The Impact of Different Crime Types on House Prices in London: Using Hedonic Model

NCG613: Data Analytics Project
Student: Chunmei Luo; 21250819

1. Introduction

The London housing market is a complex and dynamic system influenced by numerous factors, including land and property availability, supply and demand (Marsden, 2015), crime rates, economic conditions like interest rates and inflation, government policies such as taxation and regulation, and demographic changes such as migration and population growth. Given the city's global prominence in business and culture, there is high demand for property, resulting in volatile pricing that responds quickly to market and policy changes. Therefore, predicting housing prices and understanding the factors influencing them can be challenging. This is especially crucial as housing prices are an essential economic indicator for assessing overall economic health and have significant implications for policymakers, investors, and individuals.

Among the various factors contributing to housing price volatility, the impact of crime has received considerable attention in research. Crime and the associated fear impose direct and indirect costs on city residents, including the monetary value of stolen or damaged property and feelings of insecurity, which can influence their purchase willingness and therefore affect property values. Several studies have attempted to quantify the social costs of crime, such as Anderson (1999) estimating the social costs of crime in the US at $1 trillion and UK Home Office estimates suggesting a £25 billion contribution of crime against individuals and households to the overall £60 billion cost of crime (Brand and Price, 2000). Detotto and Vannini (2010) have also assessed the cost of crime in Italy, amounting to €38 billion.

Analyzing the relationship between crime and house prices is an important research topic that has gained increasing attention. Crime can significantly impact the desirability and perceived safety of a neighborhood, which in turn can influence house prices. As such, understanding the extent and nature of this relationship can provide valuable insights. Additionally, analyzing this relationship can help identify potential strategies for reducing crime and promoting safer neighborhoods, which can have positive economic and social impacts. However, identifying compensating differentials for crime data through empirical methods is challenging as crime data tend to be correlated with other factors that also impact crime rates in a region, such as unobserved variations in local amenities or distance to train. Besides, the crime rate can affect the business prosperity level nearby which leads to potential reverse causality therefore we expect there will be crime clustering.

In this study, the base model used is the hedonic model, which is widely employed in the field of housing market research. To address potential endogeneity issues and account for spatial dependence, spatial regression with instrumental variables is introduced into the hedonic function. This approach allows for the estimation of the implicit prices of intrinsic housing characteristics, such as size, house type, and the surrounding environment (e.g., distance to transit and crime rates in the neighborhood). By considering these various elements, the hedonic pricing method enables a comprehensive assessment of the relationship between a property and an explanatory variable while controlling for other factors that may affect property values.

Some previous studies have demonstrated that high crime levels can lead to decreased property value, such as (Hellman & Naroff, 1979). This decline can be interpreted in various ways, such as an increased desire to relocate, less demand, or direct lower property values. The reason for this is that buyers tend to be willing to pay a premium for homes located in safer neighborhoods, while homes in areas with higher crime rates may be subject to discounts.

The research question addressed in this study is the correlation between crime and house prices in London, primarily through spatial autocorrelation analysis and hedonic modeling with spatial regression. Previous research on this topic has primarily relied on crime rates that are aggregated at different area levels (John Paul Goncalves, 2009). In contrast, this study utilizes street-level crime data from 2018 and combines it with house price data sold in London in 2019. Specifically, three relevant types of crime data are selected: Burglary, Robbery, and Criminal Damage because they are more relevant to the property itself. This study merged crime and house data by calculating the crime count for each crime type within a 500-meter buffer around each property. This approach using street-level crime data will provide a more granular understanding of the relationship between crime and property values.

This study empirically set to test three main hypotheses regarding the relationship between crime and housing prices. Firstly, it investigates the impact of three specific types of crime on housing prices while controlling for other property attributes and neighborhood characteristics (Hypothesis 1). Secondly, the study explores how different types of crimes influence property values to varying degrees. This hypothesis expects that Criminal Damage, due to its direct destructive impact on houses, will have a stronger negative effect on housing prices compared to others (Hypothesis 2). Lastly, the study tests the hypothesis that the effect of crime on prices differs across different areas of the city (Hypothesis 3), with expectations that crime will have a greater influence on house prices in inner city areas as opposed to the outer regions. Our hypotheses are tested using the hedonic model and spatial regression to account for the high spatial autocorrelation present within both the house price and crime data.

The findings of this study highlight the significance of spatial dependence as a key factor influencing house prices. By using spatial regression models, results indicate that different types of crimes exert varying effects on housing markets, both in terms of direction and magnitude. Notably, criminal damage consistently demonstrates a negative impact on house prices, particularly in inner boroughs. However, Burglary has a positive correlation with house prices in London.

2. Literature Review

The influence of crime on individuals' inclination to purchase property varies depending on the types of crime and locations. It is crucial to recognize that the influence of crime on housing markets is highly localized, meaning that crime rates in a specific area can significantly affect local housing prices, while the effects may differ elsewhere. Therefore, the results obtained from one study cannot be applied universally to all housing markets. Consequently, findings from one study cannot be universally applied to all housing markets. By adopting this approach, more precise and context-specific findings can be obtained regarding the relationship between crime and housing prices within specific neighborhoods (George E. Tita, John, et al., 2006). In contrast to previous studies that relied on broader crime rates at the county or borough level, this paper utilizes street-level crime data and links it with all property transaction data in London for the year 2019.

The economic costs of certain crime categories are a significant concern. According to Brand and Price's (2000) estimates, acts of violence against individuals carry an average cost of £19,000, while serious wounding is associated with a total cost of £130,000. For robbery, the estimated cost per incident is £9,700. It is reasonable to assume that the increased risk of such crimes would be factored into property values. Additionally, incidents of assault and robbery may play a crucial role in individual decisions regarding when and where to walk in public spaces. On the other hand, property crimes are more linked to the selection of residential locations.

