The Economic Impact of Pocket Parks on Residential Property values: Case study-Madison Wisconsin.

By

Mbaka Mourine Mwende

A Thesis Submitted in

Partial Fulfillment of the

Requirement for the Degree of

Master of Science Applied Economics

At

The University of Wisconsin-Whitewater

December 2018

Graduate Studies

The members of the Committee approve the thesis of

Mbaka Mourine Mwende presented on 12/11/2018

____

Dr. Russ Kashian, Chair

____

Dr. Matthew Winden

____

Dr. Ran Tao

Dedication

To my father Lawrence Mbaka for believing in me and working against all odds to ensure I received an education.

Acknowledgements

First, I extend my gratitude to my husband Kevin and my daughters Sifa and Tunu. Thank you for always being on my side and believing in my dreams and walking the journey with me. To Caroline, Ann and Emma, I could not have had better friends for those days that I just needed someone to watch my little ones while I did school work. You were always a phone call away. I appreciate my professors at UW-Whitewater for your guidance and always, with enthusiasm, being ready to guide me through my education journey at UW-Whitewater. I thank all of my friends and family who helped me in one way or another with this undertaking. I had a community cheering me on. Finally, I thank my supervisor, Russ Kashian. Thank you for your guidance, I learned and accomplished more than I thought I could

The Economic Impact of Pocket Parks on Residential Property values: Case study-
Madison Wisconsin.

By

Mbaka Mourine, Mwende

The University of Wisconsin- Whitewater, Year 2018

Under supervision of Dr Russ Kashian

Abstract

When value of a property is assessed, the true determinants of the value of that property is in the attributes attached to that property. What individuals are willing to pay for is the attribute package that will give the maximum utility. Hedonic modelling has been used to appraise the value attached to this attribute. In this study, we use hedonic analysis to determine the value attached to pocket parks when valuing residential property. Our findings show that parks size does have value and it is translated in the value of property. The magnitude of the impact is biggest on parks that are not more than 1.5 acres. For this park sizes, houses near parks that are 0.1 acres bigger value 129 dollars more, above 1.5 acres the magnitude to the impact on house value decreases. Proximity to the park also does increase the value of a house. Also, the further the distance from the house to the nearest park, the lower the value of that house. Therefore, it is important to preserve open spaces even in metropolitan areas for their value on property prices. This increased value equally increases income of home owners and also this benefits the local government because as the home value increase, property tax they receive from these homes go up.

Contents

Chapter One

Introduction

National Recreation and Park Association (NRPA) defines a pocket park as a "small outdoor space that is usually not more than ¼ of an acre, usually only a few house lots in size or smaller, most often located in an urban area surrounded by commercial buildings or houses on small lots with few places for people to gather, relax, or to enjoy outdoors." Pocket parks are also known as vest parks. A pocket park is the smallest park in park classification and serves an area of approximately ¼ miles radius from the park. Unlike other parks, a pocket park is intended to be used and designed to meet the specific interest of the community nearby (Olmos 2008). According to APS 2015 pocket parks are not only important for their aesthetic value of urban beautification and recreation spaces, but these open spaces also "significantly define the layout, real estate value traffic flow, public events and the civic cultures of our communities." Such spaces are especially important breathing spaces in the increasingly congested cities.

Rapid increase in urbanization has resulted to decrease in the general access to parks and other natural areas. Van de Berg et al (2007) points out that the hectic urban life, may motivate people to more to the suburbs where they can connect more with nature. Van de Berg's study contradicts the reality because we continue to see increase rather than decrease in the urban population. The increasing urban population, and diminished land spaces for development in cities has occasioned an unprecedented penetration of businesses, companies and real estate developments into the once capacious suburb areas. The city planners and developers have resorted to creating small open spaces often referred to as pocket parks hopefully to bring at least some level of utility, if not an equivalent level of utility, which the urban residents would otherwise not get by living in a place where there is less connection to nature.

To determine the impact that pocket parks have on property value, which is the purpose of this paper, we would first need to come up with a price index which are quite complex to construct to the great variation in properties. Houses are much differentiated, that is, two houses can be very different and sell at entirely different prices because of the difference in attributes. Although hedonic analysis has been used in other fields, it is widely used in real estate to determine the market value of house characteristics such as the number of bedrooms, size of the house in square feet, number of bathrooms among other characteristics. Court (1939) who is widely considered as the founder of hedonic analysis, constructed a hedonic price index for automobiles. Rosen (1974) defines hedonic prices as “the implicit prices of attributes and are revealed to economic agents from observed prices of differentiated products and the specific amounts of characteristics associated with them.” Hedonic prices are validated by the fact that product have different attributes and each of these attributes have a value attached to them. Hedonic regression breaks down an “aggregate price” into various “attribute prices.” This paper uses hedonic analysis to the prices of homes in Madison to deconstruct these prices and determine the value, if any, that is attached to the pocket parks and consequently draw conclusions on the impact pocket parks have on property value in Madison.

Research Objectives

The goal of this study is to estimate the economic impact of pocket park on property value using Metropolitan Madison, Wisconsin, as the case study. Kitchen and Hendon (1967) research on "Land Values Adjacent to an Urban Neighborhood Park" is one of the earliest studies on the relationship between parks and house/land value. Using data from Minneapolis, they show that among the characteristics that have influence on property/land prices, is the distance from a park. The value of land decreased as the distance of the land from the park increased. McMullen (2011) found that in Gainesville Florida the proximity of the property to the park had some impact on the value of that property. He adds that hedonic analysis can be utilized to get the value of parks as an added value to residential property prices. Other past literature very consistently, with a few exceptions, agree that buyers and renters are often willing to pay a higher amount for a house/property that is near a park than one that is not close to one (Crompton et al 2000, Nicholls and Crompton 2017, Burgess et al 1988). House around a captivating view fetch a higher price than similar houses in a less captivating view. Luttik (2000) found that the presence of an open space resulted in an increased house price by approximately 6-12%.

