Measuring the Effect of Green Space on Property Value: An Application of the Hedonic Spatial Quantile Regression

Sezhu Liu

The Department of Agricultural Economics and Rural Sociology

Auburn University

Auburn, AL

Diane Hite

The Department of Agricultural Economics and Rural Sociology

Auburn University

Auburn, AL

Selected Poster prepared for presentation at the Southern Agricultural Economics Association (SAEA) Annual Meeting, Orlando, Florida, 3-5 February 2013

Abstract

Green space is an important part of environment around houses. Generally, most research focused on the economic impacts of green space on urban planning and environmental pollution cost, but ignored the impact on single family home values. Limited research was conducted in this area and few studies of green space and housing prices have incorporated spatial econometric techniques. This technique is necessary since housing value may be influenced by characteristics of nearby properties. This research attempts to quantify the impacts of green space, by using the hedonic price analysis of the relationship between property values and the green space amenities around the selected single family houses in Delaware County, Ohio. Also, by incorporating spatial-lag term, we can compare the results with and without spatial effect. Eventually, after extending the model by quantile regression, the influence of different green space characteristics on housing price may change across the conditional distribution of housing price. Substantial variation was found between the results with and without spatial effects across quantiles, which indicates that luxury house buyers may value green space differently from middle or low level house buyers.

JEL classification: C21; R20; D10

Key words: Hedonic model; Spatial-lag; Two-stage quantile regression; Generalized spatial two-stage least-square; Green space

1 Introduction

An attractive environment is likely to influence house prices. Houses with an attractive environment will have an added value over similar, less favorably attractive ones. The presence of trees and forests can make the environment a more pleasant place to live, work, and spend leisure time and thus makes substantial improvements in individual well-being, including opportunities for leisure out in the yard or in the neighborhood, reduced heating and cooling costs, privacy, and the lack of a need to construct fences or screens. Moreover, forests can strongly influence the physical/biological environment and mitigate many impacts of development by moderating climate, conserving energy, carbon dioxide and water, improving air quality, water purification, controlling rainfall runoff and flooding, and harboring wildlife thus enhancing the attractiveness of nearby parcels. Besides, field tests have shown that properly designed plantings of trees and shrubs significantly reduce noise. In sum, green space provides multiple benefits including recreational opportunities, aesthetic enjoyment, and ecological services.

As Nanette, Jeffery and Laurie (2002) represented, the effect that environmental amenities, such as forested areas and green open space, contribute to the value of real estate is often estimated using the hedonic pricing approach, a method that was based on the straightforward premise that the value of a good depends on the stream of benefits derived from that good. Using regression techniques, the hedonic pricing method identifies what portion of the differences in property value can be attributed to environmental amenities, such as green space. The sales value of real estate reflects the benefits that buyers attach to the attributes of that property, including the trees and forest resource found near the property, along the street, and in neighboring parks and greenways.

Although the hedonic model for housing was commonplace and there are a lot of studies that explored the effect of different environmental factors on house prices, few of them focus on the green space effect; even if some of them did, they only focus on urban green space and city planning issues. Besides, they are quite simple through variables and methods, since they only contain one or a few amenities and rarely contain socio-economic variables, due to a lack of using large amount of census survey data available to us.

My approach offers the potential for a richer model: first, beyond the traditional variables to explain residential values, such as housing characteristics of the parcel and distance to amenities, I also create some environment indexes to evaluate its effect on housing price more comprehensively. Second, the idea that location is an important factor in determining the property value is not new, but few people seek the factor effect, which varies with the change in housing price. This paper allows for spatial heterogeneity in estimation by introducing the spatial econometric method and combined with the quantile regression to see the location effect on the different level of housing price. This paper is organized as follows: section 2 represents some related papers to show the previous work on this topic. Section 3 outlines the basic model specifications. In Section 4, the data is described and a statistical summary is provided. The empirical results and detailed interpretations of the results are presented in Section 5. Finally, section 6 draws some conclusions from the analysis.

2 Literature Review

The study of housing price is a large field on its own, and it is impossible to cover even a small fraction of the research conducted in that field. Thus, this review concentrates mainly on the studies that looked at the relationship between the green space and housing prices.

Green spaces provide many environmental and social benefits, which are well documented in the literature (Robinette, 1972; Grey and Deneke, 1978; Laurie, 1979; Miller, 1988). Most of the values attached to the green spaces are non-priced environmental benefits. These values include those derived from pleasant landscape, clean air, peace and quiet and screening as well as potential recreational activities in wooded green spaces. Other benefits include reduced wind velocity, balanced microclimate, shading, and erosion control. But due to these non-commodity and non-priced nature, and largely intangible benefits, their contribution is usually difficult to assess and quantify, among them, various approaches have been proposed and tested. There are two ways to measure these kinds of amenity values. One is to use a survey-based method, such as travel cost or contingent valuation. The hedonic pricing approach is the other option. Hedonic methods have been gaining popularity in recent years with application of spatial analyses using geographical information system (GIS). Recently, it has been widely applied to estimate the value of environmental benefits from costs and prices of related market transactions. This method has the advantage of being based on actual transaction data, choice and purchase price.

The hedonic method can be traced back to Court (1939) and received considerable application beginning in the 1960s. However, it was not until 1974 that a theoretical model that could serve as a basis for the empirical techniques was developed by Rosen. This model considers a class of differentiated products completely described by a vector of objectively measured characteristics. Observed product prices and the specific amounts of characteristics associated with each good define a set of implicit, or "hedonic", prices.

A hedonic model of price is one that decomposes the price of an item into separate components that determine the price, since every good provides a bundle of characteristics or attributes. This theory was well explained by Brown and Rosen (1982). According to their theory, goods are valued for their utility-bearing attributes or characteristics. Each good is described by n objectively measured characteristics, represented by a vector $z = (z_1 \ldots z_i \ldots z_n)$, with z_i measuring characteristics i of the good. Each good offers buyers distinct packages of characteristics. The markets of goods implicitly reveal a price function $p(z) = p(z_1 \ldots z_i \ldots z_n)$, relating price and good characteristics. Rosen’s model has been proven to be extremely useful in many years and was cited by the majority of papers in the hedonic field. The most common application of this method is housing price. Hedonic pricing method (HPM) is based on the idea that properties are not homogenous and can differ in respect to a variety of characteristics. The method relies on the fact that house prices are affected by many factors: number of rooms, access to amenities, and so on. As Garrod (1999) represented, the most common application of the HPM is in relation to the public willingness to pay for housing. Each property is assumed to constitute a distinct combination of attributes, which determine the price or buyers’ willingness to pay. The price of a housing unit is dependent upon the availability and level of a wide range of attributes, such as structural characteristics, neighborhood characteristics, and amenity characteristics. Among them, one important factor is environment, for example, view or access to a wooded park or watercourse (Palmquist, 1991).

Theoretically, HPM can be used in calculating external benefits and costs of forests associated with housing. The price of a house reflects the people's willingness to pay in order to gain easier access to forests and to 'consume' their amenity values. In addition, HPM has been used for estimating the contribution of individual trees to property values (Darling, 1973; Morales, 1980; Morales et al., 1983; More et al., 1988). Anderson and Cordell (1985) found that tree cover increased property values by 3-5% in Athens, Georgia. A study based on the HPM was carried out in Apeldoorn, a medium-sized town in eastern Netherlands (Fennema et al., 1996). This study analyzed 106 house transactions built around a park; it demonstrated that location within 400 meter of the park attracted a premium of 60% over houses located outside this zone. This result was consistent with the expectation that green has a value-increasing effect on housing price.

There is a long history of using hedonic model to investigate the effects of amenities on sale prices of houses. The most common approach has been to include distance from property to the amenity as an explanatory variable in the model (Milon, Gressel, and Mulkey 1984; Kohlhase 1991; Mendelsohn et al. 1992; Nelson, Genereux, and Generoux 1992; Thayer, Albers, and Rahmatian 1992; Kiel 1995; Lansford and Jones 1995). In housing price research, a parcel's surroundings have a major influence on housing value. The analysis of the pattern of land use and amenities surrounding the property is important and can be captured with Geographic Information Systems (GIS) applications. Din, Hoesli and Bender (2001) argued that GIS have made possible the development of databases that can be used to better measure environmental characteristics. Their environmental parameters refer to the quality of the neighborhood and the quality of the location within a neighborhood. In the spatial model, GIS can be used to develop neighborhood characteristics that are unique to each observation, thereby allowing the examination of the impact of amenities or disamenities in proximity to the house. Another benefit of applying GIS in spatial analysis is demonstrated by Clapp (1997), he argues that GIS is a powerful tool for supporting research because of its capability of storing and manipulating large data sets on spatial relationships. GIS can quickly assemble large amounts of spatial data, link spatial features to data, and visualize spatial analysis results. Furthermore, ArcGIS10 includes a spatial statistics toolbox, with functionality for spatial autocorrelation analysis and spatial regression, but some of the functions are not available right now. Also, GeoDa is good software that can be used in spatial analysis, but it cannot deal with large dataset and the weight it generated cannot be inserted into other software.

In many instances, there may be multiple occurrences of amenities proximate to properties, and GIS can generate variables that distinguish between them. For example, in examining the influence of wetland amenities on sale prices of residential properties in Portland, Oregon, Mahan, Polasky, and Adams (2000) consider distance to, as well as size and shape of, the nearest wetland area. Similarly, Powe et al. (1997) approximate forest amenities associated with a given property with an index variable that measures the ratio of acreage to squared distance from the home, summed over all woodland areas in the Southampton and New Forest areas of Great Britain. GIS data have also been used by Geoghegan, Waiger, and Bockstael (1997) to construct variables that reflect the extent, diversity and fragmentation of land uses in various buffer sizes around residential properties in the Patuxent Watershed in Maryland. In each of these studies, GIS data have enhanced the ability of the hedonic model to explain variation in sale prices by considering both proximity and extent of environmental attributes.

When being asked, people always said that property values are determined by “location, location, location” a reasonable explanation for this is that spatial econometric techniques should be used in an analysis of housing price, therefore in research area, many hedonic price studies suggested that in a cross-sectional hedonic price analysis, the value of a property in one location may also be affected by the property value in other locations, such as in its neighboring area. Ignoring this spatial effect or spatial dependence may cause hedonic estimation result inconsistent or inefficient1. Spatial dependence among hedonic regression residuals was initially revealed by Paelinck and Klaassen (1979), who published a small volume entitled Spatial Econometrics, which arguably was the first paper in the field of spatial econometrics and its distinct methodology. Spatial analysis or spatial econometrics in hedonic analysis was introduced by Dubin (1988, 1992) and Can (1990, 1992); since then it started to be applied in many more recent studies. Those studies include: Geoghegan et al (1997) employed spatially-explicit indices in that paper, Bockstael and Bell (1997) used a simple spatial error model, He and

[1] See Anselin, 1988 for text-book treatment of spatial econometrics.

Winder (1999) demonstrated bi-directional price causality between three adjacent housing markets in Virginia, indicating the existence of spatial effect in housing markets.

Also, there are a number of studies that provide evidence of the existence of spatial effect in hedonic analysis. For example, Legget and Bockstael (2000), Gawande and Jenkin-Smith (2001) estimated a housing price hedonic model using a simple spatial autoregressive model. Bowen, Mikelbank and Prestegaard (2001) examined housing prices in Cuyahoga County, Ohio. Kim, Phipps and Anselin (2003) measured the benefits of improving air quality on housing prices in Seoul, Korea. Bransington and Hite (2004) discussed the ways to model the influence of different types of omitted variables in the spatial model. And there are still many other hedonic studies incorporate the spatial effects, such as Basu and Thibodeau (1998), Dubin, Pace, and Thibodeau (1999), Munneke and Slawson (1999), Gillen, Thibodeau, and Wachter (2001), and Irwin (2002).

Until now, we can see that housing price is affected by many factors at different perspectives, but there is still one issue we need to consider: housing characteristics may have a different effect on housing prices when we analyze it at different points of the distribution of house prices, which is referred to as quantile effects. Quantile regression is based on the minimization of weighted absolute deviations to estimate conditional quantile (percentile) functions as represented by Koenker and Bassett (1978) and Koenker and Hallock (2001). There is a large amount of literature using this model in many different topics: Eide and Showalter (1998), Knight, Bassett and Tam (2000) and Levin (2001) have addressed school quality issues. Poterba and Rueben (1995) and Mueller (2000) studied public-private wage differentials in the United States and Canada. Abadie, Angrist and Imbens (2001) considered estimation of endogenous treatment effects in program evaluation, and Koenker and Billias (2001) explored quantile regression models for unemployment duration data. A paper written by Viscusi and Hamilton (1999) considered public decision making regarding hazardous waste cleanup.