Introduced by Rosen in 1974, the concept of hedonic valuation emerged as a method to evaluate the worth of composite goods, and it has since gained significant traction in the housing market. Hedonic models decompose house prices into various components, allowing for price estimates of both the property characteristics and the attributes of its surrounding neighborhood. The underlying assumption of these models is that the value of a residential property is simply the sum of the market value of its individual characteristics. The hedonic model allows for the research topic of various factors on house price, including education (Black 1999), transportation infrastructure, air pollution (Chay and Greenstone 2005), etc. Spatial hedonic price models have been applied to investigate the relationship between transportation infrastructure and residential property values. Such as (Seo et al. 2015).

While some researchers have relied on ordinary least squares (OLS) models, disregarding the potential presence of spatial dependence, an increasing number of studies have adopted spatial regression models, considering spatial elements. This is particularly important in the context of housing properties, where spatial dependence often arises, leading to challenges related to autocorrelation and endogeneity, which can compromise the efficiency and unbiasedness of traditional hedonic models. To address these issues, the field of spatial econometrics has developed various approaches and remedies, allowing for a more accurate and robust analysis of the relationship between housing attributes and prices (LeSage and Pace, 2009). In this study, we will focus on spatial regression, but we will have a pre-analysis with the non-spatial hedonic model.

Plenty of studies in the literature have examined the impact of crime on housing prices. For instance, While Katzman (1980) contends that establishing a direct link between crime data and the relocation decisions of residences is extremely challenging, the study does provide compelling evidence that perceptions of crime significantly influence people's willingness to move. London, with its substantial foreign population, housing crisis, and high and rapidly increasing property prices, is also known as a "crime capital." Besides, Buck and colleagues (1991a, 1991b; Buck and Hakim, 1989) identified a negative correlation between crime levels and home values in New Jersey. Additionally, Lynch and Rasmussen (2001) found that the influence of crime on house prices was very small but more pronounced in areas with high crime rates. Furthermore, Gibbons (2004) revealed that crimes such as vandalism, graffiti, and arson had a more significant adverse effect on house prices in London compared to burglary. These findings collectively provide substantial evidence of the relationship between crime and housing prices. However, it is important to note the complexity and nuanced differences involved in the research topic. Factors such as the specific type of crime, location, and local market conditions can all contribute to the varying effects observed in different studies. In summary, previous literature suggests that different types of crimes may have varying impacts on housing markets, and the severity of their influence can differ across locations.

3. Data, Study Area, and Methods

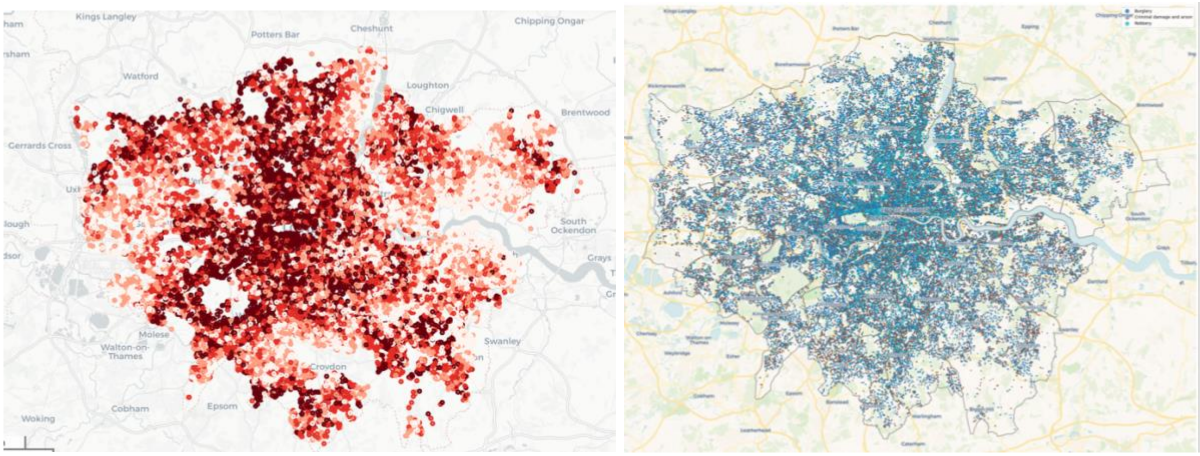
3.1 Data Source and Description

The primary data for this study consists of house price data for London in 2019 and 3 types of street-level crime data for 2018, and some socio-economic census data combined with borough boundary data for the London region.

The property data comes from the UK land registry, a government department founded in 1862 that serves as the central registry for all landowners in England and Wales. For this study, the housing transaction data with corresponding housing characteristics from London in 2019 was extracted. Since the analysis focuses on the year 2019, the influence of the COVID-19 pandemic on house prices can be disregarded.

The street-level crime data is obtained from www.plolice.uk, a website operated by the British police since December 2008 that provides monthly reports on the location of different types of crime. To ensure that the preceding year's crime data could serve as a foundation for future price trends, we opted to obtain crime data spanning the entire duration of 2018, given that our analysis relied on utilizing house transaction prices from 2019. The prevalence of crime is measured monthly by the counts of recorded offenses in several categories, such as “Antisocial behavior”, “Burglary”, “Criminal damage and Arson”, “Drugs”, “Public Disorder and Weapons”, “Robbery”, “Shoplifting”, “Vehicle crime” and “Violent crime”. To ensure the relevance of the crime data to the property itself, three specific crime types were selected for the study: (a) Criminal Damage and Arson (referred to as Criminal damage in this study following), (b) Burglary, and (c) Robbery.

The study area for this project is in the Greater London Area, United Kingdom with data shown in map Figure 1. Figure 1(a) visually represents the mapping of these housing points; deeper color means higher prices. Figure 1(b) visually represents crime points colored in three crime types, highlighting the corresponding distribution. Table II offers a comprehensive overview of the crime data present in our dataset, including descriptions and total counts for each crime category.



