ERASMUS UNIVERSITY ROTTERDAM

The effect of environmental amenities on house prices

The inner city of London

Andreas Chrysanthou (382983)

5/7/2016

Supervisor: Jeroen van Haaren

Wordcount: 9,071

The effect of environmental amenities on residential prices: The inner city of London

Research Proposal

Increased environmental awareness in the recent years has increased the demand for specific housing preferences, such as good accessibility and close proximity to green areas, water, and other physical attractions. Such features offer a variety of amenities in densely populated residential areas, including recreation, aesthetic attraction and access to clean air, creating a price premium on housing prices which is often omitted or underestimated by housing agencies. A recent study on the effect of such amenities on housing prices in the Netherlands reveals that environmental factors can create a maximum price premium of 28% (Luttik, 2000). In addition, a larger-scale case study using GIS and landscape metrics for the city of Jinan in China reveals that accessibility and distance to green areas, the size of scenery forest and the percentage land use for green areas have a strong significant effect on house prices (Fanhua, Haiwei, & Nobukazu, 2007). Furthermore, a research applying the hedonic pricing method on Joensuu town in North Carelia (Finland), with a sample of 1006 apartments reveals that proximity of watercourses and wooded recreation areas have a positive influence on housing prices (Tyrväinen, 1997).

Therefore, the research question of this study is: *What are the effects of environmental amenities on housing prices?*

The research question can be examined by using the hedonic pricing model, which explains that the presence and amount of special characteristics associated with each product can determine a set of implicit or "hedonic" prices (Rosen, 1974). In the case of residential prices, the price can be described as a function of a set of variables such as living area, number of rooms and luxury, as well as more general factors such as the location, and the presence of environmental amenities: $P = f(x_1, x_2, \ldots, x_n)$, where P is the house price, and $(x_1, x_2, \ldots, x_n)$ the set of features it embodies.

In order to examine the research question, 3 hypotheses will be tested using statistical analysis:

Hypothesis 1: Average residential prices are positively affected by the percentage coverage of green areas within the region.

The percentage of green areas (parks, gardens, forest etc.) within a region is an important indicator of the general environmental amenities which the selected area has to offer. Relevant studies have also used regional coverage of green areas as an indicator; A. B. Morancho (2003), in her research on the city of Castellón (Spain), highlights the importance of green area coverage by concluding that "large park areas should be created and planned as complements to small landscaped gardened areas" (Morancho, 2003, p. 40).

Hypothesis 2: Average residential prices are positively affected by the percentage coverage of wetlands within the region.

The coverage of wetlands (rivers, lakes, artificial sources etc.) offers physical attraction and recreational activities, and is an important indicator of environmental amenities within the selected region. A relevant study on urban wetlands shows that wetland proximity and size significantly influence property values (Mahan, Polansky, & Adams, 2000).

Hypothesis 3: Residential prices are negatively affected by the increasing distance to environmental amenities.

Measuring the distance of each residence to the closest green area and wetland is a useful indicator to accumulate the accessibility of each house to environmental amenities within the region.

In order to test the hypotheses, data will be collected from 93 different areas within a maximum proximity of 5.5 miles to the central business district of London. The dataset which is extracted from publicly available housing transactions in housing agencies (Rightmove, 2016) through UK's land registry system consists of 200 housing transactions within the selected areas, manually selected for the period 2015-2016 in order to avoid large price differences due to inflation. In addition, only single-bedroom apartments will be selected in order to avoid price differences due to differences in living space. Wetland and green area coverages will be compared to average housing price for each area, and statistical analysis will be performed using an OLS regression on aggregate basis to reveal the significance in the relationship of the explanatory variables. Lastly, a linear regression on individual basis will be performed to test the significance of the proximity of environmental amenities on property values.

Chapter 1 – Introduction

Urbanization and increased environmental awareness of the recent years in most developed economies has created a major trend in the residential market: A growth in the interaction between socio-economic and ecological factors, which created an increasing demand for specific housing preferences such as good accessibility and close proximity to green areas, water, and other physical attractions. Many recent urban policies have promoted the development of large green areas within urban regions, with the recent example of North Madrid, where approximately one million square meters have been transformed to large public green spaces, with the plantation of more than 33,000 trees and 63 fountains (Riggins, 2011). The presence of such features offers a variety of amenities in densely populated areas including aesthetic attraction, recreational activity areas and access to clean air. Such amenities could provide a price premium on residences which is often omitted or underestimated by urban developers. As a result, a general underinvestment can be observed for the development of natural areas within urban regions, which could otherwise provide a variety of private and social positive externalities. A recent study on the value of trees on housing markets reveals that, the existence of mature trees contributes approximately 2% of home values in the examined market (Dombrow, Rodriguez, & Sirmans, 2000). Another relevant research on the effect of green areas on housing prices in the Netherlands reveals that environmental factors can create a maximum price premium of 28% (Luttik, 2000). By examining the effect of environmental factors on housing prices, the development of natural areas within urban regions can be promoted through policies and decision-making processes in the markets of residential and urban development.

Therefore, the research question of this study is:

- *What are the effects of environmental amenities on housing prices?*

The research question can be examined by using the hedonic pricing model, which explains that the presence and amount of special characteristics associated with each product can determine a set of implicit or “hedonic” prices (Rosen, 1974). In the case of residential prices, the price can be described as a function of a set of variables such as living area, number of rooms and luxury, as well as more general factors such as the location, and the presence of environmental amenities: $P = f(x_1, x_2, \dots, x_n)$, where P is the house price, and $(x_1, x_2, \dots, x_n)$ the set of features it embodies.

Many relevant studies have examined the effect of green areas on the value of houses, with the hedonic pricing model being the most effective and popular method. A recent case study applying the hedonic pricing model for the city of Castellón (Spain), highlights the importance of green area coverage by concluding that “large park areas should be created and planned as complements to small landscaped gardened areas” (Morancho, 2003, p. 40). A larger-scale case study for the state of Oregon (U.S.) using more than 14,000 housing transactions for the region of Portland reveals that wetland proximity and size significantly influence property values (Mahan, Polansky, & Adams, 2000). Furthermore, hedonic pricing has also been applied in a case study for Joensuu town in North Carelia (Finland), revealing a positive influence of close proximity to watercourses and wooded recreation areas on housing prices (Tyrväinen, 1997). Other methods such as the use of GIS metrics (Geographic Information Systems) have been applied in relevant studies, such as the case study of the city of Jinan (China), which reveals that accessibility and distance to green areas, the size of scenery forest and the coverage of green areas have a strong significant effect on house prices (Fanhua, Haiwei, & Nobukazu, 2007).

Although a minority of air quality studies provided significant effects, it remains uncertain which air pollutants are mostly relevant to house prices; therefore such studies could not provide certain conclusions. In addition, air quality and its perception by residents, can be affected by industrial activities within much further distance than the location of the examined area, which includes a large variety of external factors in the relationship between regional residential prices and air-quality.

By examining relevant literature, it can be observed that the coverage and distance of environmental amenities within a region are important indicators of measuring the effects of such amenities on housing values. Therefore, several sub-questions can be formed as complementary to the research question:

- *What are the effects of green area coverage on average housing prices within an urban region?*
- *What are the effects of water area coverage on average housing prices within an urban region?*
- *To which extend does close proximity to environmental amenities affect housing prices?*

The content of this study firstly consists of an extensive literature review, through which the abovementioned and other relevant studies are examined and compared in order to derive a conclusion on the effect of environmental factors on housing values. From the literature analysis, housing price determinants are analyzed in order to distinguish environmental factors in particular. Secondly, an empirical analysis is conducted in a case study for the inner city of London, where the methodology and data is presented. Thirdly, the statistical analysis will be presented, measuring the significance in the relationship between housing values and proximity to environmental amenities, as well as the relationship between natural area coverage and average housing values within an urban region. Furthermore, the empirical analysis is compared to the literature review in order to observe any conflicting results or similarities. Lastly, the conclusion will be drawn potentially answering the research question and sub-questions of this study. Policy recommendations and the limitations of this study will also be included in the final conclusion.

Chapter 2 – Literature Review

It is important to conduct a broader analysis of residential price determinants in order to examine the potential effect of environmental factors in particular. To start with, equilibrium in a market with '*hedonic prices*' can be observed when price differences are exactly equalizing across consumers with different preferences, and the price as a function of the product's characteristics $P = f(x_1, x_2, \ldots, x_n)$ identifies the structure of demand (Rosen, 1974). In the housing market, consumer demand can be derived from the consumer's willingness of voluntary mobility, which can be separated in 2 subgroups: Adjustment mobility, which occurs due to changes in preferences for housing, neighborhood or accessibility, and induced mobility, which occurs due to changes in income or life-cycle (Clark & Onaka, 1983). Therefore in order to achieve equilibrium in the housing market, a sufficient differentiation of house and neighborhood characteristics should be implemented to reflect changes in consumers' income, life-cycle and preferences. The most important house and neighborhood characteristics can be examined by a literature review of relevant studies on urban house price determinants.