It is important to recognize that there has been a rapid increase in urbanization which has resulted to decrease in the general access to parks and other natural resources. Valuing urban open spaces/ pocket parks is very important because this acts as a motivation to the investors to ensure they create these green spaces in the urban areas and thus ensuring that despite the high demand for land in the urban areas there is the very much needed green outdoor space. More et al (1988) argues that the vulnerability of green outdoor space in urban areas can be partly explained by the fact that the green spaces have not been valued in monetary terms.

Chapter Two

Literature Review

Hedonic analysis can be traced back to Haas (1922) and Wallace (1926) who used it to value farmland. Court (1939) is however, widely referenced for using hedonic analysis in his paper "Hedonic price indexes with automotive examples, in 'The Dynamics of Automobile Demand.'" He estimates the price of automobile as a function of the automobile attributes. Court found that the price of an automobile was an aggregate of the prices of the attributes found in that automobile. Griliches (1961) expounds on how the implicit qualities of a product affect its price. He established that adjusting for quality-change is worthwhile and can greatly influence the price of to a product. Hedonic analysis is based on the idea that goods are heterogeneous rather than homogenous. That is, goods are differentiated, and they can be thought of as a package with various implicit attributes.

Lancaster (1966) introduces hedonic modelling to consumer theory. He argues that commodities do not provide any utility to the consumer. Instead, the attributes of a given product are the ones that provide the utility. He further argues that goods are, in general, a combination of many attributes and it is possible for attributes to be shared across goods. Rosen (1974) provided a framework which has been widely used to explain the impact these implicit attributes has on demand and supply of products such as houses. His work has been a significant theoretical basis for many research studies that concentrate on analyzing differentiated goods, such as houses, that have attributes that do not have observable market prices. The hypothesis for his model of differentiated products is that consumers pay for the utility they get from commodities and not the product itself. Assuming heterogeneity of products, therefore, we can say that consumers maximize their utility by consuming a commodity with the highest level of aggregate utility, which is a function of utilities from the characteristics of that commodity.

Many scholars have used hedonic modeling in their studies to understand, for instance, what brings variation in real property prices. Palmquist (1984) and Jansen et al (2001) show that the price attached to a differentiated good is implicitly inclusive of the price (hedonic price) attached to the specific attributes. Li and Brown (1978) found that the location of houses had an impact on the value attached to that house. Houses that are more accessible are valued higher than those that are not easily accessible. Furthermore, external diseconomies such as congestion and pollution decreased the value of the houses.

For instance, Goodman (1978) compares city and suburb house markets using hedonic modeling and finds city houses value at least 20% more than the suburb counterparts. Structural improvements on city houses had higher valuation than similar improvements in the suburb houses. Essentially, therefore, the price attached to a property it is not influenced just by the qualitative and quantitative attributes found specifically on that property but also by other environmental amenities and disamenities. The impact of disamenities such as air pollution and the impact this has on the property values have been studied (Harrison and Rubinfeld 1978, Small 1975, Ridker and Henning 1992, Chay et al 2005), as well as water quality (Leggett and

Bockstael 2000, Michael, Boyle and Bouchard 1996, Krysel et al 2003) among other environmental features.

Other features that have been studied in relation to property prices include school quality, plants, presence of a railway line. Past research has been based on the assumption that in the short run the supply of houses is fixed. This implies that the consumers demand for the existing houses determine the value of those houses. In modeling for new houses, their value would be determined by both the cost of including the characteristics contained in a house and other factor that bring the firm to profit optimization (Palmquist 2005).

Foundation of my Model

The foundation of my model borrows from the work by Rosen (1974) and Sheppard (1999). Considering a plane with various dimensions of equilibriums that sellers and buyers can choose from and a collection of houses with *n* measured characteristics. Any location on the plane represents a vector of coordinates such that $z = (z_1, z_2, \dots \dots \dots \dots \dots . z_n)$. z_i is the *i*th characteristic in each house. The houses in the collection are exhaustively described by the numerical values of *z* and the buyers choose from the resulting unique set of characteristics. The fact that the houses are differentiated implies that there is a variety of other options that a consumer can pick from. We also assume that the price of a good, house, is a function of the characteristics embodied in that good, house, such that P $(z) = P(z_1, z_2, \dots \dots \dots \dots \dots . z_n)$. This is a hedonic price function. This price function gives the minimum price value for the houses with similar characteristics. A seller cannot sell for more because the buyer then opts for a similar package but less expensive. The consumers—in this case, the persons purchasing the house—are price takers and cannot influence the equilibrium prices of the houses.

The hedonic price function has two main functions, first, it is used in generating the overall price indices and second as an input in analyzing consumer demand for the characteristics embodied in heterogeneous goods.

To derive consumer utility function, we assume that the consumer is deriving utility from consuming a good that contains a vector Z of N measured characteristics in addition to consumption of a composite good X. The consumer is also faced with a limited income Y and faces a price function P(Z), price of the heterogeneous good, house, which is a function of the characteristics embodied in that property.

The preferences of the household can be represented by the utility function. We assume the utility function to be strictly concave, in addition to adhering to other assumptions.

$$u = u(Z, X, \alpha) \qquad \text{equation1}$$

where Z is a vector of characteristics, X is the composite good and α is a vector of parameters that characterize the preferences of the household. We assume that the households are fully characterized by an income M and vector of parameters α.

We now can derive the expenditure a household is willing to pay using the utility function above such that:

the ‘bid house price’ function $\beta(Z, Y, u, \alpha)$ is defined implicitly as

$$u = u(Z, Y - \beta,\ \alpha) \qquad \text{equation 2}$$

We then differentiate 'bid house price' function $\frac{\partial \beta}{\partial Z_i}$ to get how a household's willingness to pay for a house changes as the characteristic *i* changes while holding the utility level constant.

To develop the budget constraint, I set the price of the composite good X to unity such that household's income Y is split between goods Z and X so that $Y \geq P(Z) + X$.