(a): House points colored by price sold (b): Crime points colored by type

Figure1: Study data

Data description: Table 1 provides a detailed description of the three types of crimes. It is evident that Criminal Damage could be the most violent one, as it involves intentional damage to properties, including buildings and vehicles. Table 2 presents a summary of key variables used in this study. It is important to note that the Crime counts here represent the number of crimes recorded within a 500-meter radius of each property across the year 2018. And the variables Price, tfarea, and the three types of crime will be log-transformed before being included in the regression model due to their highly skewed distributions.

Crime Categories	Total Count 2018	Crime Description
Criminal damage and Arson	56637	• Intentional destruction or damage of property, not necessarily to enter premises. • Deliberately setting fire to property including buildings and vehicles. • Minor types include Criminal damage to the dwelling, motor vehicle, other buildings, and other criminal damage.
Robbery	32537	• Theft with the use of force or a threat of force. Both personal and commercial robberies are included. Snatch theft is not included. • Minor types include Business property, Personal property
Burglary	80657	• Burglary is the theft, or attempted theft from a premises. • Damage to a premises that appears to have been caused by a person attempting to enter to commit a burglary, is also counted as burglary. • Minor types include Burglary in a dwelling and other buildings

Table 1: Three Types of Crime Data 2018 in London

	count	*mean*	*std*	*min*	*max*
price	42914.0	618104.83	672891.04	67500.00	30000000.00
tfarea	42914.0	93.80	49.21	14.34	959.74
numberrooms	42914.0	4.34	1.72	1.00	20.00
Flats	42914.0	0.39	0.49	0.00	1.00
New	42914.0	0.00	0.03	0.00	1.00
Detached	42914.0	0.06	0.24	0.00	1.00
Dist_Transit	42914.0	766.15	561.37	5.98	7024.69
Dist_Road	42914.0	1201.42	1116.15	10.75	7038.81
Dist_0KM	42914.0	13192.63	5888.83	163.17	30588.52
Criminal_damage_count	42914.0	10.06	17.20	2.00	319.00
Burglary_count	42914.0	13.94	28.23	2.00	653.00
Robbery_count	42914.0	6.32	15.42	2.00	976.00

Table 2: Descriptive Statistics Table of key variables in this study.

** Price, tfarea, and three types of crime will be logged before fitting into the regression model as a highly skewed distribution.*

** Flat, new, and detached are dummy variables, value 1 means Boolean yes, and 0 means no.*

3.2 Choropleth mapping

Plotting the price and relevant crime data on the London boroughs map provides more insight into the data and distributions on the map. We chose to use 'Med_priceper' because, without the influence of floor area, price per square meter provides a better representation of price level. According to Figure 2: choropleth mapping, we see a clear trend of the high level of housing prices, and high house density in the city center, gradually decreasing as we move outward. Median income, however, exhibits a different pattern. The total counts of three types of crime over London's 33 boroughs provide us with a more localized, visual representation of areas within boroughs that have the highest counts (as can be seen in Figure 2). Westminster has the highest crime count in London for all three types of crimes, even considering its small areas.

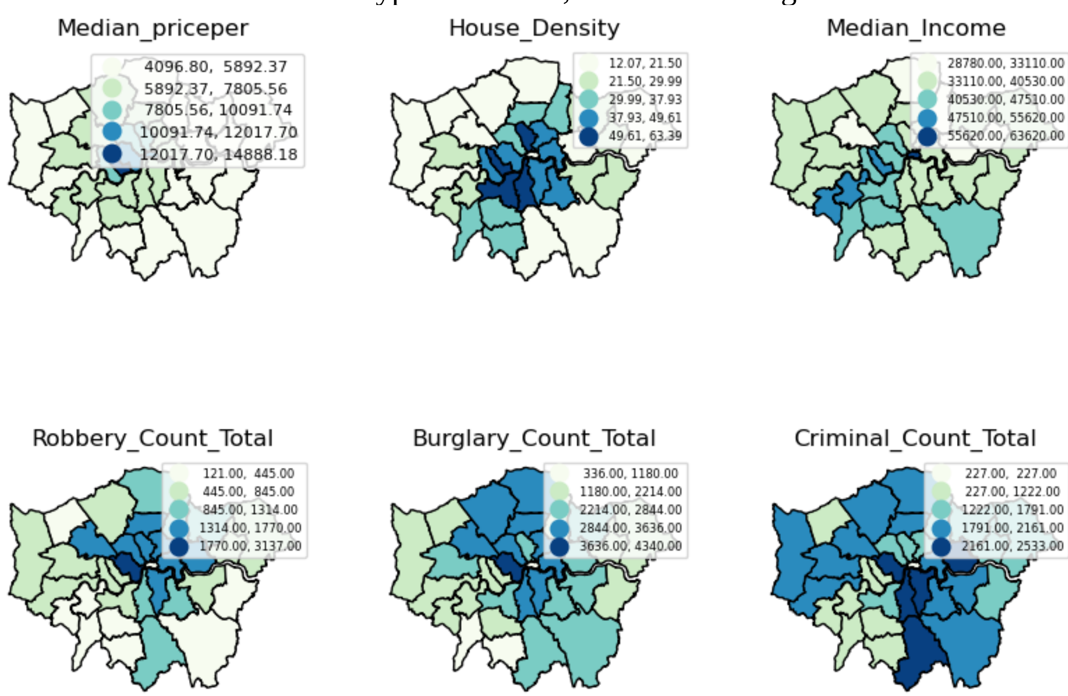


Figure 2: Choropleth mapping

3.3 Regression models

The analysis follows three stages. The first part is spatial autocorrelation analysis of house and crime variables with the objective of identifying and analyzing univariate/bivariate spatial patterns in the data and establishing the nature and direction of relationships between crime variables and house prices in the data set. The second is the application of a hedonic regression model involving different proximity distances and crime data. The third step is the spatial regression model, SAR, and SEM with details described below.

