8-2005

Is All Open Space Created Equal? A Hedonic Application within a Data-Rich GIS Environment

Bradley C, Neumann

IS ALL OPEN SPACE CREATED EQUAL?
A HEDONIC APPLICATION WITHIN A
DATA-RICH GIS ENVIRONMENT

By

Bradley C. Neumann

B.S. Northern Michigan University, 2003

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Resource Economics and Policy)

The Graduate School

The University of Maine

August, 2005

Advisory Committee:

Kevin J. Boyle, Professor and Chair of Resource Economics and Policy, Advisor

Kathleen P. Bell, Assistant Professor of Resource Economics and Policy

William A. Halteman, Associate Professor of Mathematics and Statistics

IS ALL OPEN SPACE CREATED EQUAL?

A HEDONIC APPLICATION WITHIN A

DATA-RICH GIS ENVIRONMENT

By Bradley C. Neumann

Thesis Advisor: Dr. Kevin J. Boyle

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Resource Economics and Policy)
August, 2005

Economic evidence reveals that preserved open space fosters services that are valued by members of society. However, when making the municipal decision to preserve land, communities must decide what type of open space to preserve, and must also deal with entities purchasing land and affecting tax revenues. The United States Fish and Wildlife Service (USFWS) has made efforts in recent years to expand the National Wildlife Refuge (NWR) system. This research seeks to determine if residential property owners value NWRs, and if they value NWRs differently than other types of open space, including conservation land, agricultural land, sports parks, golf courses, and cemeteries.

The hedonic method is used to estimate the benefits of each open space type that accrue to surrounding residential property owners. The hedonic models used here explain the sale price of a residential property as a function of numerous land, structure, and neighborhood characteristics, in addition to open space characteristics. The open space characteristics included in this research include measures of continuous distance from each property to the nearest open space of each type, discrete measures of distance

to the closest open space of each type, continuous measures of distance to the closest public and private open space, and an index describing the diversity of open space types evaluated at 100 and 1,000 meters around a home. As such, the hedonic method is utilized to estimate implicit prices associated with each of these open space characteristics.

The study area for this research is centered on a National Wildlife Refuge in central Middlesex County, Massachusetts called Great Meadows. The area is located approximately 20 miles northwest of Boston and is convenient for investigating the price effects of NWRs because of the abundance of residential properties adjacent to the refuge. The property sales data used in this study consists of residential transactions occurring between January, 1993 and December, 1998. Open space GIS data was obtained from the Massachusetts Office of Geographic and Environmental Information.

Results suggest that National Wildlife Refuges are valued by residential property owners. Specifically, a property located 100 meters closer to the Great Meadows NWR than a neighboring property has a price premium of $791. Further, Great Meadows is valued more highly than agricultural land, cemeteries, and conservation land but not valued significantly different than sports fields and golf courses.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Kevin Boyle, for all the guidance he has provided me throughout the course of this research and my graduate education. He has not only been an excellent advisor, but also an inspirational teacher and a great friend. Also, Dr. Kathleen Bell provided invaluable insight with respect to certain GIS operations used in this research and helped me to understand the wonderful world of spatial econometrics. Thanks also go out to Bill Halteman who was instrumental in the development of my "statistical sophistication." Also, Dr. Karin Steffens, Professor of Economics at Northern Michigan University, deserves a very grateful thank you for her enlightening lectures and talks outside of class that inspired me to pursue a degree in environmental and natural resource economics.

I would also like to thank Dr. John Vetelino, Principle Investigator of the *GK-12 Sensors!* Program at The University of Maine, for the incredible opportunities I have had as a National Science Foundation GK-12 Fellow. James Smith, at Bangor High School, deserves a very special thank you for being a mentor throughout my fellowship and a true friend. Most of all I would like to thank Melissa for her unfaltering patience, understanding, and love during the ups and downs of my graduate education.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.. iii

LIST OF TABLES.. vii

LIST OF FIGURES.. viii

Chapter

1. INTRODUCTION.. 1

Measuring the Benefits of Open Space................................ 1

The Problem.. 5

Thesis Objectives.. 6

Thesis Organization.. 9

2. THE HEDONIC METHOD.. 10

Historical Applications.. 10

Theory of Hedonic Models.. 11

Controlling for Property Attributes through Variable Selection...... 15

Functional Form.. 21

3. HEDONIC OPEN SPACE STUDIES.. 24

Capturing Open Space Effects with Measures of Proximity.......... 25

Disaggregating Measures by Open Space Type................ 26

Discrete Measures of Proximity................................ 27

Open Space Type - Public vs. Private................................ 30

Capturing Open Space Effects with Measures of

Landscape Pattern.. 31

Application to the Current Investigation................................ 32

4. DATA 34

Time Frame 36

Data Collection 38

Sales, Land, and Structural Variables 39

Neighborhood Variables 41

Environmental Variables 41

Creation of Open Space Type Variables 43

Creation of Open Space Diversity Variables 46

5. HEDONIC MODELS 48

Model Specification 48

Open Space Type Variables - Continuous Distance 51

Open Space Type Variables - Discrete Distance 52

Public/Private Open Space - Continuous Distance 55

Open Space Diversity Indices 55

Hedonic Property Value Models and Spatial Dependence 57

Accounting for Spatial Dependence 59

The Spatial Weight Matrix 61

6. RESULTS 63

Correcting for Heteroscedasticity 63

Identifying the Appropriate Structure of Spatial Dependence 66

Results of the Continuous Distance Model 70

Results of the Discrete Distance Model 76

Results of the Public/Private Continuous Distance Model 81

Open Space Categorization Results..................................... 85

7. CONCLUSIONS... 92

Limitations.. 93

Future Research.. 95

REFERENCES.. 97

APPENDIX: OLS and WLS Results.. 103

BIOGRAPHY OF THE AUTHOR.. 107

LIST OF TABLES

Table 2.1 Hedonic Property Value Studies of Open Space and Associated Structural Variables.................................. 18

Table 3.1 Hedonic Property Value Studies with Discrete Measures of Distance.. 29

Table 4.1 Adjustment of Housing Prices.. 38

Table 4.2 Descriptive Statistics for Home Characteristics........................ 41

Table 5.1 Explanatory Variables.. 49

Table 6.1 Descriptive Statistics of Hedonic Variables............................ 64

Table 6.2 Continuous Distance Model.. 72

Table 6.3 Discrete Distance Model.. 78

Table 6.4 Frequency of Sales per Open Space Zone.................................. 79

Table 6.5 Public/Private Continuous Distance Model................................ 83

Table 6.6 Distribution of Properties by Type of Nearest Public/Private Open Space... 84

Table 6.7 Comparison of Categorization Methods.................................... 86

Table 6.8 Results of the Two Open Space Categorization Methods.................... 89

Table A.1 Continuous Distance Model - OLS & WLS Results........................... 104

Table A.2 Discrete Distance Model - OLS & WLS Results............................. 105

Table A.3 Public/Private Continuous Distance Model - OLS & WLS Results... 106

LIST OF FIGURES

Figure 2.1 The Hedonic Price Function and Equilibrium........................... 14

Figure 4.1 Relative Location of Study Area.. 35

Figure 4.2 Description of Land Types in the 'Protected and Recreational Open Space' Data.. 42

Figure 6.1 Decision Criteria for Modeling Spatial Dependence.................. 67

Figure 6.2 Example of Categorization Error.. 87

Chapter 1

INTRODUCTION

In *Garden Cities of Tomorrow*, Ebenezer Howard presents the socio-economic importance of limiting the spread of urban land uses while complementarily preserving open space in and around the city-proper. Since the promotion of such ideologies in the early 20th century, planners, landscape architects, and even local government officials have strived to create the garden cities of tomorrow. However, pressures to develop exurban agricultural, grass, and forest lands, in addition to urban open space, have often been omnipotent, as these lands have been viewed as least cost locations for both public and private projects. Indeed, many of the natural jewels of the great city park era have been eroded by development pressure (Hecksher, 1977). However, over the past few decades economists have revealed significant amenity values associated with urban parks and other open space, signaling that such nonmarket services of the natural world do in fact have economic value and are beneficial to society (e.g. Knetsch, 1962; Correll et al., 1978; Beasley et al., 1986; More et al., 1988; Garrod and Willis, 1992; Bockstael, 1996; Breffle et al., 1998; Mahan et al., 2000).

Measuring the Benefits of Open Space

The majority of benefits that result from open space preservation are of the publicly provided and public good nature, such as the provision of recreational opportunities, aesthetics, and numerous ecosystem services including flood control, water purification, and habitat protection. These services sustain regional ecological processes, contribute to human psychological wellbeing, and can foster both direct and indirect benefits for surrounding property owners (Correll et al., 1978). Direct benefits result when individuals experience the positive services of open space as a result of their physical location. For instance, the associated view or ability to recreate within a particular open space often produces amenity value that is capitalized into neighboring property values. Indirect benefits of open space are somewhat less obvious; they often result from the bio-physical relationships between open space and locational attributes of a property. The association between wetlands and water quality, for instance, will produce indirect benefits for property owners located in close proximity to a body of water because of the wetland's ability to filter contaminants.

The services provided by preserved open space are not formally traded in markets. As such, we do not observe prices or values for these services or the open space from which the services are derived. Economists have developed a number of valuation techniques for quantifying such nonmarket values. The hedonic method is a valuation technique that is used to determine the implicit price of individual housing characteristics that collectively comprise a property's sale price. This thesis utilizes the hedonic method in order to derive values for a variety of open space types. For this valuation study, the housing characteristics of interest are the home's associations with surrounding environmental amenities, namely open space. In this case the hedonic property value model produces estimated implicit prices that are associated with marginal changes in the environmental amenities. Such estimates provide a basis to analyze the costs and benefits associated with a particular open space preservation project, and ultimately, allow policy makers to arrive at more informed decisions.

Historically, hedonic property value studies have utilized aggregate and rather general measures of open space for implicit price estimation. Specifically, a number of studies have used the hedonic method to estimate values for the catch-all 'open space' or 'park' (Knetsch, 1962; Correll et al., 1978; More et al., 1988; Cheshire and Sheppard, 1995). However, open space is heterogeneous and composed of a variety of different land uses and land covers that may each impose a different effect on surrounding property values. For instance, it has been revealed that agricultural land uses have negative spillover effects on surrounding residential properties (Johnston et al., 2001), while natural parks have positive effects (Lutzenhiser and Netusil, 2001; Netusil, 2005). Therefore, the inclusion of an aggregate environmental variable in a hedonic model could lead to error in the estimation of amenity values. Recent advances in Geographic Information Systems (GIS) have made the incorporation of more disaggregated open space measures within hedonic models possible. The ability to examine an area's land uses and land covers through digital aerial photography and satellite imagery in a GIS facilitates the interpretation and delineation of open space for use in hedonic models.

Regardless of the level of aggregation of the open space variables, the bulk of hedonic open space studies have attempted to capture the land capitalized value of environmental amenities or disamenities through measures of distance or access. Of course, such research focuses have been driven by the importance of location (consider the proverbial response that realtors provide when asked about the three most important factors contributing to the value of a property: 'location, location, location'). However, solely relying on measures of location restricts the spatial analysis to a single dimension. The recent advances in GIS have also set the stage for more detailed examinations of the spatial relationships among land uses. For instance, Geoghegan et al. (1997) utilized GIS to examine whether spatial patterns of land use contribute to property values by incorporating measures of land use diversity and fragmentation in the hedonic model. The inclusion of such variables adds a dimension to modeling richer spatial relationships that might exist within housing markets.