$max_{Z,X}\ u(Z, X, \alpha)$ subject to Y≥P(Z) +X equation3

Maximizing utility requires that firs order conditions be met such that

$\frac{u_i}{u_X} = P_i\ \forall$ i equation 4

where u_i and u_X are partial derivatives such that $u_i = \frac{\partial u}{\partial X_i}$ and $P_i = \frac{\partial p}{\partial X_i}$. P_i is the hedonic price of characteristic i, and P(Z) is the hedonic price function.

Combining the differentiation of the implicit function equation 2 and the first order conditions equation 4 we get

$\frac{\partial \beta}{\partial Z_i} = \frac{u_i}{u_Y} = P_i$ equation 5

The above equation (5) shows that the optimal choice of the household is achieved when the slope of the bid 'price' and the hedonic price for each characteristic are equal. This is the backbone of hedonic analysis. Equation 5 prove that we can estimate the hedonic price for an attribute and the choice made by the household. We then assume using the optimizing behavior of the observations we can get the household's willingness to pay for the attributes that border the observed choice. This makes the problem similar to one with the standard consumer behavior where the observed consumer choices and the market prices are used to determine consumer preferences.

To get a clearer understanding of the hedonic price function we bring in the production side of the equation which is characterized by the cost function.

$C(Z, J, \theta)$ equation 6

where Z is a vector of N characteristics of houses supplied, J is the houses built and θ is a vector of parameters chacterizing the producers. The producer's profit function can be written out as

$\pi = P(Z) * N - C(Z, J, \theta)$ equation 7.

The producers have a probability density. They maximize their profits

$max_{Z,J}$ P(Z)*N-C (Z, J,θ) equation 8.

The first order conditions require that

$P_i = C_i \ \forall i$

equation 9

P(Z)= C_N

At the optimal the marginal cost of every characteristic is equivalent to the corresponding hedonic price of that characteristic. Also, the producers will build houses until the point where the marginal cost of building a specific house (of type Z0 is equal to the value of such house P(Z). Thus, a competitive market with a specific type of producers such that N=1 and the cost function C(.) would be determined by the costs and technology of house remodeling and repair.

In general, we assume that at equilibrium the quantities of houses supplied is equal to the quantities of houses supplied and consequently marginal cost of a characteristic *i* in house Z is equivalent to the marginal bid value of that characteristic and the hedonic price of that characteristic.

Chapter Three

Data: Madison City Parks

This study uses data from Madison which is the capital city and the second largest city in Wisconsin to determine the impact pocket parks have on residential property value. According to DATAUSA, Madison has a population of 252,557 on an area of 94.03 square miles. The median property value is 236,100 dollars. The average age is 31.2 and there are 109, 546 households, 110, 540 housing units. Madison is known for its lakes, extensive network of parks, bike trails among other things.

The data used in this study consist of two datasets, residential properties in Madison city and parks and open spaces within Madison Boundaries. The residential properties data was acquired from Wisconsin Department of Revenue. Data on parks was created from Madison City Parks website which had detailed information on all the parks in Madison. This data consists of 237 parks within the city which are classified into various categories; mini, open space, special, neighborhood, community, sports field and conservancy parks. The parks vary in acreage starting from 0.21 acres to 916.03 acres. For this study the focus will be on the size of the park in acreage rather than the type of the park. The size of the park in acreage is more convenient since it allows us to re-classify the parks again according to their size and as a result parks that would have been ruled out as big such as neighborhood parks are put into account. This work also puts interest in the distance to the park in miles and this ranges from 0.0005 miles to 19.7716 miles.

The variables in the data set can be categorized into property and location, home and transactions characteristics. The property and location characteristics include: a dummy on the house being on water frontage, size of the home in acres, distance to the nearest park in miles. Home characteristics are number of bedrooms, number of full baths, number of half baths, and the total living area and age of the house. Transaction characteristics include: current value of the house and transaction price for the house. There are no missing observations. For all the modelling, home price is used as the dependent variable. The rest of the variable are used as independent variables. We expect number of bedrooms, full baths, half baths, total living area, size of the home in acres to have a positive correlation with current value of the house. It is also expected that the dummy variable on the house being on the water frontage to have a positive correlation with the current value of the house. Age of the house, on the other hand, is expected to have a negative correlation with the current value of the house.

The data also presents use with variables, size of the park in acres. We disintegrate this variable into five different park size variables where; park_ size _1 includes the parks that are 1.5 acres or less, park_size_2 includes the parks that have more than 1.5 acres but not more than 3.5 acres, park_size_3 are parks that have more than 3.5 acres but not more than 7.5 acres, park_size_4 are parks that have more than 7.5 acres but not more than 20 acres, park_size_5 are parks that have more than 20 acres. We also disintegrate the variable, distance to the nearest park into five different distance variables such that; park_dist_1 is for houses that are 0.5miles or less from the park, park_dist_2 is for houses that are more than 0.5 but not more than 5.5 miles from the park, park_dist_3 is for houses that are more than 5.5 but not more than 9 miles from the park, park_dist_4 is for houses that are more than 9 but not more than 12.5 miles from the park, park_dist_5 is for houses that are more than 12.5 miles from the nearest park. We then generate a park-acres-distance variable which combines the park size variable and the distance from the park. This helps us to see how the size and distance from the nearest park simultaneously affect the current value of houses.\

Findings and Results

The primary purpose of this study is to establish the impact the pocket parks have on the market value of houses if any. I have divided the parks according to different sizes to determine how size of a park impacts the value of the houses near it. The first regression is focused on park that are not more than ¼ acre. This is necessary because according to NRPA, a pocket park is not more than ¼ acre. Sorting the data according to the park size, that is not more than ¼ acre, resulted in a dataset with 8,708 houses. The average house price is 257,534 dollars. The average size of the park is 0.21 acres

Table 1.1: Descriptive statistics for house near a park less than or equal to ¼ acre.
Number of observations 8,708