House price determinants

Sirmans et al. (2005) conducted a literature review of 125 empirical studies on house price determinants using the hedonic pricing model. The authors separated the determinants in 5 generalized categories: *construction & structure* (area, number of rooms, lot size), *house internal features* (bathrooms, fireplace, air-conditioning), *external amenities* (garage, pool, porch), *environmental – natural* (lake view, ocean view, green view) *and neighborhood & location* (location, crime, distance to CBD, trees) (Sirmans, Macpherson, & Zietz, 2005, p. 11). A percentage of 11.83% out of *construction and structure* determinant appearances in all studies examined were insignificant, while 15.66% of the appearances were insignificant for house *internal features* determinants. In addition, 15.21% of *neighborhood & location* cases were insignificant, while the category of *external amenities* determinants was less accurate as 21.48% of the cases were insignificant. *Environmental – natural* determinants appearances could be observed in only 18 of the 125 empirical studies examined, however only 5.55% of the cases were insignificant. A possibility exists that the number of insignificant cases is generally underestimated, as in many cases researchers might not include insignificant variables in their results.

Furthermore, in a relevant study including observations from 93 locations within London, J. S. Wabe (1971) conducted an OLS regression analysis calculating the effect of a variety of both locational and house characteristics (Wabe, 1971). Locational variables included travelling time *(T)* and price *(P)*, socioeconomic index *(SC)* and population density *(PD)*; local employment *(J)* and greenbelt accessibility *(GB)*. House variables included living area *(A),* construction date *(D)*, and dummy variables for central heating *(CH)* and garage *(G).* The main results of the OLS regression analysis can be observed by Wabe's equation in *Figure 1*:

Figure 1: *Wabe's equation on house price determinants (Wabe, 1971)*

		standard error
Rail journey time (min.)	−20·50	(5·76)
Cost of rail fare (pence)	−18·65	(5·49)
Social class	3·44	(0·44)
Population density	−23·38	(5·10)
Green Belt	277·76	(70·95)
Floor area	3·60	(0·24)
Date of Construction	16·67	(4·96)
Central Heating	−285·95	(103·40)
Garage	26·48	(89·25)
Constant Term	3166·47	(589·88)

$R^2 = 0{\cdot}90$.
* Indicates the variable is insignificant at the 5% level.
Central Heating Classification: 1 = full C.H.; 2 = Part C.H.; 3 = No C.H.

All variables in Wabe's equation were significant at a 5% significance level, except the *garage* variable. The variables *travelling time, travelling price, population density* and *central heating* have a significant negative effect on the price; while the *socioeconomic index, greenbelt accessibility, area* and *construction date* have a significant positive effect. The existence of environmental determinants can be observed as the accessibility to green belt areas has a strong positive effect on the price. However, Wabe's calculation was simplistic with respect to green belts, as the dummy variable does not include information on the distance to green belts and green belt coverage.

Another interesting study conducted in 1968 for the city of Saint Louis, Missouri (US), 167 urban regions were examined with respect to different house and area characteristics to reveal the determinants of average property values within the selected regions (Ridker & Henning, 1968). This study included both an OLS regression analysis, and a regression analysis with residualised variables. The authors especially examine the effects of air pollution by including an air pollution index variable; other variables include accessibility to highways, median number of rooms, percentage of new houses in the area, socioeconomic (OCR) and house density (HPM) indexes. The main results are illustrated in *Figure 2*:

Figure 2: *OLS and residualised regression models on house price determinants (Ridker & Henning, 1968)*

	Ordinary regression ($)	*Regression with residualised variables* ($)
Constant	−1469	−2800
Air pollution index	−1865 (91·9)	−245·0 (88·1)
Median number of rooms per house	284·1 (46·1)	488·5 (41·1)
Percentage of recently built houses	50·1 (7·0)	48·4 (7·2)
Housing density index: $(HPM-2{\cdot}42)^2$	65·0 (21·6)	116·6 (20·4)
Travel time to CBD index: $(TIZ-3{\cdot}82)^2$	337·1 (136·1)	320·2 (138·7)
Accessibility of main highway dummy	920·0 (273)	922·5 (278·9)
Above average school quality dummy	−1468 (808)*	398·2 (302·2)*
Average/below average school dummy	−1923 (774)	—
Socioeconomic group index: $(OCR-0{\cdot}64)^2$	9847 (2889)	16940 (2840)
Persons per dwelling unit	−3385 (542)	−3210 (548·7)
Percentage non-white: $(PNW+14)^2$	0·1276 (0·066)	0·196 (0·062)
Illinois dummy	−736·3 (361·4)	—
Residualised Illinois dummy	—	−819·8 (369·1)
Mean family income	0·93 (0·1)	—
Residualised mean family income	—	0·94 (0·1)
R^2	0·94	0·94

(Standard errors in parenthesis.)
* Indicates insignificant at 5% level.

As it can be observed, all variables included in both equations are significant at 5% significance level except the dummy variable *above average school quality*. The median number of rooms, percentage of new houses, HPM and OCR indexes, travel time to CBD, accessibility to main highways , percentage of non-white population and family income have a positive effect on average property values; while air pollution, below average school quality, persons per dwelling and the *Illinois* dummy variables have a negative effect. This research differs from Wabe's research (1971) in terms of data collection and methodology, as Ridker and Henning (1968) focus on regional average property values and therefore collect locational data, omitting house-specific characteristics as price determinants.

In a more recent empirical study for the area of Utah in 2008, 1,366 properties were sampled in an OLS regression analysis including both housing and regional characteristics (Zietz, Zietz, & Sirmans, 2008). Most house-specific features had a significant positive effect on property values, such as area, number of bathrooms, and number of bedrooms, and quality of construction, while the mountain-view variable had a surprisingly significant negative effect on property values.

While the examined studies research house price determinants in general, environmental factors seem to have an important role in determining property values; in many cases, the presence of green-belts, air quality, attractive landscape and the presence of water and green areas were included as price determinants. A review of studies focusing on environmental-specific house price determinants will reveal which environmental factors have the most important effect.

Environmental house price determinants

Many empirical studies examined the effect of environmental amenities on house prices, with most of the studies focusing on specific amenities such as air quality, green areas and wetlands. Some of the studies included landscape view as an environmental determinant; however none of the studies examined was focused specifically on this variable.

- *Air quality studies*

Several air quality studies were conducted such as the example of Harrison and Rubinfeld (1978); who studied the effect of NO_2 concentration on median house values for the Boston metropolitan area, revealing a significant negative effect (at 99% confidence level), with the possibility of average annual benefits up to \$304.12 on property values (Harrison & Rubinfeld, 1978). In addition, R. Palmquist conducted an empirical study in 1982, examining the effect of several air pollutants such as NO_2, SO_2, TSP and Ozone metrics on property sales values for the area of Seattle (Washington), revealing a significant negative effect in 25 out of 70 coefficients examined (Palmquist, 1982). This research estimates a maximum negative effect of -\$479.69 on sales values due to high concentrations of O_3 pollutants. A more recent empirical study on 4 US cities in 2000 reveals a significant negative effect on 23 out of 80 coefficients including pollutants such as NO_2, SO_2 and TSP (Zabel & Kiel, 2000); however the authors did not estimate the specific value of pollution on owner-reported values. Although many studies reveal the negative effect of air pollution on house prices, the results of such studies vary widely, mainly due to the fact that air quality consists of many variables which often differ between different methodologies and data collection techniques. A recent literature review including 12 air quality studies from 1967 to 2000 concludes that, the coefficients of air quality variables are often insignificant, as the examined variables may not be relevant to homeowners (Boyle & Kiel, 2001). Therefore, empirical studies focusing on the effect of wetlands and green areas might reveal more accurate results.

- *Wetland studies*

Amenity values of wetlands have important implications on urban planning policies in deciding whether it is more beneficial to preserve existing wetlands or convert them to other uses. A relevant empirical study on the effect of wetlands on property values for the city of Portland (Oregon, US) including data for a total of 14,485 residential market sales and over 4,500 wetlands and deep-water habitats, reveals that property values are positively affected by the size and distance of the nearest wetland (Mahan, Polansky, & Adams, 2000). The authors used an OLS regression analysis of the form:

$$Ln\ P_{hi} = \beta_0 + \Sigma\beta_j S_{ji} + \Sigma\beta_k Q_{ki} + \Sigma\beta_l N_{li} + \varepsilon_i,\ (\text{for } i = 1, 2, \ldots, n)$$

Where $Ln\ P_{hi}$ is the natural logarithm of sales price, S_{ji} is the structural quality, Q_{ki} measures the k_{th} environmental amenity and N_{li} measures the l_{th} neighborhood characteristic, with ε_i measuring the error term. Through the statistical analysis it was concluded that increasing the size of the nearest wetland by one acre yields an estimate of $24.39 increase in property values, while proximity of one mile less to the nearest wetland yields an additional $436.17. In addition, the study explains examples of economic and environmental benefits provided by preserving urban wetlands, such as *"…water quality improvements, biodiversity, ground water recharge and discharge, and recreation"* (Mahan, Polansky, & Adams, 2000, p. 112).