3.3.1 Spatial autocorrelation

Spatial autocorrelation is a phenomenon where the values of a variable located within show a similar pattern. In this house price example, we are trying to quantify the degree to which similar features cluster and where such clustering occurs. Waldo Tober's first law says, "*Everything is related to everything else, but near things are more related than distant things.*"

There are many methods to analyze spatial autocorrelation, we will use Moran's I and Moran scatter plots: The Moran's I statistic is the correlation coefficient for the relationship between a variable (house price) and its surrounding values. Moran's, I generate a value between -1 and 1, where a positive value indicates positive spatial autocorrelation, a negative value indicates negative spatial autocorrelation, and a value of 0 indicates no spatial autocorrelation. Local Moran's I can decompose the global Moran's I down to its components thus constructing a localized visualization to detect "hot spots" and "cold spots" on the map. It helps to identify similar and dissimilar clusters of house price values in London and pinpoint areas with high or low spatial autocorrelation. In other words, it generates a map that highlights statistically significant clusters of high or low spatial autocorrelation. As expected, the price will have strong spatial autocorrelation, which has been proven by Figure 3.

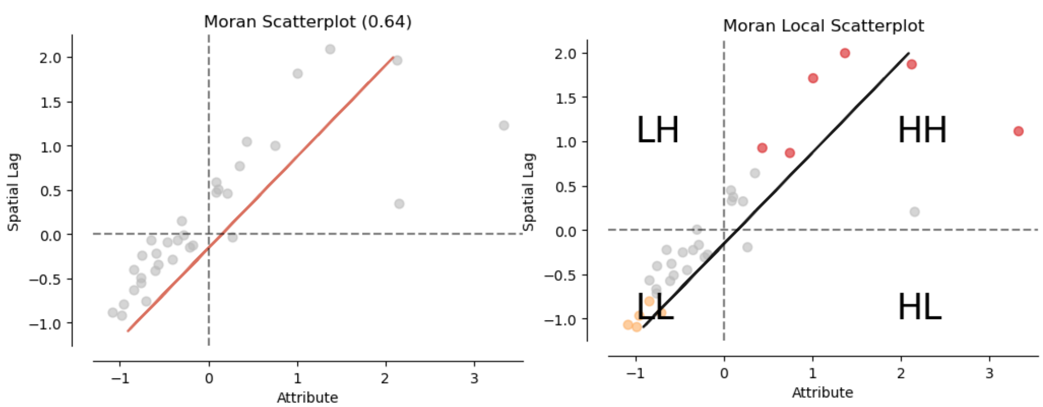


Figure 3: Moran's I stats

3.3.2 Baseline (nonspatial) regression

To lay the groundwork for our analysis, we will initially employ a non-spatial regression base model prior to introducing explicitly spatial methods following the hedonic perspective. It enables us to establish the fundamental principles of hedonic modeling and understand the correlations which is crucial as the spatial regression models will build upon this foundation. Besides, we will test and evaluate spatial dependence, errors, and their significance through diagnostic tests.

Equation 1: Base Hedonic

The basic assumption is that the price of a house is going to be significantly related to how large it is, how many rooms it has, and what kind of house it is.

$$Y = X1\beta 1 + X2\beta 2 + X3\beta 3 + \epsilon$$

where Y is the house price which has been explained by:
$X1$ is a set of housing characteristics including floor area, number of rooms, and house type which is a dummy variable (flat, detached, new).
$X2$ is a set of neighborhood characteristics, e.g., the distance to transit, distance to primary roads, and distance to the city center.
$X3$ is the crime variable measured in terms of the number of crimes within 500 meters (Criminal damage, Burglary, and Robbery)
ϵ is an error term/constant.

Group of explanatory variables	
Housing characteristic	Area(logged), number of rooms, Flat. Detached, New
Proximity Externalities	Distance to Primary Road, Distance to Transit, Distance to the city centre.
Crime (logged)	Three types of crime

The Explanatory Variables Used in Project Y (Log Price)

3.3.3 Spatial regression

In place of non-spatial regression, spatial modeling techniques will be applied in this study. There are several approaches to incorporate spatial dependence, each varying in terms of econometric sophistication. A common feature among these approaches is the formal encapsulation of spatial relationships through spatial weights matrices (n x n) and involving the spatial into statistical methods.

In this research, we examine the spatial lag model (SAR) to incorporate spatial dependence in the outcome variable and the spatial error model (SEM) to address residual spatial error. Additionally, we employ Spatial Regimes Models to analyze the data by dividing it into "spatial regimes," specifically twelve inner boroughs and outer regions. This approach allows us to compare the results of the inner and outer boroughs. The "Inner London boroughs" were defined by the London Government Act 1963, including 12 boroughs, such as Camden, Greenwich, Westminster, etc.

Equation 2: Spatial Lag Model (SAR)

$$Y = \rho WY + X1\beta 1 + X2\beta 2 + X3\beta 3 + \epsilon$$

where W is the spatial weights matrix of Y and ρ is the spatial lag coefficient. The rest are the same as the base hedonic model. The SAR model effectively captures significant spatial dependencies, including external effects and spatial interactions. It assumes that these dependencies are reflected in the spatial lag term (Wy) of the dependent variable (Y). In such cases, spillover effects are not limited to adjacent regions but extend throughout the entire regional system.

Equation 3: Spatial Error Model (SEM)

$$Y = X1\beta1 + X2\beta2 + X3\beta3 + \lambda W\upsilon + \epsilon$$

where SEM divides the complete error term into two components: a spatially structured part (incorporating W inserted into the estimation with λ as the spatial error coefficient) and a residual random portion (represented as ϵ). This partitioning allows for a more comprehensive understanding of the spatial characteristics and random variations within the model.