Variable	**Mean**	**Std Deviation**	**Minimum**	**Maximum**
Park size in acres	0.21	0	0.21	0.2
Distance to park in miles	5.9163	3.623	0.0062	13.8371
Water frontage	0.0091	0.0948	0	1
Home price	257,534	118,523.9	27000	1,600,000
Bedrooms	3.1311	0.8824	1	9
Full baths	1.8018	0.7187	1	6
Half baths	0.4821	0.5656	0	3
Total living area	1482.084	534.6803	432	6285
Age of the house	47.8529	28.8281	1	170

Without in interactions

Table 1.2: Linear and semi log regression output of houses within parks that are not more than ¼acres

	Linear regression		Semi log regression	
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**
Park size in acres	Omitted		Omitted	
Distance to park in miles	-6,956.981***	-31.63	-0.0216***	-27.55
Water frontage	359,721.9***	45.34	0.6313***	21.49
Bedrooms	-1,563.27	-1.47	0.0236***	6.34
Full baths	23,958.29***	15.88	0.1051***	19.48
Half baths	6,823.163***	4.31	0.0227***	3.97
Total living area	101.8646***	18.36	0.0008***	39.02
Total living area squared	0.0042***	3.39	-9.72e-08***	-22.10
Age	-138.6585	-1.41	0.0031***	26.58
Age squared	8.4097***	10.39	-1.81e06***	-31.96
Intercept	72,738.46***	11.39	11.207***	563.95

The linear regression analysis indicates an adjusted R-square of 0.6645. This implies that the model can explain 66.45% of the house price variation. All the variables with an exception of number of bedroom and age of the house, are found to be significant at 0.01 level. Water frontage had the biggest t-statistic, 45.34. This implies it has the highest influence in the model. A house that has a water frontage is expected to be approximately 359,722 dollars more expensive than a house with similar characteristics but not on water frontage. Distance to the nearest park had a negative coefficient implying that increasing the distance from the park by 1 mile would decrease a house value by approximately 6, 957 dollars holding all other factors constant. In this regression or variable of interest was omitted due to collinearity.

The semi log regression has explained 61.08% of the variation in the house prices. All the variables are significant at 0.01 level. Just like in the linear regression the park size variable has been omitted due to collinearity. The number of bedrooms now have a positive effect on the house price, adding one more bedroom to a house would increase its value by 2.36% holding all other factors constant. A house that has a water frontage is expected to be 63.13% more pricy than a similar house that does not have water frontage. A house that is a mile further from the park than a similar house would value at 2.16% less.

With interactions

Table1. 3: Linear and semi log regression output of houses within parks that are not more than ¼acres (interactions).

	Linear regression		Semi log regression	
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**
Park size in acres	Omitted		Omitted	
Park size squared	Omitted		Omitted	
Park size *distance	Omitted		Omitted	
Distance to park in miles	-6,956.981***	-31.63	-0.02489***	-32.32
Water frontage	359,721.9***	45.34	0.6644***	23.92
Bedrooms	-1,563.27	-1.47	0.0241***	6.48
Full baths	23,958.29***	15.88	0.0926***	17.54
Half baths	6,823.163***	4.31	0.0237***	4.26
Total living area	101.8646***	18.36	0.0007***	34.52
Total living area squared	0.0042***	3.39	-8.18e-08***	-18.66
Age	-138.6585	-1.41	-0.0012***	-3.55
Age squared	8.4097***	10.39	0.00003***	11.90
Intercept	72,738.46***	11.39	11.4278	510.86

Linear regression; *R-squared =0.6645* *Adjusted R-squared=0.6641*
Semi log regression: R-squared =0.6227 *Adjusted R-squared=0.6224*
**significant at 0.1 level* ***significant at 0.05 level* ****significant at 0.001 level*

The semi log regression has explained 62.27% of the variation in the house prices. All the variables are significant at 0.01 level. Just like in the linear regression the park size variable has been omitted due to collinearity. The number of bedrooms now have a positive effect on the house price, adding one more bedroom to a house would increase its value by 2.411% holding all other factors constant. A house that has a water frontage is expected to be 66.44% more pricy than a similar house that does not have water frontage. A house that is a mile further from the park than a similar house would value at 2.49% less.

At 1.5 Acres

Here I have further increased the size of the park to 1.5 acres. This has increased the size of the dataset from 8,708 houses to 154,630 houses. The average house price is maintained at 257,534 dollars. The average park size is now 0.8193 acres with a minimum of 0.21 acres and maximum of 1.44 acres and the distanced to the nearest park is on average 6.237 miles with a minimum of 0.0014 miles and maximum of 15.983 miles.

Table 2.1: Descriptive statistics for house near a park less than or equal to 1.5 ac

Variable	**Mean**	**Std Deviation**	**Minimum**	**Maximum**
Park size in acres	0.8193	0.3912	0.21	1.44
Distance to park in miles	6.2367	3.481	0.0014	15.983
Water frontage	0.0091	0.0948	0	1
Home price	257,534	118,517.6	27000	1,600,000
Bedrooms	3.1311	0.8824	1	9
Full baths	1.8018	0.7187	1	6
Half baths	0.4821	0.5656	0	3
Total living area	1482.084	534.6822	432	6285
Age of the house	47.8529	28.8281	1	170

<u>Without interactions</u>

Table 2.2: Linear and semi log regression output of houses within parks not more than 1.5 acres.