This model was also used in the housing value research: Gyourko and Tracy (1999) adopted the quantile regression approach to investigate changes in housing affordability between 1974 and 1997 using the American Housing Survey data. Employing housing transaction data from Chicago in 1993 through 2005, McMillen and Thornes (2006) suggested that quantile regression has advantages over the conventional mean-based approaches to estimating a housing price index. McMillen and Coulson (2007) and McMillen (2008) identified significant variations in values of physical attributes across quantiles after studied house price appreciation and constructed quantile house price indexes. Since normal quantile regression does not consider spatial autocorrelation that may be present in the data, spatial autocorrelation was incorporated into the quantile regression by adding a spatial lag variable; but adding a spatial lag into OLS regression will cause endogeneity problem (Anselin, 2001). When there are endogenous variables, the estimator of the parameter of interest is generally inconsistent. Amemiya (1982) and Powell (1983) dealt with the case of the double-stage least-absolute deviations (DSLAD) with fixed regressors, which allow researchers to focus on the median of the distribution of interest. The theoretical literature on quantile regression and LAD estimators is extensive since the seminal paper by Koenker and Basset. Other researchers have treated some endogeneity problems in quantile regressions. Kemp (1999) and Sakata (2001) studied least absolute error difference (LAED) estimators for estimating a single equation from a simultaneous equation model. Abadie and Imbens (2002) design a quantile treatment effects estimator, which is the solution to a convex programming problem with first-step non-parametric estimation of a nuisance function. MaCurdy and Timmins (2000) propose an estimator for ARMA models adapted to the quantile regression framework. Among them, Kim and Muller (2000) first introduce the Two-Stage Quantile Regression (2SQR)2. In 2004, they published another paper in 2SQR about the detailed discussion in the two stages. After that, Zietz et al. (2008) utilize quantile regression, with and without accounting for spatial autocorrelation, to identify the coefficients of a large set of diverse variables across different quantiles.

3 Data

The area covered by the data set must be sufficiently wide to ensure a representative spread of variation in the level of any external factors being investigated, and amenities in that area must be fully included in order to cover the location factor that affect the housing price. Therefore, the housing market in Delaware County, Ohio, is chosen as the case study. Delaware County has been a leader in developing a comprehensive land information

[2] See Two-Stage Quantile Regression at http://www.nottinghampublications.com/economics/documents/discussion-papers/00-01.pdf

system (DALIS/Delaware Appraisal Land Information System), and is a source of a variety of spatially explicit data with very detailed characteristics for individual houses. Also, the project provides 2010 Census Geography for Delaware County, which includes amenities and infrastructure in polygon shape files and associated tables in dbf format. At the start of this study, in an investigation of the effects of green space on housing price, variables relating to structural characteristics were designed, including the age of the house, whether it had gas, or heating and the number of bathrooms as has been done by most previous researchers (McLeod, 1984; Des Rosiers et al., 2002). This kind of information is necessary to explain those differences in price attributable to the structural characteristics, as opposed to those which are the result of amenities and socio-economic characteristics.

When making decisions, each house buyer takes the characteristics of neighboring residences into consideration. Thus, socio-economic variables that estimate the quality of neighbors were included in this study. The US Census no longer surveys long form data for its decennial census. The only demographic data are currently available from American Community Survey (ACS); it is a household survey conducted by the U.S. Census Bureau that currently has an annual sample size of about 3.5 million addresses. Socio-economic characteristics were reflected primarily by data from ACS on vacancy ratio, percentage households with medium or high income, percentage of population with different race and percentage of population with different education level. Vacancy rate is included as an indicator to capture prevailing housing market conditions. Table1 listed and defined the explanatory variables that are included in the regression. Table 2 listed the summary statistics of the variables.

Table 1 goes about here

Table 2 goes about here

As we all know that houses in an attractive location attract a premium over houses in a neutral location. Green space, ponds and lakes, smooth traffic and convenience are aspects of an attractive location. Since these factors are valued differently by residents, they will affect house prices differently. Location variables included distance to amenities and disamenities, like distance to nearest medical center, post office, railroad, police office[3], railroad,

[3] All the distance variables are in miles.

and forested amenities (amenities contain forest) and so on. These distance variables are intended to capture the effect on housing prices of the proximity to various amenities. In fact, the distance variable is an imperfect measure of this effect, for example, Strand and Vagnes (2001) represented that environmental nuisance associated with living close to the railroad. In reality noise and vibrations also depend on topographical properties, e.g. on whether the train line is elevated above the house, on level with it or sunk below it; whether there are objects (such as trees and rocks) that shield the house from noise; and whether there are other houses in between the railroad line and one's own house, and whether the unit has extra protection against noise and vibrations (such as noise-reducing windows). But since this paper is focus on green space, other variables were included just for excluding their effect on housing price. Negative distance effect was expected for the amenities since shorter distance means more convenience, and opposite effect for the disamenities because of the noise or inconvenience they brought. Figure 1-4 are the distribution of these places with the county.

Figure 1 goes about here

Figure 3 goes about here

To evaluate the convenience of parcels, except for the distance to amenities, traffic condition around the house is still important. One prime candidate for such a variable is road traffic, which will most likely reduce house price. By using number of schools near the house as representation, an index was created to examine the road traffic condition on housing price,

$$RTI_j = n^2 / \sum_{i=1}^{n} d_i \tag{1}$$

where RTI_j is road traffic index, n is the number of schools within a 1 mile buffer, d_i is the distance from the centroid of the housing to school i. If there is no school within the buffer, equation $RTI = 1/D$ would be used, where D is the straight line distance from the housing centroid to the nearest school, a negative sign of this variable is expected. Detailed interpretations of parameters in the equation are all in Table 2 and the distribution and buffer of schools are shown in Figure 2.

Figure 2 goes about here

The nature environment is also likely to affect the value of a parcel. To capture the effect associated with surrounding nature environment, accessibility to nature environment indices were created by using the shape file data of parks, ponds and lakes, and woodland. Because these places have large area, the shape of them is also important so it is inappropriate to simply change them to centroid file and calculate the distance. The cross product of distance and shape (Fanhua K., Haiwei Y., Nobukazu N., 2007) has been estimated in the regression. Taken together, these indices describe the characteristics of the surrounding environment, which should have an important effect on the value of a parcel.

(2) $$S\_DP_j = \ln{(S_p/D_{jp})}$$

Where S_DP_j is the access to park index for parcel *j*, S_p is the area of the nearest park *p*, D_{jp} is the straight line distance between the parcel centroid and the boundary of the nearest park *p*. If the parcel is within the park area, Access to Park Index = $ln\ (S_p)$. The size of nearest park is intended to capture the premium being closer to the bigger park. Park size has been found to be a significant factor on property value (Lutzenhiser and Netusil, 2001), and so did the distance to it. A positive relationship between this variable and housing price was expected, since large area and short distance were preferred by the house owner. Accordingly, variables reflecting hydrology area and distance might capture its effects on the property value.

(3) $$S\_DPL_j = \ln{(S_l/D_{jl})}$$

Where S_DPL_j the access to ponds and lakes index for parcel *j*, S_l is is the area of the nearest lake *l*, D_{jl} is the straight line distance between the parcel centroid and the boundary of the nearest lake *l*. As we know, distance to the lake may have a negative effect on the housing price, while the effect of the size of the lake is positive. For the parcel that is within the ponds and lakes area, distance for parcel to there is 0, the index equals: $S\_DPL_j = \ln(Sl)$. A positive relationship between the area and price, and negative relationship on the distance (Brown and Pollakowski, 1977) was expected. The same calculation was also used in woodland index:

(4) $$S\_DW_j = \ln{(S_w/D_{jw})}$$

Where S_DW_j the access to woodland index for parcel j, S_w is the area of the nearest woodland w, D_{jw} is the straight line distance between the parcel centroid and the boundary of the nearest woodland w. For the parcel that within the ponds and lakes area, distance for parcel to there is 0, the index equals: $S\_DW_j = \ln(S_w)$. Since the influence of woodland to house depends on both the distance and area, this index can be a better estimator. The magnitude of the index will change with both distance and area of woodland, it will increase not only if the woodland is close to the house especially if larger areas of woodland with shorter distance. Besides, another variable equaling the cross-product of area of both census tract and woodland allows the marginal effect of percentage of forest cover was also created to evaluate the green space effect on housing price. The distributions of woodland around parcels are shown in figure 4.

Figure 4 goes about here

4 Methods

This section first reviews the hedonic, spatial econometric and quantile models, and then introduces the methods to integrate them. Also, it includes specific discussion on my estimation method.

4.1. The Hedonic Housing Price Specification

In general, the hedonic equation for housing relates the sales price of a property to a set of characteristics that determine the property's value. Since this paper deals with owner-occupied housing, three groups of characteristics are included: (1) structural characteristics, (2) amenities characteristics, (3) socio-economic characteristics.

The general functional form of the hedonic price function is:

P = f (S, A, E)

Where: P = Log of housing price

S = Structural characteristics of the house

A = Amenities characteristics

E= Socio-economic characteristics

Expansion form is[4]:

$$\ln P_i = \alpha_0 + \alpha_l S_{il} + \alpha_m E_{im} + \alpha_k A_{ik} + \varepsilon_i$$

Where ε is assumed to be a normally distributed error term, with $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon') = \sigma^2 I$

P_i is the housing price in nature log form

S_{il} = Structural characteristic *l* of the house i

A_{ik} = Amenities characteristic k of house i

E_{im} = Socio-economic characteristic m of house i

The dependent variable is the natural log of the sale price. A log-linear form allows the marginal effect of each independent variable to vary with the level of the dependent variable, so the marginal effects of independent variables change as the house price changes. Because the predicted hedonic price is the result of the behavior of many different buyers and sellers, the marginal effect of independent variables are not constant for all houses regardless of differences in house price (Taylor, 2003). Therefore, the functional form of the HPM usually was not linear (Freeman, 1993). In addition, this specification was also used by Gillingham (1975), Palmquist (1979), Thibodeau (1989, 1992, 1996), and others to model the determinants of house prices.

4.2. Hedonic Analysis with Spatial Lag and Spatial Error

Whether or not any pair of houses is neighbors is based on whether or not they are located in neighboring area. Two areas are considered neighbors when they share common borders (contiguity) or when their distance to each other is below a certain level. In order to measure that in the spatial model, we need to use the spatial weight matrix (W). There are two basic types of spatial weights matrices. The first type is contiguity-based; the second

[4] Detailed information on variables in these categories are in Table 1.

type is distance-based. For both types of spatial weights matrices, we must specify two general parameters before their construction. The first is the spatial extent of the influence or the definition of the neighborhood. For a contiguity-based matrix, if two polygons are contiguous, they are considered neighbors. Two basic types of contiguity exist: rook contiguity (e.g., two polygon share a common border) and bishop contiguity (e.g., two polygons share a common vertex). Queen contiguity is a combination of these two. Specifically, a contiguity-based spatial weights matrix (W) is typically specified as

$$\text{(5)} \qquad w_{ij} = \begin{cases} 1 & \mathit{if}\ i\ \mathit{and}\ j\ \mathit{are\ contiguous} \\ 0 & \mathit{otherwise} \end{cases}$$

For a distance-based matrix, a critical value of distance must be specified within which two points are thought to be neighbors. The parameter is called the "power" of influence of two neighbors, which indicates that neighbors influence each other's housing price to different degrees, depending on the distance between them. For example, houses at different locations: the prices/error terms associated with close neighbors are more highly correlated than those of distant neighbors. The relationship in the distance-based spatial weights matrix is typically represented as an inverse function of distance, within the assumed critical value.

$$\text{(6)} \qquad w_{ij} = \begin{cases} 1/(d_{ij})^{\theta} & \mathit{if}\ i \neq j\ \mathit{and}\ d_{ij} \leq m \\ 0 & \mathit{if}\ i \neq j\ \mathit{and}\ d_{ij} > m\ \mathit{or}\ \mathit{if}\ i = j \end{cases}$$

The term d_{ij} is the distance between points *i* and *j*, and usually calculated according to their latitude and longitude (or X, Y coordinates). The parameter m is the extent of influence or critical distance value. The choice of its value is an empirical problem that depends on the scale of data and the extent of the perceived neighborhoods. Parameter θ measures the "power" of influence, whose value represents the distance decay effect within neighborhoods. As θ increases, the influence of nearby observations becomes greater than those further away. An alternative distance-based weights matrix uses linear decay. The weight corresponding to points *i* and *j* is assumed to be linearly inverse to the distance between them (d_{ij}) and equal zero at a specified distance.

The purpose of including a spatial weights matrix is to correct for potential problems due to spatial correlation and unobserved heterogeneity. Spatial autocorrelation is used to deal with the situation where the price of a house at one location is correlated with the price of neighboring houses. This dependence originates in part from the fact that each house shares with its neighbors influences from location factors that are nearly identical. In practice, parcel level distance variables, or spatial weights matrix approach, are usually used to incorporate spatial effects into hedonic regression models.