In addition, a study collecting data from 59 towns in the area of New Hampshire examines the effect of lake water clarity on house prices using the function:

$HP = f(S, L, E)$, where HP represents the house price which is a function of S structural characteristics, L locational characteristics and E environmental characteristics (Gibbs, Halstead, Boyle, & Ju-Chin, 2002). The authors' OLS regression analysis reveals that properties which are adjacent to lakes attribute a positive effect on prices as close proximity increases price by a minimum of $107.29 per foot. Furthermore, water quality and the interaction effect between water quality and lake area both have a positive effect on property values, with a minimum of $213.58 and $304.75 per additional square meter respectively, while the surface area of the lake reveals indeterminate effects. However, since the data of this study focuses on lakefront properties only, distance to wetland variables were not studied specifically.

While the abovementioned studies reveal the positive effects of wetlands on property values, another recent empirical study examines the effects from a rather controversial perspective. Harrison et al. (2001) examine the impact of flood zone status areas on property values; a negative externality of close residential proximity to wetlands. Utilizing a database of 29,887 property transactions in Alachua County (Florida, US), the researchers conducted statistical analysis revealing a negative effect of -$1,034.38 on property values within flood zone areas, significant at 90% confidence level (Harrison, Smersh, & Schwartz, 2001). The authors contribute this effect to the fact that houseowners with properties within flood areas pay increased property taxes, while they conclude that *"While a substantial portion of the increased market value discount in recent years may be attributable to corresponding increases in the nominal values of housing units, their* *contents, and thus flood insurance premiums, such factors are unable to account for the entire change in market dynamics"* (Harrison, Smersh, & Schwartz , 2001, p. 16-17).

While wetland studies provide more accurate results than air quality studies, the effects include both positive and negative externalities, which might provide controversial results on property values.

- <u>*Greenland studies*</u>

A case study for the town of Joensuu (north Carelia, Finland), using appartment sales values for 1006 transactions, measures variables such as the distance of each property to the nearest wooded recreation area, forested area, and the relative amount of forested areas within each housing district (Tyrväinen, 1997). The author uses the general hedonic function: $\boldsymbol{P = f(A_i, L_i, E_i)}$, where ***P*** represents each appartment's sales value, ***A*** apartment-specific characteristics, ***L*** is a vector for locality attributes and ***E*** describes environmental amenities within each housing district. Through a linear regression analysis, the main results show a negative effect on sales values with the increasing distance to the nearest wooded recreation area by an implicit price of -41.78FIM per 100 meter distance, while there is a strong positive effect of 471.46FIM per 100 meter distance to forested areas; green space percentage coverage attributes an implicit price of 7.36FIM. The author comments on the negative effect of close proximity to forests by concluding that: *"The negative impact of the nearby forests, however, can also be understood by the notion that dense, mature coniferous forests may not be appreciated close to a house in these latitudes"* (Tyrväinen, 1997, p. 220).

Furthermore, more than 3000 house transactions were studied in 8 regions within Netherlands, examining the value of green areas, attractive landscape view and open spaces using 2 linear regressions analyses: Firstly, a linear regression was conducted to estimate the effect of house-specific characteristics; by calculating the difference between the estimated price and the actual transaction price, the second analysis could be conducted estimating the effect of locality characteristics (Luttik, 2000). The analysis reveals that attractive landscapes could contribute a premium of 6-12% on property values, while water views contribute a premium of 8-10%. However, the author explains that *"…the impact of green areas was ambiguous; in many cases, the hypothesis that a green structure attracts a premium had to be rejected".* (Luttik, 2000, p. 163). Since the presence of greenland existed in most samples within the database and the absence of such amenities could not be compared, it was difficult to estimate the true value of green spaces on property values within the selected areas.

Another empirical study collecting a total of 810 house transactions for the city of Castellón (Spain), provides a linear OLS regression, a double-logarithmic model and a reciprocal model to assess the effect of urban green areas on property values (Morancho, 2003). The author includes several environmental variables such as a dummy variable for *green view*, *green distance* and *green size*; most of house-specific characteristics had a signifcant effect on price, however only *green distance* has a significant negative effect: property values decreased by €1,800 for every 100 meters distance to a green area.

Lastly, a recent larger-scale case study for Jinan city (China) in 2007, included more variables than previous studies recorded; distance and size of green areas, type (plaza, park, scenery forest), accesssibility, and percentage coverage of green areas were some of the independent variables included (Fanhua, Haiwei, & Nobukazu, 2007). *Figure 3* illustrates the main results of the linear and semi-log regression analysis:

Figure 3

Regression results after eliminating collinear variables, dependent variable: in price *(Fanhua, Haiwei, & Nobukazu, 2007)*

Independent variable	Linear model ($R^2 = 0.648$, adjusted $R^2 = 0.620$)			Independent variable	Semi-log model ($R^2 = 0.673$, adjusted $R^2 = 0.647$)		
	Coefficients	*t*-Ratio	*P*-value		Coefficients	*t*-Ratio	*P*-value
Constant	3735.850	4.666	0.000	Constant	8.172	37.838	0.000
S-DSCEN**	141.094	3.275	0.001	S-DSCEN**	4.919E−02	4.094	0.000
PRLA**	−417.878	−5.319	0.000	PRLA**	−0.123	−5.084	0.000
PLGR**	63.546	4.997	0.000	PLGR**	2.055E−02	5.702	0.000
ACPLAZA*	−51.631	−1.737	0.085	ACPLAZA*	−1.849E−02	−2.571	0.011
EE**	71.813	2.964	0.004	DLOC*	6.295E−02	2.529	0.013
DLOC*	155.359	1.746	0.084	ACPARK	−1.642E−02	−2.405	0.018
ACPARK	−32.087	−1.379	0.171	EE**	1.841E−02	2.648	0.009
TYPEGR**	370.239	2.313	0.022	NPGR**	0.139	3.283	0.001
S-DPLAZA*	188.579	1.760	0.081	PRGR**	−0.192	−2.780	0.006

* Indicates statistical significance at the 10% level.
** Indicates statistical significance at the 5% level.

As it can be observed in the linear model, the size-distance index for scenery forest and plaza (*S-DSCEN & S-DPLAZA*), percentage coverage of greenland (*PRLA*), type of green (*TYPEGR*), location (*DLOC*) and educational environment (*EE*) have a significant positive effect on house prices; while patch richness landscape (*PRLA*) and distance to plaza (*ACPLAZA*) have a significant negative impact. The semi-log model shows minor differences with respect to the coverage of green areas in the housing cluster and fragmentation of public green areas (*NPGR & PRGR*) having a positive and negative impact respectively.

Effects &Mechanisms

Through the abovementioned studies, it can be summarized that wetland and greenland coverage and proximity were the environmental variables which had the most significant effects on prices. *Table 1* summarizes the main findings of the studies examined:

Authors	Year	Area	Dependent variable	Independent variables	Sign (+/-) & significance	Value of coeffecient	Methodology
Harrison & Rubinfeld	1978	Boston (US)	Median house value	NO_2 concentration	negative, 1% s. level	$-304.12	OLS regression, Semi-log model
Palmquist	1982	Seattle (US)	Property sales value	NO_2, SO_2, TSP, O_3 and Ozone metrics	negative, 25 out of 70 coefficients significant	$-479.69 (for O_3 variables)	OLS & GLS regression models
Zabel & Kiel	2000	US	Property sales value	NO_2, SO_2 and TSP	negative, 23 out of 80 coefficients significant	not reported	Log - Linear & Log - Log regression models
Mahan, Polansky & Adams	2000	Oregon (US)	Property sales value	Wetland size, type and proximity	positive, wetland type not significant	$24.39 per acre (size) $436.17 per mile (proximity)	OLS regression
(Gibbs, Halstead, Boyle, & Ju-Chin)	2002	New Hampshire (US)	Property sales value	Lake proximity, water quality, lake area	positive for proximity and water quality, indeterminate for area	$107.29 per foot (proximity) minimum of $213.58 per sq. m. for water quality	OLS regression
Harrison, Smersh & Schwartz	2001	Florida (US)	Property sales value	flood zone areas	negative, 10% s. level	$-1034.38 if property is located in flood zone area	OLS regression

Tyrväinen	1997	Joensuu (North Carelia, Finland)	Property sales value	distance to parks, forests & greenland coverage	negative for distance to parks, positive for distance to forests and green land coverage	(-)41.78FIM per 100m distance to park 471.46 per 100m distance to forest 7.36 per sq. m. coverage	OLS regression
Luttik	2000	Netherlands	Property sales value	green view, water view, presence of green & water within each area	positive for green & water view, insignificant for presence of green & water	6-12 % premium for green view, 8-10% premium for water view	OLS regression
Morancho	2003	Castellón (Spain)	Property sales value	green view, distance and size	negative for distance to green areas insignificant for size & view	€-1800 per 100m distance to green areas	OLS regression, double - logarithmic, Reciprocal model
Fanhua, Haiwei & Nobukazu	2007	Jinan (China)	Property sales value	distance and size of green areas, type (plaza, park, scenery forest), green view, and percentage coverage	positive for size, type, coverage, green view, negative for distance	¥141.094 & ¥ 188.58 for size-distance index to forest and plaza, ¥417.78 for green coverage and ¥-51.6 for distance to nearest plaza	OLS regression & Semi-log model

Through the summary of the main findings, it can be derived that distance to environmental amenities such as green and water sources as well as the size and coverage of these features are some of the most important and significant environmental variables. According to the studies examined, such amenities offer additional value to properties through access to recreational facilities, aesthetics and healthy atmosphere and provide positive environmental externalities by improving biodiversity. In order to conduct an empirical research, this study will focus on 3 variables, which according to previous studies appear to have the largest impact on property values: distance to environmental amenities (water & green), and their natural coverage within each examined region. The size of green areas appears to be less important; the relative size can be derived by measuring the greenland coverage within each region.