4. Results and Discussion

4.1 Spatial clustering analysis

While analyzing spatial autocorrelation clustering, the Modifiable Area Unit Problem (MAUP) becomes a concern due to the sensitivity of measurements in cross-sectional data to the levels of aggregation and the configurations of contiguous units (Anselin, 1988). To mitigate the challenges posed by the MAUP, we have opted to utilize Lower Layer Super Output Areas (LSOAs) in London for the Local Moran cluster Map. By adopting LSOAs as the spatial units, we expect to reduce the impact of the MAUP, thereby improving the accuracy of our results.

The Local Moran I Cluster analysis categorizes the data into four distinct classes: HH (High-high), LH (Low-High), HL (High-low), and LL (Low-low). Shown in Figure 4, HH (red) signifies a clustering of high values with other high values, indicating regions with higher prices or a greater likelihood of crimes. Interestingly, the LL class (blue) is observed only in the price cluster map, indicating regions with lower prices that are clustered together. The clustering patterns observed for crimes exhibit some similarities. There is a region in the middle and another in the northeast that show high clustering for all crime types, implying that these areas may be less safe. Overall, the Local Moran I Cluster analysis provides insights into the spatial patterns of housing prices and crime rates, highlighting clustering tendencies.

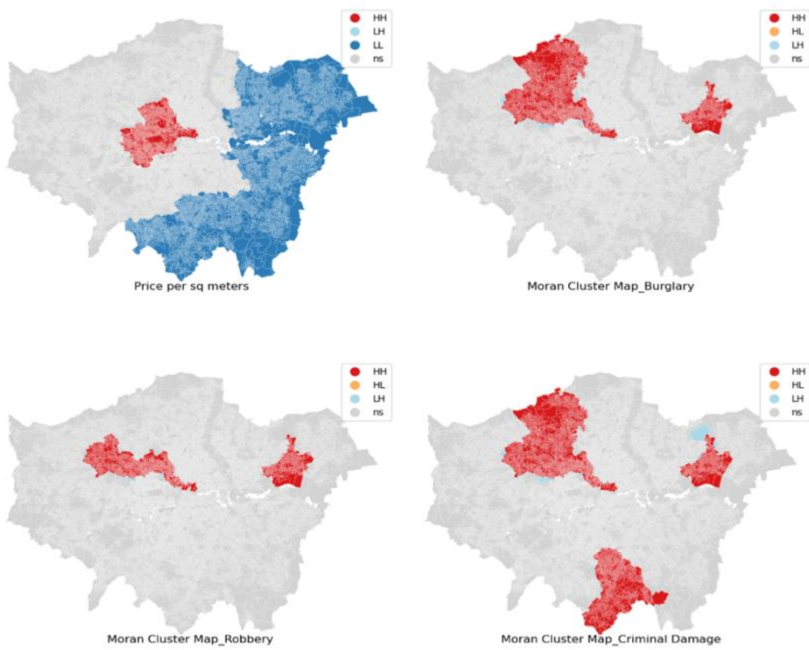


Figure 4: Moran Cluster Map

4.2 Hedonic Model

Hedonic base mode: Table 3 presents the results of the hedonic base model with equations (1). Both models (1) and (2) in this table use the same equation but with different variables. The inclusion of all variables in the study significantly improves the explanatory power, as indicated by the higher R-squared value of 0.73 compared to the initial value of 0.57. The coefficients can be interpreted as follows:

- The 8% increase in housing prices is associated with a 10% increase in housing size.
- The addition of an extra room leads to a 1.4% increase in price, holding all other variables constant.
- Flats are priced approximately 10% lower compared to other types of houses, controlling for other factors in the model.
- Distance to transit and the city center has a negative coefficient, implying that as the distance increases, the housing price decreases. Coefficient of distance to primary road is positive.
- 10% increase in Burglary count around a property is associated with 0.7% house price increase.
- The coefficient of the Robbery count is not significant.
- 10% increase in Criminal damage around a property is associated with 1% house price decrease.

	Dependent variable:	
	(1)	(2)
$\log_a rea$	1.062***	0.838***
	(0.008)	(0.006)
numberrooms	0.001	0.014***
	(0.002)	(0.002)
Flats	0.161***	-0.105***
	(0.005)	(0.004)
Detached	0.005	0.175***
	(0.008)	(0.006)
New	0.014	0.091**
	(0.055)	(0.043)
$\text{Dist}_T ransit$		-0.000***
		(0.000)
$\text{Dist}_R oad$		0.000***
		(0.000)
$\text{Dist}_0 KM$		-0.000***
		(0.000)
$\log_b ur$		0.072***
		(0.005)
$\log_r ob$		0.004
		(0.003)
$\log_c ri$		-0.117***
		(0.006)
Observations	42,914	42,914
R^2	0.574	0.736
Adjusted R^2	0.574	0.735
Residual Std. Error	0.369(df = 42908)	0.291(df = 42902)
F Statistic	11542.334*** (df = 5.0; 42908.0)	10847.538*** (df = 11.0; 42902.0)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 3: Hedonic results of two OLS model.

TEST	MI/DF	Value	Probability
Moran's I (error)	0.5420	682	0.0000
Lagrange Multiplier (lag)	1	30,336	0.0000
Robust LM (lag)	1	485	0.0000
Lagrange Multiplier (error)	1	464,636	0.0000
Robust LM (error)	1	434,784	0.0000
Lagrange Multiplier (SARMA)	2	465,120	0.0000

Table 4: Diagnostics test for spatial dependence.

Diagnostics: These diagnostics tests in Table 4 evaluate whether the residuals of the regression display spatial correlation, examining the assumption of random distribution across space. The diagnostics revealed substantial spatial dependence and spatial error in the model specifications, suggesting that the null hypothesis of spatial randomness in the residuals can be rejected. Consequently, additional spatial regression techniques should be employed to further investigate spatial relationships.