	Linear regression		Semi log regression	
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**
Park size in acres	1,287.36**	2.47	0.0045**	2.46
Distance to park in miles	-4,172.619***	-66.27	-0.0146***	-66.08
Water frontage	352,303.7***	160.30	0.6379***	82.70
Bedrooms	-3,591.762***	-12.23	0.0168***	16.28
Full baths	26,209.97***	62.79	0.1008***	68.76
Half baths	8,680.335***	19.79	0.0303***	19.70
Total living area	117.3095***	76.75	0.0007***	135.39
Total living area squared	0.0018***	5.33	-9.04e-08***	-74.50
Age	363.3364***	13.60	0.0006***	6.44
Age squared	4.8651***	22.02	0.00002***	26.92
Intercept	28,136.02***	15.85	11.2649***	18.8.50

Linear regression: R-squared =0.6389 Adjusted R-squared=0.6389
Semi log regression: R-squared =0.5921 Adjusted R-squared=0.5920
**significant at 0.1 level **significant at 0.05 level ***significant at 0.001 level*

The adjusted R-square for the second linear regression is lower than the first one. The linear model has explained 63.89% of the variations in the house prices. The semi log regression has a lower adjusted R-square (0.5920) in comparison with the earlier semi log model. In the linear model water frontage has the highest t-statistics while in the semi log model total living area has the highest t-statistics. All variables are significant in both models. Park size variable is significant at 0.05 level in linear and the semi log model. The linear model indicates that a house near a park that 1 acre bigger in size will value 1,287 dollars more than a similar house located near a park that 1 acre smaller. The semi log model shows if we have two houses that are similar in all aspects, but one is located near a park that is 1 acre bigger, that house will value at 0.45% more that its counterpart.

The distance from the nearest park has an inverse relationship with the house value. A house that is a mile closer to the park will value 4,173 dollars (linear regression) than a similar house that is further from the park. According to the semi log regression increasing the total living area by 1 square foot would cause the price of the house to go up by 0.07% holding all other factors constant. The number of bedrooms continue to exhibit a negative coefficient, suggesting that the more the bedrooms the lower the house value. This could be an indication that the model might not be completely reliable. Age variable is also not consistent with the expected result. I would expect that house that are older to value less than newer houses, the model however depict the older houses to be more expensive.

<u>With interactions</u>

The adjusted R-square for the second linear regression is lower than the first one. The linear model has explained 63.89% of the variations in the house prices. The semi log regression has a lower adjusted R-square (0.5920) in comparison with the earlier semi log model. In the linear model water frontage has the highest t-statistics while in the semi log model total living area has the highest t-statistics.

Table 2.3: Linear and semi log regression output of houses within parks not more than 1.5 acres.

	Linear regression		Semi log regression	
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**
Park size in acres	4,374.267*	1.50	0.0196*	1.91
Park Size squared	2,670.104*	1.63	0.0098*	1.71
Park size *distance	-1211.385***	-8.63	-0.0051***	-10.26
Distance to park in miles	-3116.087***	-22.60	-0.0102***	-21.07
Water frontage	352,526***	160.44	0.6389***	82.85
Bedrooms	-3,557.439***	-12.12	0.0169***	16.43
Full baths	26,141.71***	62.63	0.1005***	68.59
Half baths	8,618.665***	19.66	0.0301***	19.53
Total living area	117.0818***	76.61	0.0007***	135.25
Total living area squared	0.0019***	5.44	-9.03e-08***	-74.39
Age	35.3524***	13.67	0.0006***	6.54
Age squared	4.8966***	22.16	0.00002***	27.09
Intercept	23,157.02***	10.89	11.2433***	1505.98

Linear regression: R-squared =0.6389 *Adjusted R-squared=0.6389*

Semi log regression: R-squared =0.5921 *Adjusted R-squared=0.5920*

**significant at 0.1 level* ***significant at 0.05 level* ****significant at 0.001 level*

All variables are significant in both models with an exception of park size variables which are insignificant both in the linear and the semi log model. Distance from the nearest park has an inverse relationship with the house value. A house that is a mile closer to the park will value 4,173 dollars (linear regression) than a similar house that is further from the park holding all other factors constant. According to the semi log regression increasing the total living area by 1 square foot would cause the price of the house to go up by 0.07% holding all other factors constant. The number of bedrooms continue to exhibit a negative coefficient, suggesting that the more the bedrooms the lower the house value. This could be an indication that the model might not be completely reliable. Age variable is also not consistent with the expected result. I would expect that house that are older to value less than newer houses, the model however depict the older houses to be more expensive.

<u>At 2.5 acres</u>

The model has further been expended to include parks sizes up to 2.5 acres. The data set now includes 235,116 houses with house prices ranging from 27,000 dollars to 1,600,000 dollars and an average price of 257,534 dollars. The park size ranges from 0.21 acres to 2.47 with a mean of 1.35 acres.

Table 3.1: Descriptive statistics for house near a park less than or equal to 2.5 acres.

Variable	**Mean**	**Std Deviation**	**Minimum**	**Maximum**
Park size in acres	1.3533	0.6560	0.21	2.47
Distance to park in miles	6.4518	3.5609	0.0005	18. 69
Water frontage	0.0091	0.0948	0	1
Home price	257,534	118,517.6	27000	1,600,000
Bedrooms	3.1311	0.8824	1	9
Full baths	1.8018	0.7187	1	6
Half baths	0.4821	0.5656	0	3
Total living area	1482.084	534.6822	432	6285
Age of the house	47.8529	28.8281	1	170

Without interactions

Table 3.2: Linear and semi log regression output of houses within parks not more than 2.5 acres.

	Linear regression		Semi log regression	
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**
Park size in acres	831.8921***	3.65	0.0027***	3.36
Distance to park in miles	-1,857.137***	-40.94	-0.0060***	-37.70
Water frontage	351,565.3***	218.98	0.6357***	82.70
Bedrooms	-4,390.463***	-20.49	0.0140***	18.54
Full baths	26,940.73***	88.39	0.1033***	96.49
Half baths	9,182.222***	28.67	0.0320***	28.47
Total living area	122.9447***	110.34	0.0007***	190.73
Total living area squared	0.0010***	4.05	-9.33e-08***	-105.28
Age	611.7526***	31.62	0.0015***	22.24
Age squared	3.4806***	21.68	0.00002***	28.20
Intercept	1,089.411***	0.86	11.1659***	2509.95

Linear regression: R-squared =0.6285 Adjusted R-squared=0.6285
Semi log regression: R-squared =0.5800 Adjusted R-squared=0.5800
**significant at 0.1 level **significant at 0.05 level ***significant at 0.001 level*

The adjusted R-square for the linear regression is 0.6285. The linear model has explained 62.85% of the variations in the house prices. The semi log regression has a lower adjusted R-square (0.5800) in comparison with the earlier semi log model. In the linear model water frontage has the highest t-statistics while in the semi log model total living area has the highest t-statistics. All variables are significant in both models at 0.01 level. The linear model indicates that a house near a park that is 1 acre bigger in size will value approximately 832 dollars more than a similar house located near a park that 1 acre smaller. The semi log model shows if we have two houses that are similar in all aspects, but one is located near a park that is 1 acre bigger, that house will value at 0.27% more that its counterpart. Distance from the nearest park has an inverse relationship with the house value. A house that is a mile closer to the park will value approximately 1,857 dollars (linear regression) than a similar house that is further from the park.