There are two kinds of weighting methods: the first one is spatial-lag model, which is weighting the sum of neighboring observations on the dependent variable (y), which is generally accomplished by creating a spatial lag term *Wy* weighted by neighbors' proximities to each observation. The spatial-lag model implicitly assumes that the spatially weighted average of housing prices in a neighborhood affects the price of each house (indirect effects) in addition to the standard explanatory variables of housing and neighborhood characteristics (direct effects). It assumes that the spatially weighted sum of neighborhood housing prices (the spatial lag) enters as an explanatory variable in the specification of housing price formation.

$$\tilde{\mathbf{P}} = \rho \mathbf{W} \tilde{\mathbf{P}} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{7}$$

Where ρ is the spatial dependence parameter and W is an n×n standardized spatial weight matrix (n is the number of observations). The spatial weight matrix, W, tells us whether any pair of observations are neighbors, that is, if house *i* and house *j* are neighbor, then $W_{ij} = 1$ and 0 otherwise. The spatial weight matrix was standardized here, meaning that every row of the matrix sums to 1. This enables us to interpret the spatial lag term in a model as simply a spatially-weighted average of neighboring house prices, for example:

$P_1 = \rho(w_{12}P_2 + w_{13}P_3 + w_{16}P_6) + \sum_{j=1}^{k} \beta_j x_j + \varepsilon_1$, where observation 2, 3 and 6 are neighbors of observation 1. The spatial lag model more or less resembled the autoregressive (AR) model in time-series econometrics. However, unlike the AR model, OLS estimation in the presence of spatial dependence will be inconsistent, because of the endogeneity problem. The spatial lag model will be estimated using maximum likelihood estimation.

The second one is the spatial error model, which is done by creating a proximity-weighted error term Wε, where ε is the weighted sum of neighboring errors. Compared to the spatial lag model, the spatial error model does not include indirect effects but is based on the assumption that there is one or more omitted variables in the hedonic price equation and that the omitted variables vary spatially. Due to this spatial pattern in the omitted variables, the error term of the hedonic price equation tends to be spatially autocorrelated.

The spatial error model takes the following form:

(8) $$\tilde{\mathbf{P}} = \mathbf{X\beta} + \boldsymbol{\varepsilon} \qquad \boldsymbol{\varepsilon} = \lambda \mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u} \tag{5}$$

Where u is an i.i.d error term, and λ is spatial error parameter. The spatial error model resembles more or less the moving average (MA) model in time series econometrics, in which error of certain observations is affected by errors of other observation. OLS estimation of spatial error model will be inefficient because it violates the assumption of independence among the disturbance term.

Generalized Spatial Two-stage Least-Square (GS2SLS) model is the combined spatial-autoregressive model with spatial-autoregressive disturbances. The basic functional form is:

(9) $$\mathbf{y} = \boldsymbol{\lambda}\mathbf{Wy} + \mathbf{X\beta} + \boldsymbol{u}$$

(10) $$\boldsymbol{u} = \boldsymbol{\rho}\mathbf{M}\boldsymbol{u} + \boldsymbol{\varepsilon}$$

Where y is an n $\times$1 vector of observations on the dependent variable, W and M are n $\times$n spatial-weighting matrices (with zero diagonal elements), Wy and Mu[5] are n$\times$1 vectors typically referred to as spatial lags, and λ and ρ are the corresponding scalar parameters typically referred to as spatial-autoregressive parameters, X is an n $\times$k matrix of observations on k right-hand-side exogenous variables (where some of the variables may be spatial lags of exogenous variables), and β is the corresponding k $\times$1 parameter vector, ε is an n $\times$1 vector of innovations.

4.3. Quantile Regression and Spatial Autocorrelation

[5] In this paper, I use the same W in both the lag term and the error term, in other words, W=M

While ordinary least-squares regression models the relationship between one or more covariates X and the conditional mean of a response variable Y given X = x, quantile regression can be employed to explain the determinants of the dependent variable at any point of the distribution of the dependent variable. It is preferred to the approach that quantile regression first subdivides the sample according to the unconditional distribution of the response variable and subsequently performs OLS for each subsample, we can see from this process that quantile regression uses the full sample and avoids the truncation problem that the alternative approach usually encounters. Other important advantages of quantile regression include its superior capability in handling heteroscedasticity, outliers, and unobserved heterogeneity.

In practice, Koenker and Hallock (2001) and Koenker (2005) review this econometric method thoroughly. For hedonic price functions, quantile regression makes it possible to statistically examine the extent to which housing characteristics are valued differently across the distribution of housing prices.

The detailed equations for quantile regression are as follows:

For a random variable Y with probability distribution function

(11) $$F(y) = \text{Prob } (Y \leq y)$$

The τ th quantile of Y is defined as the inverse function

(12) $$Q(\tau) = \inf \{y : F(y) \geq \tau\}$$

Where $0 < \tau < 1$. In particular, the median is $Q(1/2)$.

The mechanism to carry out quantile regression is similar to ordinary regression. The difference is, instead of searching for the argmin of sums of squared residuals, quantile regression looks for the argmin of weighted sums of absolute residuals.

Least squares minimizes the sum of the squared residuals,

(13) $$\min_{\{\beta_j\}^k{}_{j=0}} \left[\sum_i (y_i - \sum_{j=0}^{k} x_{i,j}\beta_j)^2 \right]$$

where y_i is the dependent variable at observation i, $x_{j,i}$ is the jth regressor variable at observation i, and β_j is an estimate of the model's jth regression coefficient. By contrast, quantile regression minimizes a weighted sum of the absolute deviations,

(14) $$\min_{\{\beta_j\}^k{}_{j=0}} \left[\sum_i \left| y_i - \sum_{j=0}^{k} x_{i,j}\beta_j \right| h_i \right]$$

Where the weight h_i is defined as $h_i = 2\tau$, if the residual for the ith observation is strictly positive or as $h_i = 2-2\tau$, if the residual for the ith observation is negative or zero. The variable τ ($0< \tau <1$) is the quantile to be estimated or predicted.

The general quantile functional form used in this paper is:

(15) $$Q_\tau(P) = \beta_\tau + \sum_{i=1}^{k} \beta_{\tau,k} X_k$$

Where $Q_\tau(P)$ is the log housing price in τth percentile. X_k is all the variables used in this paper, and for the 2SQR model, it includes the predicted value of spatial lag of housing price after regression.

As mentioned above, spatial autocorrelation is a special problem must be considered in the housing data. Therefore, in this paper, a spatial lag variable was incorporated into the quantile model. But the presence of the spatial lag term in the right-hand side introduces endogeneity in the model, which will make biased and inconsistent estimators. There are two commonly used alternative estimation procedures: instrumental variable (IV) estimation and maximum likelihood estimation. The former is more robust than the latter in the sense that it does not require the error term to be normally distributed. In this study, the IV estimation, or more specifically the two stage least squares (2SLS) was used, this mixed method is called the two-stage quantile regression.

As the name implied, the Two-Stage Quantile Regression (2SQR) includes two steps, but actually I did three steps to finish this model: in the first step, I created spatial lag variable of housing price and all the independent variables. Then, I regressed the spatially lagged independent variables as well as the independent variables themselves against the spatial lag of housing price, and got the predicted value of the spatial lag of housing price. Finally, I ran the quantile regression of the housing price against all the characteristics and the predicted value I got from the last step. The reason why I used the predicted value instead of the true value is that it can eliminate correlation between the spatially lagged endogenous variable and the error term.

5 Results and Discussion

The results of the variables estimated using OLS, QR, GS2SLS and 2SQR were reported in Table 3 and Table 4. The first columns in Table 3 and Table 4 were OLS and GS2SLS model separately, the remaining columns in the two tables were QR and 2SQR models; the numbers in the parentheses were the bootstrapped standard errors.

Table 3 goes about here

Table 4 goes about here

The OLS and GS2SLS estimates were presented in the first column of Tables 3 and 4[6]. They both estimated on the entire data set. These regression analyses went beyond simple correlations and allowed us to separate the various effects of green space, house quality and location, and socio-economic characteristics, yielding a better picture of the impact of green space on sales price.

As mentioned above, in a spatial-lag model, a characteristic change of one parcel affects not only this parcel's price, but also the prices of the neighboring parcels, which may further influence some units far away. Therefore, the coefficients of the spatial model (in this paper, it is the GS2SLS and 2SQR model) does not represent the marginal effects that measure how changes in the exogenous variables affect the endogenous variable, which

[6] Coefficients for X_Sphat are all scaled.

means that the estimated β vector does not have the same interpretation as in a simple linear model because including a spatial lag of the dependent variable implies that the outcomes are determined simultaneously. The Average Total Direct Impact (ATDI) of the variables was calculated as the marginal effects (LeSage and Pace, 2009) in GS2SLS and 2SQR. The ATDI is the average over $i = \{1, . . . , n\}$ of the changes in the $\hat{y}_i$ attributable to the changes in the corresponding x_{ik}; it measures the average change in $\hat{y}_i$ attributable to sequentially changing x_{ik} for a given k. We can calculate the reduced-form predictor $\hat{y}$ = E [y|X,W,M] = $(I - \lambda W)^{-1} X\beta$. This expression for the predictor shows that change in a single observation on an exogenous variable will typically affect the values of the endogenous variable for all n units because the spatial model forms a system of simultaneous equations. ATDI can be calculated by computing $\hat{y}(x_k)$, $\hat{y}(x_k + \delta_i)$, and the average of the difference of these vectors of predicted values, where δ is the magnitude by which x_{ik} is changed. Since the result is not changed a lot when compared to the corresponding coefficients in 2SQR columns, I just the ATDI for GS2SLS model in Table 5.

Table 5 goes about here

Therefore, now, we could compare the signs and magnitudes of the coefficients between the GS2SLS and OLS model: for the non-green variables, most of the coefficients have the same signs and levels of significances, except for some structural characteristics variables, such as remodel and number of full bathrooms, which indicates that the structural characteristics are highly spatial-dependent. In other words, correcting spatial autocorrelation has greater impact on the structural characteristics. And, we can see that all the coefficients' magnitudes are changed. In general, the absolute value for most coefficients in GS2SLS are smaller than in OLS, which indicates that after correcting the spatial autocorrelation, the characteristics' effect on housing price become weak.

However, it was not done, as general comparison between models is not this paper's focus. Variables that represented green space are distance and area index of woodland, percentage of forest area in each census tract, and nearest distance to forested amenities area. Besides, there are also some other good indicators for green space; for example, some research use golf courses to represent green space (Bolitzer and Netusil 2000;

Lutzenhiser and Netusil 2001). In this paper, I simply use the distance to nearest golf course since its data are point shape file. Both sign and statistical significance are the same in OLS and GS2SLS for woodland distance area index. Most importantly, it shows a fairly strong positive correlation with selling price, meaning that shorter distance and larger area of woodland around the house are associated with houses that sell for more money. The intuitively expectation for the effect of percentage of the census tract area occupied by forest and effect of distance to forested amenity on housing price were both quite ambiguous, no a priori expectations for the sign of the coefficients. Because although forested amenities are the major means to carry out outdoor activities and provide enjoyable green views and a greater green coverage of area around the house could result in higher house price, a portion of the value was actually also reduced by location and traffic condition. This is because there was a statistical tendency for areas far from urban area to have more trees and traffic condition around forested amenities to be bad. The regression result shows that the coefficient estimate of distance to nearest forested amenity was positive and statistically significant suggests that the negative effect dominates the positive effect. The coefficient estimate for the forest percentage in census tract is negative in both OLS and GS2SLS and only significant in the first model, which indicates that there is a probability that this effect is zero.

The main difference between QR and 2SQR models is that they evaluate the characteristics' effect at the different housing price point before and after including the spatial effect. In order to better see the tendency of the coefficients change across quantiles, figures for each variable was created for both QR and 2SQR models in Figure 5 and Figure 6. It is the graph version of the results in Table 3 and Table 4.

Figure 5 goes about here

Figure 6 goes about here

The results from the QR and 2SQR show that most of the structural variables are not statistically significant except for age, number of bedrooms, garage and basement. Coefficients' signs of the significant structural variables are as intuitively expected for most quantiles. Coefficients' signs for the distance variables are as expected. The coefficient for the distance to railroad variable is positive and statistically significant in almost all quantiles, and the magnitude gets larger with the increasing quantile level. This suggests that house price increases with increasing distance from railroad, and the increasing amount grows for more expensive houses. This may be explained by the fact that railroad is likely to be associated with a noise disamenity or other inconvenience. Also, the road traffic variable is negative and statistically significant for all quantile levels as expected. The coefficients for the distances to fire district, medical center and police office are statistically significant and the signs are as expected. The variable for the cross product of distance and shape of park and lake are both positive as expected, and are statistically significant at medium or high quantile level. Besides, the coefficients of socio-economic variables from the census tract level, such as percentage of population who travels more than 90 minutes to work, percentage of population with college undergraduate or graduate or professional school degree, percentage households with high income are of the predicted sign with statistical significance in most quantiles.