Therefore, in order to answer the research question, in addition to the literature review, 3 hypotheses can be examined through statistical analysis:

Hypothesis 1: Average residential prices are positively affected by the percentage coverage of green areas within the region.

Hypothesis 2: Average residential prices are positively affected by the percentage coverage of wetlands within the region.

Hypothesis 3: Residential prices are negatively affected by the increasing distance to environmental amenities.

By conducting a statistical analysis and comparing the results to the main findings of the literature review, it will be possible to extract safe conclusions on the effect of the mentioned environmental amenities on house prices.

Chapter 3 –Methodology & Data

By examining the previous studies, it can be observed that the most popular method of estimating the effects of environmental amenities on property values is the OLS regression model; on individual basis, where the dependent variable consists of individual property sales values, and on aggregate basis, where the average or median house value within a region is used as the dependent variable.

The hedonic pricing model $P = f(x_1, x_2, \ldots, x_n)$, where the property sales value P is determined by a set of characteristics $(x_1, x_2, \ldots, x_n)$ (Rosen, 1974), can form the following equation, assuming a linear relationship between the dependent and the explanatory variables:

$P_i = \beta_0 + \Sigma\beta_k x_{ik} + \varepsilon_i$, where P represents the property sales value in location i, β_0 is a positive constant, $\Sigma\beta_k x_{ik}$ represents the sum of explanatory variables and ε_i is the error term.

However, some of the previous researches have used logarithmic and GLS models (Harrison & Rubinfeld, 1978; Palmquist, 1982; Zabel & Kiel, 2000; Fanhua, Haiwei, & Nobukazu, 2007) and reciprocal models (Morancho, 2003) in order to eliminate heteroskedasticity. A relevant study on the effect of green and water areas on property values using a log-linear model, presents the equation:

$Ln(y_i) = \beta_0 + \Sigma\beta_k x_{ik} + \varepsilon_i$, where $Ln(y_i)$ represents the natural logarithm of the property sales value in location i (Cho, Bowker, & Park, 2006). The authors explain that the logarithmic transformation of the dependent variable is essential in order to eliminate heteroskedasticity, which occurs when there is a wide range in the explanatory variables.

The statistical analysis will occur in two models: Firstly, on individual basis where the dependent variable consists of individual property values, and secondly on aggregate basis, where the dependent variable consists of average sales values within each location. In order to choose the most appropriate methodology for each model of the statistical analysis, the data will firstly be presented and operationalized to specify the dependent and explanatory variables included in this study.

Data

The dataset consists of 200 house transactions, collected from 19 districts within a maximum radius of 5.5 miles from the central business district of London; the data was extracted from publicly available websites through UK's land registry system (Rightmove, 2016). Since this study mainly focuses on environmental determinants, only single-bedroom apartments were included in the dataset, in order to eliminate large price differences due to house-specific determinants such as the total living area, number of rooms etc. In addition, in order to avoid inflationary changes in house prices, transactions were selected only for the period 2015 – 2016. On individual basis (model 1), the explanatory variables (*Greendis, Waterdis, CBDdis and PTdis)* represent the distance of each property to the closer green area (park, forest, etc.) and wetland (lakes, rivers), and the distance of each property to the central business district (CBD) and public transport (PT), as observed in the map of each district, measured in miles.

For the aggregate analysis (model 2), average sales values are provided for 93 selected districts within 5.5 miles from the city center (Rightmove, 2016); while the map of each area was analyzed in an image color extract software which recognizes the percentage coverage of green and water within each selected area (PHPTools)[1]. Apart from the percentage coverage of green and wetlands for each region, the explanatory variables *education* (number of educational facilities within each region) and *CBDdis* (distance to the Central Business District, in miles) are included in the analysis to improve the explanatory power of the equation. It is important to note that, for the aggregate analysis, the average price of all transactions within the chosen period has been included, independently of the number of rooms of each house.

The selection of data was strictly selected within a close proximity (5.5 miles) to the central business district, as the research focuses in an urban environment where natural resources such as green and water are theoretically scarce. Since the data sources did not provide further house-specific details apart from the number of rooms, relevant determinants such house quality and the floor which the apartment is located are omitted from the dataset. *Table 2* and *Table 3* include descriptive statistics for the dependent and independent variables, for the individual and aggregate models respectively:

Table 2: Descriptive statistics, dependent variable - individual price sold (model 1)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Pice sold	200	175000,0	3350000,0	560974,365	278600,5695
Greendis (miles)	200	,0	,4	,127	,1030
Waterdis (miles)	200	,0	1,7	,421	,3300
CBDdis (miles)	200	,3	5,2	1,678	1,0291
PTdis (miles)	200	,0	,5	,137	,0937
Valid N (listwise)	200				

Table 3: Descriptive statistics, dependent variable - Average Sales Price (model 2)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AV. PRICE	93	357440,0	3300516,0	847972,677	580907,2550
GREENLAND %	93	0,0%	34,1%	8,841%	5,7960%
WETLAND %	93	0,0%	26,0%	4,721%	5,2990%
EDUCATION	93	1	11	7,16	2,397
CBDdis	93	,0	4,9	2,628	1,2941
Valid N (listwise)	93				

[1] PHPTools: Map images of each district are imported in online software which analyzes the percentage of various colors within the image. Therefore the percentages of green and blue colors indicate the percentage of green areas and wetlands within each map respectively.

Methodology

It is essential to identify the distribution of data, signs of heteroskedasticity (variability of residuals) and signs of collinearity between the explanatory variables in order to choose the most appropriate method for the statistical analysis.

Firstly, in order to assume a linear relationship between the independent and explanatory variables, it must be assured that the data follow a normal distribution; where the value of the independent variable *Y* can be calculated as follows:

$\boldsymbol{Y = (1/\sigma * (2\pi)^2) * e -(x - \mu)^{2/2\sigma 2}}$, where *X* is a normal random variable, μ is the mean, σ is the standard deviation, π is approximately 3.14159, and *e* is approximately 2.71828 (Stat-Trek, 2016). By a observing the histogram and a P-P plot of standardized residuals created using SPSS statistics software, the normality of the data distribution can be observed for the individual analysis (model 1):

Figure 4: Histogram of data distribution on individual basis (model 1)

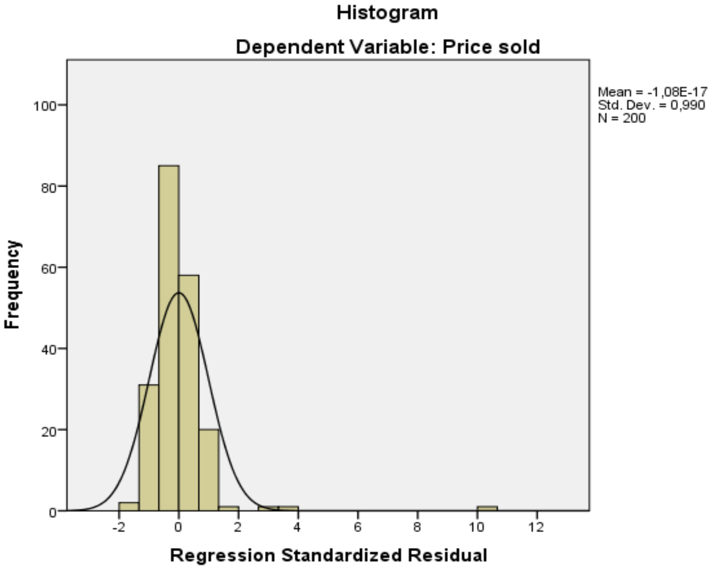
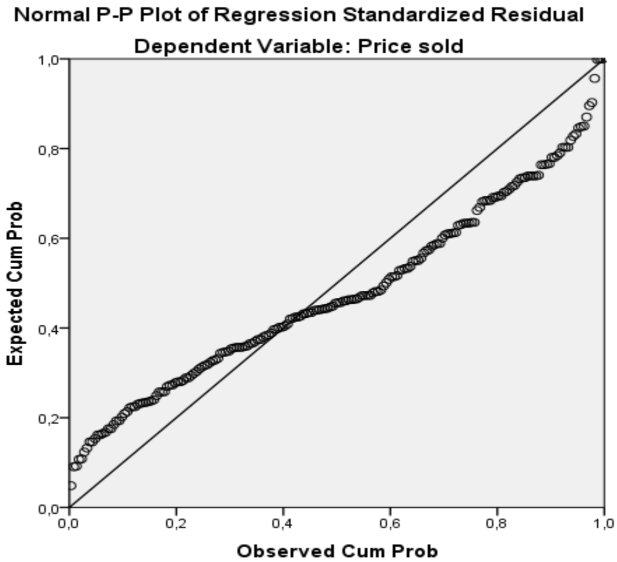


Figure 5: Normal P-P plot of Regression on individual basis (model 1)



As it can be observed, the distribution of the data slightly differ from a normal distribution as the observations are mainly concentrated towards the mean (short-tailed), however this does not indicate a violation of linearity assumptions: *Figure 4* illustrates the expected bell-shaped histogram with a median slightly below the average standardized residual, while *Figure 5* shows that observed cumulative probabilities do not vary significantly from the least squares regression line.