Whether or not richer measures of spatial location are indeed important in explaining property values remains an empirical question. The value of a residential property's location relative to environmental amenities may be better explained within a hedonic model under simpler spatial relationships, in which case data richness offers little benefit. If the latter situation is true, at the very least, the incorporation of GIS and associated spatial data within a hedonic property value study will only contribute to the accuracy and ease of generating environmental variables.

In a similar vein, the power that GIS offers for interpreting and disaggregating open space into its component types for use in a hedonic analysis is of little use if property owners are indifferent between certain types of open space. It may in fact be the case that open space is valued simply because it is – open space. If so, property owners would likely perceive open space not for what it is, but instead, for what it is not. The associated value of the environmental amenity would then likely be derived by its existence as undeveloped land, in which case the incorporation of alternative open space types does little to increase the accuracy of implicit price estimates. If such a situation does prevail, property owners would essentially perceive certain open space types as substitutes, willing to trade one for the other when interacting in the housing market as long as some amount of land in the surrounding neighborhood remains undeveloped.

As the previous discussion indicates, the price effects of preserved open space that accrue to residential property owners can be estimated with the hedonic method. Additionally, techniques and technologies for modeling the price effects of open space have improved since the first open space application of the hedonic method. However, irrespective of the method of valuation or empirical advances, the commitment of open space for preservation remains a contentious issue in public policy.

The Problem

Committing land as preserved open space has two important issues with which citizens and local governments must cope. First, purchasing land for the purpose of preservation is costly and communities must decide what type of open space to preserve and how much of it to preserve. For instance, development rights may simply be purchased leaving land in a more natural state, or land may be purchased outright and maintained as a particular type of open space such as an urban park or complex of sports fields. However, it has been revealed that not all open space has the same effect on the surrounding community – the decision to commit land as a particular type of open space could have either positive or negative spillover effects on surrounding properties.

Second, communities must deal with internal and external groups purchasing land for preservation and, as a result, affecting tax revenues. Preservation can result in reduced productivity of farm or forestland and reduced development potential in the community by restricting certain uses of land and/or preventing development altogether. The preservation efforts by entities such as The Nature Conservancy, the Trust for Public Land, or regional land conservancies and watershed councils provide example.

Another entity with goals of open space preservation is the United States Fish and Wildlife Service (USFWS), specifically the National Wildlife Refuge (NWR) system. There have been efforts in recent years to expand the land holdings of the system. However, it has been argued by community officials that NWRs reduce tax revenue because of the commitment of land to the federal government. As previously mentioned, economic evidence suggests there are positive price effects associated with naturally vegetated open space, in which the amenities of open space become capitalized into land values of neighboring properties. While NWRs make payments to communities in lieu of taxes, the effect that land preserved as NWR has on property values and in turn, the tax role, remains unexplored.

This thesis investigates if residential property owners value being proximate to NWRs, and if they value NWRs differently than other open space types. Implicit prices are estimated for a variety of open space types through application of the hedonic property value model to residential transactions surrounding the Great Meadows NWR. The results of this research provide new information for land use policy and aid communities in determining which types of open space have the greatest associated benefits. In turn, these results can be used to determine if the benefits of preserving certain open space types outweigh the costs.

Thesis Objectives

It is the goal of this thesis to apply the hedonic property value model within a data-rich GIS environment in order to estimate the implicit prices of a variety of different open space types. As previously described, this goal is multidimensional; the following objectives express the specific hypotheses to be examined by this work:

1) Investigate if residential property owners value National Wildlife Refuges using a continuous measure of distance from each property sale as the environmental variable.

2) Investigate if property owners value NWRs differently than five other types of open space using continuous measures of distance from each property sale.

3) Investigate if the hedonic price function exhibits discontinuities with respect to distance to each open space type using discrete measures of distance.

4) Investigate if publicly and privately accessible open space is valued by residential property owners using continuous measures of distance from each property sale.

5) Investigate if property owners value open space diversity.

6) Investigate if alternative methods of categorizing open space into type produce substantially different results.

First of all, it is an objective of this research to determine if residential property owners value being proximate to National Wildlife Refuges. The price effects that NWRs have on surrounding properties are not well known. Additionally, this research will determine if NWRs are valued differently than other types of open space, including conservation land, agricultural land, sports parks, golf courses, and cemeteries. This research will derive marginal values for NWRs and the five additional open space types by utilizing measures of continuous distance from each property to the closest open space of each type. Comparison of marginal values will indicate relative levels of amenity (disamenity) across open space types.

The use of continuous measures of distance for valuation eliminates the possibility for discontinuities in the hedonic price function with respect to distance to each open space type. It is possible for a particular open space type to be an amenity (disamenity) to surrounding properties at one distance while being a disamenity (amenity) to properties at some other distance. Therefore, it is also an objective of this research to estimate open space values based on the location of open space within various concentric rings or zones around each property using discrete variables.

This thesis will also investigate if residential property owners value open space accessibility. There are additional characteristics other than land use that may be important to property owners, one of which is whether open space is accessible to the public or not. This objective will be investigated by creating a set of continuous distance variables that measure the distance from each property to the closest publicly and privately accessible open space.

Additionally, recent advances in GIS have set the stage for more detailed examinations of the spatial relationships among land uses. Therefore, it is also an objective of this research to investigate if residents value open space diversity using an index calculated at 100 meters and 1,000 meters around a property.

Pursuit of these objectives for valuing open space provides an opportunity for also incorporating a methodological research objective. With the increasing use of GIS for scientific inquiry, the quality and completeness of secondary GIS data used for spatial analyses is often unknown or neglected. Therefore, it is also an objective of this research to investigate if analysis with secondary open space GIS data produces substantially different results than data that has been ground-truthed and referenced against other sources. Specifically, this objective seeks to determine if the results of an automatic method of categorizing open space are substantially different than the results of an interactive method of categorizing open space.

Thesis Organization

Chapter 2 continues with a description of the hedonic method and a detailed explanation of the conceptual framework used for valuation. Chapter 3 provides a review of the hedonic open space literature. Chapter 4 describes the study area and the types of data used in this analysis, including specific manipulations used in order to prepare the data for inquiry. Chapter 5 includes a description of the hedonic models used for valuation. Chapter 6 presents the results and discussions of the analysis, and Chapter 7 concludes with policy implications and recommendations for future research.

Chapter 2

THE HEDONIC METHOD

Residential properties are inherently heterogeneous goods. Housing is comprised of distinct characteristics that vary by quality and quantity to create differentiated product varieties within a single market. These housing characteristics are often viewed by hedonic theory as falling into attribute bundles, including land, structural, neighborhood, and environmental characteristics (Freeman, 2003). The presence of product variety in the market gives rise to price variation across the differentiated commodity. When a consumer chooses between alternative residential properties, it is revealed that the purchased property is overall comprised of more desirable characteristics, and therefore offers a greater level of utility for the consumer, *ceteris paribus*. Thus, if two otherwise identical residential properties differ only by a particular environmental characteristic, such as the presence of open space, the price differential between the two properties can be interpreted as the marginal implicit price of this environmental characteristic. The hedonic method is a valuation technique that relies on the observation of such market transactions to attach prices to the characteristics of a heterogeneous good, such as housing. Therefore, the market for housing can also function as a market for environmental quality.

Historical Applications

Hedonic property value studies have measured everything from the amenity value of a home's proximity to greenbelts (Correll et al., 1978), to the disamenity associated with a home's proximity to landfills (Nelson et al., 1992). However, historical use of the hedonic method for empirical research is more diverse. One of the earliest applications of the hedonic method was Waugh's (1928) analysis of vegetable prices. Based on specific perceivable quality characteristics, Waugh systematically graded vegetables in the Boston marketplace and used a hedonic model to estimate the premium that consumers were willing to pay for each characteristic of quality. However, it was Griliches (1961) who popularized the hedonic method and formed the basis for modern applications through his analysis of the automobile market. Using market auto prices, Griliches estimated the marginal implicit prices of the options and characteristics that together comprise the price of an automobile. The first application of the hedonic method to estimate the value of environmental quality on property values was conducted by Ridker and Henning (1967) who estimated the effects of air pollution on property values in St. Louis, Missouri. The study concluded that the marginal value of a change in the city's air quality could be used to estimate the benefits (costs) of improvements (degradations) in the city's air quality. The use of the hedonic method for nonmarket valuation was formalized by Rosen (1974) whose seminal article developed a framework of consumer utility theory explaining a hedonic equilibrium and its underlying market processes. More specifically, this work documented the linkages between consumer preferences for the characteristics of a heterogeneous good and the relevant equilibrium price function.

Theory of Hedonic Models

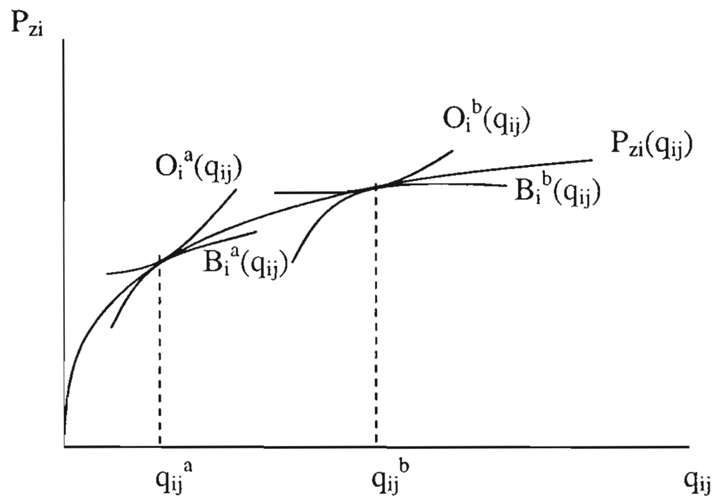
Application of the hedonic method for valuing the characteristics of a differentiated product relies on the establishment of a relationship between the overall price of a good and the quantity and quality of the good's characteristics. This relationship is referred to as the hedonic price function. It is a reduced form statistical model that represents the locus of equilibrium points resulting from the interaction of many consumers and producers in a perfectly competitive market.

Let Z represent the product class housing. In the market for housing, a particular home, z_i, can be represented by a vector of differentiated characteristics Q, such that $z_i = z_i(q_{i1}, q_{i2}, \ldots, q_{in})$. It follows that the price of the home, z_i, is a function of its characteristics, represented by the hedonic price function $P_{zi} = P_{zi}(q_{i1}, q_{i2}, \ldots, q_{in})$. Therefore, the price an individual consumer pays for a house is affected by the housing characteristics that they choose. In most instances consumers cannot pick and choose individual housing characteristics to repackage them as they please. Instead, consumers must settle for bundles of attributes that have already been assembled as a particular home. Thus, it has been suggested that the costs associated with reassembling or repackaging certain housing characteristics fosters a nonlinear hedonic price function (Rosen, 1974).

