research and literature on segregation have focused on population groups whose spatial segregation causes political and social problems. In the United States, the focus of segregation studies has been on racial and ethnic groups, where they remain a major concern. This focus is related to the conditions of American cities at the beginning of the twentieth century, in terms of the legacy of the slave system and discriminatory practices against racial and ethnic groups, including residential segregation. Although this situation gradually changed after World War II, the established patterns of separation on racial-ethnic bases did not remodel as the segregation index for African Americans is still very high.

Whereas in European cities that are more homogeneous in racial and ethnic terms, studies in the first decades after the Second World War focused on the social class as an identifier of spatial separation, especially with labor migration (Fujita and Maloutas, 2016). In the last decades, with the strong waves of international immigration to European countries, residential segregation has become more visible and influential, where there is a strong relationship between residential segregation and immigration (Murdie and Borgegård, 1998; Fujita and Maloutas, 2016).

In Sweden, the segregation issues and problems were identified as segregation by age due to the primarily young adults moving to multi-family housing in the 1940s and 1950s. The residential segregation by the end of the 1960s identified as a social problem derives from differences between socioeconomic groups. Finally, with international immigration to Sweden, the beginning of the 1970s was recognized as a `social problem' based on ethnic-racial segregation (Murdie and Borgegård, 1998). Segregation has increased over time according to statistics based on country of birth from Statistics Sweden. On average, the segregation index is equal to 23 in 2017, which can be interpreted as the fact that 23% of the population needs to move for even distribution (segregation index equal to 0). We can further note that the variation between the municipalities is significant. The maximum and minimum segregation index of municipalities is 46.8 and 1.6, respectively (Statistic Sweden (SCB), 2017). Part of this variation was undoubtedly explained by the variation in income, the proportion of migrants, the level of housing prices, and the general access to housing.

Sweden faces the challenge of creating sustainable diversity by facilitating the arrival and integration of new immigrants into the labor market while facing residential segregation, as social exclusion leads to an increase in crime and other problems with a significant negative impact on society. This leads to the creating of parallel societies, poses a threat to social cohesion (Papillon, 2002; Böhlmark and Willén, 2017). As Sweden has transformed in recent decades from one of the most ethnically homogenous countries in the world to one where those born abroad or parents born abroad represent 22% of the population living in certain areas, which makes the study of residential segregation critical (Böhlmark and Willén, 2017).

The way of the spatial distribution for different types of tenures, whether concentrated in different neighborhoods or mixed, determines the extent to which the housing market segmentation affects segregation (Andersen et al., 2016). The residential segregation by race and income has spillover effects on housing values. High-income households are willing to pay a more significant premium for high-income neighbors, where municipalities will be able to collect higher taxes. Hence, the quality of schools and services is better, in addition to social standards, this will lead to higher prices for homes in these neighborhoods more, and as a result, the income segregation will increase (Fogli, and Guerrieri, 2019).

The specific purpose of the paper is to analyze the trends regarding housing segregation over the past 10-20 years and analyze the relationship between housing segregation and its spillover effect on neighboring housing prices in Stockholm metropolitan.

The main research questions we seek to answer:

- Has the segregation increased over time in some of the housing markets?
- Is there a relationship between housing segregation and neighboring housing prices ─ spillover effect? In other words, does proximity to a specific type of segregated housing market have a negative impact, while another type of segregation has a positive impact on nearby housing markets?

The research questions will be addressed by estimating housing segregation concerning income at a low disaggregated level in Stockholm, then analyzing the causal effect on housing prices in neighboring residential areas.

The main contribution of our paper is going beyond the traditional studies of segregation that mainly emphasize residential segregations based on income levels, i.e., low-income or high-income households. We have analyzed the spillover effect of proximity to hot spots (high income) and cold spots (low income) of segregation areas in the municipality of Stockholm in 2013 on the housing values of nearby condominiums or single-family homes.

The remainder of the paper is divided into four major sections. Section 2, we provide a theoretical and conceptual framework for the residential segregation, the spillover effect of housing segregation on neighboring housing areas with a review of previous studies. Section 3 describes the different data sets we used and the research methodology. Section 4 presents the empirical analysis by analyzing the housing segregation estimated concerning income, and the casual effect on housing prices in neighboring areas. Finally, the paper concludes with a discussion on the implications of the study, the limitations of the study, and possible routes for future studies.

2 Theoretical Framework

Segregation Mechanism and Measurement

Residential segregation grew in Europe between socioeconomic groups in the last decades (Fujita & Maloutas, 2016; Musterd & Ostendorf, 1998; Tammaru, Marcińczak, van Ham, & Musterd, 2016). Growing Inequalities in Europe are an essential challenge in threatening the sustainability of European urban communities and cities (Tammaru et al., 2016). Empowerment and promotion of social inclusion irrespective of ethnic-racial, socioeconomic, or another status represent the primary purpose for the Sustainable Development Goals (SDGs). In 2015, the United Nations General Assembly adopted the 2030 Agenda for Sustainable Development, which includes 17 Sustainable Development Goals (SDGs) as the agenda emphasizes a comprehensive approach to achieving sustainable development for all. Where four of the (SDGs) Goals which are 8, 10, 11, and 16 emphases on inclusion (Silver, 2015; Sustainable Development Goals (SDGs), 2015). Residential segregation and social inclusion occupy an essential place in discussions about the sustainability of cities (Papillon, 2002), where the Stockholm region has the most substantial fraction of immigrants compared to other regions of Sweden (Hårsman, 2006).