The respective histogram and P-P plot are expected to have significant differences for the aggregate analysis (model 2), mainly due to the fact that the sample size is less (93 areas) than on individual basis (200 house transactions). Although a significantly large sample population is needed in order to identify normality in the data distribution, a similar approach can be used by observing a histogram of standardized residuals and P-P plot of Regression Standardized Residuals. The following figures illustrate the histogram and P-P plot for the aggregate analysis:

Figure 6: Histogram of data distribution for aggregate analysis (model 2)

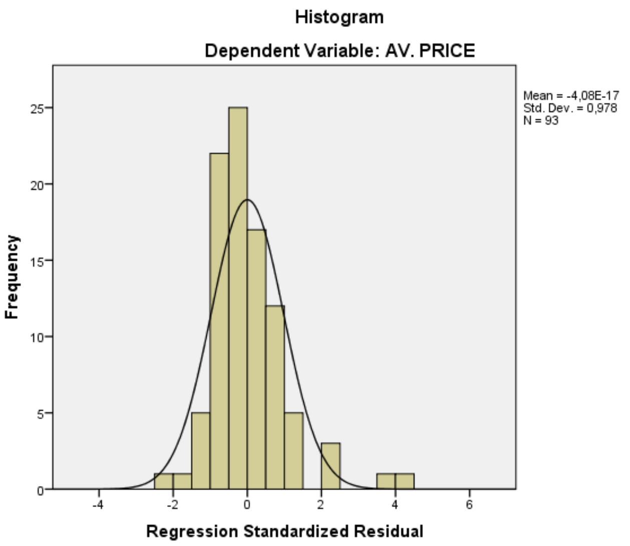
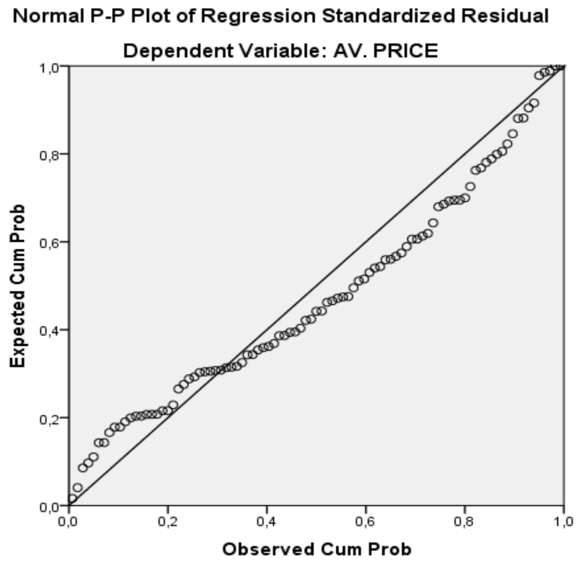


Figure 7: Normal P-P plot of Regression for aggregate analysis (model 2)



While the histogram for the aggregate analysis shows a slight difference from normality of the distribution (short-tailed), the P-P plot shows that the observed cumulative probability has less deviation from the least squares line in model 2. By observing the relevant figures it can be assumed that in both models, the dataset follows a normal distribution, and therefore the data does not violate the normality assumption of linear regression analysis.

Another concern in linear regression models as observed through relevant literature is the possibility of multicollinearity between two or more explanatory variables, which occurs when the independent variables are linearly related (Cho, Bowker, & Park, 2006). There can be a violation of the linear regression assumptions if two or more explanatory variables have a correlation coefficient greater than 0.8 (Judge, Hill, Griffiths, Lütkepohl, & Lee, 1982, p. 620). The *multicollinearity diagnostics* process of SPSS software can be used to test the existence of multicollinearity between the explanatory variables.

Furthermore, the threat of heteroskedasticity is reduced due to the method of data collection; the wide range of explanatory variables is avoided by choosing only single-bedroom apartments, therefore large price differences due to house-specific characteristics are eliminated. Heteroskedasticity might be a larger threat in model 2, where aggregate sales values are used, as all house transactions were included independently of differences in house-specific characteristics. A scatterplot of the standardized residuals will help observe the variability of the residuals and whether heteroskedasticity exists in the dataset. In addition, a White test will reveal the possible existence of heteroskedasticity, where the unstandardized squared residuals of each regression are tested for a linear relationship with the explanatory variables.

Overall, as factors which lead to non-linear relationships between the variables such as non-normality, multicollinearity and heteroskedasticity can be examined and eliminated, the linear OLS regression method is proven to be the most appropriate for statistical analysis of both models; on individual and on aggregate basis. Therefore the additional use of GLS and Log-linear approaches as observed in previous studies can be avoided. It is important to note that for the aggregate analysis, the results might not be representative for the behavior of prices at individual level, as in accord with the ecological fallacy theorem (Robinson, 2009). Therefore, the coefficients of variables used only for the aggregate analysis (model 2) such as *Greenland %* and *Wetland %,* represent only the behavior of average sales prices, and might not represent the behavior of individual house prices, as used in the individual analysis (model 1).

By examining the abovementioned dataset and methodologies, 2 generalized functions can be formed for the individual and aggregate models respectively; considering the hedonic pricing model where the property value is a function of the product's characteristics:

- **Price function for individual analysis (model 1):**

$SP_i = f$ (Greendis, Waterdis, CBDdis, PTdis)

SP_i = Sales price of single-bedroom apartment *(i)* measured in UK currency (£)
Greendis = Distance of the individual property to the nearest green area, in miles.
Waterdis = Distance of the individual property to the nearest wetland, in miles.
CBDdis = Distance of the individual property to London's Central Business District, in miles.
PTdis = Distance of the individual property to the nearest public transport, in miles.

- **Price function for aggregate analysis (model 2):**

$AP_k = f$ (Greenland %, Wetland %, Education, CBDdis)

AP_k = Average residential sales price of location *(k)* measured in UK currency (£)
Greenland % = Percentage coverage of green areas for location *(k)*
Wetland % = Percentage coverage of wetlands for location *(k)*
Education = Number of educational facilities within location *(k)*
CBDdis = Distance of location *(k)* to London's Central Business District, in miles.

Table 4 shows analytically the variables included in each model and the expected sign of each coefficient, which shows the impact (positive or negative) of each explanatory variable on the independent variable:

Table 4: Expected sign of coefficients for each explanatory variable for models 1 & 2

Locational & Environmental characteristic (model 1)	Expected sign	Locational & Environmental characteristic (model 2)	Expected sign
Distance to Green	-	Greenland coverage	+
Distance to Water	-	Wetland coverage	+
Distance to CBD	-	Number of educational facilities	+
Distance to Public Transport	-	Distance to CBD	-

The following results will reveal if the observed sign of each coefficient is in accord with the expected results, and will additionally provide the value of the effect of each explanatory variable.

Chapter 4 – Results

Model 1:

Table 5 provides the coefficient results for the individual analysis, by conducting a linear OLS regression:

Table 5: Coefficient results for individual analysis (model 1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	600467,717	60388,422		9,943	,000
	Greendis (miles)	-594228,209	196700,705	-,220	-3,021	,003
	Waterdis (miles)	-161317,970	63905,986	-,191	-2,524	,012
	CBDdis (miles)	40844,334	20389,417	,151	2,003	,047
	PTdis (miles)	255872,923	214559,527	,086	1,193	,234

a. Dependent Variable: Pice sold

By examining the coefficients of the individual model, it can be observed that distance to green (*Greendis*) and water (*Waterdis*) has a significant negative impact on individual property values at 5% significance level. A very strong impact can be observed for the distance to green areas variable, as 0.1 additional mile shows a negative effect of approximately £-59,423 on individual property price; however the coefficient appears to have a high standard error, and thus a large variability of the effect between individual properties. Distance to water also shows a negative effect of approximately £-16,132 with less variability on the results. An unexpected positive effect can be observed for the distance to CBD variable, where each 0.1 additional mile has a significant positive effect of approximately £4,084 on property values, while the distance to public transport appears to be insignificant.