According to the semi log regression increasing the total living area by 1 square foot would cause the price of the house to go up by 0.07% holding all other factors constant which is similar with the earlier regression. The number of bedrooms continue to exhibit a negative coefficient only on the linear regression. The semi log regression indicates a positive relationship suggesting that the more the bedrooms the lower the house value, increasing the number of bedrooms in a house by one would result in 1.4% increase in the value of the house holding all other factors constant. Age variable is also not consistent with the expected result. I would expect that house that are older to value less than newer houses, the model however depict the older houses to be more expensive.

<u>With interactions</u>

Table 3.3: Linear and semi log regression output of houses within parks not more than 2.5 acres.

	Linear regression		Semi log regression		
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**	
Park size in acres	-14,274.77***	-13.33	-0.0518***	-13.78	
Park size squared	-583.1608	-1.55	-0.0023*	-1.75	
Park acre*distance	2569.646	40.56	0.0094***	42.07	
Distance to park in miles	-5,532.571***	-54.65	-0.0194***	-54.55	
Water frontage	352,089.7***	220.07	0.6377***	113.53	
Bedrooms	-4,192.557***	-19.63	0.0147***	19.57	
Full baths	26,703.63***	87.90	0.1024***	96.02	
Half baths	8,989.935***	28.16	0.0313***	27.85	
Total living area	121.3657***	109.23	0.0007***	189.86	
Total living area squared	0.0013***	5.05	-9.24e-08***	-104.63	
Age	571.5202***	29.61	0.0014***	20.14	
Age squared	3.7495***	23.42	0.00002***	30.02	
Intercept	26,341.41***	17.70	11.2576***	2155.00	

Linear regression: R-squared =0.6285 *Adjusted R-squared=0.6285*
Semi log regression: R-squared =0.5800 *Adjusted R-squared=0.5800*
**significant at 0.1 level* ***significant at 0.05 level* ****significant at 0.001 level*

The adjusted R-square for the linear regression is 0.6285. The linear model has explained 62.85% of the variations in the house prices. The semi log regression has a lower adjusted R-square (0.5800) in comparison with the earlier semi log model. In the linear model water frontage has the highest t-statistics while in the semi log model total living area has the highest t-statistics. Park size variable is still insignificant in both linear and semi log models which implies the size of the park does not impact the price or the value of a house, at least not for parks sizes 2.5 acres or less. All other variables are significant in both models at 0.01 level. The linear model indicates that a house near a park that is 1 acre bigger in size will value approximately 832 dollars more than a similar house located near a park that 1 acre smaller. The semi log model shows if we have two houses that are similar in all aspects, but one is located near a park that is 1 acre bigger, that house will value at 0.27% more that its counterpart.

On the other hand, the distance from the nearest park has an inverse relationship with the house value. A house that is a mile closer to the park will value approximately 1,857 dollars (linear regression) than a similar house that is further from the park. According to the semi log regression increasing the total living area by 1 square foot would cause the price of the house to go up by 0.07% holding all other factors constant which is similar with the earlier regression. The number of bedrooms continue to exhibit a negative coefficient only on the linear regression. The semi log regression indicates a positive relationship suggesting that the more the bedrooms the lower the house value, increasing the number of bedrooms in a house by one would result in 1.4% increase in the value of the house holding all other factors constant. Age variable is also not consistent with the expected result. I would expect that house that are older to value less than newer houses, the model however depict the older houses to be more expensive.

<u>All park sizes</u>

I work with the whole dataset without putting restrictions on how big the size is to find out how the effect reflects in the results and how different the impact is from the other regressions. The data set now includes 809,844 houses with house prices ranging from 27,000 dollars to 1,600,000 dollars and an average price of 257,534 dollars. The park size ranges from 0.21 acres to 916.03 acres with a mean of 23.237 acres.

Table 4.1: Descriptive statistics for with data for all park sizes

Variable	**Mean**	**Std Deviation**	**Minimum**	**Maximum**
Park size in acres	23.237	97.0139	0.21	916.03
Distance to park in miles	7.088	4.2129	0.0005	19.7716
Water frontage	0.0091	0.0948	0	1
Home price	257,534	118,517.6	27000	1,600,000
Bedrooms	3.1311	0.8824	1	9
Full baths	1.8018	0.7187	1	6
Half baths	0.4821	0.5656	0	3
Total living area	1482.084	534.6822	432	6285
Age of the house	47.8529	28.8281	1	170

Without interactions

Table 4.2: Linear and semi log regression output of houses within all parks (without interactions)

	Linear regression		Semi log regression	
Variable	**Coefficient**	**t-statistics**	**Coefficient**	**t-statistics**
Park size in acres	1.4579*	1.76	4.60e-06	1.58
Distance to park in miles	-767.4249***	-38.76	-0.0024***	-34.86
Water frontage	352,499.3***	406.49	0.6388***	209.86
Bedrooms	-4,501.059***	-38.89	0.0136***	33.46
Full baths	26,869.65***	163.18	0.1030***	178.27
Half baths	9,049.338***	52.30	0.0316***	52.00
Total living area	123.5242***	205.23	0.0007***	354.12
Total living area squared	0.0010***	7.13	-9.35e-08***	-195.30
Age	703.2564***	67.88	0.0018***	49.78
Age squared	3.1521***	36.41	0.00001***	48.76
Intercept	-7919.316	-12.09	11.1363***	4841.53