In establishing an equilibrium point on the hedonic price function, it is assumed that a consumer purchases two goods: a particular variety of housing, and the numeraire good, X, comprised of all other consumer goods. Therefore, a consumer seeks to maximize utility, defined as $U = U(X, q_{i1}, q_{i2}, \ldots, q_{in})$, subject to the budget constraint $M - P_{zi} - X = 0$, where M is income. The necessary first order condition for utility maximization requires that the marginal rate of substitution between any characteristic, q_{ij}, and the composite good must equal the ratio of the marginal prices. The consumer's desire to attain a particular residential property can be represented by a bid function, which is simply the inverse of the consumer's indirect utility function. The bid function establishes the relationship between the consumer's willingness to pay for z_i as one or more of its component characteristics change with a given level of the consumer's utility and income. Thus, the bid function, represented as $B_i = B_i(M - P_{zi}, q_{ij}, Q^*, U^*)$ where Q^* is a vector of the optimally chosen other characteristics, and U^* is the optimal level of consumer utility, indicates how the consumer's optimal bid must change as q_{ij} changes in order to maintain the optimal level of utility. This bid function is concave, exhibiting a diminishing marginal rate of substitution between q_{ij} and X (Rosen, 1974).

Equilibrium in the housing market also requires the presence of producers of housing. The goal of the producer is to determine the quality and quantity of the housing product to sell in order to maximize profits. The necessary first order condition for profit maximization requires that the producer supply a particular housing characteristic up to the point where the marginal revenue of that characteristic equals the marginal cost. In turn, the output level of the composite good is chosen such that the price of the residential property is just equal to the marginal cost of producing the housing unit. The inverse of the firm's profit function is called the offer function and is described by $O_i = O_i(q_{ij}, Q^*, \pi^*)$, where q_{ij} is the offered characteristic, and Q^* and π^* are the optimal other characteristics and profit, respectfully. Therefore, equilibrium in the housing market results from tangency between the producer's offer function and the consumer's bid function (Figure 2.1). Such points of equilibrium define the hedonic price function for a particular housing characteristic (Rosen, 1974). Along this hedonic price function the marginal utility of the characteristic to the consumer is equal to the marginal cost to the producer of providing that characteristic. Thus, the marginal implicit price of any characteristic, q_{ij}, can be estimated by calculating the first partial derivative of the hedonic price function evaluated at the desired quantity or quality of the characteristic, holding all other attributes constant. Mathematically, the marginal implicit price (MIP) of housing characteristic q_{ij}, is $MIPq_{ij} = \partial P_{zi}/\partial q_{ij}$.

Figure 2.1: The Hedonic Price Function and Equilibrium.



The marginal implicit price derived from the hedonic price function represents the cost of experiencing a marginal increase in a particular characteristic of housing. For example, the marginal implicit price of an environmental attribute, such as proximity to a natural park, represents the additional amount that must be paid to be located an additional unit closer to the park. In the market, such a marginal implicit price represents the equilibrium price and quantity combination on a particular individual's demand (willingness to pay) function.

This combination of price and quantity represents only a single point on the consumer's demand function, with the function itself remaining unidentified (Freeman, 2003). Therefore, estimation of willingness to pay measures for non-marginal changes of a particular characteristic is not possible with only marginal implicit prices derived from a first stage hedonic analysis. Instead, the demand function for the characteristic must be identified, requiring either multi-market estimation or the creation of restrictions on functional form within single market estimation (Freeman, 1974; Brown and Rosen, 1982; Palmquist, 1984). In this second stage of the hedonic method, socio-economic data about consumers must be combined with information on the quantities of characteristics purchased and the marginal implicit prices derived from the hedonic price function in order to identify inverse demand functions for the characteristics. As a result of such informational requirements, most applications of the hedonic method are solely concerned with the estimation of first stage marginal implicit prices of characteristics. This study is also solely focused on the first stage of the hedonic method.

Controlling for Property Attributes through Variable Selection

In a hedonic model, estimating the value of an environmental characteristic requires the major characteristics that determine the value of a property to be controlled. In addition to environmental variables, land, structural, and neighborhood variables must also be included in the model as determinants of the value of a residential property. Unfortunately, alternative specifications of these variables within the hedonic function can produce substantially different coefficient estimates for the variables. Economic theory does not suggest the variables to include in the hedonic equation and the researcher must consider the tradeoff between increased variance, resulting from irrelevant variable inclusion, and increased bias, resulting from relevant variable omission (Freeman, 2003). This tradeoff results from the econometric phenomenon known as multicollinearity, in which certain variables are correlated with other housing variables and/or environmental variables. Multicollinearity causes large standard errors if variables have a near exact linear relation; therefore, researchers may be inclined to omit certain collinear variables. However, if these variables are relevant for explaining some of the variation in the dependent variable, their removal can cause omitted variable bias. As a result, the hedonic function's specification is highly sensitive and the possibility exists for inducing error in the estimates of environmental variables (Graves et al., 1988). Furthermore, Michael et al. (2000) revealed that even the method of measurement of environmental variables included in a hedonic function can induce variation in coefficient estimates. Michael et al. (2000) suggest that the selection of environmental variables should be "...based on conceptually and theoretically sound logic and should reflect the public's perceptions of environmental quality" (p. 296).

Despite the inherent difficulty in the selection of variables, a well established group of structural characteristics used in hedonic models exists throughout the literature. This vector of structural characteristics often includes measures such as house age, interior square footage, number of bathrooms, and lot size (Lupi et al., 1991; Do and Grudnitski, 1995; Doss and Taff, 1996; Mahan et al., 2000; Bolitzer and Netusil, 2000; Lutzenhiser and Netusil, 2001; Irwin, 2002; Thorsnes, 2002). Additional structural characteristics have been included in hedonic models such as number of rooms, number of bedrooms, number of fireplaces, view quality, slope of property, elevation of property, month of sale, building materials, heating system, presence of a garage, presence of a pool, and presence of a basement. It is assumed that these attributes also affect property values, but are not always included in studies because of the presence of multicollinearity with other structural variables. The structural variables included in numerous hedonic open space studies are summarized in Table 2.1.

Additionally, neighborhood variables are important in hedonic models because of their role in determining the value of a residential property. Neighborhood variables comprise a vector of locational and socio-economic characteristics. These variables are often selected from a national census and utilized at the tract level. Typical neighborhood variables included in numerous hedonic models include median income, percentage nonwhite, percentage of residents older than 65 years of age, and percentage of residents over the age of 18 with some college education. Other neighborhood variables that are less frequently included in the hedonic model include traffic noise, location relative to CBD, tax rate, population density, and percentage of surrounding lands in commercial and industrial use. The inclusion of neighborhood variables depends on the characteristics of the study area and therefore, the signs of the coefficients are often unknown prior to estimation. Furthermore, the coefficients and signs associated with these variables are also very sensitive to the specification of the hedonic function.

Table 2.1: Hedonic Property Value Studies of Open Space and Associated Structural Variables.

Author	Weicher & Zerbst	Correll et al.	More et al.	Lupi et al.	Garrod & Willis	Do & Grundinski	Cheshire & Sheppard	Doss & Taff
Date	1973	1978	1988	1991	1992	1995	1995	1996
Location	Columbus, OH	Boulder, CO	Worcester, MA	Ramsey County, MN	UK	Rancho Bernardo, CA	Reading & Darlington, UK	Ramsey County, MN
Func. Form	linear	linear	linear	linear	Box-Cox	log-linear	Box-Cox	linear
Enviro. Variables	neighborhood parks	greenbelts	urban parks	wetlands	broadleaf & conifer forests	golf course	open land closed land	wetlands
Structural Variables	age house	age house	age house	age house	-	age house	-	age house
	-	dwelling size	-	dwelling size	dwelling size	dwelling size	dwelling size	dwelling size
	# rooms	# rooms	# rooms	# rooms	# rooms	-	-	-
	-	-	# bedrooms	# bedrooms	-	# bedrooms	# bedrooms	-
	-	-	# baths	# baths	# baths	# baths	# baths	# baths
	-	-	-	-	-	-	-	-
	-	-	# fireplaces	# fireplaces	-	# fireplaces	-	-
	-	-	-	# stories	-	-	# stories	-
	-	-	-	constr. type	-	-	-	-
	-	-	-	basement	-	-	-	-
	-	-	-	porch size	-	-	-	-
	-	-	heat source	-	heat source	-	heat source	-
	-	-	-	-	-	-	-	-
	-	-	-	central air	-	-	-	-
	-	-	-	pool	-	-	-	-
	-	-	garage	garage size	garage	-	garage	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	dwelling quality	dwelling quality	-	-	-	-
	-	-	-	-	dwelling type	-	dwelling type	-
	lot size	lot size	lot size	lot size	-	lot size	lot size	lot size
	-	-	-	lot topography	-	-	-	-
	time sale	-	time sale	-	time sale	time sale	-	-

Table 2.1 Continued.

Author	Geoghegan et al.	Tyrvainen	Powe et al.	Bolizer & Netusil	Mahan et al.	Tyrvainen & Miettinen	Luttik	Shultz & King
Date	1997	1997	1997	2000	2000	2000	2000	2001
Location	Patuxent Watershed, MD	Joensuu, Finland	Southhampton & New Forest, UK	Portland, OR	Portland, OR	Salo, Finland	Netherlands	Tucson, AZ
Func. Form	double-log	linear & semi-log	semi-log	linear & semi-log	semi-log	Box-Cox	linear	semi-log
Enviro. Variables	openspace land use - diversity & fragmentation	wooded park wooded rec area	woodland access	public/private parks cemeteries golf courses	wetlands	urban forests	parks	parks golf courses wildlife habitats
Structural Variables	age house	age house	age house	age house	age house	age house	-	-
	-	dwelling size	dwelling size	dwelling size	dwelling size	dwelling size	-	-
	-	# rooms	-	-	-	# rooms	# rooms	# rooms
	-	-	# bedrooms	-	-	-	-	-
	-	-	# baths	# baths	# baths	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	# fireplaces	# fireplaces	-	-	-
	-	-	-	-	-	-	-	-
	constr. type	constr. type	-	-	-	constr. type	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	heat source	-	heat source	-	-	-
	-	-	-	-	hardwood floor	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	pool	-	-	-
	-	-	garage	-	garage size	-	-	-
	-	flat roof	-	-	-	flat roof	-	-
	-	-	-	-	-	kitchenette	-	-
	-	-	-	-	-	sauna	-	-
	-	-	-	-	-	-	-	-
	-	-	dwelling type	-	-	-	dwelling type	-
	lot size	-	-	lot size	lot size	lot size	lot size	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-

Table 2.1 Continued.