Table 6 shows analytically the correlations between explanatory variables. As it can be observed, none of the explanatory variables have a correlation coefficient greater than 0.8, which eliminates the case of multicollinearity (Cho, Bowker, & Park, 2006)(see methodology):

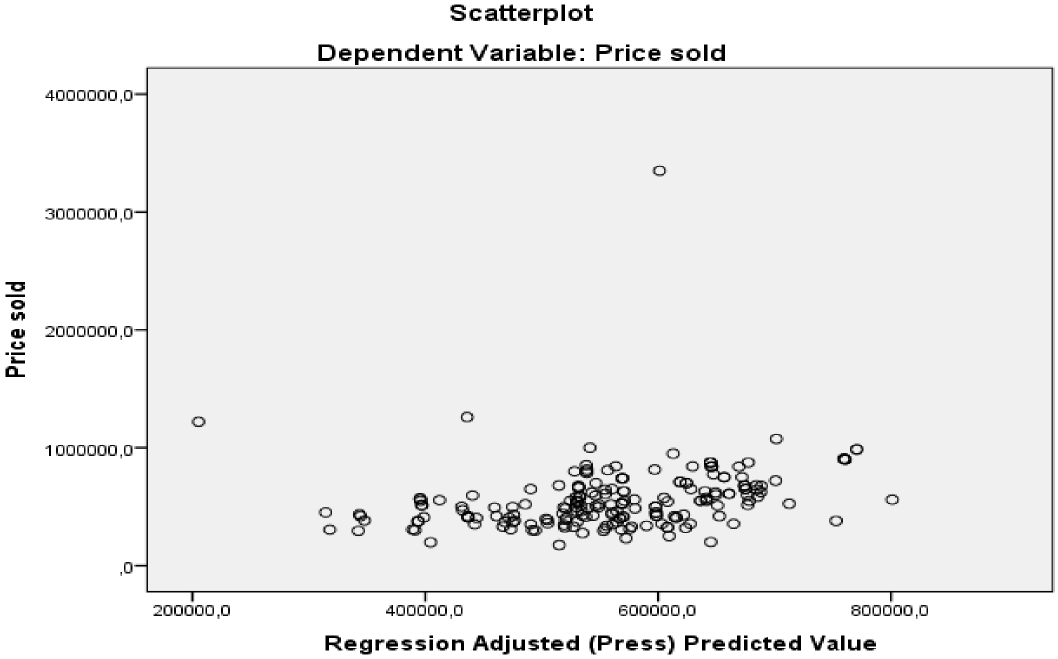
Table 6: Collinearity diagnostics (model 1)

Correlations

		Pice sold	Greendis (miles)	Waterdis (miles)	CBDdis (miles)	PTdis (miles)
Pearson Correlation	Pice sold	1,000	-,293	-,224	,054	,086
	Greendis (miles)	-,293	1,000	,338	,036	-,159
	Waterdis (miles)	-,224	,338	1,000	,317	-,072
	CBDdis (miles)	,054	,036	,317	1,000	-,325
	PTdis (miles)	,086	-,159	-,072	-,325	1,000
Sig. (1-tailed)	Pice sold	.	,000	,001	,222	,113
	Greendis (miles)	,000	.	,000	,307	,012
	Waterdis (miles)	,001	,000	.	,000	,154
	CBDdis (miles)	,222	,307	,000	.	,000
	PTdis (miles)	,113	,012	,154	,000	.
N	Pice sold	200	200	200	200	200
	Greendis (miles)	200	200	200	200	200
	Waterdis (miles)	200	200	200	200	200
	CBDdis (miles)	200	200	200	200	200
	PTdis (miles)	200	200	200	200	200

In addition, *Table 7* provides a scatterplot for the dependent variable *price sold*, using adjusted predicted values on the x-axis, in order to avoid outliers which cause anomalies on the graph:

Table 7: Scatterplot using Adjusted Predicted values on X-axis for individual analysis (model 1)



With the exception of minor outliers, residuals appear to have similar width along the x-axis, with a slight reduction in variability for values less than £400,000. As the dataset does not show a large variability in residuals, no clear signs of heteroskedasticity exist. *Table 7.1* shows that no significant relationship exists between the unstandardized residuals and the explanatory variables; therefore the possibility of heteroskedasticity can be eliminated:

Table 7.1: White-test for heteroskedasticity for individual analysis (model 1)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-6,180E+10	1,163E+11		-,531	,596		
	Greendis (miles)	-5,420E+11	3,789E+11	-,109	-1,431	,154	,851	1,174
	Waterdis (miles)	1,459E+11	1,231E+11	,094	1,186	,237	,785	1,274
	CBDdis (miles)	49344250622	39274183179	,099	1,256	,210	,793	1,261
	PTdis (miles)	3,933E+11	4,133E+11	,072	,952	,342	,865	1,157

a. Dependent Variable: UnsResiduals

The extreme impact of distance to green on house prices can be explained by observing that in general, properties with direct access to green spaces (and thus 0 distance) have a much larger sales price than properties without direct access to green areas. Therefore the difference in price from the first 0.1 additional miles might have a larger impact than increasing the distance from 0.1 to 0.2 miles. The significant positive effect of distance to the central business district can be explained by the fact that areas further away from the City of London (CBD area) have larger opportunities in developing alternative uses for urban planning, such as the creation of parks and artificial water sources. Finally, the effect of distance to the nearest public transport appears to be insignificant, as most of the property samples within the selected regions have access to public transport within less than 0.1 miles.

By analyzing the results of model 1, the following equation can be formed to predict the value of single-bedroom apartments within the selected areas, considering the environmental factors which have a significant effect:

$$SP_i = 600{,}467 - 594{,}228\, x_1 - 161{,}318\, x_2 + 40{,}844\, x_3 + 255{,}873 x_4$$

Where x_1 represents the distance to the nearest green area, x_2 the distance to the nearest water source, x_3 the distance to the central business district, and x_4 the distance to public transport.

Model 2:

Table 8 provides the coefficient results for the aggregate analysis, by conducting a linear OLS regression:

Table 8: Coefficient results for aggregate analysis (model 2)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	926981,593	192983,150		4,803	,000
	GREENLAND %	64841,341	10052,408	,647	6,450	,000
	WETLAND %	11382,244	9885,285	,104	1,151	,253
	EDUCATION	-58717,338	22555,774	-,242	-2,603	,011
	CBDdis	-108643,464	46329,970	-,242	-2,345	,021

a. Dependent Variable: AV. PRICE

The coefficient results for the aggregate analysis show a positive significant effect of regional green coverage on average house prices, with an additional 1% of green areas having an impact of approximately £64,841. In contrast with the results of model 1, the distance to the City of London (CBD area) show a significant negative effect of approximately £-108,643 per mile on average house prices; while the number of educational facilities within each area has a surprising negative effect on the price by approximately £-58,717. Wetland coverage appears to have insignificant positive effects.

Table 9 provides collinearity diagnostics for the aggregate analysis. A minor correlation can be observed between the explanatory variables *Greenland %* and *CBDdis (0.500)*, which enforces the possibility that, increased distance to the city center provides more opportunities for alternative uses of urban development, such as the creation of green areas:

Table 9: Collinearity diagnostics (model 2)

Correlations

		AV. PRICE	GREENLAND %	WETLAND %	EDUCATION	CBDdis
Pearson Correlation	AV. PRICE	1,000	,459	,041	-,246	-,019
	GREENLAND %	,459	1,000	-,252	,167	,500
	WETLAND %	,041	-,252	1,000	-,259	-,153
	EDUCATION	-,246	,167	-,259	1,000	,349
	CBDdis	-,019	,500	-,153	,349	1,000
Sig. (1-tailed)	AV. PRICE	.	,000	,349	,009	,427
	GREENLAND %	,000	.	,007	,055	,000
	WETLAND %	,349	,007	.	,006	,071
	EDUCATION	,009	,055	,006	.	,000
	CBDdis	,427	,000	,071	,000	.
N	AV. PRICE	93	93	93	93	93
	GREENLAND %	93	93	93	93	93
	WETLAND %	93	93	93	93	93
	EDUCATION	93	93	93	93	93
	CBDdis	93	93	93	93	93

However, as in the case of model 1, none of the explanatory variables have a correlation coefficient greater than 0.8, and thus the diagnostics do not show a clear violation of linearity assumptions (Cho, Bowker, & Park, 2006) (see methodology).