Linear regression: R-squared =0.6266 Adjusted R-squared=0.6266
Semi log regression: R-squared =0.5781 Adjusted R-squared=0.5781
**significant at 0.1 level **significant at 0.05 level ***significant at 0.001 level*

The linear model has explained 62.66% of the variations in the house prices. The semi log regression has an adjusted R-square of 0.5781, that is, the semi log regression has explained 57.81% 0f the variation in the house prices. In the linear model water frontage has the highest t-statistics followed by the total living area while in the semi log model total living area has the highest t-statistics. All variables are significant in both models with an exception of park size variable is significant at 0.1 level on the linear regression and insignificant in the semi log regression. The linear model indicate that the value of a house will increase by 1.46 dollars when the size of the park nearest to it is increased by 1 acre holding all other factors constant. The semi log model has an insignificant coefficient for the park size variable which implies that holding all other factors constant we cannot say that the size of the park has any impact on the value of houses. Distance from the nearest park has an inverse relationship with the house value. A house that is a mile closer to the park will value 4,173 dollars (linear regression) than a similar house that is further from the park. According to the semi log regression increasing the total living area by 1 square foot would cause the price of the house to go up by 0.07% holding all other factors constant. The number of bedrooms continue to exhibit a negative coefficient, suggesting that the more the bedrooms the lower the house value. This could be an indication that the model might not be completely reliable. Age variable is also not consistent with the expected result. I would expect that house that are older to value less than newer houses, the model however depict the older houses to be more expensive.

<u>With Interactions</u>

Table 4.3: Linear and semi log regression output of houses within all park sizes

	Linear		Semi log	
Variable	Coefficient	t-statistic	Coefficient	t-statistic
Park size in acres	4.0619	1.13	0.00001	0.96
Park acres squared	-0.0038	-0.96	-1.08e-08	-0.78
Distance to park in miles	1501.524***	20.21	0.0038***	14.41
Distance squared	-143.7915***	-31.69	-0.0004***	24.59
Water frontage	351,995.8***	406.09	0.6374***	209.45
Bedrooms	-4497.243***	-38.88	0.0136***	33.49
Full baths	26,920.73***	163.58	0.1032***	178.57
Half baths	9,107.321***	52.66	0.0317***	52.28
Total living area	123.4716***	205.26	0.0007***	354.78
Total living area squared	0.0010***	7.09	-9.35e-08***	-195.41
Age	677.6562***	65.25	0.0017***	47.73
Age squared	3.2696***	37.76	0.00002***	49.79
Intercept	-13421.83***	-19.78	11.1213***	4667.89

Linear regression: R-squared =0.6270 Adjusted R-squared=0.6270
Semi log regression: R-squared =0.5784 Adjusted R-squared=0.5784
**significant at 0.1 level **significant at 0.05 level ***significant at 0.001 level*

The linear model has explained 62.70% of the variations in the house prices. The semi log regression has an adjusted R-square of 0.5784, that is, the semi log regression has explained 57.84% 0f the variation in the house prices. In the linear model water frontage has the highest t-statistics followed by the total living area while in the semi log model total living area has the highest t-statistics. All variables are significant in both models with an exception of park size variable is insignificant in both regressions. The linear model shows the park size variable to have a concave relationship with house value. If this variable was significant it could be interpreted that holding all other factors constant, for houses near a park that is small, we would expect that the bigger park sizes the higher the value of the house, however as we increase the size and approach 534.46 acres the magnitude of this positive effect diminishes. For houses near parks that are above 534.46 acres in size, we would expect a bigger park size to decrease the value of the house. The distance from the park to the house also has an impact on the value of the house and according to the linear regression the impact is concave. Holding all other factors constant, for houses close to a park increasing the distance from the park would result in higher house value, however as we approach 10.44 miles from the park the magnitude of the positive effect diminishes. For houses that are far from the park (beyond 10.44 miles), we would expect the further the park is from the house the lower the value of the house, the magnitude of this effect will decrease as the distance from the park increases.

The semi log model shows that a house that has a water frontage will value 63.74% more than a similar house that does not have water frontage ceteris palibus. The number of full bathrooms depict a positive relationship with the value of the house, holding all other factors constant a house with one more bathroom value 10.32% more than a house that is similar in all aspects but is 1 bathroom less. The number of bedrooms continue to exhibit a negative coefficient, suggesting that the more the bedrooms the lower the house value. This could be an indication that the model might not be completely reliable. Age variable is also not consistent with the expected result. I would expect that house that are older to value less than newer houses, the model however depict the older houses to be more expensive.

Conclusion

We could not arrive at any useful information for houses near parks that are 0.25 acres or less. The park size coefficient was omitted due to collinearity. We find parks that are not more than 1.5 acres to have the biggest impacts on house value. For houses within this range increasing the park size by 0.1 acres would result in house value increasing by approximately 129 dollars holding all other factors constant. We also find the distance to the park to significantly impact the value houses. For houses within this range (1.5 acres park sizes) increasing the distance from the park by 0.1 miles would reduce the value of the house approximately 417 dollars. We can conclude that for houses near parks that are 1.5 acres house near parks that are bigger in size value more and the shorter the distance to the park the higher the value of the house.

Increasing the range of the size of the park still give is significant positive impact of park size on house prices /value, however, the impact is lesser. For houses near parks that are not more than 2.5 acres, increasing the size of the parks by 0.1 acres would result in the value of the house increasing by 83 dollars and if we have two houses that are similar in every way, but one is further from the park by 0.1 miles we would expect it to value approximately 186 dollars less. In this case the impact of the increase in the distance from the park has decreased too.