For the green space variables, when we first see numbers in Table 3 and Table 4, as well as lines in Figure 5 and Figure 6, we can find substantial variation of the coefficient of area and distance index of woodland across quantiles, there is more than a 400 percent difference between the coefficient for the 0.2 quantile and the 0.9 quantile. The positive sign reported in Tables 3 and 4 in this quantile range suggests that houses closer to larger green space sell for relatively more amount, and the statistical significance for high quantile coefficients also confirm this effect. Moreover, the increasing magnitudes coefficients reveal that there is a higher positive effect for green space in higher-priced homes. Both Figure 5 and Figure 6 show a sudden increase of the coefficient in the 0.9 quantile, the positive, significant and large magnitude imply the strong preference to green space of richest people. As we all know an obvious interrelation, which is difficult to disentangle, occurs between social status and attractive location. People who can afford to do so have a tendency to choose attractive, green settings for their homes. As a consequence, certain towns or districts in attractive, green settings have become known as places for the rich. Buyers in these areas are willing to pay more premiums for the attractive environmental setting, such as green space. Consequently, houses in these areas are the highest priced.

The result shows an opposite effect for the nearest distance to forested amenity area: the regression coefficients are positive and not changed a lot across quantiles, indicating that the negative location effect as discussed previously dominates the positive forest effect, and this effect does not change a lot with increasing house price. The coefficients are statistically significant primarily in the middle and higher price ranges and are not significant in the lowest price range. This means that forested amenity may has no effect on low-priced houses. The percentage of forest area in each census tract has a negative effect on housing price in all quantiles, but the magnitudes of coefficients experiences a wave in the middle percentiles and they are statistically significant at this range. This reveals that both cheapest and luxury houses have no obvious relationship with green space around, but it has a significant negative influence on the middle level housing price. The underlying economic reason for this result may be tied to the fact that area with a higher green percentage is always far away from working places and downtown area. Furthermore, if buyers of more expensive houses bear more taxes or fees of provision of green space while all home owners nearby equally benefit from it, then households who purchase lower-priced units would have a premium by enjoying a higher green percentage but those who buy higher-priced units would not. This will make middle-priced houses in higher green percentage area less competitive, which will decrease the house price in this kind of area. Usually, the people who buy luxury houses are the richest people, they may not care about this amount, and so does the poorest people since the amount they pay are quite low, they actually are "free-riding", that's why coefficients are not significant at both highest and lowest quantiles.

6 Conclusion

Intuitively, we felt that houses in attractive locations will have an added value over similar, less favorably located houses. But, the definition of attractive location is quite complicated; it depends on many factors such as people's income and their preference. By excluding the effect of different characteristics and factors on housing price, this study finds that green space was not always a positive factor that can make a house sell in a higher price; location, traffic, tax and other factors may affect the housing price as green space accessories. As the result shows that the overall impact of green space was ambiguous. Furthermore, separating houses in different levels makes the relationship more complex. People with different income level may have huge difference on valuation to green space, this partly reflect house price. Generally speaking, the results show that green space effect was only significant for middle and high priced houses, but it has no significant effect on the bottom level home prices. Nearby woodland has increasing positive effect with increasing housing price. But house near forest amenities suffer from this location factor no matter what level the house is, thus price reduced by this factor, but the negative effect was almost the same for houses in all quantiles. Forest coverage percentage in census tract is a negative factor in housing price, especially for medium level houses.

Methodologically, this paper first demonstrates the importance of incorporating spatial effects in hedonic housing price models when assessing the effect of green space on housing prices. The incorporation of spatial dependence into the hedonic model (GS2SLS) illustrates how OLS estimates from traditional hedonic housing price models tend to overestimate the parameters on explanatory variables. Also, as shown in the result section, we can see more detailed relationships between the green space variables and housing price by using quantile regression, since this method helps us to see the relationship changes over different price ranges. The quantile regression results confirm that the effect of nearest distance to forested amenity area remains relatively stable across different price ranges. For other variables that are not statistically significant in the GS2SLS estimation such as the percentage of forest area in each census tract, the quantile regression results suggest that they are not insignificant over all price ranges. This study also uses the Average Total Direct Impact (ATDI) to estimate the marginal prices. All the coefficients derived from the spatial adjustments were different than those resulting from an OLS model. This suggests that using spatial information in hedonic studies is necessary.

In sum, the general hypothesis that an attractive, wooded landscape attracts a premium on the house price had to be rejected. An attractive environment with more green space around is a perfect place for a high-priced house, but this green space did not include the green space in public amenities, and attractive environment means the percentage of green area in whole census tract should not be too large. Finally, green space has no effect or negatively affects the housing price for median or low level houses.

References

Abadie, A., J. Angrist, and G. Imbens. 2002. "Instrumental Variables Estimates of The Effect of Subsidized Training on The Quantiles of Trainee Earnings." *Econometrica* 70:91–117.

Abadie, A., J. Angrist and G. Imbens. 2001. "Instrumental Variables Estimation of Quantile Treatment Effects." Econometrica. Forthcoming.

Amemiya, T. 1982. "Two stage least absolute deviations estimators." Econometrica 50: 689–711.

Anselin, L. 1992. SpaceStat Software for Spatial Data Analysis.

A.T. Court. 1939. "Hedonic Price Indexes with Automotive Examples, in the Dynamics of Automobile Demand," General Motors, New York. pp. 98-119.

Anderson, L.; Cordell, H. 1985. "Residential Property Values Improved by Landscaping with Trees." Southern Journal of Applied Forestry 9: 163-166.

Anderson, L.M., Cordell, H.K. 1988. "Influence of Trees on Residential Property Values in Athens, Georgia (USA): A Survey Based on Actual Sales Price." Landscape Urban Plann. 15: 153–164.

Anselin L. 2001. Spatial econometrics, in Baltagi B. (ed.) A Companion to Theoretical Econometrics, Blackwell, Oxford, 310–330.

Basu, S. and Thibodeau, T. G. 1998. "Analysis of Spatial Autocorrelation in House Prices." *Journal of Real Estate Finance and Economics* 17: 61-85.

Bell, K. P. and Bockstael. N. E. 2000. "Applying the Generalized-Moments Estimation Approach to Spatial Problems involving Micro-level Data." *Review of Economics and Statistics* 82: 72-82.

Bockstael, N.E. and Bell, K.P. 1997. "Land Use Patterns and Water Quality: The Effect of Differential Land Management Controls. Conflict and Cooperation in Trans-Boundary Water Resources. Just, R.E., Netanyahu, S., eds. Boston, USA: *Kluwer Academic Publishers*: 169– 191.

Bowen, W. M., Brian Mikelbank, and D. M. Prestegaard. 2001. “Theoretical and Empirical Considerations Regarding Space in Hedonic Housing Price Model Applications.” *Growth and Change* 32: 466–490.

Bransington D. M. and D. Hite. 2005. “Demand for Environmental Quality: A Spatial Hedonic Analysis.” *Regional Science and Urban Economics* 35: 57– 82.

Brown, J. and H. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models" *Econonmetrica* 50: 765-768.

Bolitzer, B., and N. Netusil. 2000. “The Impact of Open Space on Property Value in Portland, Oregon.” J. Environ. Manage. 59:185-93.

Biazee.1992. "Measuring Hazardous Waste Damages with Panel Models." Journal of Environmental Economics and Management 22: 259-71.

Can. A. 1990. “The Measurement of Neighborhood Dynamics in Urban House Prices,” *Economic Geography* 66: 254–272.

Clapp, J. M., M. Rodriguez, and G. Thrall. 1997. “How GIS Can Put Urban Economic Analysis on the Map”. *Journal of Housing Economics* 6: 368-386.

Conley, T. and D. Galenson. 1998. “Nativity and Wealth in Mid-Nineteenth-Century Cities.” *Journal of Economic History* 58: 468-93.

Deaton, A. 1997. The Analysis of Household Surveys. Baltimore: Johns Hopkins.

Din Garrod, G. and G.W. Kenneth. 1999. “Economic Valuation of the Environment: Methods and Case Studies.” Massachusetts: Edward Elgar Publishing Inc.

Dubin, R. A. 1988. “Estimation of Regression Coefficients In the Presence of Spatially Auto-Correlated Error Terms.” *Review of Economics and Statistics* 70: 466-474.

Dubin, R. A. 1992. “Spatial autocorrelation and neighborhood quality.” *Regional Science and Urban Economics, 22*: 433-452.

Dubin, R., R.K. Pace, and T. Thibodeau. 1999. “Spatial Autoregression Techniques for Real Estate Data,” *Journal of Real Estate Literature* 7: 79-95.

Darling, A.H. 1973. “Measuring Benefits Generated by Urban Water Parks.” Land Econ XLlX (I): 22-34.

Eide, E. and Mark S. 1998. “The Effect of School Quality on Student Performance: A Quantile Regression Approach.” *Economics Letters* 58: 345-50.

Fennema, A.T., Veeneklaas, F.R., Vreke, J. 1996. “Meerwaarde woningen door nabijheid van groen (Surplus Value of Dwellings in the Vicinity of Green Areas).” *Stedebouw en Ruimtelijke Ordening* 3: 33-35.

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: 240–252

Garrod, G. and G.W. Kenneth. 1999. “Economic Valuation of the Environment: Methods and Case Studies.” Massachusetts: Edward Elgar Publishing Inc.

Gawande, K., and Jenkins-Smith, H. 2001. “Nuclear Waste Transport and Residential Property Values: Estimating The Effects Of Perceived Risks.” *Journal of Environmental Economics and Management 42*: 207-233.

Geoghegan, J, L. A. Wainger, and N. E. Bockstael. 1997. “Spatial Landscape Indices in a Hedonic Framework: an Ecological Economic Analysis Using GIS”. *Ecological Economics* 23: 251-264.

Geoghegan, J., Wainger, L. A., and Bockstael, N. E. 1997. Spatial landscape indices in a hedonic framework: An ecological economics analysis using GIS. *Ecological Economics*, *23*: 251-264.

Gillen, K., Thomas T., and S. Wachter. 2001. “Anisotropic Autocorrelation in House Prices.” *Journal of Real Estate Finance and Economics* 23: 5-30.

Gould, W.W. 1992. “Quantile Regression with Bootstrapped Standard Errors.” *Stata Technical Bulletin* 9: 19-21.

Grey. C. and Dcneke. F. 1978. Urban Forestry. *John Wiley, Chichester* 279.

Gyourko, J. and Tracy, J. 1999. “A Look at Real Housing Prices and Incomes: Some Implications for Housing Affordability and Quality.” *Federal Reserve Bank of New York Economic Policy Review*, September, pp. 63–77.

He, L.T. and Winder. R.C. 1999. “Price Causality between Adjacent Housing Markets within a Metropolitan Area: A Case Study.” *Journal of Real Estate Portfolio Management* 5 (1): 47–58.

Irwin, Elena G. 2002. “The Effects of Open Space on Residential Property Values.” *Land Economics* 78: 465-480.

Kelejian H.H. and Prucha I.R. 1998. “A Generalized Spatial Two Stage Least Squares Procedure For Estimating A Spatial Autoregressive Model With Autoregressive Disturbances.” *Journal of Real Estate Finance and Economics* 17: 99-121.

Kemp, G. 1999. “Least Absolute Error Difference Estimation of a Single Equation From A Simultaneous Equations System.” Mimeo, University of Essex.

Kestens, Y., Thériault, M. and Des Rosiers. F. 2004. “The Impact of Surrounding Land Use and Vegetation On Single-Family House Prices.” *Environment and Planning B: Planning and Design* 31: 539-567.

Kim, C. W. and Phipps, T. T. and Anselin, L. 2003. “Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach.” *Journal of Environmental Economics and Management 45*(1): 24-39.

Kim, Chong Won, Tim T. Phipps, and Luc Anselin. 2003. “Measuring the Benefits of Air Quality Improvement: A Spatial Hedonic Approach,” *Journal of Environmental Economics and Management* 45: 24–39.

Kim, T.H., Muller, C., 2004. “Two-Stage Quantile Regression when the First Stage is Based on Quantile Regression.” *Econometric Journal* 7: 218-231.

Knight, Keith, Gilbert Bassett and Mo-Yin S. Tam. 2000. “Comparing Quantile Estimators for the Linear Model.” Preprint.

Koenker, Roger and Yannis Bilias. 2001. “Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments.” *Empirical Economics* 26: 199-220.