Spatial segregation is one of the most significant factors that reflect the specific conditions of an urban housing market. Depending on the householders' preferences for a specific set of housing characteristics, and demographic features, similar dwellings are concentrated spatially, and these preference options are bound by financial ability, where the price/rent level distinguishes between housing selection of households. Thus, the dynamics of the housing market reinforce the income-based focus of similar groups of households, which leads to spatial segregation based on income (Giffinger, 1998; Owens, 2019). In addition to racial and ethnic factors (Schill and Wachter, 1995; King and Mieszkowski, 1973).

Income inequality represents the most important and crucial reason for residential segregation between socioeconomic groups (Musterd & Ostendorf, 1998; Quillian & Lagrange, 2016). The causal mechanism of income inequality includes three-part—the first part where the level of income inequality generated by the labor market and ethnic-racial discrimination. In the second part, income inequality and discrimination translate into social and racial segregation in the form of residential segregation. In the last part, residential segregation feeds back inequality and discrimination (Fujita & Maloutas, 2016). Stockholm has a very rapid rise in residential segregation between the top and bottom socioeconomic groups between the 1990s and 2000s (Tammaru et al., 2016).

The differences in income are explained by the fact that a significantly larger proportion of people born abroad have unemployment compared to persons born in Sweden. According to recent studies (OECD, 2020), Sweden is among the top three countries with the highest foreign-born unemployment rate. The unemployment rate for foreign-born was 15.7% compared to 3.9% for native-born. Based on Statistics Sweden, the proportion of long-term unemployed in the group born abroad is as much as 4%, compared to group born in Sweden with 1% of unemployment. Another part can also be explained by income differences where the median income of persons born in Sweden amounts to SEK 393,000, while that of persons born abroad amounts to SEK 303,000, almost 30% gap or difference between the two groups (Statistics Sweden (SCB), 2017).

It has become challenging for foreign-born to find jobs in Sweden during the last decades (Hårsman, 2006). This is due to several reasons, the most important of which is the attitude towards immigrants that leads to spatial segregation, especially towards who came in recent decades, and who are considered from the employers' perspective less attractive due to the lower educational level and skills compared to migrants who arrived during the period of labor migration. On the other hand, many immigrants have higher levels of education than the Swedish-born population, but they do not have equal opportunities to find employment in their specialty. In order to achieve equality in the housing market, the crucial prerequisite is equal access to employment opportunities. In addition to the apparent differences in racial and ethnic background and lifestyle, which contributed to the increasing social gap between many immigrants' groups and Swedes, migrants also tend to live with local groups, which promotes residential segregation (Murdie and Borgegård, 1998).

Furthermore, residential segregation impedes education and the performance of the labor market, as it affects long and short-term education and the results of the labor market, especially for those born abroad. The PISA (Programme for International Student Assessment) results show variance in the results of mathematics, science, and reading between natives and foreign-born (Böhlmark and Willén, 2017). These factors make entry to the labor market more difficult, which leads to a contrast in the level of income.

Spillover effect on the housing market

Income is one of the driving forces behind segregation and income segregation can be a stable equilibrium since high-income households have a higher willingness to pay to live near other high-income households than low-income households have (see, e.g., O'Sullivan, 2007, and the empirical findings by Dickerson et al. 2015). This will have several consequences, such as the level of housing prices in segregated areas and non-segregated areas as well as house prices near segregated areas. Analysis of spillover effects from segregated residential areas is rare. Some older studies analyze the question, but the question has not been empirically analyzed with Swedish data.

A related issue that has been analyzed in the literature is the issue of proximity to affordable housing detrimentally affecting housing values. Nguyen (2005) has done a literature review on the subject. The results indicate that not only is the existence of affordable housing sufficient to have an impact on housing values, but the effect depends mainly on how affordable housing is designed and the management of the buildings, the characteristics of the host neighborhood and its compatibility with affordable housing, and the concentration of affordable housing. Similar results could be considered more generally for the proximity to segregated areas. The effect of affordable housing, social housing, rental housing, or segregated neighborhoods can be seen as a *not in my backyard* (NIMBY) effect (see, e.g., Brunes et al., 2020). Lyons and Loveridge's (1993) results indicated that locating affordable housing near residential property has a statistically significant negative effect on its value, which diminishes with greater distance, and this effect tends to be quite small. Cummings and Landis (1993) argued that the proximity of affordable housing has no significant effect on property values if located within 1/8 or 1/4 mile, but there was a negative effect when located within a 1/2-mile radius,

Neighborhood composition is a crucial factor. Clustering and situating the affordable housing in disadvantaged and declining neighborhoods have adverse effects on property values are more likely to occur. Galster, Tatian, and Smith (1999) examined the effect of (Section 8 sites) dispersed housing subsidy programs "are programs that are designed to provide greater locational opportunities to households receiving housing assistance" on the sale prices of housing situated surrounding it between 1991 and 1995 in Baltimore County, located at three different distances: 500 feet, 501-1,000 feet, and 1,001-2,000 feet. The results indicated to strong positive impact related to higher property values If only a few of Section 8 households located within 500 feet, Section 8 sites had a significant negative impact on prices in the 2000-foot range, with the effect declining after 500-feet. So, policies should be put in place, directing the households of Section 8 away from vulnerable neighborhoods and avoiding a clustering of affordable housing.

The proximity to affordable housing not only has negative impacts on property values. It is possible the affordable housing with good quality and design, well-maintained, and rehabilitation has positive impacts on property values and raises their value. Alternatively, at least it did not have any harmful effects on the property values, and it does not cause property values to decrease. (Briggs, Darden, and Aidala 1999; Goetz, Lam, and Heitlinger, 1996: and Santiago, Galster, and Tatian, 2001).