When we consider all park sizes, we found the impact to be significant, but the magnitude is low. Increasing the size by 1 acre would result in the value increasing by only approximately 1.5 dollars. Including the interactions (squaring the park size variable gave insignificant result). The distance from the park was significant and the relationship is concave. The distance from the park to the house also has an impact on the value of the house and according to the linear regression the impact is concave. Holding all other factors constant, for houses close to a park increasing the distance from the park would result in higher house value, however as we approach 10.44 miles from the park the magnitude of the positive effect diminishes. For houses that are far from the park (beyond 10.44 miles), we would expect the further the park is from the house the lower the value of the house, the magnitude of this effect will decrease as the distance from the park increases.

In conclusion we can say that the size of a park nearest to a house has a positive impact on the value of that house. The most significant impact, however, would be expected on parks not too big in size. The best impact is felt on parks that are not more than 1.5 acres in size, beyond this size the magnitude of the impact reduces.

Bibliography

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, *58*(304), 933–942.

Burgess, J., Harrison, C. M., & Limb, M. (1988). People, Parks and the Urban Green: A Study of Popular Meanings and Values for Open Spaces in the City. *Urban Studies*, *25*(6), 455–473.

Crompton, J. L., National Recreation and Park Association, & National Recreation Foundation. (2004). *The proximate principle: the impact of parks, open space and water features on residential property values and the property tax base*. Ashburn, Va.: National Recreation and Park Association.

Fabozzi, F. J., Francis, J. C., & Lee, C. F. (1980). Generalized Functional Form for Mutual Fund Returns. *Journal of Financial and Quantitative Analysis*, *15*(5), 1107–1120.

Fletcher, M., Gallimore, P., & Mangan, J. (2000). Heteroscedasticity in hedonic house price models. *Journal of Property Research*, *17*(2), 93–108

Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, *5*(4), 471–484.

Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric of Quality Change. *The Price Statistics of the Federal Goverment*, 173–196.

Gujarati, D. N. (2009). *Basic econometrics*. Fifth edition. Boston: McGraw-Hill Irwin.

Haas, G. C. (1922). *Sale Prices as a Basis for Farmland Appraisal*. University Farm.

Halvorsen, R., & Pollakowski, H. O. (1981). Choice of functional form for hedonic price equations. *Journal of Urban Economics*, *10*(1), 37–49

Hulten, C. R., & Wykoff, F. C. (1981). The Estimation of Economic Depreciation Using Vintage Asset Prices: An Application of the Box-Cox Power Transformation. *Journal of Econometrics*, *15*(3), 367–396

Kitchen, J. W., & Hendon, W. S. (1967). Land Values Adjacent to an Urban Neighborhood Park. *Land Economics*, *43*(3), 357–360.

Krysel, C., Minnesota, Legislature, Legislative Commission on Minnesota Resources, Bemidji State University, Minnesota, & Mississippi Headwaters Board. (2003). *Lakeshore property values and water quality: evidence from property sales in the Mississippi Headwaters region*. Minnesota: Mississippi Headwaters Board.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, *74*(2), 132–157.

Leggett, C. G., & Bockstael, N. E. (2000). Evidence of the Effects of Water Quality on Residential Land Prices. *Journal of Environmental Economics and Management*, *39*(2), 121–144.

Li, M. M. (1978). *Micro-neighborhood externalities and hedonic housing prices*. Cambridge, MA: Department of City and Regional Planning, Harvard University.

Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the Netherlands. *Journal of Planning Literature*, *15*(2).

McMullen, B. A. (2011). *The economic impact of parks on residential property values: evidence from Gainesville, Florida*. Gainesville, Fla.: University of Florida.

Michael, H. J., Boyle, K. J., & Bouchard, R. (1996). *Water quality affects property prices: a case study of selected Maine lakes*. Orono: Maine Agricultural and Forest Experiment Station, University of Maine.

More, T. A., Stevens, T., & Allen, P. G. (1988). Valuation of urban parks. *LAND Landscape and Urban Planning*, *15*(1), 139–152.

Nicholls, S., & Crompton, J. (2005a). The Impact of Greenways on Property Values: Evidence from Austin, Texas. *Journal of Leisure Research*, *37*.

Nicholls, S., & Crompton, J. L. (2005b). The Impact of Greenways on Property Values: Evidence from Austin, Texas. *Journal of Leisure Research; Urbana*, *37*(3), 321–341.

Olmos, M. (2008). Pocket Parks Development Standard. Retrieved from http://www.ci.visalia.ca.us/civicax/filebank/blobdload.aspx?blobid=4542

Palmquist, R. B. (2006). *Property Value Models* (Handbook of Environmental Economics) (pp. 763–819). Elsevier.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, *82*(1), 34–55.

Sheppard, S. (1999). Chapter 41 Hedonic analysis of housing markets. In *Handbook of Regional and Urban Economics* (Vol. 3, pp. 1595–1635). Elsevier.

Small, K. A. (1975). Air Pollution and Property Values: Further Comment. *The Review of Economics and Statistics*, *57*(1), 105–107.

Söderberg, B., Zhou, J., & Janssen, C. (2001). Robust estimation of hedonic models of price and income for investment property. *Journal of Property Investment & Finance*, *19*(4), 342–360.

Sopranzetti, B. J. (2015). Hedonic Regression Models. In C.-F. Lee & J. C. Lee (Eds.), *Handbook of Financial Econometrics and Statistics* (pp. 2119–2134). New York, NY: Springer New York.

Spitzer, J. J. (1982). A Primer on Box-Cox Estimation. *The Review of Economics and Statistics*, *64*(2), 307–313.

U.S. Census Bureau QuickFacts: Madison city, Wisconsin. (n.d.). Retrieved December 10, 2018, from https://www.census.gov/quickfacts/fact/table/madisoncitywisconsin/PST045217

Wallace, H. A. (1926). Comparative Farm-Land Values in Iowa. *The Journal of Land & Public Utility Economics*, *2*(4), 385–392.

Zabel, J. (2015). The hedonic model and the housing cycle. *Regional Science and Urban Economics*, *54*, 74–86.