Table 10 provides the scatterplot as a test for heteroskedasticity. The limited number of observations, as expected, is not sufficient to provide clear results on the variability of the residuals:

Table 10: Scatterplot using Regression Standardized Residuals on x-axis for aggregate analysis (model 2)

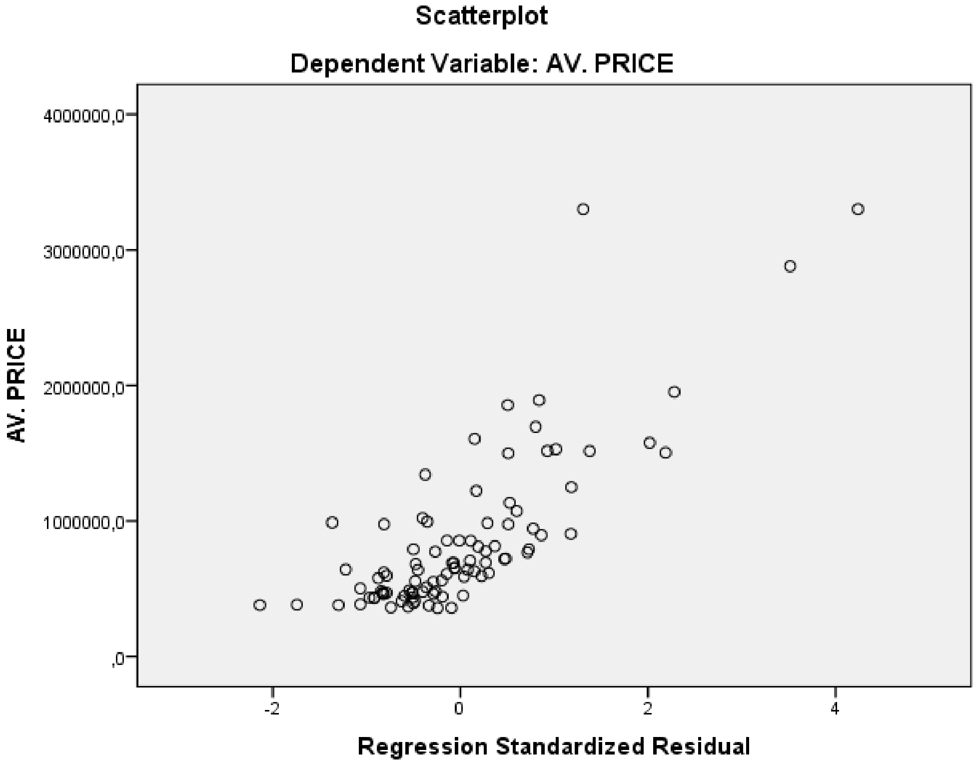


Table 10.1 shows that no significant relationship exists between the unstandardized residuals and the explanatory variables; therefore the possibility of heteroskedasticity can be eliminated:

Table 10.1: White-test for heteroskedasticity for aggregate analysis (model 2)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	72695272498	51899883716		1,401	,183		
	GREENLAND %	-3968091849	3078901548	-,456	-1,289	,218	,463	2,161
	WETLAND %	-2910934846	2861654858	-,273	-1,017	,326	,803	1,245
	EDUCATION	-7091070721	6228915707	-,331	-1,138	,274	,686	1,458
	CBDdis	35228399060	23202026449	,569	1,518	,151	,413	2,420

a. Dependent Variable: UnsResiduals

An interesting observation is that the area of Mayfair, which was included in the study area, appears to be an outlier with respect to its average house price, with an amount of £3,300,435 by far exceeding the mean of all 93 areas (£847,973). In addition, the specific area is covered by 34.1% with large parks, which shows the large impact of green area coverage on average house prices. However, house-specific characteristics such as luxury and average living area have been omitted from model 2, which might be a drawback in the explanatory power of this model. Furthermore, the distance to CBD appears to have contrasting results with model 1, as the variable has a positive effect on individual property values, while it has negative effects on average property values; this result comes in accord with the ecological fallacy theorem, as explained in the methodology (Robinson, 2009).

By analyzing the results of model 2, the following equation can be formed, explaining the relationship between the dependent variable of average house prices and the explanatory variables for the distance to CBD (x_4), green area coverage (x_1), wetland coverage (x_2) and distance to public transport (x_3):

$$AP_k = 926{,}982 + 64{,}841\, x_1 + 11{,}382\, x_2 - 58{,}717\, x_3 - 108{,}643\, x_4$$

Therefore, the **1st hypothesis** of this statistical analysis: "*Average residential prices are positively affected by the percentage coverage of green areas within the region"* cannot be rejected, as by observing *model 2* it can be concluded that green area coverage has a significant positive effect on average residential prices. The statistical analysis does not provide significant results for the **2nd hypothesis:** ***"****Average residential prices are positively affected by the percentage coverage of wetlands within the region"* as the coefficient has a significance of 0.616 (with 0.05 significance level); while for the **3rd hypothesis**: *"Residential prices are negatively affected by the increasing distance to environmental amenities", model 1* provides significant results to reveal a significant negative coefficient.

Chapter 5 – Synthesis

The real effects of the environmental variables included in the statistical analysis should be compared to the results of previously examined relevant researches, in order to draw clear conclusions. The effects of variables which are excluded from environmental determinants (distance to CBD, distance to public transport, Education) are not discussed and compared to previous studies, since this study is clearly focused in environmental variables.

In general, most of the previous studies examined used *property sales value* as the dependent variable for their statistical analysis; with only 2 studies using average and median property values to determine the effect of air quality variables (Harrison & Rubinfeld, 1978; Ridker & Henning, 1968). Therefore, *model 1* as presented in the statistical analysis of this study is more representative to the results of previous researches. In addition, comparisons of *model 2* to previous researches can be implemented on the basis of similarities between the selected explanatory variables, as many studies have used green and water area coverage and size in order to determine house prices; however using individual property values as the dependent variable.

To start with, the coefficient of distance to the nearest green area (£-59,423 per 0.1 mile) as described in the 1st model is in accord with the expected sign (see *Table 4*), and is mostly representative of the results from previous studies which included this variable. Significant negative results have been revealed in studies conducted in Finland and Spain, with coefficients of -41.78FIM and €-1,800 per 100 meters distance respectively (Tyrväinen, 1997; Morancho, 2003); while the study conducted in 2007 in China reveals a negative coefficient of ¥-51.60 (Fanhua, Haiwei, & Nobukazu, 2007).

In addition, the effect of increasing distance to the nearest water area (£-16,132 per 0.1 mile), is in accord with previous relevant studies, such as the study from *Mahan, Polansky & Adams* which reveals a significant positive coefficient of $436.17 per mile of decreasing distance to the nearest wetland (Mahan, Polansky, & Adams, 2000). Futhermore, the study conducted in the area of *New Hampshire* (US) in 2002 which included only lakefront properties, revealing a positive effect of $107.29 per foot of decreasing distance to the lake (Gibbs, Halstead, Boyle, & Ju-Chin, 2002). Only one from the reviewed studies examines the presence of wetlands from a different perspective, revealing a negative effect of $-1034.38 for properties located in flood zone areas (Harrison, Smersh, & Schwartz, 2001); although the effect of flood zone areas cannot be generalized for the effect of wetlands on property values, as not all water areas induce the risk of flood.

However, the relevant studies provide much lower coefficients than *model 1*; a possible explanation could be the large difference in house prices and consumer preferences between different countries. For instance, residents of an economically developed urban area such as London might have a higher willingness to pay for an a cleaner environment than residents of less developed areas. This possibility is in accord with the environmental Kuznet's curve (EKC) theorem, which explains that as income increases, environmental emmisions firstly increase, reach a maximum point and then start decreasing; thus forming an inverse U-shaped curve which explains the effect of economic development on environmental preferences (Stern, 2004). The fact that the distance coefficients have such large differences between different areas questions the external validity of relevant statistical analyses.

The results for the 2nd model reveal a significant positive coefficient of £64,841 on average property values for an additional 1% of greenland coverage within each region. As already mentioned, the reviewed studies have not included average property values to test the effect of green coverage; individual property values have been used instead. The relevant study by *Tyrväinen* in 1997 examines the effect of the size of the nearest green area on individual property values, also revealing a significant positive effect of 7.36FIM per square meter of green area (Tyrväinen, 1997); while the study conducted in Jinan City (China), implicitely selecting green area coverage as one of the explanatory variables, reveals a positive effect of ¥417.78 per 1% greenland coverage (Fanhua, Haiwei, & Nobukazu, 2007).

The variable indicating wetland coverage show a strongly insignificant positive effect of £11,382 on average property values per 1% increase in water areas within each region. While the results of this study do not show a significant relationship between regional water coverage and average property values, 2 relevant studies show a significant positive coefficent of $24.39 per acre and $213.58 per square meter of additional water coverage on nearby individual property values (Mahan, Polansky, & Adams, 2000; Gibbs, Halstead, Boyle, & Ju-Chin, 2002).

As in *model 1*, the 2nd model provides much higher coefficients with respect to environmental amenities than relevant studies in different areas. However, the results of *model 2* cannot be easily compared to previous studies due to differences in the dependent variables. Overall, by comparing the results of this study to previous researches it can be observed that the proximity to environmental amenities such as water and green areas, and the relative size of such amenities as examined, have both positive effects on property values. The methodology of this study (OLS regression) is in accord with the methodologies used in the previous studies examined, while the possibilities of non-linear relationships are eliminated through heteroskedasticity tests and collinearity diagnostics. It can be assumed that the results of relevant studies have weak external validity, due to the great variability of the coefficients. Differences in the sampled population and settings often act as a threat to the external validity of the results; which is important in terms of forecasting. Minimizing the number of explanatory variables and identyfing the most important determinants could improve the external validity of the studies examined, and therefore such models can be used in forecasting equations to explain the behavior of property values with respect to environmental determinants.