3 Data and Descriptive Statistics

To answer the purpose of the paper, we must combine information on segregation within a city with information on property values in the city. We have, therefore, used data on the income of the population and data on housing values via housing transactions. The case study utilized is the city of Stockholm, the capital of Sweden. The following sections present the two data sets with descriptive statistics and illustrate how we have estimated the concentration of segregated areas.

The disaggregated income data and the measurement of segregation

To analyze the housing segregation, we have used information about the population's income in the Stockholm municipality presented in Table 1. The data is spatial to such an extent that it provides information at a disaggregated level of 250 meters by 250 meters squares. The total number of squares is 2421, with information on the population and their incomes.

Table 1. Population in the municipality of Stockholm in 2000, 2007, and 2013.

	2000		**2007**		**2013**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
Total number of people with income	280.645	343.412	288.337	346.53	315.57	369.48
Number of people with low income in the included area	70.189	97.0784	73.7791	101.308	77.9654	105.152
Number of people with medium low income in the included area	56.1881	68.4725	56.2398	67.5937	60.9717	71.502
Number of people with medium high income in the included area	58.8557	69.5556	59.378	69.3397	66.3452	75.5665
Number of people with high income in the included area	95.4118	132.904	98.9399	141.137	110.287	155.319
N	2421		2421		2421	

Note: Different income levels measured in TSEK are low income (<150), medium-low (150-250), medium-high (250-360), and high income (>360). Only income levels of persons older than 20 years are considered.

As Table 1 demonstrates, the average number of the total population living in the included area from 2003 to 2013 has increased by around 5% (log change of 0.39% per year). Similarly, the average number of people with different levels of incomes also increased almost the same rate as the population over the same period. In the year 2000, the share of the number of high income in the included area was 26% and 40% higher than the shares of low income and medium low income, respectively. The gap of the shares of different income levels living in the included has slightly changed over the study period.

The data refers to almost the last two decades, which enables us to analyze trends over time. The following Table 2 shows the income segregation in Stockholm between 2000 and 2013. Income segregation measured as, for example, the Dissimilarity and the Information Theory Index have increased over time. These measures are the most widely used to measure residential segregation. The Dissimilarity Index measures the extent to which the population of each group is distributed evenly across the region. It also used to measure inequality, known as mean relative deviation. Information Theory Index used to analyze segregation within and between communities, and as a measure of segregation, the index measures the diversity of local areas compare to the overall diversity of a region, instead of measuring the diversity between the local and overall proportions of each group. Indices values typically range between 0 and 1 (Roberto, 2015).

Table 2. Diversity measures. Income.

Year	**Dissimilarity index**	**Information theory index**
2000	0.1777	0.0457
2007	0.2003	0.0568
2013	0.2047	0.0582

Note Dissimilarity index = a measure of how different the social composition of neighborhoods is, on average, from the social composition of the study area, where 0 is no segregation. Information theory index = a measure of how much less socially diverse neighborhoods are, on average, than is the study area, where 0 is no segregation.

Getis-Ord statistics G_i^* given a set of weighted features, identifies statistically significant hot spots and cold spots. The index used to study evidence of identifiable spatial patterns. The standard definition of Getis-Ord G_i^* statistic is:

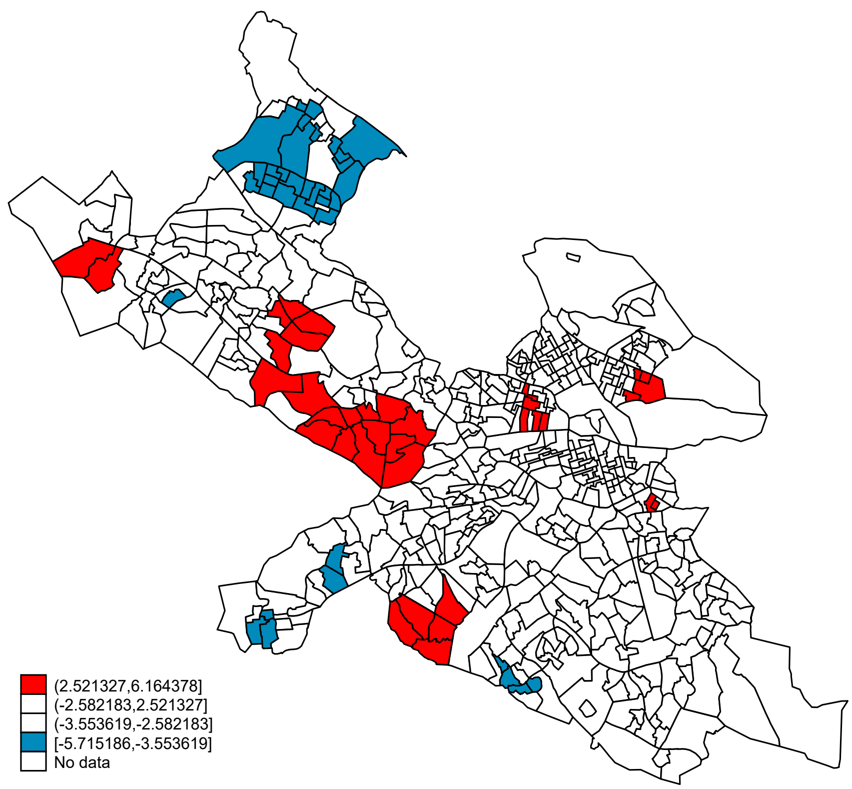
$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j}x_j - \bar{X}\sum_{j=1}^{n} w_{i,j}}{S\sqrt{\frac{n\Sigma_{j=1}^{n} w_{i,j}^2 - (\Sigma_{j=1}^{n} w_{i,j})^2}{n-1}}} \quad (1)$$

Consider a study area subdivided into n regions $i= 1,2,.....,n,$ and each region is identified with known Cartesian coordinates point with measurements $X=[x_1,...,x_n]$. Each i value has associated with it a value x_i.