Koenker, R. and G. Bassett. 1978. “Regression Quantiles.” *Econometrica* 46: 33-50.

Koenker, R. and K. F. Hallock. 2001. “Quantile Regression.” *Journal of Economic Perspectives* 15: 143-56.

Leggett, C. G., and Bockstael, N. E. 2000. “Evidence of the Effects of Water Quality on Residential Land Prices.” *Journal of Environmental Economics and Management 39*: 121-144.

Levin, Jesse. 2001. “For Whom the Reductions Count: A Quantile Regression Analysis of Class Size on Scholastic Achievement.” *Empirical Economics* 26: 221-46.

LeSage, J. P., and R. K. Pace. 2009. Introduction to Spatial Econometrics. Boca Raton, FL: CRC Press Taylor & Francis Group.

Lutzenhiser, M., and N. Netusil. 2001. “The Effect of Open Space on a Home’s Sale Price.” *Contemporary Econ. Pol.* 19:291-98.

Laurie, I. 1979. “The Natural Environment in the Design and Development of Urban Green” John Wiley, Chicbester. 428.

Lindsey, G., Man, J., Payton, S., and Dickson, K. 2004. “Property Values, Recreation Values and Urban Greenways.” *Journal of Parks and Recreation Administration* 22: 69-90.

MaCurdy, T. and C. Timmins 2000. “Bounding The Influence of Attrition on Intertemporal Wage Variation in the NLSY.” Mimeo Stanford University.

McMillen, D. P. and Thornes, P. 2006. “Housing Renovation and The Quantile Repeated-Sales Price Index.” *Real Estate Economics* 34: 567–584.

McMillen, D.P., 2008. “Changes in the Distribution of House Prices over Time: Structural Characteristics, Neighborhood, or Coefficients.” *Journal of Urban Economics* 64: 573-589.

McMillen, D.P., and Coulson N.E. 2007. “The Dynamics of Intra-urban Quantile House Price Indexes.” *Urban Studies* 44: 1517-1537.

Morillo, Daniel. 2000. “Income Mobility with Nonparametric Quantiles: A Comparison of the U.S. and Germany.” Preprint.

Mueller, R. 2000. “Public- and Private-Sector Wage Differentials in Canada Revisited.” *Industrial Relations* 39: 375-400.

Munneke, Henry J. and V.Carlos Slawson. 1999. “A Housing Price Model with Endogenous Externality Location: A Study of Mobile Home Parks.” *Journal of Real Estate Finance and Economics* 19: 113-131.

Morales, D.J. 1980. “The Contribution of Trees to Residential Property Value.” *J. Arboricol* 7: 109- 12

Morales, D.J., Micha, F.R. and Weber. R.L. 1983. “Two Methods of Valuating Trees on Residential Sites.” *J. Arboricol* 9: 21-24.

More, T.A., Stevens, T., Allen, P.G. 1988. “Valuation of urban parks.” *Landscape Urban Plann.* 15: 139–152.

Milon. J. Walter, Jonathan Gressel, and David Muikey. 1984. "Hedonic Amenity Valuation and Functional Form Specification." *Land Economics* 60: 378-87.

Mendelsohn, Robert O., Daniel Hellerstcin, Michael Huguenin, Robert E. Unsworth, and R. Nelson, Arthur C, John Genereux, and Michelle G.1992. "Price Effects of Landfills on House Values." *Land Economics* 68: 359-65.

Thayer, and Don Waldman.1988. "The Robustness of Hedonic Price Estimation: Urban Air Quality." *Land Economics* 64: 220-33.

Kiel, Katherine A.1995. "Measuring the Impact of the Discovery and Cleaning of Identified Hazardous Waste Sites on House Values." *Land Economics* 71:428-35.

Lansford, Notie H, and Lonnie L. Jones. 1995."Marginal Price of Lake Recreation and Aesthetics: A Hedonic Approach." *Journal of Agricultural and Applied Economics* 27: 212-23.

Mahan, Brent L., Stephen Polasky, and Richard M. Adams. 2000. "Valuing Urban Wetlands: A Property Price Approach." *Land Economics* 76: 100-13.

McLeod, P.B. 1984. “The demand for local amenity: a hedonic price analysis.” *Environ. Plann.* 16: 389–400.

Pebesma, E., Wesseling, C. G. 1998. “Gstat: A Program for Geostatistical Modeling, Prediction, and Simulation.” *Computers and Geosciences* 24: 17-31.

Poterba, James and Kim Rueben. 1995. “The Distribution of Public Sector Wage Premia: New Evidence Using Quantile Regression Methods.” *NBER Working Paper No.* 4734.

Powell, J. 1983. “The Asymptotic Normality of Two-Stage Least Absolute Deviations Estimators.” *Econometrica* 51: 1569–75.

Paterson, R., and K. Boyle. ‘‘Out of Sight, Out of Mind? 2002. Using GIS to Incorporate Visibility in Hedonic Property Value Models.’’ *Land Economics* 78:417–25.

Paelinck, J.H.P. and Klaassen, L.H. 1979. Spatial Econometrics, Saxon House Farnborough

Powe, Neil A., Guy Garrod, Chris Brunsdon, and Ken Willis. 1997. "Using a Geographic Information System to Estimate a Hedonic Price Model of the Benefits of Woodland Access." *Forestry* 70: 139-49.

Palmquist. R. 1991. “Measuring the Demand for Environmental Quality.” *Elsevier, North-Holland.* pp 77-120.

Rogers, W. H. 1993. “Calculation of Quantile Regression Standard Errors.” *Stata Technical Bulletin* 13: 18-19.

Rosen, S.1974. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition.” *Journal of Political Economy* 82: 34-55.

Robinette, C. 1972. “Plants, People and Environmental Quality.” Department of the Interior. National Fark Service, Washington DC, 40

Sakata, S. 2001. “Instrumental Variable Estimation Based on Mean Absolute Deviation.” Mimeo University of Michigan.

Strand, J., Vagnes, M., 2001. “The relationship between property values and railroad proximity: a study based on hedonic prices and real estate brokers appraisals.” *Transportation* 28, 137–156.

Viscusi, W. Kip and James Hamilton. 1999. “Are Risk Regulators Rational Evidence from Hazardous Waste Cleanup Decisions.” *American Economic Review* 89: 1010-27.

Zietz, J., Zietz, E.N. and Sirmans, G.S. 2008. “Determinant of Housing Prices: A Quantile Regression Approach.” *Journal of Real Estate Finance and Economics* 37: 317–333.

Table 1. Variable Descriptions and Expected Signs

Variable	Description	GIS	Expected Sign
Structural characteristics			
BASE_DUMMY	Availability of Basement (0 or 1)	NO	+
BDROOMS	Number of Bedrooms	NO	+
FAMROOMS	Number of Family-Rooms	NO	+
DINROOMS	Number of Dining-Rooms	NO	+
GARAGE_CAP	Number of Garages	NO	+
FULBATHS	Number of Full Bathrooms	NO	+
FIREPL_STA	Number of Fireplaces	NO	+
AGE	The age of the house	NO	-
AVERAGE_PRICE	The average price of parcels in each census tract	YES	+
Amenities characteristics			
NEAR_DISTFIRE	Closest Distance to Fire Districts	YES	?
NEAR_DISTGOLF	Closest Distance to Golf Course	YES	-
NEAR_DISTMEDICAL	Closest Distance to Medical Center	YES	+
NEAR_DISTPOLICE	Closest Distance to Police Office	YES	-
NEAR_DISTPOSTOFFICE	Closest Distance to Post Office	YES	-
NEAR_DIST_RAILROAD	Closest Distance to Railroad	YES	+
RTI	A measure of road traffic, RTI $=n^2/\sum_{i=1}^{n} d_i$,where n is the number of schools within a 2 mile buffer, and d_i is the distance from the centroid of the housing cluster to school i; or RTI=1/D, where D is the nearest distance from the centroid of the housing cluster to school when there is no school within 2 mile buffer	YES	-
AD_PARK	A measure of the integrated impact of the nearest park area and straight-line distance from the housing centroid to the boundary of the nearest park	YES	-
AD_PL	A measure of the integrated impact of the nearest ponds and lakes area and straight-line distance from the housing centroid to the boundary of the nearest ponds and lakes	YES	-
AD_WOODLAND	A measure of the integrated impact of the nearest woodland area and straight-line distance from the housing centroid to the boundary of the nearest woodland	YES	+
NEAR_DIST_FORESTAM	The nearest distance to forested amenity area	YES	?
PERCENT_WOOD	The percentage of forest area in each census tract	YES	?
CITY	Whether the parcel in city area or not	YES	+
Socio-economic characteristics			
VAC_RATIO	Vacancy ratio = Vacant Housing Unites/Total Housing Unites	YES	-
HH_SIZE	The number of people living in the Respondent's Household	YES	?
WHITE	Percentage of White Population in Census Tract	YES	+
BLACK	Percentage of Black Population in Census Tract	YES	-
ASIAN	Percentage of Asian Population in Census Tract	YES	-
TMR	Percentage of Two or More Races Population in Census Tract	YES	?
TTW	Percentage of Population Who Travels More Than 90 Minutes to Work	YES	-
WOR	Percentage of Population Worked Outside Place of Residence	YES	-
PNS	Percentage of Population not Enrolled In School	YES	-
PAC	Percentage of Population with College Undergraduate or Graduate or Professional School Degree	YES	+
PMI	Percentage Households with Medium Income (%)US $ 50,000 - US $ 99,000)	YES	+

PHI	Percentage Households with High Income (%)US $ 100,000 and over	YES	+
PGH	Percentage of House Use Gas as Heating Fuel	YES	?
PEH	Percentage of House Use Electricity as Heating Fuel	YES	?
X_Sphat	Predicted Value of Spatial Lag of the Housing Price	NO	?
SALE_AMNT	Price of the Single-Family House Sale	YES	

Table 2. Summary Statistics of the Variables

Variable	Mean	Std. Dev.	Min	Max
LSALEPRICE	12.3568	0.6086	9.4335	14.8088
AGE	21.4584	27.6927	1.0000	171.0000
REMOD	0.0512	0.2204	0.0000	1.0000
BDROOMS	3.5826	0.6871	0.0000	8.0000
FAMROOMS	0.6320	0.5058	0.0000	3.0000
DINROOMS	0.5977	0.4923	0.0000	2.0000
FULBATHS	2.1718	0.7877	0.0000	7.0000
FIREPL_STA	0.8451	0.4766	0.0000	4.0000
BASE_DUMMY	0.9132	0.2816	0.0000	1.0000
GARAGE_CAP	1.3737	1.2228	0.0000	6.0000
NEAR_DISTFIRE	1.5027	0.8903	0.1123	6.3006
NEAR_DISTGOLF	1.4780	1.0966	0.0730	10.6318
NEAR_DISTMEDICAL	2.8951	2.2154	0.1265	13.9460
NEAR_DISTPOLICE	2.3560	1.4841	0.0843	12.0236
NEAR_DISTPOSTOFFICE	2.2236	1.2599	0.0469	7.6603
AD_WOODLAND	0.0071	0.0032	0.0009	0.0194
NEAR_DIST_FORESTAM	0.5733	0.6334	0.0000	8.6952
PERCENT_WOOD	0.2485	0.1109	0.0428	0.7072
AD_PARK	0.0075	0.0030	0.0007	0.0158
NEAR_DIST_RAILROAD	0.0211	0.0243	0.0019	0.2950
AD_PL	3.8284	1.3567	-0.2376	11.0881
RTI	0.0009	0.0006	0.0000	0.0029
AVERAGE_PRICE	269962.5340	101761.6433	86019.2778	451089.5000
CITY	0.3200	0.4666	0.0000	1.0000
VAC_RATIO	6.9839	4.3493	0.0000	17.3700
HH_SIZE	2.1041	0.4987	1.0000	4.0000
WHITE	0.8925	0.0628	0.7818	1.4334
BLACK	0.0413	0.0349	0.0000	0.1689
ASIAN	0.0401	0.0305	0.0000	0.1219
TMR	0.0198	0.0106	0.0000	0.0491
TTW	0.0146	0.0108	0.0000	0.0448
WOR	0.2508	0.2512	0.0000	0.7060
PNS	0.7071	0.0463	0.6286	1.1193
PAC	0.0506	0.0173	0.0262	0.1010
PMI	0.3182	0.0780	0.1864	0.5215
PHI	0.4497	0.1915	0.0813	0.9414
PGH	0.7176	0.2179	0.1169	1.3122
PEH	0.1506	0.0789	0.0430	0.3098
N	2247			