Author	Acharya & Bennett	Espey & Owusu-Edusei	Lutzenhisher & Netusil	Irwin	Smith et al.	Geoghegan	Thorsnes	Geoghegan et al.
Date	2001	2001	2001	2002	2002	2002	2002	2003
Location	New Haven, CT	Greenville, SC	Portland, OR	Central MD	Wake County, NC	Howard County, MD	Grand Rapids, MI	Central MD
Func. Form	semi-log	log-linear	Box-Cox	semi-log	Box-Cox & semi-log	semi-log	linear & log-linear	double-log
Enviro. Variables	openspace land use - diversity & fragmentation	parks	parks natural areas golf courses cemeteries	cropland pastureland forest	parks agriculture forests golf course right-of-ways	permanent & developable open space	forests	permanent & developable open space
Structural Variables	-	-	age house	age house	age house	-	age house	age house
	dwelling size	dwelling size	dwelling size	dwelling size	-	dwelling size	dwelling size	dwelling size
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	# baths	# baths	# baths	# baths	-	# baths	-
	-	-	-	# half-baths	-	-	-	-
	-	-	# fireplaces	-	-	-	-	-
	-	-	-	-	-	# stories	-	# stories
	-	-	-	-	-	-	-	constr. type
	-	-	-	-	-	-	-	basement
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	central air	-	-	-	-	-	-
	-	garage	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	-	-	-	-	-	-	-
	-	dwelling quality	-	dwelling quality	-	dwelling quality	-	dwelling quality
	-	-	-	dwelling type	-	-	dwelling type	-
	lot size	lot size	lot size	lot size	lot size	lot size	lot size	lot size
	-	-	-	-	-	-	-	-
	-	-	-	time sale	-	-	time sale	-

Functional Form

It is not clear from economic theory the ideal functional form of the hedonic equation (Rosen, 1974). Cassel and Mendelson (1985) have suggested that best fit criteria be used among alternative hedonic functional forms. Rosen (1974) has provided an additional criterion for the functional form of the hedonic equation noting that housing attributes cannot typically be purchased independently and therefore, the hedonic price function should be of a different form than linear.[1] Among recent hedonic studies, the most common functional forms are the log-linear (exponential, in which only the dependent variable is logarithmic) (Espey and Owusu-Edusei, 2001; Acharya and Bennett, 2001; Thorsnes, 2002) and the semi-log (where some of explanatory variables are logarithmic) (e.g. Mahan et al., 2000; Shultz and King, 2001; Smith et al., 2002; Geoghegan et al., 2003). A fewer number of recent studies have used the double-log functional form (also known as log-log, in which all continuous and unbounded variables are logarithmic) (Geoghegan, 1997; Irwin, 2002) or the flexible Box-Cox form of the hedonic equation to allow the nature of the data to determine the exact functional form (Tyrvainen and Miettinen, 2000; Lutzenhiser and Netusil, 2001).

The applicability of Box-Cox functional forms to the diversity of explanatory variable specifications within hedonic studies has been debated. Cropper et al. (1988) have suggested that more complex functional forms, such as the linear Box-Cox, or simpler forms, such as linear, log-linear, semi-log, or double-log, be used when certain variables are missing or are instead replaced by proxies. Indeed, controlling for every

---

[1] Additionally, a linear functional form is indicative of constant implicit prices across the range of the variable. However, economic theory suggests that individuals are willing to pay more for a marginal increase in a particular good (environmental amenity) when endowed with little of the good and pay less for the good when endowed with more.

housing characteristic that comprises the sale price of a home is extremely difficult considering the cross-sectional nature of hedonic property value studies. Also, to avoid problems of multicollinearity, a specification of variables that consists of only the primary drivers of house prices is often used. However, criticism of Box-Cox functional forms has been voiced because of the difficulty in calculating implicit prices of the attributes of interest when the form is other than linear. In most cases then, calculation of the marginal implicit price of any particular variable not only depends on the focus variable's level, but also depends on the level of all other attributes, a complicating factor that may compromise the estimates' use for policy implementation (Cassel and Mendelsohn, 1985). One advantage of the Box-Cox form is the ability to test the estimated restrictions against the more typical forms, such as linear and log-linear, using the asymptotic likelihood ratio statistic (Halvorsen and Pollakowski, 1981). Although, Box-Cox functional forms necessarily have more coefficient estimates than other functional forms with the same number of variables, so the use of best fit criteria among alternative forms may be at the expense of parameter estimate accuracy and resulting implicit prices (Cassel and Mendelsohn, 1985).

While constraints of data availability are a limiting factor in any hedonic study, the models estimated in this thesis intentionally assume a parsimonious specification. Additionally, based on the difficulty controlling for every housing characteristic that comprises the sale price of a home and the likelihood that the hedonic price function contains proxies, it seems that either more complex functional forms, or simpler forms are the most appropriate (Cropper et al., 1988). Further, flexible Box-Cox forms, such as those employed by Tyrvainen and Miettinen (2000), and Lutzenhiser and Netusil (2001) offer the nicety that the nature of the data will determine the exact functional form. Thus, based on the current status of the literature, this research employs a flexible Box-Cox specification in order to shed light on the most appropriate functional form among simpler specifications. Box-Cox estimation suggests a log-linear specification of the hedonic price function ($\lambda = 0.25$).[2]

---

[2] $$P^{(\lambda)} = \left\{ \begin{array}{ll} \dfrac{P^{\lambda} - 1}{\lambda} & if\ \lambda \neq 0 \\ \ln P & if\ \lambda = 0 \end{array} \right\}$$

Chapter 3

HEDONIC OPEN SPACE STUDIES

Applications of the hedonic method to examine the relationship between open space and a home's sale price have become increasingly more numerous in recent years. One reason for such growth is the increased number of land use disputes as society has become increasingly more urban. With this growth in urban and suburban populations has come the expansion of cities and an increase in the urban-rural interface. Additionally, suburbanization has given rise to the decline of many city centers, raising questions about the relative amenity and disamenity levels of land uses. Therefore, land conversion at the exurbs, and transformations of cities at their centers have necessitated an understanding of the price effects inherent among differing land uses. The hedonic method has been used extensively for generating new knowledge about such land use relationships.

In addition to frequency, land use applications of the hedonic method have also increased in quality. The increased availability of data produced from the U.S. Census of Population and Housing and the growing number of multiple listing services have together made housing and population data more readily available (Shultz and King, 2001). Additionally, the increasing use of GIS and the ability to spatially reference home sales for automated calculations of open space proximity and quantity has contributed to the quality and frequency of hedonic property value studies of open space price effects.

Capturing Open Space Effects with Measures of Proximity

Even before the widespread availability of census data and GIS, the hedonic method had been utilized in studies to value open space. It was Knetsch (1962) who first suggested that a residential property located closer to open space could be expected to command a premium over a property located farther away, *ceteris paribus*. Much of the literature that immediately followed sought to establish this connection between a home's price and its relation to surrounding urban parks and/or greenbelts, for which the hedonic method was often the statistical vehicle of choice. Correll et al. (1978) examined the effect of a home's proximity to greenbelts (natural buffers of open space in and around a city) on property values in Boulder, Colorado. The environmental variable was measured as the distance in feet using the most direct public access to a greenbelt. The authors used a relatively small set of land, structural, and neighborhood variables in their study. The hedonic price function used in the study was linear in the sale price of a single-family residential property, and therefore, assumed a constant marginal implicit price for proximity to open space. The model revealed that the presence of greenbelts throughout the city added $4.20 to the value of a residence for every foot closer the home was located to a greenbelt (Correll et al., 1978).

More et al. (1988) examined the price effects of proximity to urban parks in Worcester, Massachusetts. Two measures of proximity to four different urban parks were used: Euclidean distance from each house to the park, and the distance via the network of roads from each house to the closest park entrance. The hedonic price function was of the semi-log form. Results indicate that on average, a house located 20 feet from a park sold for approximately $2,675 more than a house 2,000 feet from a park (More et al., 1988). The authors concluded that the positive amenity effect of living in the proximity of an urban park extended to properties as far as 2,000 feet away (More et al., 1988).

Disaggregating Measures by Open Space Type

Disaggregating the catch-all 'open space' into its specific land use types has been applied in surprisingly few hedonic property value studies. An exception is the work of Lutzenhiser and Netusil (2001) which disaggregated the general characteristic 'park' into the component parts of urban park, natural area park, and specialty park/facility. In addition, the authors included open space measures for cemeteries and golf courses, as well as typical land, structural, and neighborhood characteristics in the hedonic equation. The study, which focused on Portland, Oregon, utilized discrete measures to determine if distance to each of the open space types affects a home's sale price. Results revealed that the highest capitalized values of open space were present in homes located adjacent (within 200 feet) to golf courses. Location between 601 and 800 feet of natural area parks was found to have the second highest affect on residential property values. However, averaged across all of the discrete zones included in the model (see Table 3.1 on page 29) natural area parks were found to have the largest average positive effect on surrounding property values.

This research follows the approach taken by Lutzenhiser and Netusil (2001) and disaggregates open space based on land use type. Here, open space is disaggregated into six types including NWR, conservation land, agricultural land, sports park, golf course, and cemetery. It remains unexamined in the literature what effect NWRs have on neighboring properties, and how the effect (if any) compares to other types of open space.

Mahan et al. (2000) examined the effect of proximity to urban wetlands on surrounding property values in Portland, Oregon. Numerous wetland proxies were used including, distance to, and size of open water, emergent vegetation, scrub-shrub, and forested wetland. In order to control for other amenity generating features, the authors also included as environmental variables distance measures to parks, streams, rivers, lakes and commercial and industrial areas, in addition to measures of view quality, property slope, and elevation. Results reveal that a 1,000 foot reduction in the distance of a home from wetland increases property values by $436.17. Comparing this result with the marginal implicit price of proximity to urban parks, in which a 1,000 foot reduction in the distance to a park increases property values by $33.24, suggests that wetland is valued differently than other urban open space (Mahan et al., 2000).

The results of Mahan et al. (2000) provide support for the first two objectives of this thesis. Not only do the results suggest that National Wildlife Refuges, which are often dominated by wetland, are valued by residential property owners, they also suggest that NWRs are valued differently than other types of open space.

Discrete Measures of Proximity

While the previous studies have identified amenity effects associated with location relative to open space, the studies have incorporated only continuous measures of proximity to the environmental amenities. Valuing open space proximity with only a continuous measure of distance assumes that the hedonic price function with respect to this distance is constant over the entire geographic extent of the study area. However, it seems quite possible for a particular open space type to be an amenity (disamenity) to surrounding properties at one scale while being a disamenity (amenity) to properties at some other scale.[3] Consider agricultural land. To neighboring properties the open space of a farm has been revealed to have downward pressure on value because of nuisance sights, smells, and/or sounds (Johnston et al., 2001). However, at a certain scale, property owners may value farm land as contributing to a rural sense of place that has less traffic and congestion than urban locations. Therefore, in order to examine the hedonic price function for discontinuities and gain more understanding about the effects of open space on residential property values, a handful of researchers have incorporated in the hedonic equation discrete measures of distance to open space (Johnston, 1998; Bolitzer and Netusil, 2000; Tyrvainen and Miettinen, 2000; Espey and Owusu-Edusei, 2001; Lutzenhiser and Netusil, 2001). Using dummy variables, a property is assigned a value of 1 if the nearest open space is located within a particular distance or zone, and 0 otherwise. Review of the threshold distances or zones used in the studies in Table 3.1 suggests that little guidance exists in the literature regarding the specification of zones.

In addition to using continuous measures of distance, this research also includes discrete measures of distance for valuing open space. However, this research incorporates an alternative specification of discrete distance than the measures used in previous studies. Following Geoghegan et al. (1997), the zones are based on what can be seen from a property versus what would be encountered on a walk. The specification has conceptual grounding and avoids the bias associated with manipulating delineations in an iterative process based on the significance of parameters. The zones are described in more detail in Chapter 5.

[3] Of course, a quadratic specification of the hedonic price function with respect to distance is capable of capturing such changes in sign while remaining continuous. However, few researchers have taken this approach.

Table 3.1: Hedonic Property Value Studies with Discrete Measures of Distance.