Moreover, the weights w_{ij} are defined between all pairs of points i and j (for all $i, j \in \{1,\ldots,n\}$). The Gi^* statistic for each feature is a Z score and, therefore, can be attached to the statistical significance. Z score has a positive statistical significance; the larger the Z score is, the more intense the clustering of high values (hotspot). While the Z score has a negative statistical significance, the smaller the z-score is, the more intense the clustering of low values (cold spot). (Ord and Getis, 1995; Songchitruksa and Zeng, 2010).

We can observe in Map 1 for the Stockholm municipality in 2013 that the concentration and size measured with the Getis-Ord statistics (hot and cold spots) in the segregated areas are substantial. Areas marked in blue refer to areas with a more significant proportion of lower-income households (cold spots), and red refers to areas with a more significant proportion of high-income households (hot spots).

Map 1. Hot (high income) and cold (low income) spots segregation in Stockholm municipality 2013 (p=0.05).



The methods that we used are partly to calculate different types of segregation measures and analyze them over time or in space such as the dissimilarity index and the information theory index as well as the Lorenz curve Figure (2 and 3) and Getis-Ord statistics, and partly to estimate different types of econometric models.

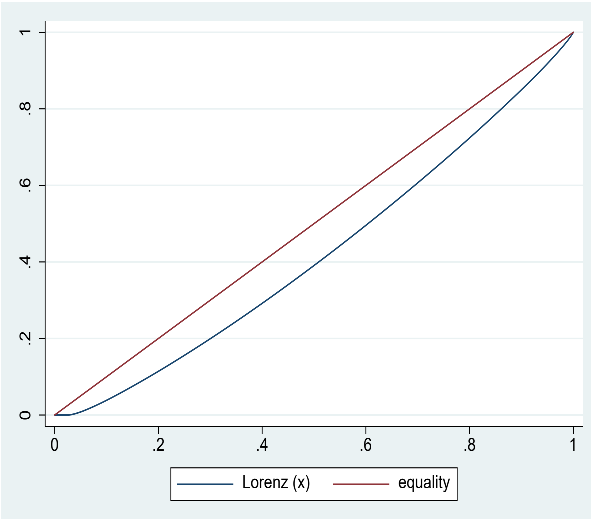


Figure 2. Lorenz curve, Median income, 2000.

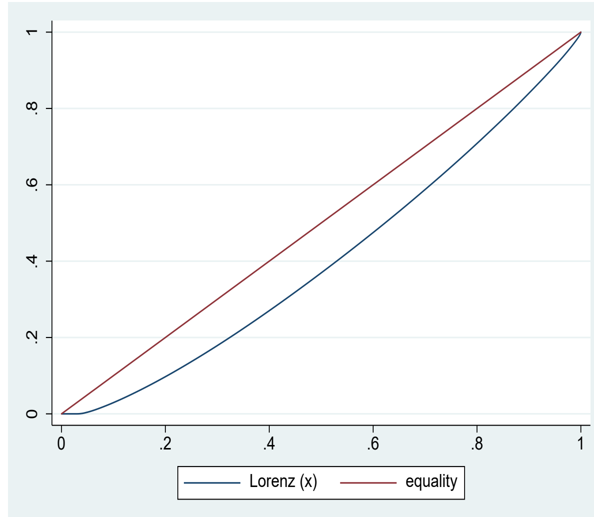


Figure 3. Lorenz curve, Median income, 2013.

The Lorenz curve of income inequality shows a slight increase between 2000-2013 figures 1 and 2. The horizontal axis represents the percentages of the population according to income, and the vertical axis the cumulative income. A straight line with a slope of 1 refers to the ideal equality in income distribution. A Lorenz curve is located below it, showing the actual distribution. The area between a straight and a curved line represents the Gini coefficient and is a measure of inequality.

The transaction data and the measurement of segregation spillover

The empirical analysis will be the estimation of the traditional hedonic pricing model. It will be estimated for the condominium market, where we have the most observations, and for single-family houses with ownership. Data refer to the year 2014 in the City of Stockholm. The source of the data is Mäklarstatistik AB.

In the hedonic price equation, the value of housing will be related to the nearest segregated housing area. The measured distance is the shortest Euclidian distance in meters. We have included one variable that measures the distance to a segregated area with high income and one that measures the distance to a segregated area with high income in the hedonic price equation. The calculation of the distance variables has been made possible by merging the data regarding disaggregated household income in the city with all real estate transactions in the city.

Other variables included in the price equation are living area, number of rooms, and monthly fee in the apartment model and living area and plot size in the single-family model. In both models, the distance to

CBD will also be included as well as the distance to four different sub-centers in Stockholm, namely Farsta center, Vällingby center, Skärholmen center, and Ringen center. Also, the variable's longitude and latitude are included to minimize the incidence of spatial dependence. Descriptive statistics regarding the included variables in the hedonic model are illustrated in Table 3.

Table 3. Descriptive statistics.

	Mean	Median	Standard deviation
Condominiums			
Price	3,356,004		1,935,522
Living area	62,1		27.1
No. of rooms	2.5		23.5
Monthly fees	3229.3		1341.5
Distance to CBD	664.2		504.0
Distance hot spot	1597.0		1272.7
Distance cold spot	1333.0		6.41
Single-family houses			
Price	5,655,744		2,411,516
Living area	129.9		41.6
No. of rooms	5.7		1.4
Plot size	572.2		331.5
Distance to CBD	1529.5		645.9
Distance hot spot	771.9		8.2
Distance cold spot	1857.1		22.6

The table presents descriptive statistics regarding the data on condominium sales and single-family housing sales. The number of observations of apartments amounts to just under 20,000. The average price is just over 3 million SEK, with a standard deviation of as much as 2 million SEK. The average size of flats is 62 square meters, with a standard deviation of 27 square meters. As the owner of a condominium, you pay a monthly fee to the association to cover the cost of common areas and ongoing maintenance. On average, this fee amounts to SEK 3,200, with a variation corresponding to one-third of that amount. The vast majority of observations have a city-close location with an average of only 66o meters from the central business district (CBD), which means that most apartments are relatively far from the various sub-centers around

Stockholm that are included in the model. The distance to a segregated residential area, whether it is one with low or high incomes, amounts to about 1500 meters; however, the variation is significant.