Chapter 6 – Conclusion

This research has included a critical analysis of past studies concerning the effects of environmental amenities on residential prices, and created a separate statistical analysis using the most appropriate methodologies according to the reviewed studies. Through an evaluative literature review and an objective statistical analysis, this study aimed to answer its proposed research question:

- *What are the effects of environmental amenities on housing prices?*

By examining several environmental variables included in relevant studies, it can be concluded that the most important determinants which have a significant effect on property values include the proximity of properties to wetlands and green areas, as well as the relative sizes of such amenities. Studies which focused exclusively on the presence of green areas have statistically proved a negative effect of increasing distance to green areas on property values (Luttik, 2000; Morancho, 2003; Tyrväinen, 1997), while others have additionally proved a positive effect of their relative regional coverage (Fanhua, Haiwei, & Nobukazu, 2007). Studies focusing on the presence of wetlands have also proved a negative effect of increasing distance to the nearest water source, and a positive effect of the increasing coverage of wetlands (Mahan, Polansky, & Adams, 2000; Gibbs, Halstead, Boyle, & Ju-Chin, 2002); while other studies have proved contradicting results, revealing a negative effect of the presence of wetlands on property values due to flood risks (Harrison, Smersh, & Schwartz, 2001).

By collecting and classifying data for a strictly urban environment such as the city of London, this study aims to create a representative model in which the real effects of environmental amenities on property values can be examined. Similarly, the statistical analysis of this research studies the effect of distance to environmental amenities such as green areas and wetlands on individual property values (*model 1*), as well as the effect of green and water area coverage on regional average property values (*model 2*). As in accord with the reviewed studies, the results show a significant negative effect of increasing distance to the nearest green and water source (see page 17) on individual property values. In addition, the aggregate analysis reveals a positive effect of green area coverage on regional average property values, while the effect of wetland coverage reveals insignificant results (see page 19).

However, the results of the statistical analysis of this research differ from previous results, as distance and coverage coefficients appear to have much larger values for the city of London than in relevant past researches conducted in Europe, China and the US. In addition, it can be observed that studies conducted in different areas and within a different time-frame, show a large variability in the coefficients of the selected environmental variables. This phenomenon could occur for 2 reasons:

1. The implicit value of environmental amenities increases over time, as environmental awareness and the increasing scarcity of natural resources in urban environments has caused residents to have an increasing willingness to pay for such amenities over

time. This reasoning can be justified by observing the behavior of environmental coefficients included in relevant studies and comparing the coefficients by the time-frame in which each study has been conducted.

2. The implicit value of environmental amenities differs significantly between different areas. As each of the studies examined has been conducted in a different area, it can be assumed that the results of each study are not externally valid to different populations and settings. For instance, residents of an economically developed country/area may have a relatively higher willingness to pay for such environmental amenities, while residents of developing or less economically developed countries might have in contrast a lower willingness to pay. External invalidity could additionally act as a threat to the forecasting power of such statistical models.

Furthermore, as the data used for the statistical analysis did not provide details on house-specific features apart from the number of rooms, the effect of such variables was partly eliminated by including only single-bedroom apartments for the individual analysis. However this method did not eliminate additional variables which can affect property values such as the level of luxury, or each property's construction date.

Overall, it can be concluded that urban residents distribute a rapidly increasing implicit value on the presence of environmental amenities in close proximity to their properties. Therefore, by considering the main results of this research and the previous studies examined, several policies can be recommended to urban developers. For instance, the presence of at least one public green area (park) including a water source (pond/fountain) for each residential block could better match the preferences of urban residents and therefore capture consumer surplus, as residents have a higher willingness to pay for properties with such amenities. Additionally, direct access of apartment complexes to public green areas can create a rapid increase on property values. The presence of at least one large public green area in each region can also improve property markets and provide residents with additional values such as access to recreational areas, improved air-quality and aesthetics. The promotion of natural areas within urban regions could also improve biodiversity, and therefore creating a social positive externality to which residents could positively respond by fulfilling their demand for a cleaner environment. By investing in such alternative land uses within urban regions, developers can improve both the living standards of residents, and stimulate the urban economy by means of creating more environmentally-friendly residential areas.

Bibliography

Boyle, M. A., & Kiel, K. A. (2001). A survey of house price hedonic studies of the impact of environmental externalities. *Journal of Real Estate Literature; 9, 2*, 117-144.

Cho, S.-H., Bowker, J. M., & Park, W. M. (2006). Measuring the Contribution of Water and Green Space Amenities to Housing Values: An Application and Comparison of Spatially Weighted Hedonic Models. *Journal of Agricultural and Resource Economics, Vol. 31, No. 3*, 485-507.

Clark, W. A., & Onaka, J. L. (1983). Life Cycle and Housing Adjustment as Explanations of Residential Mobility. *Urban Studies*, 50.

Dombrow, J., Rodriguez, M., & Sirmans, C. F. (2000). The Market Value of Mature Trees in Single-Family Housing Markets. *Appraisal Journal 68(1)*, 39-43.

Fanhua, K., Haiwei, Y., & Nobukazu, N. (2007). Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning 79 (2007) 240–252*, 240 - 252.

Gibbs, J. P., Halstead, J. M., Boyle, K. J., & Ju-Chin, H. (2002). An hedonic analysis of the effects of lake water clarity on New Hampshire lakefront properties. *Agricultural and Resource Economics Review; Apr 2002; 31, 1*, 39-46.

Harrison, D., & Rubinfeld, D. L. (1978). Hedonic housing prices and the demand for clean air. *Journal of environmental economics and management, 5(1)*, 81-102.

Harrison, D., Smersh, G. T., & Schwartz, A. (2001). Environmental determinants of housing prices: the impact of flood zone status. *Journal of Real Estate Research, 21(1-2)*, 3-20.

Judge, G., Hill, C., Griffiths, W., Lütkepohl, H., & Lee, T. (1982). Introduction to the Theory and Practice of Econometrics. New York: John Wiley & Sons,.

Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the Netherlands. *Landscape and Urban Planning 48 (2000) 161-167*, 161-167.

Mahan, B. L., Polansky, S., & Adams, R. M. (2000). Valuing Urban Wetlands: A Property Price Approach. *Land Economics, Vol. 76, No. 1 (Feb., 2000), pp. 100-113*, 100 - 113.

Morancho, A. B. (2003). A hedonic valuation of urban green areas. *Landscape and Urban Planning 66 (2003) 35–41*, 35 - 41.

Palmquist, R. B. (1982). Measuring environmental effects on property values without hedonic regressions. *Journal of Urban Economics, 11(3)*, 333-347.

PHPTools. (n.d.). *Image Color Extract*. Retrieved May 13, 2016, from Tools To Make Your Life Easier: http://www.coolphptools.com/color_extract

Ridker, R. G., & Henning, A. J. (1968). The Determinants of Residential Property Values with Special Reference to Air Pollution. . *The Review of Economics and Statistics, 49(2)*, 246–257.

Riggins, J. (2011, October 12). *ZD Net*. Retrieved April 28, 2016, from From highway to green space in Madrid: http://www.zdnet.com/article/from-highway-to-green-space-in-madrid/

Rightmove. (2016). *Find your happy Search properties for sale and to rent in the UK*. Retrieved April 18, 2016, from House Prices in City Of London: http://www.rightmove.co.uk/house-prices-in-City-Of-London.html

Robinson, W. S. (2009). Ecological correlations and the behavior of individuals. *International journal of epidemiology 38(2)*, 337-341.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy, Vol. 82, No. 1 (Jan. - Feb., 1974), pp. 34-55*, 34 - 55.

Sirmans, S., Macpherson, D., & Zietz, E. (2005). The composition of hedonic pricing models. *Journal of real estate literature, 13(1)*, 1-44.

Stat-Trek. (2016). *Stat Trek Teach yourself statistics*. Retrieved May 18, 2016, from Statistics and Probability Dictionary: http://stattrek.com/statistics/dictionary.aspx?definition=Normal%20distribution

Stern, D. I. (2004). The Rise and Fall of the Environmental Kuznet's Curve. *World Development Vol. 32, No. 8*, 1419–1439.

Tyrväinen, L. (1997). The amenity value of the urban forest: an application of the hedonic pricing method. *Landscape and Urban Planning 37 (1997) 211 - 222*, 211 - 222.

Wabe, S. J. (1971). A study of house prices as a means of establishing the value of journey time, the rate of time preference and the valuation of some aspects of environment in the London metropolitan region. . *Applied Economics, 3(4)*, 247-255.

Zabel, J. E., & Kiel, K. A. (2000). Estimating the demand for air quality in four US cities. *Land Economics (2000)*, 174-194.

Zietz, J., Zietz, E. N., & Sirmans, S. G. (2008). Determinants of house prices: a quantile regression approach. *The Journal of Real Estate Finance and Economics, 37(4)*, 317-333.