Table 3. Estimates and Statistical Significance of the Parameters in OLS and QR

	OLS	QR 0.1	QR 0.2	QR 0.3	QR 0.4	QR 0.5	QR 0.6	QR 0.7	QR 0.8	QR 0.9
AGE	0.0012**	0.0003	0.0005	0.0008	0.0007	0.0005	0.0008*	0.0008***	0.0006	0.0004
	(2.72)	(0.49)	(1.45)	(1.56)	(1.48)	(1.73)	(2.20)	(3.75)	(1.43)	(0.97)
REMOD	-0.0078	0.0431	0.0506	0.0201	0.0048	-0.0073	-0.0167	-0.0126	0.0015	-0.0058
	(-0.17)	(0.56)	(0.97)	(0.37)	(0.18)	(-0.17)	(-0.45)	(-0.33)	(0.04)	(-0.11)
BDROOMS	0.0089	0.0202	0.0167	0.0138	0.0252	0.0244	0.0260	0.0160	0.0093	-0.0112
	(0.54)	(0.79)	(0.90)	(0.64)	(1.91)	(1.74)	(1.76)	(1.05)	(0.57)	(-0.48)
FAMROOMS	-0.0049	0.0164	0.0015	-0.0000	0.0101	0.0061	-0.0016	-0.0120	-0.0276	-0.0418*
	(-0.24)	(0.43)	(0.06)	(-0.00)	(0.71)	(0.39)	(-0.08)	(-0.91)	(-1.09)	(-2.14)
DINROOMS	0.0252	-0.0285	-0.0090	0.0085	0.0024	0.0077	0.0291	0.0536**	0.0549	0.0751***
	(1.13)	(-0.60)	(-0.33)	(0.41)	(0.12)	(0.33)	(1.26)	(2.95)	(1.95)	(3.49)
FULBATHS	0.0039	-0.0075	-0.0138	-0.0123	-0.0068	-0.0119	-0.0004	-0.0005	0.0045	0.0167
	(0.25)	(-0.34)	(-0.65)	(-0.77)	(-0.43)	(-0.95)	(-0.03)	(-0.04)	(0.33)	(0.83)

FIREPL_STA	0.0072	-0.0167	-0.0068	-0.0068	-0.0052	0.0056	0.0153	0.0105	0.0172	-0.0011
	(0.32)	(-0.40)	(-0.24)	(-0.23)	(-0.32)	(0.36)	(0.92)	(0.64)	(0.85)	(-0.06)
BASE_DUMMY	0.5281***	0.7146***	0.6128***	0.5974***	0.4674***	0.4482***	0.4409***	0.4401***	0.4281***	0.4669***
	(15.60)	(7.47)	(8.13)	(9.34)	(7.61)	(10.14)	(13.98)	(14.90)	(13.04)	(10.26)
GARAGE_CAP	-0.0134	0.0000	-0.0197	-0.0214*	-0.0214	-0.0243**	-0.0240***	-0.0275***	-0.0323***	-0.0264**
	(-1.64)	(0.00)	(-1.75)	(-2.11)	(-1.68)	(-3.14)	(-3.33)	(-3.74)	(-3.54)	(-2.90)
NEAR_DISTFIRE	-0.0359*	-0.0289	-0.0490	-0.0340	-0.0308	-0.0243	-0.0223	-0.0064	-0.0142	-0.0098
	(-2.43)	(-0.64)	(-1.47)	(-1.41)	(-1.30)	(-1.02)	(-1.13)	(-0.27)	(-0.72)	(-0.70)
NEAR_DISTGOLF	-0.0840***	-0.1491***	-0.0946***	-0.0892***	-0.0869***	-0.0805***	-0.0688***	-0.0608***	-0.0617***	-0.0506**
	(-7.18)	(-4.67)	(-5.07)	(-5.15)	(-4.19)	(-4.51)	(-5.60)	(-4.70)	(-3.98)	(-2.93)
NEAR_DISTMEDICAL	-0.0571***	-0.0742***	-0.0561***	-0.0459***	-0.0485***	-0.0448***	-0.0443***	-0.0515***	-0.0699***	-0.0980***
	(-6.85)	(-4.09)	(-3.43)	(-3.69)	(-6.40)	(-3.30)	(-3.52)	(-6.88)	(-4.97)	(-10.84)
NEAR_DISTPOLICE	-0.0357***	-0.0030	-0.0479***	-0.0411***	-0.0433***	-0.0445***	-0.0398***	-0.0342**	-0.0302*	-0.0306
	(-3.93)	(-0.18)	(-3.31)	(-4.80)	(-3.35)	(-5.58)	(-4.14)	(-2.77)	(-2.20)	(-1.96)
NEAR_DISTPOSTOFFICE	0.0344**	0.0024	0.0071	0.0040	0.0137	0.0168	0.0360*	0.0469**	0.0646***	0.0867***
	(2.60)	(0.08)	(0.32)	(0.18)	(0.70)	(1.13)	(2.08)	(2.86)	(4.37)	(4.64)
AD_WOODLAND	11.4348***	-3.0553	6.8793	9.6117**	11.8365***	14.9006***	17.2668***	18.9499***	19.8774***	28.9386***
	(3.64)	(-0.32)	(1.43)	(2.72)	(4.69)	(4.88)	(4.38)	(5.94)	(4.91)	(5.84)
NEAR_DIST_FORESTAM	0.1220***	0.1187	0.1210**	0.1365***	0.1550***	0.1337***	0.1376***	0.1351***	0.1361***	0.1641***
	(5.78)	(1.73)	(3.03)	(4.86)	(5.77)	(3.32)	(4.86)	(4.30)	(4.15)	(4.42)
PERCENT_WOOD	-0.3461*	-0.2311	-0.6078**	-0.7126***	-0.6078***	-0.7063***	-0.7609***	-0.6194***	-0.4232	-0.2192
	(-2.27)	(-0.65)	(-2.64)	(-3.68)	(-3.67)	(-5.64)	(-6.08)	(-4.47)	(-1.83)	(-0.81)
AD_PARK	17.9940***	11.3074	18.5506***	19.2374***	20.2074***	20.1022***	21.1818***	17.9619***	17.5527***	20.5434***
	(4.40)	(1.00)	(3.91)	(3.76)	(4.19)	(6.79)	(4.22)	(5.49)	(4.05)	(4.52)
NEAR_DIST_RAILROAD	3.6158***	0.5194	1.4327*	1.9015*	2.2162**	3.3517***	3.9807***	3.7717***	5.2682***	7.1305***
	(8.20)	(0.44)	(2.07)	(2.54)	(3.23)	(4.01)	(5.29)	(5.07)	(4.03)	(8.33)
AD_PL	0.0278***	0.0307	0.0315***	0.0259***	0.0219***	0.0195**	0.0151*	0.0146*	0.0112	0.0053
	(3.91)	(1.59)	(3.79)	(4.35)	(3.36)	(2.88)	(2.03)	(2.45)	(1.57)	(0.53)
RTI	-191.8306***	-312.5770***	-233.4290***	-181.5194***	-168.1121***	-184.9356***	-159.4514***	-161.1920***	-136.7645***	-122.5995***
	(-6.57)	(-3.67)	(-6.17)	(-4.55)	(-4.57)	(-6.15)	(-5.27)	(-6.12)	(-5.03)	(-3.30)
AVERAGE_PRICE	0.1330	-0.1929	0.3394	0.6582***	0.4753***	0.3504**	0.2135	0.2002	0.1145	-0.0923
	(0.92)	(-0.47)	(1.33)	(4.24)	(3.34)	(2.67)	(1.28)	(1.13)	(0.90)	(-0.46)
CITY	0.0177	-0.0351	0.0086	0.0188	0.0714**	0.0613**	0.0678**	0.0790***	0.0809***	0.0130
	(0.70)	(-0.78)	(0.28)	(0.58)	(2.75)	(2.61)	(2.61)	(3.68)	(3.37)	(0.46)
VAC_RATIO	-0.0008	0.0033	0.0054	-0.0009	-0.0051	-0.0046	-0.0076*	-0.0072	-0.0100*	-0.0093
	(-0.21)	(0.33)	(1.07)	(-0.15)	(-1.34)	(-1.29)	(-2.23)	(-1.33)	(-2.37)	(-1.45)
HH_SIZE	0.0532	0.0128	0.0684	0.1125**	0.0964***	0.0411	0.0145	0.0258	0.0378	0.0266
	(1.88)	(0.17)	(1.28)	(3.19)	(3.88)	(1.85)	(0.52)	(0.69)	(1.00)	(0.69)
WHITE	1.2321*	-0.6517	-0.0525	0.1000	0.3081	1.3521	1.6204	2.3011	2.3325*	4.3270***
	(2.28)	(-0.39)	(-0.05)	(0.09)	(0.45)	(1.56)	(1.47)	(1.79)	(2.05)	(3.50)
BLACK	1.8214**	0.8126	0.3414	0.9363	1.8315*	2.3543*	2.2268*	2.7789	3.1792**	5.5262***
	(2.59)	(0.43)	(0.30)	(0.80)	(2.04)	(2.07)	(1.96)	(1.96)	(2.84)	(3.94)
ASIAN	-0.5748	-4.3135*	-4.1411**	-3.1153	-2.1417	0.0945	0.7154	2.4798	2.5801	3.5084*
	(-0.68)	(-2.04)	(-3.27)	(-1.95)	(-1.86)	(0.08)	(0.50)	(1.61)	(1.53)	(2.44)
TMR	2.7072	-4.5907	-1.4027	-1.2931	-0.0349	1.5356	3.2446	5.2490*	6.6266**	9.5390***
	(1.86)	(-0.97)	(-0.64)	(-0.55)	(-0.02)	(0.96)	(1.82)	(2.32)	(3.01)	(3.64)
TTW	-3.8364**	-1.6235	-4.8431*	-4.3232**	-4.1987**	-4.0237**	-3.7622**	-3.2890	-2.9261	-4.2393***
	(-2.71)	(-0.67)	(-2.39)	(-2.58)	(-2.68)	(-3.15)	(-2.88)	(-1.58)	(-1.64)	(-3.32)
WOR	0.0754	0.0179	-0.0873	-0.1011	-0.0528	0.0094	0.0572	0.1526	0.2571**	0.2654**
	(1.13)	(0.10)	(-0.94)	(-1.15)	(-0.63)	(0.15)	(0.78)	(1.74)	(3.07)	(2.75)
PNS	-1.0522*	-0.6293	-1.3108	-0.7847	-1.0262**	-1.1488***	-1.3227***	-1.1896	-1.3594	-1.7521*
	(-2.42)	(-0.61)	(-1.65)	(-1.55)	(-2.94)	(-3.36)	(-3.45)	(-1.75)	(-1.88)	(-2.38)
PAC	9.8716***	8.8742*	12.0132***	9.8322***	9.7678***	7.7664***	8.4573***	7.9402***	9.3666***	11.9148***
	(6.54)	(2.36)	(5.37)	(4.46)	(7.30)	(4.15)	(5.22)	(3.76)	(4.98)	(5.45)
PMI	-0.5498	1.2939	0.7922	0.3090	-0.2441	-0.4071	-0.8957***	-1.4084***	-2.1574***	-2.5305***
	(-1.90)	(1.39)	(1.61)	(0.70)	(-0.87)	(-1.70)	(-4.03)	(-3.78)	(-4.75)	(-6.48)
PHI	1.7993***	2.3908***	2.5856***	2.2318***	1.9464***	1.7548***	1.5797***	1.3442***	1.1748***	1.1564***

	(12.06)	(6.65)	(12.62)	(13.84)	(13.48)	(12.98)	(13.30)	(9.17)	(7.99)	(6.91)
PGH	-1.0739***	-0.5861	-0.9533**	-0.7209*	-0.9640***	-0.8579***	-0.9456***	-1.1693***	-1.6560***	-2.0586***
	(-6.10)	(-1.19)	(-2.98)	(-2.34)	(-5.73)	(-3.38)	(-4.27)	(-5.48)	(-6.60)	(-10.18)
PEH	-1.4932***	-0.9464	-1.3128*	-0.8677	-1.0587**	-0.9914**	-1.3398***	-1.8201***	-2.4094***	-3.0018***
	(-3.86)	(-0.99)	(-1.98)	(-1.64)	(-2.68)	(-2.85)	(-3.73)	(-3.65)	(-5.58)	(-5.44)
_CONS	11.2340***	11.6407***	11.7690***	11.2547***	11.8158***	11.2105***	11.3799***	11.0685***	11.8397***	10.6361***
	(29.98)	(14.13)	(15.03)	(12.00)	(14.83)	(11.94)	(10.28)	(11.76)	(12.53)	(8.21)

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. Estimates and Statistical Significance of the Parameters in GS2SLS and 2SQR