Author	Date	Publication	Resource Application	Zones						
Johnston	1998	URI: Coastal Resources Center		≤ 400 m of a 10 acre tract						
			open space	**2.73%**						
				≤ 1,000 m of a 50 acre tract						
			open space	**9.07%**						
Bolitzer & Netusil	2000	J of Environmental Management		≤ 30 m*	31-121*	122-213*	214-304*	305-396*	397-457*	
			open space	5.34%	**4.17%**	**3.00%**	**2.31%**	**2.20%**	1.52%	
Tyrvainen & Miettinen	2000	J of Enviro. Econ. & Management		5-99 m	100-299	300-599	600-999			
			forest park	**7.57%**	**5.87%**	**4.71%**	4.19%			
Espey & Owusu-Edusei	2001	J of Agricultural & Applied Econ.		≤ 91 m*	91-152*	152-457*				
			small basic park	**-13.9%**	**15.0%**	**6.18%**				
				≤ 183 m*	183-457*					
			small attractive park	**11.6%**	-0.10					
				≤ 61 m*	61-457*					
			medium attractive park	3.05%	**6.18%**					
				≤ 183 m*	183-366*					
			medium basic park	**-51.3%**	-1.00%					
Lutzenhiser & Netusil	2001	Contemporary Economic Policy		≤ 60 m*	61-121*	122-182*	183-243*	244-304*	305-365*	366-457*
			urban park	2.96%	**3.11%**	1.80%	1.23%	1.42%	**2.55%**	0.52%
			natural area park	**16.9%**	**15.4%**	**19.1%**	**17.0%**	**13.6%**	**12.3%**	**15.1%**
			golf course	**21.0%**	**11.9%**	4.25%	**13.4%**	**13.4%**	6.63%	**6.60%**
			specialty park	**11.2%**	**8.68%**	**15.5%**	**8.55%**	**7.51%**	**6.89%**	**5.80%**
bold indicates significance at the 0.05 level							* indicates conversion from feet to meters (divide by 3.28)			

Open Space Type - Public vs. Private

Similar to the approach taken by Lutzenhiser and Netusil (2001), Geoghegan (2002) also utilized the inherent characteristics of open space to produce more specific estimates of how different types of open space affect residential property values. Geoghegan (2002) approached the situation by disaggregating the catch-all 'open space' into categories that represent the presence of or lack of development rights, by applying the hedonic method to estimate the marginal implicit prices of developable open space and permanent open space. Developable open space includes privately owned forested land and agricultural crop and pasture land, while permanent open space includes parks and land that has conservation easements or has had the development rights sold. The two variables were calculated using a 1,600 meter buffer or neighborhood around properties in Howard County, Maryland. Results suggest that permanent open space increases residential property values over three times as much as developable open space (Geoghegan, 2002).

Geoghegan's (2002) classification incorporates additional characteristics of open space, besides land use, into the hedonic price function. The classification, developable versus permanent, is essentially an underlying measure of ownership, either private or public, respectfully. However, there exists an alternative set of open space types based on the public/private dichotomy that are distinguished by accessibility. This research expands on the work of Geoghegan (2002) by incorporating variables for private and public open space that measure the continuous distance from properties to the nearest publicly and privately accessible open space.

Capturing Open Space Effects with Measures of Landscape Pattern

As previously mentioned, technical advances have paved the way for the incorporation of more complex measures of the effects of open space on property values. For instance, Geoghegan et al. (1997) incorporated in the hedonic price function measures of the diversity and fragmentation of land uses around residential properties in a Maryland watershed, in addition to measures of the percentage of open space in a home's neighborhood. Originally developed by landscape ecologists, the diversity index measured the heterogeneity of land uses by describing whether there were relatively few or many land use categories in a given neighborhood, whereas the fragmentation index used in the model was a ratio of perimeter to area that increases as land is subdivided. Both landscape indices and the open space index were calculated at two distinct scales (0.1 kilometer and 1.0 km) in order to capture any differential effects between what can be seen from a home and what would be encountered on a walk through the surrounding neighborhood, respectfully. Based on a double-log model estimated with Ordinary Least Squares (OLS), the coefficients for the diversity and fragmentation indices were not significantly different from zero. However, taking into account the possibility that the hedonic function exhibited spatial variation, Geoghegan et al. (1997) also estimated a spatial expansion model in which the parameters varied linearly and quadratically with distance from Washington, D.C. Results of this second specification suggest that increases in land use diversity and fragmentation are only valued in the immediate proximity of Washington, D.C. and at the outermost edge of the sample (Geoghegan et al., 1997).

In the only other hedonic property value study to estimate the value of land use patterns in tandem with the value of open space, Acharya and Bennett (2001) applied the hedonic method to residential property sales in an urban watershed of New Haven County, Connecticut. The environmental variables in the hedonic price function included a land use diversity index and the percentage of open space around each home. In a similar fashion to that of Geoghegan et al. (1997), the landscape variables were measured at both a 0.25 mile and 1.0 mi radius around each home. Results suggest that the diversity of land uses around a home has a negative effect on property values at both the 0.25 and 1.0 mi scale (Acharya and Bennett, 2001).

Similar to Geoghegan et al. (1997) and Acharya and Bennett (2001), this research also estimates the effects that diversity has on property values. However, this research seeks to determine if property owners value open space diversity, as opposed to land use diversity, a landscape characteristic that remains unexplored in the literature. Therefore, the indices used in this thesis are restricted to only include open space types, as opposed to land use types. The indices are fully described in Chapter 5.

Application to the Current Investigation

Based on previous hedonic open space studies and other hedonic literature, the hedonic price functions estimated in this thesis are log-linear and take the following general form:

$$\ln(P_i) = \beta_0 + \sum\beta_j L_{ji} + \sum\beta_k S_{ki} + \sum\beta_l N_{li} + \sum\beta_m E_{mi} + \varepsilon_i$$

where $\ln(P_i)$ is the natural logarithm of the sale price of property i, L_{ji} is a vector of land characteristics of property i, S_{ki} is a vector of structural characteristics of property i, N_{li} is a vector of neighborhood characteristics of property i, E_{mi} is a vector of environmental characteristics of property i, and ε_i is the observation specific error term. Specific descriptions of the explanatory variables used in this research and an overview of the study area are the subjects of the next chapter.

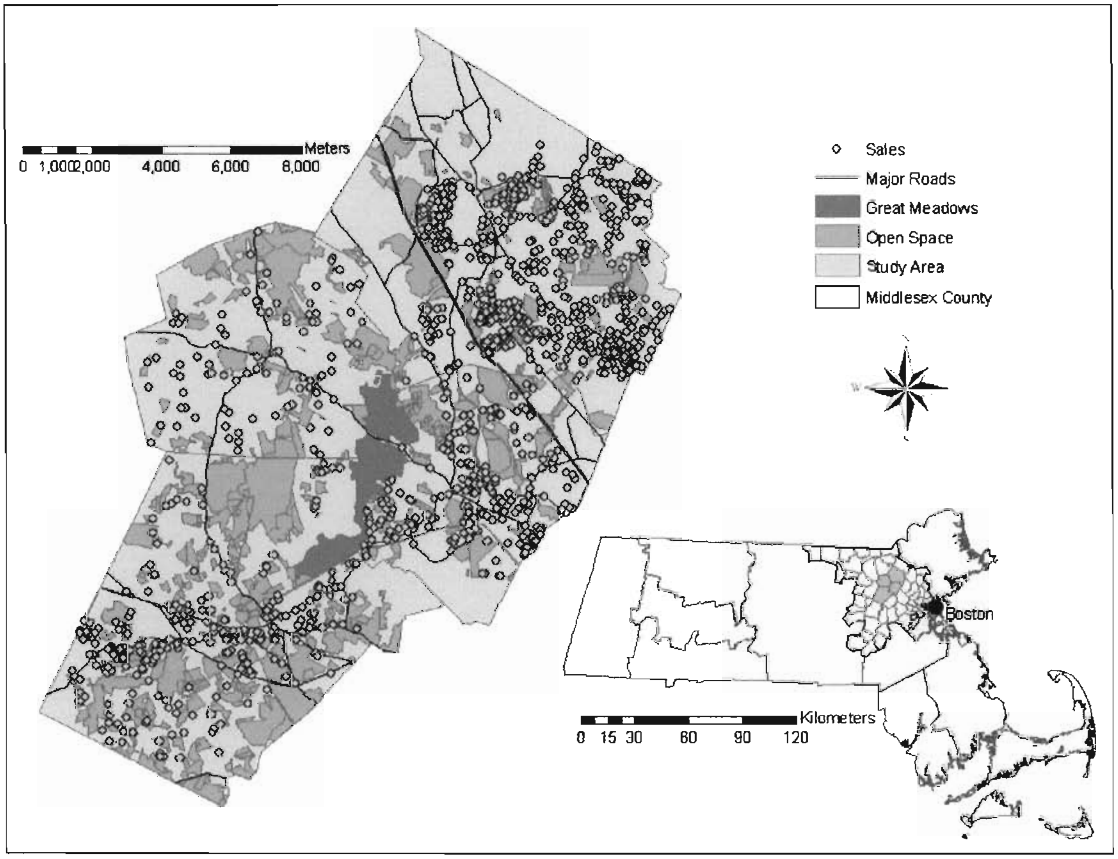
Chapter 4

DATA

The study area for this research is centered on a National Wildlife Refuge (NWR) in central Middlesex County, Massachusetts called Great Meadows (see Figure 4.1). This protected and federally owned, natural open space is located approximately 20 miles northwest of Boston and consists of approximately 3,626 acres managed for the protection of migratory waterfowl.[4] The area is convenient for investigating the price effects of NWRs because of the abundance of property sales adjacent to the refuge. Established in 1944, the refuge is located along the Atlantic Flyway and has had over 220 bird species recorded on site. Approximately 90 percent of Great Meadows is comprised of wetlands that serve as a transient home to waterfowl such as mallards, black ducks, wood ducks, and blue-winged teal. Other animals such as white-tailed deer, muskrats, red fox, raccoons, cottontail rabbits, weasels, squirrels, and various other small mammals, amphibians, and reptiles can also be found in the refuge. In addition to the protection of wildlife and wildlife habitat, Great Meadows also serves as a natural environment for wildlife viewing and recreation. The refuge has multiple hiking trails throughout the wetland and an observation tower for wildlife viewing. Based on the area's quality of wildlife habitat and accessible recreational amenities, ornithologists have called Great Meadows one of the best inland birding areas in Massachusetts (U.S. Fish and Wildlife Service, 2000). As a large tract of open space in a predominantly developed area, the refuge is also simply undeveloped land, reducing the density of development in the area.

[4] This statistic includes both the Concord and Sudbury divisions of the NWR. However, Figure 4.1 depicts only the Concord division of the NWR.

Figure 4.1: Relative Location of Study Area.



Great Meadows is surrounded by the four towns Billerica, Bedford, Concord, and Carlisle (clockwise from the northern most town Billerica). Much of the land closest to Great Meadows was developed after the creation of the refuge in 1944 when the area experienced its first major residential development pressure in the 1950s and 1960s. Today, the towns Billerica, Bedford, Concord, and Carlisle are largely developed and characterized by low-density residential development scattered along curvilinear roads. This residential form or morphology is, however, heavily vegetated with natural cover that contributes to a surprisingly rural-feeling sense of place. Such character is in sharp contrast to the adjacent Route 128 corridor to the east that has been home to the region's most rapid development in recent years (Gittell and Flynn, 1995). The 128 corridor has become a significant source of employment for high-tech, manufacturing, and commercial jobs in the suburban Boston region.