The number of observations in the sample of single-family house sales is significantly fewer, only 1,200. This means that our results regarding the single-family housing market will not be as robust and evident as for the condominium market. The average price is significantly higher, closer to SEK 5.6 million, with a standard deviation of SEK 2.4 million. Also, the size is more than twice as large, with its 130 square meters and with a standard deviation of 42 square meters. Unsurprisingly, single-family houses are located longer out of the CBD than the average apartments. The distance to the nearest segregated area is shorter when it comes to the high-income area and longer when it comes to the low-income area.

4 The Econometric Analysis

The econometric analysis will intend to estimate the hedonic equation. We will use the Box-Cox transformation to find the empirically best form of function. The proximity to hot and cold spots will be included as a continuous variable in a so-called gravity model, and as a binary variable in a so-called treatment effect model. We will also test for parameter heterogeneity in space and price distribution.

Gravity model

Here we will estimate the gravity type model. However, first, we will test which functional form is most suitable. We have performed a Box-Cox transformation. The result of the analysis is that we cannot reject a natural logarithm transformation of the price, but we reject the same transformation of the independent variables. Hence, a semi-logarithmic specification form of the hedonic pricing model is preferred. This applies to both the condominium model and the model for the single-family houses. Hence, all models in this analysis are using a semi-logarithmic form of the hedonic price equation.

In step 2, proximity to Hot and Cold spots are included as a continuous variable in the hedonic price equation where we analyze the entire material (all of Stockholm), but also where we analyze a more limited sample where only transactions within 1.5 kilometers of a segregated area are included. Table 4 shows the results of these models for both apartments and single-family houses. A hot spot is defined as a high-income area and cold spot as a low-income segregated area. All model has been estimated using a method controlling for outliers in the data.

The impact of outliers on estimated parameters is a complex issue. We are following the process laid out in Rousseuw (1997) concerning detecting outliers. First, we are estimating a hedonic price equation and detect outliers with Cook's D, and then we analyze the absolute residuals. The most influential observations are excluded, and observations with large absolute residuals are weighted down by an iterative process where observation weights are recalculated until convergence (see Huber (1964) and Beaton and Tukey (1974)). Berk (1990) provides a full description of the methodology.

In the models regarding condominiums, we can note that the goodness-of-fit in the model is high 84%. All estimated parameters have expected value and reasonable values. For example, increasing the living area by one square meter increases the price by about 1.4%. The further away from the CBD the apartment is located, the lower the price. When it comes to distances to segregated areas, we can see a negative spillover effect from a segregated area with low incomes. The parameter estimate is statistically significant. However, the spillover effect from high-income areas is not statistically significant. If, on the other hand, we analyze the effects within a distance of 1500 meters from the segregated areas, we can find a statistically significant positive spillover effect in the condominium apartment market.

Table 4. Gravity model

Variables	Apartment		Single-family	
	Full sample	Restricted sample	Full sample	Restricted sample
Living area	0.0136	0.0136	0.0023	0.0025
	(114.45)	(119.66)	(11.34)	(13.39)
Plot area	-	-	0.0003	0.0003
			(13.12)	(13.27)
Monthly fee	-0.0001	-0.0001	-	-
	(-31.25)	(-47.84)		
No. of rooms	0.0300	0.0322	0.0454	0.0467
	(10.66)	(11.24)	(7.27)	(7.52)
Distance CBD	-0.0009	-0.0010	-0.0014	-0.0011
	(47.37)	(-51.12)	(-4.32)	(-3.18)
Distance hot spot (10^{-6})	1.82	-12.7	-6.99	-1.67
	(0.68)	(-4.39)	(-0.68)	(-0.14)
Distance cold spot (10^{-6})	42.2	32.3	134.4	133.9
	(15.06)	(10.89)	(13.79)	(13.12)
Constant	269.8	306.1	383.6	366.3
	(25.69)	(29.00)	(5.02)	(4.76)
R^2_{adj}	0.8354	0.8263	0.7246	0.7114
No. of observations	19,121	17,942	1,168	1,193

Note. t-values within brackets. Distance to sub-center and longitude and latitude are included in the model.

The degree of explanation is somewhat lower in the model with single-family houses. One explanation for this is that this market is significantly more heterogeneous. Included independent variables can explain about 70% of the variation in price. All estimated parameters have expected signs; for example, the value increases by 2.5% for each additional square meter of living space, and each additional room the price increases by about 5%. For each additional square meter of land, the price is 0.3%. The distance variable has an adverse effect on the price; that is, the longer the distance from the CBD, the lower the price. The effect of proximity to the segregated area is significant, but only if the segregated area has a lower income level. Proximity to high-income segregated areas has no spillover effect.