	GS2SLS	2SQR 0.1	2SQR 0.2	2SQR 0.3	2SQR 0.4	2SQR 0.5	2SQR 0.6	2SQR 0.7	2SQR 0.8	2SQR 0.9
main										
AGE	0.0010*	0.0005	0.0005	0.0008	0.0007	0.0005*	0.0008*	0.0009***	0.0005	0.0003
	(2.29)	(0.70)	(1.58)	(1.53)	(1.47)	(1.98)	(2.18)	(3.67)	(1.13)	(0.76)
REMOD	0.0040	0.0311	0.0406	0.0169	0.0055	-0.0093	-0.0099	-0.0199	-0.0003	-0.0003
	(0.09)	(0.40)	(0.84)	(0.30)	(0.21)	(-0.21)	(-0.27)	(-0.50)	(-0.01)	(-0.01)
BDROOMS	0.0026	0.0280	0.0187	0.0140	0.0254	0.0236	0.0261	0.0137	0.0122	-0.0110
	(0.16)	(0.95)	(1.01)	(0.66)	(1.87)	(1.66)	(1.71)	(0.93)	(0.72)	(-0.48)
FAMROOMS	-0.0061	0.0145	0.0010	0.0009	0.0103	0.0056	-0.0029	-0.0110	-0.0286	-0.0340
	(-0.31)	(0.40)	(0.03)	(0.05)	(0.73)	(0.36)	(-0.16)	(-0.78)	(-1.07)	(-1.83)
DINROOMS	0.0271	-0.0412	-0.0088	0.0087	0.0029	0.0098	0.0305	0.0521**	0.0553	0.0765**
	(1.25)	(-0.89)	(-0.28)	(0.42)	(0.13)	(0.41)	(1.29)	(3.15)	(1.90)	(3.10)
FULBATHS	-0.0001	-0.0044	-0.0150	-0.0113	-0.0105	-0.0099	-0.0001	0.0017	-0.0055	0.0199
	(-0.01)	(-0.19)	(-0.67)	(-0.65)	(-0.63)	(-0.76)	(-0.01)	(0.12)	(-0.39)	(0.91)
FIREPL_STA	0.0030	-0.0245	-0.0054	-0.0089	0.0001	0.0051	0.0145	0.0076	0.0200	-0.0030
	(0.14)	(-0.59)	(-0.17)	(-0.30)	(0.01)	(0.34)	(0.87)	(0.44)	(0.99)	(-0.15)
BASE_DUMMY	0.5171***	0.7062***	0.6121***	0.5959***	0.4704***	0.4514***	0.4383***	0.4378***	0.4525***	0.4619***
	(15.52)	(7.27)	(8.07)	(8.91)	(7.65)	(10.08)	(13.58)	(14.01)	(11.84)	(8.58)
GARAGE_CAP	-0.0038	0.0067	-0.0207	-0.0214*	-0.0220	-0.0239**	-0.0243***	-0.0286***	-0.0323***	-0.0264**
	(-0.46)	(0.22)	(-1.76)	(-2.01)	(-1.72)	(-3.06)	(-3.36)	(-3.82)	(-3.68)	(-2.85)
NEAR_DISTFIRE	-0.0280	-0.0376	-0.0452	-0.0343	-0.0326	-0.0238	-0.0255	-0.0065	-0.0157	-0.0079
	(-1.88)	(-0.78)	(-1.33)	(-1.45)	(-1.43)	(-0.97)	(-1.36)	(-0.27)	(-0.87)	(-0.58)
NEAR_DISTGOLF	-0.0829***	-0.1380***	-0.0917***	-0.0887***	-0.0881***	-0.0808***	-0.0665***	-0.0606***	-0.0631***	-0.0497**
	(-6.99)	(-4.39)	(-6.05)	(-4.90)	(-4.37)	(-4.58)	(-5.31)	(-4.73)	(-4.15)	(-2.64)
NEAR_DISTMEDICAL	-0.0556***	-0.0704***	-0.0550***	-0.0457***	-0.0485***	-0.0461***	-0.0446***	-0.0516***	-0.0675***	-0.0985***
	(-6.67)	(-3.76)	(-3.45)	(-3.70)	(-6.60)	(-3.35)	(-3.62)	(-6.43)	(-4.56)	(-9.26)
NEAR_DISTPOLICE	-0.0322***	-0.0117	-0.0530***	-0.0417***	-0.0419**	-0.0451***	-0.0400***	-0.0351**	-0.0311*	-0.0329*
	(-3.52)	(-0.68)	(-3.84)	(-4.89)	(-3.29)	(-5.67)	(-4.00)	(-2.86)	(-2.52)	(-2.20)
NEAR_DISTPOSTOFFICE	0.0288*	0.0128	0.0060	0.0037	0.0157	0.0130	0.0358*	0.0467**	0.0593***	0.0870***
	(2.16)	(0.41)	(0.27)	(0.15)	(0.82)	(0.87)	(2.10)	(2.84)	(4.03)	(4.59)
AD_WOODLAND	11.4697***	-5.9191	6.1859	9.5477**	12.1874***	14.7218***	16.9329***	19.2793***	19.6233***	29.2489***
	(3.66)	(-0.64)	(1.32)	(2.62)	(4.64)	(4.88)	(4.30)	(6.33)	(4.73)	(6.32)
NEAR_DIST_FORESTAM	0.1168***	0.1092	0.1220**	0.1370***	0.1545***	0.1388***	0.1398***	0.1358***	0.1489***	0.1579***
	(5.53)	(1.82)	(3.04)	(5.05)	(5.74)	(3.50)	(4.82)	(4.62)	(4.71)	(4.13)
PERCENT_WOOD	-0.2921	-0.1768	-0.5891*	-0.7206***	-0.5941***	-0.7268***	-0.7458***	-0.6498***	-0.4143	-0.2825
	(-1.88)	(-0.47)	(-2.44)	(-3.84)	(-3.71)	(-6.00)	(-6.11)	(-4.86)	(-1.78)	(-1.05)
AD_PARK	17.4727***	12.6674	17.7628**	19.4013***	20.4283***	19.6037***	21.2405***	17.5916***	19.0043***	20.0835***
	(4.21)	(1.34)	(3.18)	(3.84)	(4.42)	(6.97)	(4.38)	(5.69)	(4.54)	(4.17)
NEAR_DIST_RAILROAD	3.7256***	1.1250	1.5267*	1.9130*	2.0987**	3.4008***	3.9803***	3.7903***	5.0970***	7.2866***
	(8.55)	(0.96)	(2.10)	(2.41)	(3.19)	(4.15)	(5.26)	(5.35)	(3.93)	(8.87)
AD_PL	0.0281***	0.0281	0.0319***	0.0259***	0.0210**	0.0196**	0.0156*	0.0143*	0.0117	0.0058
	(3.94)	(1.53)	(4.00)	(4.27)	(3.17)	(3.10)	(2.12)	(2.33)	(1.42)	(0.59)
RTI	-174.6670***	-303.2014***	-231.8411***	-181.6696***	-168.7579***	-183.6917***	-160.1728***	-163.9568***	-132.7023***	-120.7114**
	(-5.94)	(-3.51)	(-6.13)	(-4.47)	(-4.46)	(-6.31)	(-5.17)	(-6.33)	(-5.46)	(-3.21)

AVERAGE_PRICE	0. 1544	-0. 1139	0. 3356	0. 6663***	0.4594***	0.3663**	0.2043	0.1756	0.0757	-0.0947
	(1.07)	(-0.28)	(1.39)	(4.16)	(3.44)	(2.77)	(1.31)	(1.05)	(0.54)	(-0.48)
CITY	0.0069	-0.0409	0.0089	0.0201	0.0753**	0.0662**	0.0672**	0.0757***	0.0682**	0.0105
	(0.27)	(-0.73)	(0.26)	(0.64)	(2.99)	(2.81)	(2.82)	(3.57)	(2.67)	(0.38)
VAC_RATIO	-0.0007	0.0034	0.0064	-0.0010	-0.0050	-0.0055	-0.0078*	-0.0068	-0.0101*	-0.0102
	(-0.19)	(0.35)	(1.21)	(-0.18)	(-1.36)	(-1.64)	(-2.30)	(-1.30)	(-2.22)	(-1.66)
HH_SIZE	0.0563*	0.0132	0.0634	0.1155**	0.0940***	0.0457*	0.0150	0.0237	0.0307	0.0308
	(1.98)	(0.18)	(1.18)	(3.18)	(3.69)	(2.06)	(0.52)	(0.61)	(0.82)	(0.82)
WHITE	1.1774*	-0.4319	-0.1642	0.0745	0.3709	1.2468	1.6500	2.2851	2.3486*	4.1584***
	(2.20)	(-0.26)	(-0.16)	(0.07)	(0.56)	(1.45)	(1.50)	(1.75)	(2.13)	(3.77)
BLACK	1.6696*	1.2635	0.0600	0.9215	1.9037*	2.2826*	2.3072*	2.6282	3.2685***	5.3610***
	(2.38)	(0.70)	(0.05)	(0.81)	(2.25)	(2.02)	(2.05)	(1.83)	(3.30)	(4.36)
ASIAN	-0.8054	-4.4312*	-4.3121***	-3.1688	-2.0234	0.1526	0.7714	2.5842	2.3389	3.1625*
	(-0.95)	(-2.09)	(-3.57)	(-1.96)	(-1.79)	(0.13)	(0.54)	(1.64)	(1.43)	(2.34)
TMR	2.3069	-3.8001	-1.9539	-1.4116	0.0282	1.5523	3.2742	5.6319*	6.7465**	9.3762***
	(1.58)	(-0.78)	(-0.83)	(-0.59)	(0.02)	(1.00)	(1.86)	(2.45)	(3.01)	(3.65)
TTW	-3.8363**	-1.2065	-5.3041*	-4.4589**	-4.1078*	-3.9897**	-3.6576**	-3.3584	-2.9623	-3.9644**
	(-2.68)	(-0.46)	(-2.49)	(-2.70)	(-2.54)	(-3.15)	(-2.86)	(-1.57)	(-1.79)	(-2.74)
WOR	0.0249	-0.0094	-0.1138	-0.0988	-0.0435	0.0143	0.0575	0.1613	0.2486**	0.2428**
	(0.37)	(-0.05)	(-1.10)	(-1.16)	(-0.51)	(0.23)	(0.79)	(1.86)	(3.22)	(2.68)
PNS	-0.9380*	-0.9048	-1.3205	-0.7818	-1.0947**	-1.0236**	-1.3556***	-1.1739	-1.4246*	-1.6258*
	(-2.15)	(-0.91)	(-1.51)	(-1.50)	(-2.99)	(-2.79)	(-3.52)	(-1.68)	(-2.03)	(-2.28)
PAC	9.8548***	9.7070**	11.4980***	9.8800***	9.7640***	7.9150***	8.6074***	7.8026***	9.4445***	12.0814***
	(6.50)	(2.76)	(5.05)	(4.77)	(6.91)	(4.20)	(5.24)	(3.98)	(4.96)	(5.64)
PMI	-0.5492	1.3110	0.9328*	0.3292	-0.2849	-0.4159	-0.9077***	-1.4297***	-2.0612***	-2.4809***
	(-1.89)	(1.35)	(2.05)	(0.82)	(-1.02)	(-1.66)	(-4.04)	(-3.68)	(-4.64)	(-6.01)
PHI	1.7856***	2.4051***	2.5961***	2.2424***	1.9245***	1.7851***	1.5712***	1.3452***	1.2016***	1.1771***
	(11.76)	(7.65)	(12.56)	(13.81)	(13.95)	(12.55)	(13.28)	(9.02)	(7.81)	(7.97)
PGH	-0.9959***	-0.6514	-0.8488**	-0.7126*	-0.9762***	-0.8980***	-0.9629***	-1.1662***	-1.6081***	-2.0452***
	(-5.65)	(-1.22)	(-2.59)	(-2.36)	(-5.84)	(-3.51)	(-4.28)	(-5.60)	(-6.92)	(-9.30)
PEH	-1.3628***	-1.0771	-1.1462	-0.8505	-1.0648**	-1.0191**	-1.3545***	-1.8185***	-2.3650***	-3.0740***
	(-3.51)	(-1.08)	(-1.73)	(-1.68)	(-2.64)	(-2.98)	(-3.70)	(-3.71)	(-5.45)	(-5.52)
X_SPHAT		-0.8188	-0.4384	-0.2073	0.0843	0.0966	0.0393	0.0995	0.3140	0.1911
		(-1.45)	(-1.35)	(-1.09)	(0.45)	(0.67)	(0.27)	(0.41)	(1.17)	(0.69)
_CONS	11.1476***	11.5696***	11.8008***	11.2539***	11.8324***	11.2300***	11.3860***	11.1071***	11.7917***	10.6968***
	(30.75)	(14.43)	(15.56)	(11.40)	(15.33)	(12.04)	(10.16)	(11.75)	(13.03)	(9.05)
lambda										
_cons	0.0128									
	(1.65)									
rho										
_cons	0.9749***									
	(6.30)									

t statistics in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001
GS2SLS is spatial two-stage least-square model
In the parentheses under 2SQR estimates are bootstrapping standard errors with setting seed 1001 in Stata
Lambda is the coefficient of the spatial lag term and rho is for error term in GS2SLS