In determining the appropriate geographic extent of the study area from which to collect property sales data, it is important to consider the spatial extent of the price effect that Great Meadows might have on nearby properties. It is apparent that at least some properties in the adjacent towns Billerica, Bedford, Concord, and Carlisle will experience the price effect, if one in fact exists, and should therefore be included in the analysis. It is less certain if this effect extends to properties in outlying towns. However, as the geographic extent of the study area grows, one must be increasingly concerned with whether or not the study area constitutes a single housing market. The hedonic price function is assumed to be an equilibrium function describing a specific market; as such, the definition of the study area is important for estimating implicit prices. If the geographic extent of properties included in a sample reaches into more than one housing market, estimating a single hedonic price function will no longer be appropriate. For this reason, the towns Billerica, Bedford, Concord, and Carlisle are assumed to constitute the appropriate geographic extent for examining the price effect of the NWR. The property sales data based on this geographic extent that are used in this thesis have a relatively uniform spatial distribution, with a slight concentration of sales in closest proximity to Boston (see Figure 4.1).

Time Frame

The property sales used in this study occurred between January of 1993 and December of 1998. The selection of this duration of market activity was based on the necessity to obtain a sufficient number of property sales for statistical analysis and a representative sample of housing sales and associated open space relations. Analysis with a larger sample will benefit from the increased likelihood of observing similar properties near different open space of the same type, making control of property characteristics easier.

The selected duration of market activity was also based on conditions of the Greater Boston housing market. The presence of relatively stable housing prices in a market is necessary for the assumption of equilibrium to be made in a hedonic property value model. Significant changes in housing prices within the study area can eclipse any influence open space has on sale prices. In the Greater Boston housing market, home prices decreased slightly over the first half of the 1990s, a reflection of the area's recession between 1991 and 1992 (Allen et al., 2002). However, after 1995 housing prices began to climb with prices skyrocketing after 1998 (Allen et al., 2002). Therefore, based on the Greater Boston housing market, the six year time frame from 1993 to 1998 was selected as the most recent and stable duration for the collection of property sales data.

Property sale prices for the time frame 1993 to 1998 were adjusted to constant 1990 dollars using the Consumer Price Index (CPI) for housing costs within the Boston-Brockton-Nashua area.[5] The nominal mean sale price, deflation factor, and real mean sale price for each year of property sales used for analysis are presented in Table 4.1.

[5] Prices were adjusted to 1990 dollars because the census variables used in the analysis were also based on the year 1990.

Table 4.1: Adjustment of Housing Prices.

Year of Sale	Nominal Mean Sale Price for Study Area	Deflation Factor	Real Mean Sale Price for Study Area
1993	$145,283	1.069	$135,905
1994	$200,876	1.080	$185,996
1995	$244,398	1.107	$220,775
1996	$256,944	1.148	$223,819
1997	$261,231	1.186	$220,262
1998	$301,280	1.210	$248,992

Data Collection

Hedonic property value studies require numerous types of data that are often collected from multiple sources. Property sale data alone can be obtained from many sources; however, as Freeman (2003) has noted, actual market transactions are preferred because the hedonic price function is assumed to be in equilibrium. Therefore, the "arm's-length" transaction occurring between a willing buyer and a willing seller is the ideal form of property sale data for hedonic studies. Such data is often made proprietary by multiple listing services (MLS) that collect and compile data from tax assessment inventories at the city and/or county level. The majority of recent hedonic studies have utilized actual market transactions (e.g. Lupi et al., 1991; Garrod and Willis, 1992; Do and Grudnitski, 1995; Geoghegan, 1997; Mahan et al., 2000; Tyrvainen and Miettinen, 2000; Bolitzer and Netusil, 2000; Lutzenhiser and Netusil, 2001; Espey and Owusu-Edusei, 2001; Acharya and Bennett, 2001; Thorsnes, 2002; Smith et al., 2002; Irwin, 2002; Geoghegan et al., 2003).

For a hedonic study, one must also obtain data that describe the characteristics of the property, including both the land and the structure. This data is maintained by most town or county tax assessors and may also be available through MLS. For neighborhood or locational characteristics, the researcher typically turns to the national census for group or tract level aggregate data. Additionally, neighborhood characteristics, such as distance to employment centers or transportation routes may be created by the researcher in a GIS using spatially referenced property sales and regional GIS data. Finally, the focus variables of most hedonic studies, the environmental characteristics, are almost always created by the researcher in a GIS, or at the very least, created by appending existing environmental GIS data to fit the application of interest.

Sales, Land, and Structural Variables

The property sale data used in this analysis consists of 1,597 residential, market transactions occurring between 1993 and 1998 (inclusive), purchased from Warren Information Services of Boston. The original data of 2,983 observations consisted of both housing sales and vacant land sales that were zoned residential at the time of sale. The sales data was complete with land and structural characteristics for most properties. Of the 2,983 observations received, 1,957 were complete with individual latitude and longitude coordinates, making spatial reference in the GIS straightforward.[6] An effort to spatially reference the remaining 1,026 observations without latitude and longitude coordinates was conducted in the GIS by performing an interactive address match. The sales were matched to a 2000 U.S. Census Bureau Topologically Integrated Geographic Encoding and Referencing system (TIGER) line file covering the entire State of Massachusetts. The address matching process produced an additional 251 spatially

[6] Projection: Massachusetts State Plane (mainland zone); Datum: NAD 83 (meters).

referenced property sale observations.[7] This rather small success rate was a function of the interactive process used in order to ensure a high rate of accurate address matches. During the matching process, potential matches were simultaneously ground-truthed using large-scale, detailed street atlases of the study area. The two sets of data were then merged, producing 2,208 spatially referenced observations.

After talking to tax assessors from the four towns, the determination was made to remove observations from the data that had lot sizes less than 0.05 acres (2,178 ft^2). These properties are not conducive for building based on land use code in the study area. Additionally, observations with adjusted sale prices less than $17,000 were removed from the data because the unadjusted prices of these properties were less than 15 percent of the unadjusted assessed prices. The hedonic method assumes that the sale of a residential property is the product of an "arm's length" transaction in which a willing buyer and a willing seller interact in a perfectly competitive market. Therefore, such undercut prices are not consistent with the conceptual framework used for valuation. Alternatively, extremely high, outlying adjusted sale prices of $4.2 million and $17.6 million were removed from the data by truncating the property sales at $2.03 million. Finally, 11 homes with negative ages, and 302 homes with missing values for interior square-footage were removed from the data, as these structural characteristics are important determinants of price and will enter the hedonic price function. The final property sales data used for analysis consists of 1,597 observations. The descriptive statistics of the land and structural characteristics of these observations appear in Table 4.2.

[7] A dummy variable indicating if observations were provided with latitude and longitude coordinates was included in early models. The dummy variable was significant at the five percent level; however, it was omitted from final models because of the robustness of the remaining variables regardless of its presence.

Table 4.2: Descriptive Statistics for Home Characteristics.

Variable	Mean	Std. Dev.	Minimum	Maximum
Real sale price (1990 dollars)	$214,651	$186,034	$18,067	$2,024,793
Lot acreage	0.86	1.25	0.05	19.5
Age (years at time of sale)	44	35	0.0	313
Interior square-footage	1,823	965	372	9,483

Neighborhood Variables

Neighborhood variables used in this study consist of data obtained from two different sources. The 1990 U.S. Census of Population and Housing was utilized for two of the neighborhood variables – the percentage of people in each census tract with at least some college education (P_EDUC), and the percentage of people in each census tract over the age of 65 (P_AGE65). Additional neighborhood variables were created to capture the proximity of each property sale to the closest Massachusetts Bay Transportation Authority (MBTA) commuter rail station (T-stop), major road, and commercial land use. Theses variables were created in the GIS by performing spatial joins between the property sales data and GIS data obtained from the Massachusetts Office of Geographic and Environmental Information. The variables are measures of the linear distance in meters from each property sale to the closet T-stop, major road, and commercial land use. All five neighborhood variables were then joined to the sales, land and structural variables using a unique identifier assigned to each property sale and imported into a statistical software package.

Environmental Variables

Open space data for this study was obtained from the Office of Geographic and Environmental Information (also know as the Massachusetts Geographic Information

System), a division within the state's Executive Office of Environmental Affairs. The 'Protected and Recreational Open Space' data was not only obtained for the four towns of Billerica, Bedford, Carlisle, and Concord, but also was acquired for the concentric ring of towns surrounding the four-town study area to ensure that peripheral sales within the study area had accurate measures of open space within surrounding neighborhoods. The 'Protected and Recreational Open Space' data was compiled, and is continually updated, on a volunteer basis. The combined efforts of state environmental agencies, regional planning commissions, municipal planning and engineering departments, town conservation commissions, local watershed associations, local and regional nonprofits, and open space planning committees has produced this vector GIS data consisting of open space throughout Massachusetts. The specific land uses included in the open space data are presented in Figure 4.2.

Figure 4.2: Description of Land Types in the 'Protected and Recreational Open Space' Data.[8]

Conservation land
Habitat protection with minimal recreation (walking trails).

Recreation land
Privately or publicly owned outdoor facilities including town parks, commons, playing fields, school fields, golf courses, bike paths, scout camps, and fish and game clubs.

Town forest parkways
Natural buffers along roads.

Agricultural land
Land protected under an Agricultural Preservation Restriction and administered by the state Department of Food and Agriculture.

Aquifer protection land
Excluding zoning overlay districts.

Watershed protection land
Excluding zoning overlay districts.

Cemeteries
If recognized as a conservation or recreation resource.

[8] Source: Massachusetts Office of Geographic and Environmental Information, 2004.

Creation of Open Space Type Variables

The original open space data acquired from the Office of Geographic and Environmental Information was appended to meet the objectives of this research. The data was simply assigned an additional field in the attribute table in order to create a variable indicating the particular type of each open space. Open space was assigned one of six different codes based on whether the open space is in use as a NWR, natural park or conservation land, agricultural operation, urban park or sports facility, golf course, or cemetery. Additionally, a small number of open spaces were excluded from the study because of failure to fall within any of the aforementioned categories (for instance, an indoor skating rink was omitted because this site lacked any vegetated open space and consisted of only a structure and asphalt).

Open space types were assigned using two different methods: an interactive categorization process and an automatic categorization process. The decision to use two different methods of categorizing open space was made after first attempting to classify observations using only the 'Protected and Recreational Open Space' GIS data. As previously described, this data was compiled on a volunteer basis and lacked consistency within each attribute field across like observations. Therefore, the task of categorizing observations in the data into open space type was an arduous one using only the original open space GIS data. Digital aerial photographs of the area were then obtained and the interactive process of simultaneously examining the attributes of the open space GIS data and the land cover associated with each open space in order to assign the appropriate code began. However, this process involved a substantial time burden, and an automatic categorization method was developed with the intention of being able to produce similar results to those of the interactive method at a fraction of the cost of time.[9] The two methods are described in more detail in what follows.

The interactive process consisted of assigning user defined codes within the GIS. The existing fields: site name (where applicable) and primary purpose were examined in order to determine the appropriate open space type for each observation to be categorized. In addition, a visual inspection of each open space was performed within the GIS through acquisition of digital orthophotos for the entire study area. The 0.5-meter resolution, color images were originally collected on film in April 2001 by Keystone Aerial Surveys, Inc. of Philadelphia, Pennsylvania, and provided exceptional detail for assisting with the categorization of open space. Furthermore, municipal comprehensive plans, open space plans, and plat maps for the four-town study area were examined (where available) in order to establish greater confidence in the interactive open space categorization and ensure that the assignment of open space types was similar to that perceived by town officials and citizens.