Treatment effect model

In step 3, the relationship between segregation and housing values is analyzed by creating a binary treatment variable, where one equals that the house is close to a segregated area and 0 elsewhere. The control group thus consists of housing further away from the segregated areas where there is no expected spillover effect. This is done in order to reduce the problem of spatially dependent and omitted variable bias, i.e., ultimately, to reduce the problem of endogeneity. Several distance measures will be tested empirically. The treatment group will consist of 500 and 800 meters from the segregated area, and the control group will be 1000 and 1500 meters from the segregated area, respectively. The 500/1000 is labeled *smaller* in the result table 5, and the 800/1500 is labeled *wider*. When we measure the spillover effect, there is a risk that some housing is located in the segregated area, and that is not desirable. We have, therefore, included a buffer zone. Any housing located within 125 meters of the segregated area may in reliance be within the area. Thus, we will only include housing that is at least 125 meters from a segregated area. However, this means that some housing located near the segregated area, but outside will not be included. This fact means that we are introducing bias, but this is expected to be lower than the bias that occurs if we include all housing with less than 125 meters from a segregated area in the model.

As stated, we have estimated two models where we have varied treatment and control groups. In the first model, the treatment group is between 125 and 500 meters, and the control group is between 500 and 1000 meters. In the second model, the treatment group is between 125 and 800 meters, and the control group is between 800 and 1500 meters. We have also separated the estimates of proximity to a segregated area with high incomes and one with low incomes. The results from these models are presented in Table 5.

In treatment model 1 (M1), we can see that the degree of explanation is remarkably high. About 84% of the variation in the price can be explained by the variables used. Compared to the gravity models, significantly fewer observations are used in the calculation. As mentioned earlier, this has done to reduce the risk of omitted variable bias and spatial dependency. All coefficients have expected signs. The estimate for treatment is positive and statistically significant. The parameter estimates amount to 0.0303, which corresponds to 3.1% of the value of the apartment. For an average apartment, this would equal SEK 102,000. In Model 2 (M2), the treatment area has been extended to 800 meters, and the control area has been extended to 1500 meters. Again, the effect is positive and slightly higher than in M1.

Table 5. Treatment effect model.

Variables	Apartment				Single-family			
	Hot spot Smaller	Hot spot Wider	Cold spot Smaller	Cold spot Wider	Hot spot Smaller	Hot spot Wider	Cold spot Smaller	Cold spot Wider
	M1	M2	M3	M4	M5	M6	M7	M8
Living area	0.0136	0.0132	0.0144	0.0137	0.0026	0.0023	0.0013	0.0018
	(60.20)	(69.29)	(74.14)	(84.36)	(7.18)	(7.73)	(2.50)	(4.76)
Plot area	-	-	-	-	0.0003	0.0004	0.0004	0.0004
					(7.61)	(11.45)	(9.02)	(11.40)
Monthly fee	-0.0001	-0.0001	-0.0001	-0.0001	-	-	-	-
	(-13.90)	(-17.56)	(-25.90)	(-25.65)				
No. of rooms	0.0371	0.0340	0.0126	0.0196	0.0461	0.0438	0.0499	0.0411
	(7.43)	(8.03)	(2.72)	(4.99)	(4.31)	(4.85)	(3.56)	(3.69)
Distance CBD	-0.0012	-0.0011	-0.0008	-0.0008	0.0018	-0.010	-0.0011	-0.0004
	(-21.53)	(-29.57)	(-25.88)	(-33.51)	(2.60)	(1.67)	(-1.31)	(-0.62)
Treatment	0.0202	0.0303	-0.0276	-0.0599	0.0458	0.0753	-0.0676	-0.1556
	(4.76)	(8.01)	(-6.24)	(-15.59)	(2.04)	(3.82)	(-2.14)	(-6.29)
Constant	280.1	207.1	386.6	355.6	-56.63	144.9	392.7	100.3
	(12.07)	(10.49)	(26.53)	(27.48)	(-0.38)	(1.19)	(2.35)	(0.72)
R^2_{adj}	0.8449	0.8311	0.8368	0.8383	0.6990	0.7088	0.6775	0.6394
No. of obs.	4,862	7,651	8,115	10,965	439	570	215	354

Note. t-values within brackets. Distance to sub-center and longitude and latitude are included in the model.

The proximity to a segregated area with low incomes has been estimated in models 3 and 4 (M3 and M4). Like M1 and M2, the goodness-of-fit is high, around 83%. The number of observations is also slightly higher, i.e., more transactions have taken place in a location near segregated areas with low incomes. The estimates of area, monthly fee, and distance to CBD and sub-centers all have expected signs. Treatment here means that the home is close to a segregated area with low incomes, and we expect a negative capitalization in housing values. It is also something we find. The effect is statistically significant and negative in both M3 and M4. It is slightly higher in M4, where the treatment area is more widespread than in M3. The effect corresponds to a negative price effect of up to 6%.

In models 5 to 8 (M5-M8), the results regarding the single-family model are presented. Estimates of living area, plot size, and the number of rooms are relatively constant in the different models. However, this does not apply to the coefficient in terms of distance to the city. When it comes to the effects of being close to a segregated area, we can see that the price effect is stronger if we allow treatment and the control area to be wider. The treatment effects are all statistically significant, with t-values above 2. The effects have also expected signs to such an extent that proximity to a low-income area spills over into lower prices for the properties that are close by (M7 and M8), while the effect is positive if the property is near a high-income area (M5 and M6).

Parameter heterogeneity

In a third step, we test the parameters, whether they are constant in space. We do this by testing whether the effect is different north and south of the CBD. We also test parameter heterogeneity in the price distribution by estimating a quantile regression model. The latter makes it possible to analyze whether the effect is the same throughout the distribution of prices. Table 6 presents the results where we have estimated a model where we have included an interaction variable between treatment and north location, and in Table 7, the results regarding quantile regression are reported.

Table 6. Parameter heterogeneity in space.