Table 5. ATDI of the GS2SLS Model

Variable	Value	Variable	Value	Variable	Value
BDROOMS	0.0008	NEAR_DISTPOLICE	-0.0693	BLACK	-145.9284
FAMROOMS	0.0034	NEAR_DISTPOSTOFFICE	-0.0465	ASIAN	1.2898
DINROOMS	0.0021	AD_WOODLAND	-0.0269	TMR	0.0057
FULBATHS	-0.0051	AD_PARK	0.0240	TTW	-0.0006
FIREPL_STA	0.0226	NEAR_DIST_RAILROAD	9.5826	WOR	0.0471
BASE_DUMMY	-0.0001	AD_PL	0.0976	PNS	0.9836

GARAGE_CAP	0.0025	RTI	-0.2440	PAC	1.3948
NEAR_DISTFIRE	0.4320	VAC_RATIO	14.5979	PMI	-0.6729
NEAR_DISTGOLF	-0.0031	HH_SIZE	3.1127	PHI	1.9274
NEAR_DISTMEDICAL	-0.0234	WHITE	0.0235	PGH	-3.2050
NEAR_DIST_FORESTAM	-0.7836	AGE	-0.8320	PEH	0.0207
PERCENT_WOOD	8.2333	REMOD	-1.1386		
CITY	1.4918	AVERAGE_PRICE	-0.4589		

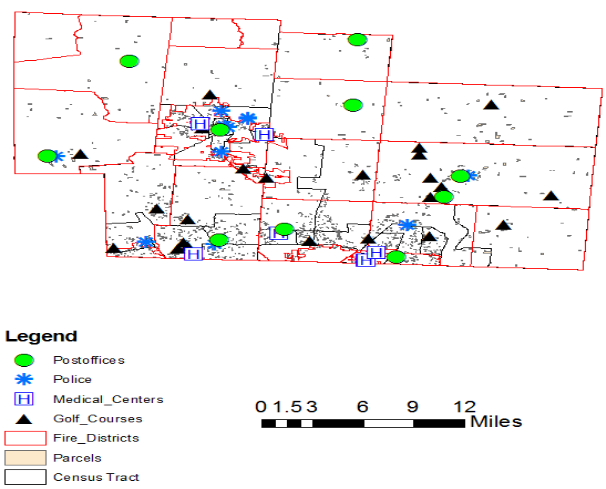

Figure 1. Places of Interest and Parcels

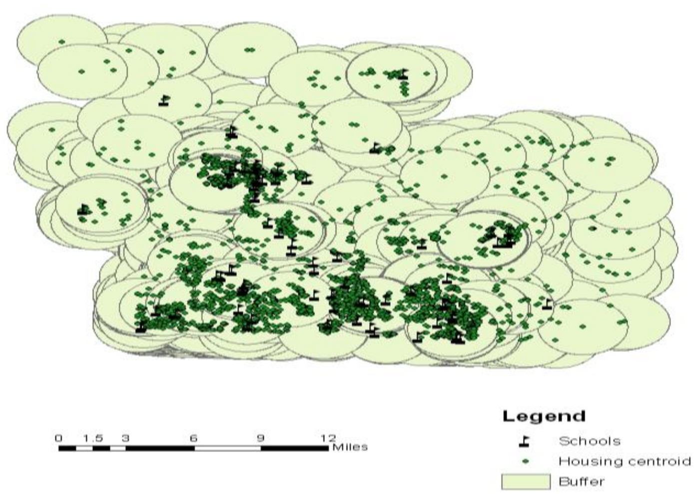

Figure 2. Parcel Centroid and Schools and within Buffer 2 miles

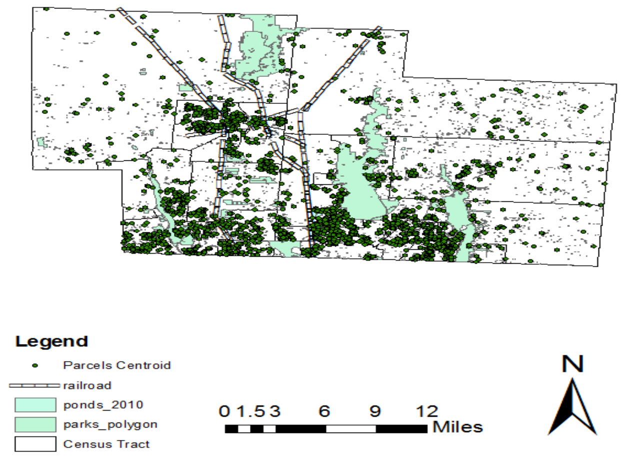

Figure 3. Railroad, Ponds and Parks

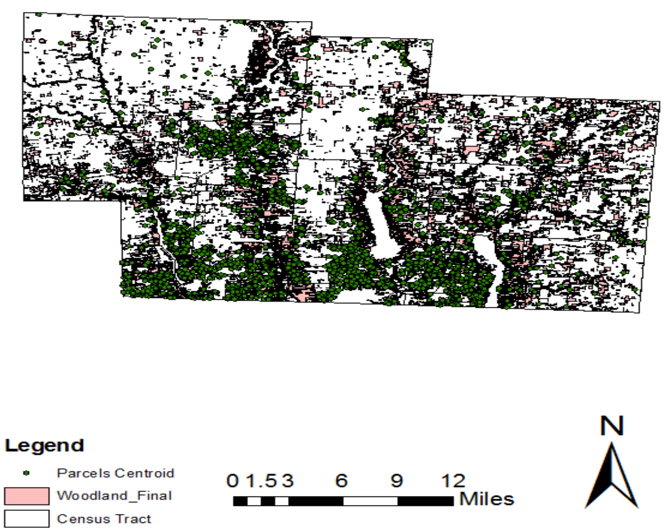

Figure 4. Woodland and Parcels

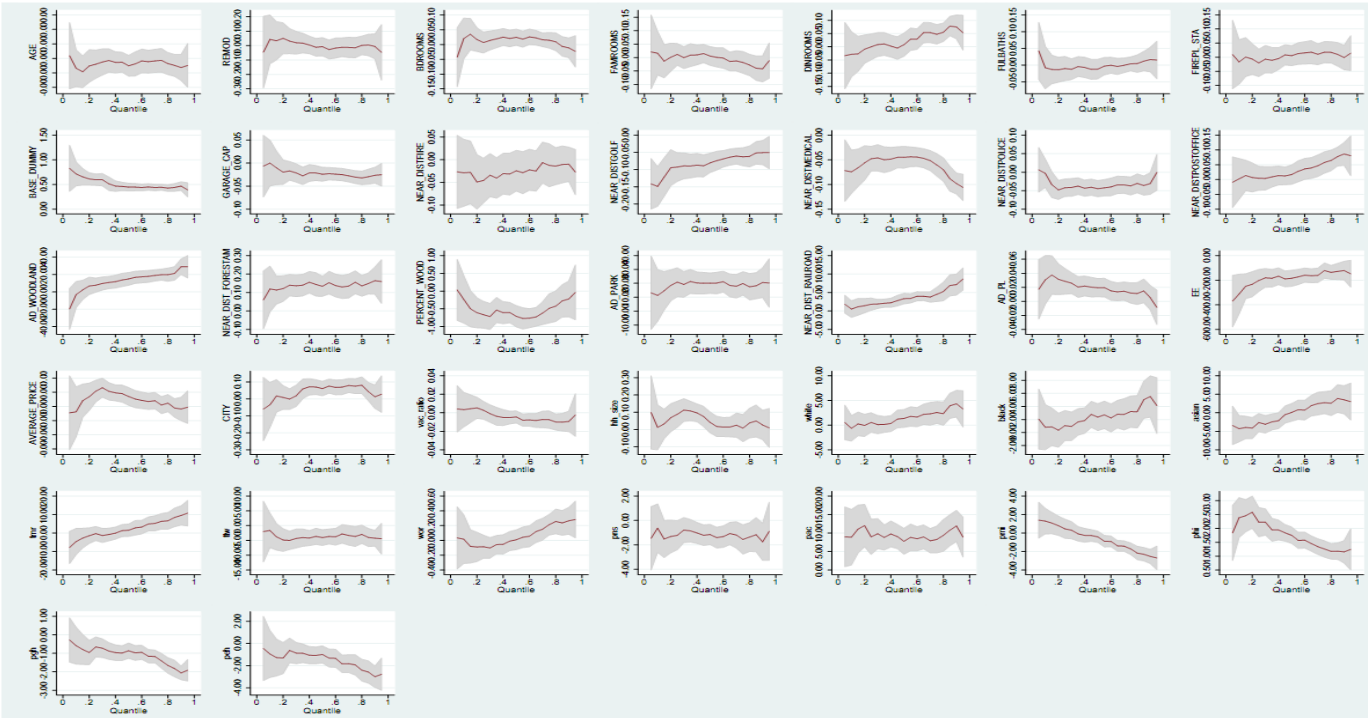

Figure 5. Plot of Coefficients in Different Quantile in QR

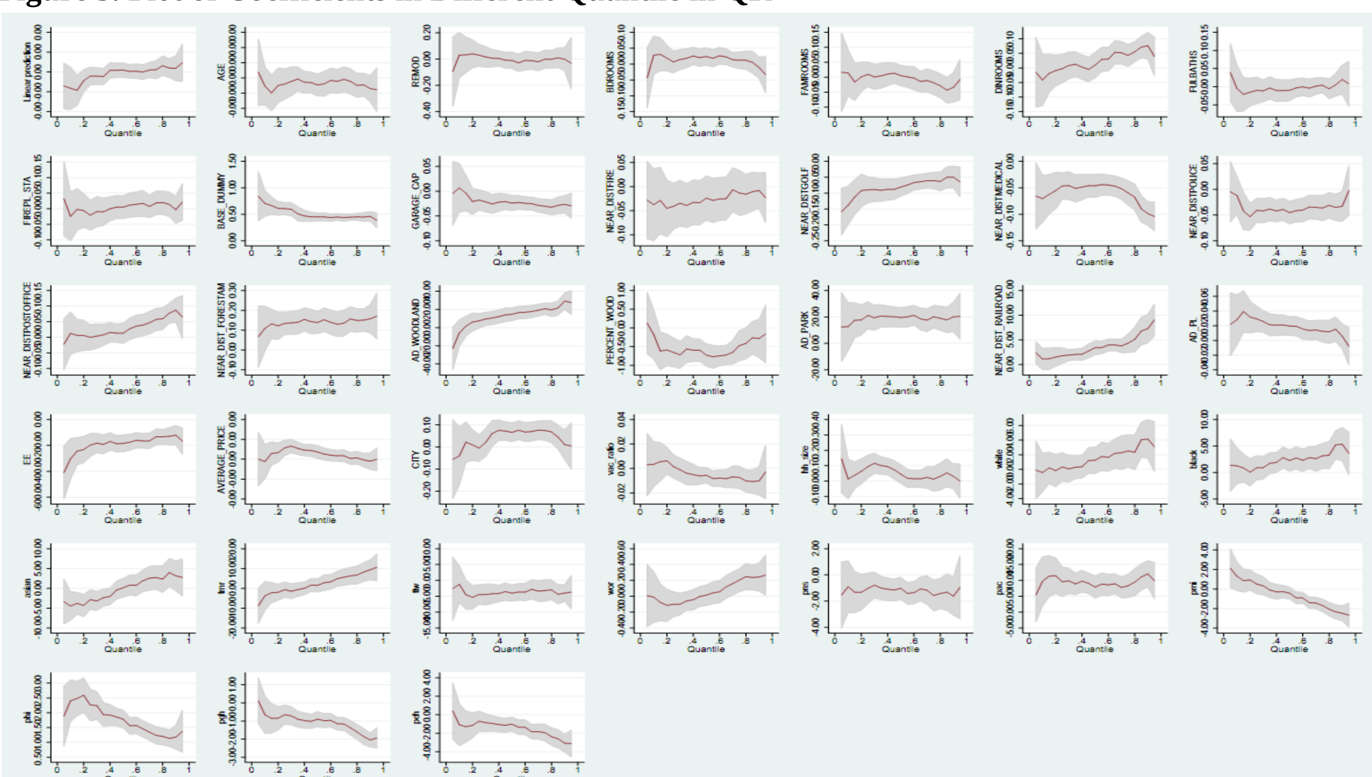

Figure 6. Plot of Coefficients in Different Quantile in 2SQR