The second method used to disaggregate open space into land use type was an automatic process developed within statistical analysis software. This method relied on the development of a programming code that would objectively assign a type to each open space. Using existing fields in the original open space data, such as primary purpose, public access, level of protection, and ownership type, a code was written that iteratively selected observations from the data and then assigned the associated open space code based on the attributes of each open space. As was the case with the

[9] Of course, the success of an automatic categorization method at reproducing the results of an interactive method should be examined with sensitivity analysis.

interactive coding process, a small number of observations were omitted from analysis because of failure to associate with any of the aforementioned open space types.

After the assignment of open space types, both of the open space data sets were utilized in conjunction with the property sales data in the GIS in order to generate environmental variables for the hedonic model. The first variable generated in the GIS was a measure of the linear distance from each property sale to the closet open space of each type. First, separate spatial data sets for each of the six open space types were created in the GIS. Spatial joins were then performed between the property sales data and each open space type to produce distance variables. The distance variables are linear measures in meters from each property sale to the closest open space of each type. Once imported into the statistical analysis software, discrete measures of distance to the closest open space of each type were created based on the continuous distance variables generated in the GIS.

In a similar fashion, distance measures to the closest publicly accessible open space and privately accessible open space were calculated. Separate open space data sets were first generated based on a field provided describing public access in the original open space data. Publicly accessible open spaces were defined as those with public access equal to 1 (public), 2 (public, residents only), 4 (private, public welcome), and Y (yes, open to public), while privately accessible open spaces were defined as those with public access equal to 5 (private, members only), 6 (none), and N (no, not open to public). Spatial joins were then performed between the property sales data and the publicly accessible open space data and the privately accessible open space data to produce distance variables. The variables were then imported into statistical analysis software and joined to the land, structural, and neighborhood data.

Creation of Open Space Diversity Variables

The appended open space GIS data was used to derive indices describing the diversity of open space types in neighborhood zones around each property sale. Two indices were created, each at a different scale: 100 meters and 1,000 meters around each property in the data. To derive the indices, buffers were first created in the GIS at the two scales. The open space data was then intersected with the buffers which produced clipped polygons of the six open space types contained within the two distinct neighborhoods around each property sale. The area of each intersected open space was then recalculated in the GIS and the total area of each open space type contained within each buffer was created by summarizing the area of each polygon of a particular open space type across each property sale. After adding a new field to the attribute table of each intersection, a variable depicting the proportion of each open space within the associated neighborhood of each property sale was created by dividing the areas of the newly clipped open space by the area of each of the buffers.[10] These variables were also joined into a single table using a unique identifier assigned to each property sale and imported into the statistical analysis software for analysis, where the open space diversity index for each scale was calculated as:

$$D = -\sum_k (P_k) \ln(P_k)$$

[10] The area of each circular buffer was calculated as: $AREA = \prod r^2$. Thus, the area of the 100m buffer is $31{,}416\text{m}^2$ and the area of the 1,000m buffer is $3{,}141{,}592\text{m}^2$.

where P_k is the proportion of the surrounding landscape in open space type k, and ln is the natural logarithm.

The data requirements of hedonic models are numerous and the researcher must attain sales, land, structural, neighborhood, and environmental characteristics for each property in the study, often from multiple sources. As such, compiling data for analysis becomes a process of joining the numerous attributes of each property into a single set of data. This composite set of data becomes the hedonic model of the composite good, housing.

Chapter 5

HEDONIC MODELS

Model Specification

The hedonic price function relates the sale price of a residential property to a vector of land characteristics, a vector of structural characteristics, a vector of neighborhood characteristics, and vector of environmental characteristics, such that:

$$\ln(P_i) = \beta_0 + \sum\beta_j L_{ji} + \sum\beta_k S_{ki} + \sum\beta_l N_{li} + \sum\beta_m E_{mi} + \varepsilon_i$$

where ln is the natural logarithm, and ε_i is the observation specific error term.

The land, structural, and neighborhood variables selected for the hedonic model were based on the literature, availability of property data, and geographic characteristics of the study area (see Table 5.1). The selection of structural variables requires consideration of which attributes of a home are most representative of the quality and value of the property. However, at the same time, one must keep in mind the fact that many structural variables of a residential property are likely to be multicollinear. It is apparent that the size of the land under ownership (ACRES) contributes to the value of a residential property and is therefore included in the model. Additionally, the size of the structure on the land is important in determining the value of a property, so the interior square-footage of a home (INTSF) is also included. Another structural variable included in the final specification is the age of the house at the time of sale (AGE) and is intended to be a proxy for quality. Other structural variables typically included in hedonic property value models, such as the number of total rooms, the number of bedrooms, and the number of bathrooms, were not included in the final specification because of the likelihood that these variables are multicollinear; a common result of the near linear relation such variables have with the size (INTSF) of residential properties. The signs of the coefficients for the size variables are expected to be positive, indicating that a unit increase in lot size or house size would be associated with an increase in the sale price of the residential property. The age of a home is expected to have a negative coefficient, indicating that as a home ages, the value of the home falls.

Table 5.1: Explanatory Variables.[11]

Name	Description
Land	
ACRES	the size of the lot measured in acres.
Structural	
INTSF	the interior size of the house measured in square feet.
AGE	the age of the house at the time of sale.
Neighborhood	
P_EDUC	the percentage of people in each census tract with at least some college education.
P_AGE65	the percentage of people in each census tract over the age of 65.
DIST_TSTOP	the Euclidian distance in meters to the closest commuter train station.
DIST_MJRD	the Euclidian distance in meters to the closest major road.
DIST_COMRC	the Euclidian distance in meters to the closest commercial land use.
Environmental	
OPEN_SPACE	various measures of open space.

Neighborhood attributes of a property are included in the hedonic model in order to control for the socio-economic characteristics that have amenity (disamenity) value and contribute to (depress) the sale price of a residential property. Census variables at the tract level are often included as neighborhood variables in hedonic models. The percentage of people in each census tract with at least some college education (P_EDUC) and the percentage of people in each census tract over the age of 65 (P_AGE65) are included in this analysis and were obtained from the 1990 U.S. Census of Population and Housing. The signs of the coefficients for P_EDUC and P_AGE65 are expected to be positive, indicating that as the percentage of people in any census tract with at least some

[11] The descriptive statistics of these variables and all the environmental variables appear in Table 6.1.

college education (over the age of 65) increases, the sale prices of residential properties within that tract also increase. Distance measures from each residential property to the closest Massachusetts Bay Transportation Authority (MBTA) commuter rail station (DIST_TSTOP) and major road (DIST_MJRD) are included as neighborhood variables in order to account for accessibility and/or nuisance effects of residential location in the proximity to the area's transportation networks. Additionally, the distance from each residential property to the closest commercial land use (DIST_COMRC) is included in the model in order to account for either convenience effects of being located in proximity to shopping or nuisance associated with increased congestion. Therefore, because of the potential for differential price effects, *a priori*, it is unclear the signs that the coefficients for these neighborhood distance variables will assume.

As previously mentioned, open space can enter the hedonic model as an environmental variable in the form of numerous measures. In selecting among the alternative measures of open space to include in the hedonic price function, it is important to consider both how residents perceive open space, and the ease with which results of a particular measure can be applied to policy. The measures of open space included in this analysis include continuous distance from each property to the closest open space of each type, discrete measures of distance to the closest open space of each type, continuous measures of distance to the closest public and private open space, and an index describing the diversity of open space types evaluated at 100 and 1,000 meters around a home.

The bulk of hedonic property value studies have attempted to capture the land capitalized value of environmental amenities through measures of distance or access (e.g.

Knetsch, 1962; Weicher and Zerbst, 1973; Correll et al., 1978; More et al., 1988; Do and Grudnitski, 1995; Doss and Taff, 1996; Tyrvainen, 1997; Powe et al., 1997; Johnston, 1998; Kluge and White, 1999; Mahan et al., 2000; Tyrvainen and Miettinen, 2000; Bolitzer and Netusil, 2000; Luttik, 2000; Lutzenhiser and Netusil, 2001; Espey and Owusu-Edusei, 2001; Johnston et al., 2001; Shultz and King, 2001; Thorsnes, 2002; Smith et al., 2002; Netusil, 2005). Such measures provide the researcher and policy maker with easily interpreted marginal values of the environmental amenity that describe the increase (decrease) in the price of a home for a one unit increase in the distance of the property from the environmental amenity (disamenity). This approach to valuing open space is also used for the current research. Therefore, a subset of the environmental variables in this study consist of distance measures from each residential property to the closest NWR land, conservation land, agricultural land, sports park, golf course, and cemetery in the study area.

Open Space Type Variables - Continuous Distance

The six open space distance variables are measures of Euclidian distance in meters to the nearest open space of each type. The coefficients for the variables are estimated within a log-linear hedonic price function of the form:

$$\ln(P_i) = \beta_0 + \beta L_i + \sum\beta_k S_{ki} + \sum\beta_l N_{li} + \sum\beta_m DIST\_OS_{mi} + \varepsilon_i$$

For the log-linear functional form, the marginal effect or marginal implicit price (MIP) of the each open space distance variable is calculated as:

$$\frac{\partial \ln(P)}{\partial DIST\_OS_1} = \beta_1 * \overline{P} = MIP$$

where $\overline{P}$ is the mean sale price in the study area.

It is anticipated that the signs of these variables will vary across open space type, indicating differences between whether or not each open space is considered an environmental amenity or disamenity within the housing market. Negative signs on the coefficients will indicate that as the distance from a residential property to a particular open space increases, the sale price of that property decreases, suggesting that the open space is an amenity. Positive coefficients will indicate the opposite distance price relationship and suggest that an open space is a disamenity.

Open Space Type Variables - Discrete Distance

While the inclusion of continuous, linear measures of distance in the hedonic price function allows one to compare the relative levels of amenity (disamenity) across open space types, it also restricts the hedonic price function to assume a constant sign over the range of each distance variable. In order to examine discontinuities in the hedonic price function with respect to distance to open space, the continuous measures of distance to the closest open space of each type are also disaggregated into discrete zones and estimated. Little guidance exists in the literature regarding the appropriate delineation of zones. Review of the few studies that have taken this approach to valuing open space with discrete measures of distance suggests that researchers either made arbitrary delineations in the data or manipulated the delineations in an iterative process based on the significance of parameters. Indeed, the existence of unique delineations for each type of open space included in a model suggests *ex post* justification (see Espey and Owusu-Edusei (2001) in Table 3.1). However, borrowing from Geoghegan et al. (1997), there exist a distinct set of urban geographies that provide strong conceptual grounding for disaggregating continuous measures of distance into discrete zones. As described below, the zones enter a log-linear model as dummy variables in the following manner:

$$\ln(P_i) = \beta_0 + \beta L_i + \sum\beta_k S_{ki} + \sum\beta_l N_{li} +$$

$$\sum\beta_m ZONE1_{mi} + \sum\beta_n ZONE2_{ni} + \sum\beta_o ZONE3_{oi} + \varepsilon_i$$

where: *ZONE1* = 1 if the closest open space of type *z* exists within 50 meters of property *i*, 0 otherwise,
ZONE2 = 1 if the closest open space of type *z* exists between 51 and 100 meters of property *i*, 0 otherwise,
ZONE3 = 1 if the closest open space of type *z* exists between 101 and 1,000 meters of property *i*, 0 otherwise, and
ZONE4 is the omitted case in which the closest open space of type *z* exists beyond 1,001 meters of property *i*.