Variables	Apartment		Single-family	
	Hot spot Wider	Cold spot Wider	Hot spot Wider	Cold spot Wider
Treatment	0.0447	-0.0595	0.0748	-0.0508
	(9.80)	(-12.58)	(2.16)	(-1.30)
Treatment * North	-0.0334	-0.0012	-0.0439	-0.0606
	(-5.60)	(-0.15)	(-1.10)	(-0.96)
No. of observations	7,651	10,964	439	215

Note: All independent variables are included in the estimations, as in previous models.

When we test the hypothesis that treatment is different depending on whether you live north or south of CBD, we find that dwellings located near a low-income area have a constant impact in both north and south of CBD. From the models where we analyze a possible positive effect from high-income areas, we find that the effect is higher in the southern parts of the city.

If an interaction variable between treatment and north is included in the single-family model, it can be noted that none of the parameter estimates regarding treatment are statistically significant except to be near a high-income segregated area. The low significance is probably an effect of the fact that the number of observations is very limited. The estimates are jointly statistically significant. However, the conclusion is that we cannot reject the hypothesis that the parameters are constant in space, but caution in interpretation is required.

The result from quantile regression, presented in Table 7, is evident as we can observe that the effect is more significant in the lower price ranges than in the higher ones on the condominiums housing market. However, since we have estimated a percentage impact, one must keep in mind that in absolute terms, price impact is more noteworthy in the higher price ranges. The same pattern can be observed both in proximity to a segregated area with low incomes and with high incomes.

Table 7. Quantile regression.

	Apartments		Single-Family	
	Low-income	High-income	Low-income	High-income
.1	-0.0752	0.0725	-0.1897	0.0646
.2	-0.0656	0.0530	-0.1645	0.0675
.3	-0.0685	0.0431	-0.1286	0.0474
.4	-0.0680	0.0350	-0.1342	0.0759
.5	-0.0585	0.0263	-0.1467	0.0761
.6	-0.0474	0.0204	-0.1555	0.0809
.7	-0.0377	0.0122	-0.1646	0.0640
.8	-0.0338	0.0166	-0.1615	0.0786
.9	-0.0368	0.0150	-0.1452	0.1064

In the case of the spillover effect on the housing value of single-family homes, the price effect is generally higher, higher for proximity to low-income areas, and more stable over the entire price distribution. It is possible to see a slightly lower effect in the higher price ranges, but that difference is not statistically significant.

In summary, we can conclude that there is a spillover effect from segregated housing areas. In segregated areas with low incomes, this spillover effect is negative, while for areas with high incomes, the spillover effect is positive on nearby housing values.

5 Conclusion and Policy Implications

The purpose of the following study was to identify residential areas that were segregated from either low-income or high-income households. We use Getis-Ord Statistics to identify so-called hot spots and cold spots. By identifying areas of the city where there is a concentration of households with similar incomes, it was possible to answer our second purpose, which was to investigate whether the proximity to these areas spills over into the housing values of nearby condominiums or single-family homes.

Why is it of interest to investigate this issue, what policy implications does it have? Whether segregation is a stable equilibrium or not, segregation brings with it socioeconomic costs. Not only for those living in the segregated residential areas but also for nearby areas and the city. Measuring the spillover effect is part of these costs and, therefore, essential to analyze. Of course, segregated areas are not just areas with lower-income households, but there are, of course, areas that are segregated but where income is high. We have chosen to identify segregated areas with both high and low incomes in our case study, which is Stockholm, Sweden.

We can see that segregation has increased between the years 2000 and 2013 that we have investigated. The difference, however, is small. With the help of Getis-Ord Statistics, we have identified hot and cold spots in Stockholm for 2013. These segregated areas we have then matched with sales data regarding condominiums and single-family houses. This has made it possible to calculate the nearest distance from each transaction to any segregated residential area with high and low income. Then we have estimated a traditional hedonic price equation. Whether we estimate a gravity model or a treatment effect model, we can observe an apparent positive effect of being near segregated areas with high income and a negative capitalization effect in the vicinity of segregated areas with low incomes. The effect is stable whether we analyze it in space or the distribution of prices, even if we can observe a higher percentage impact in low-priced houses than in the higher-price segment.

The policy implications of the findings of this research are manifold. Since the identified spillover of house prices emanating from the nearness to segregated areas with high income is different from segregated areas with low-income, policies that address socioeconomic cost and benefits, as well as property assessment levels, should reflect this pronounced differences.

On a property level, positive spillover on house prices near high-income segregated areas will cause an increase in the number of higher-income groups and exacerbate segregation based income. Contrary, negative spillover on house prices near low-income areas might discourage high-income households from moving to near low-income segregated areas. Local government should be aware of these spillover effects on house prices in order to ensure that policies intended to reduce socioeconomic segregation such as residential and income segregation produce desirable results.

A better understanding of the different possible spillover effects on house prices in relation to the spatial distribution of various income groups is beneficial in determining appropriate property assessment levels. In other words, this spillover effect awareness could improve exiting property assessment methods and provide local governments with extra information to make an informed decision on policies and services needed in different neighborhoods.

Furthermore, a good estimation of these spillover effects on house prices would allow local governments to carry out cost-benefit analysis on policies intended to combat segregations and invest deprived communities.

References

Beaton, A.E., J.W. Tukey. (1974). The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16: 147–185.

Berk, R. A. (1990). *A primer on robust regression. In Modern Methods of Data Analysis*, ed. J. Fox and J. S. Long, 292–324. Newbury Park, CA: Sage. 1990.

Böhlmark, A., Willén, A. (2017). Tipping and the effects of segregation. The Institute for Evaluation of Labor Market and Education Policy (IFAU).

Brunes, F., Hermansson, C., Song, H. S., & Wilhelmsson, M. (2020). NIMBYs for the rich and YIMBYs for the poor: analyzing the property price effects of infill development. *Journal of European Real Estate Research*. 13(1), 55-81.