Spatial Regression Results. SAR and SEM:

According to Table 5, the statistical significance of the additional spatial parameters, w_log_price (0.1155) and lambda (0.8865), in the SAR and SEM models provides further evidence of spatial dependence. The magnitude of 'w_log_price' suggests a substantial spillover effect from neighboring properties, indicating that the price of a residential property is not only influenced by its own characteristics but also by the prices of nearby properties. This implies that the original OLS estimators may be biased and inconsistent in the presence of spatial correlation. Accounting for spatial effects is crucial for understanding the dynamics of housing prices and the impact of neighboring properties. Therefore, the interpretation of the research question will focus on the spatial models.

In the SAR model, the partial coefficient for criminal damage counts around a house is -0.09, indicating that a 10% increase in crime-criminal counts is associated with a 0.9% decrease in housing prices. The coefficients for the other two types of crime remain positive but with a smaller magnitude. The complete regression results are shown in the Appendix.

	SAR			SEM		
	Coeff.	Std. Error	P-Value	Coeff.	Std. Error	P-Value
CONSTANT	8.520600	0.211200	0.000000	10.733100	0.034100	0.000000
log_area	0.813200	0.009000	0.000000	0.672000	0.006300	0.000000
numberrooms	0.014000	0.001900	0.000000	0.023500	0.001500	0.000000
Flats	-0.114400	0.004600	0.000000	-0.222600	0.003500	0.000000
New	0.089300	0.029000	0.002100	0.143400	0.031900	0.000000
Detached	0.162600	0.006200	0.000000	0.126500	0.004500	0.000000
Dist_Transit	-0.000000	0.000000	0.000000	-0.000000	0.000000	0.216700
Dist_Road	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Dist_0KM	-0.000000	0.000000	0.000000	-0.000000	0.000000	0.000000
log_bur	0.052500	0.005100	0.000000	0.027600	0.006900	0.000100
log_rob	0.008500	0.002900	0.003600	-0.006500	0.004800	0.170300
log_cri	-0.095200	0.006300	0.000000	-0.041100	0.007600	0.000000
W_log_price	0.115500	0.016800	0.000000	nan	nan	nan
lambda	nan	nan	nan	0.886500	0.004600	0.000000

Table 5: Coefficient summary of SAR and SEM

Summary of research question:
Because of the slightly higher Pseudo R-squared value in SAR, we split the data to compare inner boroughs and outer boroughs using SAR regression. The results in summary Table 6 revealed that the impact of Criminal damage and Burglary on housing prices in Inner boroughs is greater than the effect found in outside boroughs, while the impact of robbery in the inner region is not significant. This suggests that Criminal damage has the most significant negative impact on housing prices.

Specifically, if crime counts of criminal damage within 500 meters of a property increase by 10%, house prices located in twelve inner boroughs will expect to fall by 4%, whereas the corresponding decrease in outer boroughs is only 1%. This finding emphasizes the importance of considering the location-specific effects of different types of crimes on housing prices. The higher sensitivity of inner boroughs to criminal damage suggests that potential homebuyers in these areas prioritize safety and security.

Contrary to expectations, results consistently indicate a positive effect of burglaries on property values across all models and regions. A 10% increase in burglary counts nearby is associated with a 3% increase in house prices in inner boroughs and a 0.5% increase in outer regions. However, it is important to acknowledge that this unexpected outcome may be influenced by unobserved factors related to property characteristics and neighborhoods. Factors such as the presence of higher returns to burglaries in higher-priced dwellings and the higher likelihood of wealthier households reporting crimes may introduce bias into these estimates. Burglary tends to focus on properties that offer the highest expected return in terms of the value of stolen goods. As high-priced neighborhoods typically have a greater proportion of high-income residents, burglars are likely to target these areas, where the potential rewards for burglary are greater. Therefore, it is reasonable to assume that high burglary rates will be found in these areas.

Interestingly, the study reveals that the impact of robbery on property prices in inner boroughs is not statistically significant, while it does have a negative effect on prices in outer boroughs (0.18%). Besides, when considering all of London city, robbery is not significant in the spatial error and ordinary least squares (OLS) models. Overall, it suggests that robbery may have a limited influence on housing prices in the context of this study.

	Burglary	**Robbery**	**Criminal Damage**	**Pseudo R-squared**
Inner boroughs (SAR; Obs:13942)	+3 %	+0.4 % (Not significant)	- 4%	0.7690
Outer boroughs (SAR, obs:28972)	+ 0.5%	-0.18%	- 1%	0.7471
London (SAR, Obs: 42914)	+0.5%	+0.08%	-0.95%	0.7613
London (SEM, Obs: 42914)	+0.27%	-0.06 % (Not significant)	-0.4%	0.7226

Table 6: Summary results and compare inner and outer boroughs.

** Impact of 1o% increase in crime counts nearby a house on the percentage change in house price*

** The rest P-value is significant*

5. Conclusion

In this paper, we have investigated the relationships between house prices and three types of crimes that may impact people's safety concerns and purchase decisions using hedonic spatial regression. The findings of this analysis reveal that different types of crimes have varying effects on house prices, with some crime types leading to lower prices, such as Criminal Damage, which aligns with previous research. Contrary to initial expectations, the results indicate a positive correlation between Burglary and house prices.

The study's main contribution lies in highlighting the nuances of various types of crimes on house prices in different regions. It demonstrates that the influence of neighborhood and location on housing prices is less important compared to property characteristics and crime variables, aligning with findings from other studies. Furthermore, distance to transit and the city center has a negative coefficient significantly, indicating that being far away from transit and the city center leads to lower housing prices, while proximity to major roads was a negative externality for housing prices.

Addressing our hypothesis 1, it can be concluded that not all crime has a negative impact on house prices. Criminal damage has the most significant negative correlation with price as predicted in Hypothesis 1. However, Burglary shows a positive correlation with price contrary to the initial expectation. The impact of robbery, overall, was not significant.