Cummings, P. M., Landis, J. D. (1993). Relationships between affordable housing developments and neighboring property values. *Berkeley: University of California at Berkeley*.

de Souza Briggs, X., Darden, J. T., Aidala, A. (1999). In the wake of desegregation: Early impacts of scattered-site public housing on neighborhoods in Yonkers, New York. *Journal of the American Planning Association*, *65*(1), 27-49.

Dickerson, A., Brooks, R., Rose, C. (2015). Sorting the Neighborhood. *Journal of Affordable Housing & Community Development Law*, 23(3-4), 311-340.

Fogli, A., Guerrieri, V. (2019). The end of the American dream? Inequality and segregation in us cities (No. w26143). National Bureau of Economic Research.

Fujita, K., Maloutas, T. (2016). *Residential segregation in comparative perspective: Making sense of contextual diversity.* London: Routledge.

Galster, G. C., Tatian, P., Smith, R. (1999). The impact of neighbors who use Section 8 certificates on property values. *Housing Policy Debate*, 10(4), 879-917.

Giffinger, R. (1998). Segregation in Vienna: impacts of market barriers and rent regulations. *Urban Studies,* 35(10), 1791-1812. https://doi.org/10.1080/0042098984196

Goetz, E. G., Lam, H. K., & Heitlinger, A. (1996). There Goes the Neighborhood? The Impact of Subsidized Multi-Family Housing on Urban Neighborhoods.

Greene, W.H. (2008). *Econometric Analysis*. Pearson International Edition. Pearson Prentice Hall, Upper Saddle River, New Jersey, USA.

Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35: 73–101.

Hårsman, B. (2006). Ethnic diversity and spatial segregation in the Stockholm region. *Urban Studies*, 43(8), 1341-1364.

King, A. T., Mieszkowski, P. (1973). Racial discrimination, segregation, and the price of housing. *Journal of Political Economy, 81(3),* 590-606.

Lyons, R. F., Loveridge, S. (1993). A hedonic estimation of the effect of federally subsidized housing on nearby residential property values (No. 1701-2016-138670).

Murdie, R. A., Borgegård, L.-E. (1998). Immigration, Spatial Segregation and Housing Segmentation of Immigrants in Metropolitan Stockholm, 1960-95. *Urban Studies*, 35(10), 1869-1888. https://doi.org/10.1080/0042098984196

Musterd, S., Ostendorf, W. (1998). *Urban segregation and the welfare state: Inequality and exclusion in Western cities*. London: Routledge. https://www.academia.edu/19294872

Nguyen, M.T. (2005). Does affordable housing detrimentally affect property values? A review of the literature, *Journal of Planning Literature*, 20 (1), 15–26.

OECD (2020), The Organization for Economic Co-operation and Development, Foreign-born unemployment (indicator). DOI: 10.1787/ba5d2ce0-en (Accessed on 23 April 2020)

Ord, J. K., & Getis, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical analysis*, *27*(4), 286-306.

O'Sullivan, A. (2007). *Urban Economics*, McGraw-Hill International Edition, ISBN 007-124471-9.

Owens, A. (2019). Building Inequality: Housing Segregation and Income Segregation. *Sociological Science*, 6, 497.

Papillon, M. (2002). Immigration, diversity and social inclusion in Canada's cities. Ottawa: Canadian Policy Research Networks.

Quillian, L., Lagrange, H. (2016). Socioeconomic segregation in large cities in France and the United States. Demography, 53(4), 1051–1084. DOI:10.1007/s13524-016-0491-9

Roberto, E. (2015). The Divergence Index: A Decomposable Measure of Segregation and Inequality. arXiv preprint arXiv:1508.01167.

Rousseeuw, P.J., Leroy, A.M. (1987). *Robust regression and outlier detection*. New York: Wiley Interscience, 1987.

Rosenbaum, P.R., Rubin, D.B. (1983). The central role of the propensity score on observational studies for causal effects. *Biometrica* 1983, 70:1,41-55

Santiago, A. M., Galster, G. C., & Tatian, P. (2001). Assessing the property value impacts of the dispersed subsidy housing program in Denver. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management, 20(1), 65-88.*

Schill, M. H., Wachter, S. M. (1995). Housing market constraints and spatial stratification by income and race. *Housing Policy Debate*, 6(1), 141-167.

Silver, H. (2015). The contexts of social inclusion. *Available at SSRN 2641272*.

Skifter Andersen, H., Andersson, R., Wessel, T., Vilkama, K. (2016). The impact of housing policies and housing markets on ethnic spatial segregation: Comparing the capital cities of four Nordic welfare states. *International Journal of Housing Policy*, 16(1), 1-30.

Songchitruksa, P., & Zeng, X. (2010). Getis–Ord spatial statistics to identify hot spots by using incident management data. *Transportation research record*, *2165*(1), 42-51.

Statistic Sweden, SCB, (2017). https://www.scb.se/en/

Stock, J.H., Watson, M.W. (2012). *Introduction to Econometrics*, 8th ed. Addison Wesley, Boston.

Sustainable Development Goals (SDGs), Knowledge Platform, U.N., 2015. (SDGs)

Tammaru, T., Marcińczak, S., van Ham, M., Musterd, S. (2016). *Socioeconomic segregation in European capital cities: East meets West*. Regional Studies Association. London: Routledge.

Wilhelmsson, M. (2002). Spatial models in real estate economics, *Housing, Theory and Society*, 19 (2), 92–101.

Wooldridge, J.M. (2006). *Introductory Econometrics: A Modern Approach*, 3rd ed. Thomson South-Western, Mason, OH.