Regarding the range of effects in hypothesis 2, within London, the study found that a 10% increase in crime counts of Criminal Damage within 500 meters of property results in a 1% on house prices in London. If the Robbery count increases by 10%, house prices will expect to increase by 0.5%. Criminal Damage is the most influential crime variable on house price in the context of the study.

One of our key findings is hypothesis 3. notable differences in the impact of criminal damage, burglary, and robbery on housing prices between inner boroughs and outside regions. The study found that a 10% increase in crime counts of Criminal results in a 4% decrease in house prices in inner boroughs, whereas the corresponding decrease in outer boroughs is only 1%. Additionally, a 10% increase in burglary counts nearby is associated with a 3% increase in house prices in inner boroughs and a 0.5% increase in outer regions, confirming our hypothesis that crime impact in the inner city is greater than in the outer regions.

One of the key findings of this study is related to hypothesis 3, which suggests notable differences in the impact of criminal damage, burglary, and robbery on housing prices between inner boroughs and outer regions. Specifically, a 10% increase in crime counts of Criminal Damage results in a 4% decrease in house prices in inner boroughs, whereas the corresponding decrease in outer boroughs is only 1%. Additionally, a 10% increase in burglary counts leads to a 3% increase in house prices in inner boroughs and a 0.5% increase in outer regions, confirming our hypothesis that the influence of crime is greater in the inner city compared to the outer regions.

In conclusion, this study provides extra information into the existing studies as it highlights the importance of considering the effects of different crimes on housing prices in different regions. The findings contribute to a deeper understanding of the factors influencing housing prices and emphasize the significance of safety and security considerations for potential homebuyers. However, one of the limitations of this paper is that it did not consider multicollinearity.

References:

Cullen, J.B. & Levitt, S.D. (1999) Crime, urban flight, and the consequences for cities. Review of Economics and Statistics. 81 (2).

Black SE (1999) Do better schools' matter? Parental valuation of elementary education. Q J Econ.

Dugan, L. (1999) The effect of criminal victimization on a household's moving decision. Criminology. 37 (4).

Gibbons, S. (2004) The costs of urban property crime. Economic Journal.

Hellman DA, Naroff JL (1979) The impact of crime on urban residential property values. Urban Study.

Ihlanfeldt, K. & Mayock, T. (2010) Panel data estimates of the effects of different types of crime on housing prices. Regional Science and Urban Economics.

Lynch, A. K. & Rasmussen, D. W., 2001. Measuring the impact of crime on house prices. Applied Economics, 33(15), pp. 1981-1989

Polinsky, A. M. & Shavell, S., 1976. Amenities and property values in a model of an urban area. Journal of Public Economics, 5(1-2), pp. 119-129.

Morenoff, J.D., Sampson, R.J. & Raudenbush, S.W. (2001) Neighborhood inequality, collective efficacy, and the spatial dynamics of urban violence.

Rizzo, Mario J. (1979), 'The cost of crime to victims: An empirical analysis', Journal of Legal Studies 8,177-205.

Rosen, S., 1974. Hedonic prices and implicit markets: product differentiation in pure competition, Journal of Political Economy 82(1), 34-55.

Taylor, R. B. (1995). The impact of crime on communities. *Annals of the American*

Ceccato, V. and Wilhelmsson, M. (2020) Do crime hot spots affect housing prices? *Nordic Journal of Criminology*, 21(1), 84-102.
Troy, A., & Grove, J. M. (2008). Property values, parks, and crime: A hedonic analysis in Baltimore, MD. Landscape and Urban Planning, 87(3), 233–245.

Nunziata L (2011) Crime perception and victimization in Europe: does immigration matter? CSEA. Working Paper 04/2011

MacDonald Z (2002) Official crime statistics: their use and interpretation. Econ J 112: F85–F106

Malpezzi, S., 2003. Hedonic pricing models: selective and applied research, in O'Sullivan, T., Gibb, K. (Eds.) Housing Economics and Public Policy, Oxford: Blackwell Science.

Ceccato, V. and Wilhelmsson, M. (2011). The impact of Crime on apartment prices: evidence from Stockholm, Sweden. Geografiska Annaier: Series B, Human Geography 93(1):81-103

Appendix:

London (SAR):

```
----------
SUMMARY OF OUTPUT: SPATIAL TWO STAGE LEAST SQUARES
--------------------------------------------------
Data set            :     unknown
Weights matrix      :     unknown
Dependent Variable  :   log_price                Number of Observations:       42914
Mean dependent var  :     13.1318                Number of Variables   :          13
S.D. dependent var  :      0.5657                Degrees of Freedom    :       42901
Pseudo R-squared    :      0.7613
Spatial Pseudo R-squared:  0.7355

White Standard Errors
------------------------------------------------------------------------------------
            Variable     Coefficient       Std.Error     z-Statistic     Probability
------------------------------------------------------------------------------------
            CONSTANT       8.5205757       0.2112072      40.3422511       0.0000000
            log_area       0.8132231       0.0089590      90.7715939       0.0000000
         numberrooms       0.0140224       0.0019405       7.2262580       0.0000000
               Flats      -0.1144151       0.0046332     -24.6946971       0.0000000
                 New       0.0892766       0.0289867       3.0799169       0.0020706
            Detached       0.1626377       0.0062456      26.0402934       0.0000000
        Dist_Transit      -0.0000243       0.0000026      -9.1743781       0.0000000
           Dist_Road       0.0000194       0.0000013      15.3190306       0.0000000
            Dist_0KM      -0.0000382       0.0000006     -59.1392155       0.0000000
             log_bur       0.0525129       0.0051448      10.2069693       0.0000000
             log_rob       0.0084676       0.0029078       2.9120208       0.0035910
             log_cri      -0.0951930       0.0063055     -15.0968382       0.0000000
         W_log_price       0.1154752       0.0167583       6.8906136       0.0000000
------------------------------------------------------------------------------------
Instrumented: W_log_price
Instruments: W_Detached, W_Dist_0KM, W_Dist_Road, W_Dist_Transit, W_Flats,
             W_New, W_log_area, W_log_bur, W_log_cri, W_log_rob,
             W_numberrooms
================================ END OF REPORT =====================================
```

London (SEM):

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIALLY WEIGHTED LEAST SQUARES (HET)
---------------------------------------------------------
Data set            :     unknown
Weights matrix      :     unknown
Dependent Variable  :   log_price                Number of Observations:       42914
Mean dependent var  :     13.1318                Number of Variables   :          12
S.D. dependent var  :      0.5657                Degrees of Freedom    :       42902
Pseudo R-squared    :      0.7226
N. of iterations    :           1                Step1c computed       :          No

------------------------------------------------------------------------------------
            Variable     Coefficient       Std.Error     z-Statistic     Probability
------------------------------------------------------------------------------------
            CONSTANT      10.7330864       0.0341023     314.7323282       0.0000000
            log_area       0.6719551       0.0062508     107.4994041       0.0000000
         numberrooms       0.0235140       0.0014745      15.9468004       0.0000000
               Flats      -0.2226343       0.0035365     -62.9532913       0.0000000
                 New       0.1434433       0.0318979       4.4969505       0.0000069
            Detached       0.1265232       0.0044788      28.2494666       0.0000000
        Dist_Transit      -0.0000091       0.0000074      -1.2353286       0.2167082
           Dist_Road       0.0000371       0.0000062       5.9949526       0.0000000
            Dist_0KM      -0.0000470       0.0000018     -25.7734248       0.0000000
             log_bur       0.0275938       0.0069035       3.9970840       0.0000641
             log_rob      -0.0065426       0.0047709      -1.3713588       0.1702631
             log_cri      -0.0411366       0.0075774      -5.4288531       0.0000001
              lambda       0.8864595       0.0046383     191.1190050       0.0000000
------------------------------------------------------------------------------------
================================ END OF REPORT =====================================
```

Inner boroughs (SAR):

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL TWO STAGE LEAST SQUARES
--------------------------------------------------
Data set            :     unknown
Weights matrix      :     unknown
Dependent Variable  :   log_price                Number of Observations:       13942
Mean dependent var  :     13.3370                Number of Variables   :          13
S.D. dependent var  :      0.6489                Degrees of Freedom    :       13929
Pseudo R-squared    :      0.7690
Spatial Pseudo R-squared:  0.7672

------------------------------------------------------------------------------------
            Variable     Coefficient       Std.Error     z-Statistic     Probability
------------------------------------------------------------------------------------
            CONSTANT       9.7006465       0.0449703     215.7123287       0.0000000
            log_area       0.9355860       0.0109833      85.1825160       0.0000000
         numberrooms       0.0071776       0.0031654       2.2675002       0.0233597
               Flats      -0.1350952       0.0077237     -17.4909445       0.0000000
                 New       0.1864402       0.0867453       2.1492815       0.0316121
            Detached       0.1724392       0.0222341       7.7556089       0.0000000
        Dist_Transit      -0.0000696       0.0000093      -7.4746567       0.0000000
           Dist_Road       0.0000436       0.0000055       7.8787739       0.0000000
            Dist_0KM      -0.0000794       0.0000011     -74.9089644       0.0000000
             log_bur       0.3503711       0.0232209      15.0886171       0.0000000
             log_rob       0.0423134       0.0241620       1.7512358       0.0799053
             log_cri      -0.4247768       0.0384031     -11.0609974       0.0000000
         W_log_price       0.0000706       0.0000034      20.8258235       0.0000000
------------------------------------------------------------------------------------
Instrumented: W_log_price
Instruments: W_Detached, W_Dist_0KM, W_Dist_Road, W_Dist_Transit, W_Flats,
             W_New, W_log_area, W_log_bur, W_log_cri, W_log_rob,
             W_numberrooms
================================ END OF REPORT =====================================
```

Outer boroughs (SAR):

```
REGRESSION
----------
SUMMARY OF OUTPUT: SPATIAL TWO STAGE LEAST SQUARES
--------------------------------------------------
Data set            :     unknown
Weights matrix      :     unknown
Dependent Variable  :   log_price                Number of Observations:       28972
Mean dependent var  :     13.0331                Number of Variables   :          13
S.D. dependent var  :      0.4913                Degrees of Freedom    :       28959
Pseudo R-squared    :      0.7471
Spatial Pseudo R-squared:  0.7462

------------------------------------------------------------------------------------
            Variable     Coefficient       Std.Error     z-Statistic     Probability
------------------------------------------------------------------------------------
            CONSTANT      10.2918441       0.0289361     355.6749868       0.0000000
            log_area       0.7161192       0.0071431     100.2537007       0.0000000
         numberrooms       0.0244973       0.0017794      13.7671772       0.0000000
               Flats      -0.1371002       0.0042608     -32.1768283       0.0000000
                 New       0.1333597       0.0438565       3.0408163       0.0023594
            Detached       0.1972508       0.0057496      34.3068390       0.0000000
        Dist_Transit      -0.0000311       0.0000026     -12.1037449       0.0000000
           Dist_Road       0.0000229       0.0000013      18.0465600       0.0000000
            Dist_0KM      -0.0000324       0.0000004     -85.3594326       0.0000000
             log_bur       0.0594434       0.0040804      14.5681130       0.0000000
             log_rob      -0.0183211       0.0031374      -5.8394843       0.0000000
             log_cri      -0.1044991       0.0049493     -21.1139214       0.0000000
         W_log_price       0.0000404       0.0000027      15.0885250       0.0000000
------------------------------------------------------------------------------------
Instrumented: W_log_price
Instruments: W_Detached, W_Dist_0KM, W_Dist_Road, W_Dist_Transit, W_Flats,
             W_New, W_log_area, W_log_bur, W_log_cri, W_log_rob,
             W_numberrooms
================================ END OF REPORT =====================================
```