The discrete zones above have conceptual grounding. *ZONE1* was selected based on the premise that the types of land uses bordering any residential property *i* are likely to affect the value of property *i*. This zone is theorized to capture 'neighbor effects'.[12] The existence of zoned buffers between incompatible land uses in contemporary land use regulation provides evidence of such effects. Additionally, Do and Grudnitski (1995) found that golf courses in San Diego County have positive effects on values of adjacent residential properties. *ZONE2* was selected based on the premise that the land use types that are within the same neighborhood as any residential property *i* are likely to affect the value of property *i*. This zone is conceptualized to capture 'neighborhood effects'. Geoghegan et al. (1997) used 100 meters as the threshold distance for what can be seen from properties in a Maryland watershed and found a positive and significant estimate for the marginal contribution of open space. *ZONE3* was selected based on the premise that the types of land uses within walking distance of any residential property *i* are likely to affect the value of property *i*. Geoghegan et al. (1997) used 1,000 meters as the distance

[12] It is important to note that 50 meters was selected as the threshold distance instead of a lower value because two of the six open space types have too few observations in this lower range for model estimation.

for what land uses would be encountered on foot within the bounds of a comfortable walk from one's home. The authors again found a positive estimate for the marginal contribution of open space at the 1,000 meter scale. At a similar scale, a study by Bowes and Ihlanfeldt (2001) found a negative estimate for the presence of Metropolitan Atlanta Rapid Transit Authority (MARTA) rail stations located between one-half mile (535 meters) and one mile (1,069 meters) from residential properties. The threshold distance between *ZONE3* and *ZONE4* was also selected based on the premise that the types of land uses that require an automobile (or other form of transportation) to access from any residential property *i* are likely to be considered part of a distinct geography from that of *ZONE 1*, *2*, or *3*.

The marginal effect of each open space zone is calculated in an identical manner to that of the continuous distance variables. However, the marginal effects are interpreted as the difference in house sale price between the homes located in the particular zone of interest (β_{zj}) and the omitted zone as a result of the closest open space of type z being located in zone j. For the most part, it is expected that the signs of the coefficients for the discrete zones will tell a similar story to the signs of the continuous distance variables, with the magnitudes of the coefficients decreasing from *ZONE1* to *ZONE3*. Using dummy variables an environmental amenity would have positive coefficients and be decreasing in magnitude, while a disamenity would have negative coefficients and be decreasing in magnitude.

Public/Private Open Space - Continuous Distance

In addition to open space type as defined by land use, the literature has revealed that the ownership of open space, either public or private, is important in explaining sale prices of residential properties (Bolitzer and Netusil, 2000; Geoghegan, 2002; Irwin, 2002). However, it remains unknown whether the accessibility of a particular open space, either public or private, affects surrounding property values. It is an objective of this thesis to estimate marginal values for an additional set of environmental focus variables that measure distance to publicly accessible and privately accessible open space, regardless of land use type. As described in Chapter 4, these variables were derived from existing fields in the 'Protected and Recreational Open Space' GIS data obtained from the Massachusetts Office of Geographic and Environmental Information. The private and public open space variables enter a log-linear model as continuous distance measures to the closest publicly accessible open space (DIST_PUBOS) and the closest privately accessible open space (DIST_PRIOS) from each property sale. Based on the results of previous studies, the signs of these coefficients are expected to both be negative.

Open Space Diversity Indices

While proximity to open space has been repeatedly shown to be an important determinant in the price of a residential property, solely relying on such measures of location limits the explanation of open space valuation. Again, recent advances in GIS have set the stage for more detailed examinations of the spatial relationships among land uses. Expanding on the work of Geoghegan et al., (1997) and Acharya and Bennett

(2001), this research also incorporates, as environmental variables, landscape indices that describe the diversity of open space surrounding a home, calculated at 100 and 1,000 meters.

Ecologists have used indices to measure landscape pattern and determine the capacity of various landscapes to support specific ecological processes. Indeed, animals have preferred habitats which govern how they inhabit the landscape in terms of spatial extent and population sizes. In a similar vein, it has been suggested that landscape pattern may also affect human settlement. Therefore, it is logical to examine whether landscape heterogeneity helps to determine residential location. Specifically, it remains unexplored whether or not heterogeneity of open space is valued in the housing market when selecting a residence.

Borrowing from the landscape ecology literature, an index is used which measures open space heterogeneity and describes whether a few concentrated open spaces, or a distribution of many open spaces, dominate the surrounding landscape of a home (Turner, 1990).[13] Each index is calculated as:

$$D = -\sum_k (P_k) \ln(P_k)$$

where P_k is the proportion of the surrounding landscape in open space type k. This variable depends on both the diversity of open space and the similarity of proportions of the open space in a given area. The greater the diversity index associated with a specific observation, the greater the number of open space categories and the more similar the proportions of the categories. Following the approach taken by Geoghegan et al. (1997), the diversity indices enter the hedonic price function at two distinct scales of urban

[13] This diversity index is based on information theory developed by Shannon and Weaver (1962). A version of the index was first applied to landscape ecology by O'Neil et al. (1988).

geography: a 100 meter scale and 1,000 meter scale. The two scales are intended to distinguish between what open space is in the surrounding neighborhood of a property versus what open space would be encountered on a walk from a property. The indices enter both the log-linear continuous distance model and the log-linear discrete distance model. Additionally, the marginal effects of open space diversity on sale price are calculated in an identical manner to that previously described for the continuous distance variables.

A priori, it is difficult to know if the coefficients for the diversity indices will be positive or negative. A positive coefficient would suggest that the more diverse the selection of open space in a given neighborhood, the more expensive the residential properties in that neighborhood, whereas a negative coefficient would suggest just the opposite. While greater diversity of open space in a neighborhood increases the choice set of individuals seeking recreation and potentially limits the extent of the residential-commercial/industrial interface, it also may be accompanied by greater nuisance as a result of outsiders seeking to exploit the recreational diversity of the area, or less convenience associated with a lack of other land uses such as neighborhood commercial centers. Therefore, it is unclear the expected signs of the coefficients for the diversity indices.

Hedonic Property Value Models and Spatial Dependence

When estimating hedonic models econometric complications are likely to arise. Particular econometric concerns common in hedonic property value models result from spatial dependencies among observations. Essentially, spatial dependence is the lack of autonomy between observations within cross-sectional data (Anselin, 1988). In a hedonic property value model when relative locations are important (for the estimation of land use spillovers), failure to correct for spatial dependencies between observations can not only affect the magnitude and significance of parameter estimates, but will also affect standard diagnostic tests and resulting inferences (Anselin, 1988).

Spatial dependence was first addressed in hedonic property value models by Dubin (1988) and Can (1990). Using a data of residential property sales in Baltimore, Maryland, Dubin (1988) utilized a maximum likelihood (ML) procedure for estimating the covariance matrix of the error terms in order to obtain efficient parameter estimates and unbiased standard errors. Estimation results indicated that the cost of ignoring spatial dependence within the data and pursuing estimation with ordinary least squares (OLS) was the incorrect interpretation of two of the 13 effects of housing characteristics on sale price. Can (1990) utilized a spatial lag (spatially weighted dependent variable) in addition to a varying parameters approach in order to account for both spatial dependence and spatial heterogeneity.[14] Estimation results based on residential properties in Columbus, Ohio indicated that the housing market was replete with significant spatial dynamics, factors that necessitated correction in order to obtain unbiased and consistent parameter estimates.

Application of spatial econometrics to hedonic property value models concerned with the valuation of environmental variables has received little attention in the literature. Two of the notable exceptions pertain to studies of air pollution conducted by Beron et al. (2002) and Kim et al. (2003). In the only hedonic open space study to account for spatial

[14] Spatial heterogeneity refers to the lack of stability in the behavior of events across space and is often exhibited in a model in the form of a heteroscedastic error term (Anselin, 1988).

dependence, Bell and Bockstael (2000) estimated a spatial error model for a housing market in a Maryland watershed. Acharya and Bennett (2001) considered the presence of spatial dependence in their hedonic study of open space in Connecticut, but rejected the hypothesis that perceptible spatial dependence due to omitted spatially correlated variables exists. Therefore, to the best of the author's knowledge, modeling spatial dependence in a hedonic open space study remains largely unexplored in the literature.

Accounting for Spatial Dependence

Spatial dependence is accounted for in econometric models with two distinct modeling techniques. One technique of correcting for spatial dependence assumes a structural dependence exists between observations of the dependent variable. In this case, what is observed at one location is determined, at least in part, by what happens at other locations throughout the system. In the housing market, structural dependence exists if the sale price of a particular residential property is influenced by the sale prices of other properties in the surrounding neighborhood. This form of spatial dependence, which warrants the application of a spatial lag model, is comparable to the ordered dependence within time-series data; however, the spatial version is multidimensional. In time-series data, only events from the past can affect current events, a unidirectional relationship, whereas in cross-sectional spatial data, an observation (property) is likely to be related to surrounding observations (properties) in multiple dimensions. Furthermore, as opposed to time-series data, dependence between cross-sectional data is void of time's natural ordering and is therefore bi-directional.

The second technique of accounting for spatial dependence assumes that there exists dependence across error terms, a result of the omission of variables from the hedonic function that follow a spatial pattern. Since much of the cross-sectional data that exists is only available at an aggregate scale, which may have little correspondence with the scope of underlying spatial phenomena, there is likely to be error in the measurement of variables. Such measurement error can spillover into other spatial units, making the errors themselves related. If these variables are omitted from the hedonic specification, the model will exhibit spatial autocorrelation. In other words, the price of any property will not only be a function of the associated land, structural, neighborhood, and environmental variables, but will also be a function of any omitted variables associated with the observation and/or neighborhood. Remediation of this form of spatial dependence requires estimation of a spatial error model.

As indicated by the previous discussion, correcting for spatial dependence requires the assumption of an underlying structure of spatial dependence. A parameter of the structure is then estimated in unison with the parameters of the focus variables of the model. Cliff and Ord (1973) provide a framework for the spatial lag model in which $Y = \rho WY + X\beta + \varepsilon$ where ρ is an unknown scalar parameter, and WY is a vector of the spatially lagged dependent variable, which implies that $Y = (I - \rho W)^{-1}X\beta + (I - \rho W)^{-1}\varepsilon$, where Y is an n × 1 vector of dependent variable observations, W is an N × N spatial weight matrix, X is an n × k matrix of explanatory variables, β is an n × k matrix of unknown parameters, and ε is an n × 1 vector of random error terms with expected value 0 and variance-covariance matrix $\sigma^2 I$, or $\varepsilon \sim N(0, \sigma^2 I)$.

In the spatial error model, outlined by Cliff and Ord (1973), spatial dependence takes the form: $Y = X\beta + \varepsilon$ where $\varepsilon = \rho W\varepsilon + \mu$, implying that $Y = X\beta + (I - \rho W)^{-1}\mu$, where Y is an n × 1 vector of dependent variable observations, X is an n × k matrix of explanatory variables, β is an n × k matrix of unknown parameters, W is an n × n spatial weight matrix, ρ is an unknown scalar parameter, μ is an n × 1 vector of random error terms with expected value 0 and variance-covariance matrix $\sigma^2 I$ (i.e. $\mu \sim N(0, \sigma^2 I)$), and ε is an n × 1 vector of random error terms with expected value 0 and nonspherical variance-covariance matrix, such that $E[\varepsilon\varepsilon'] = \sigma^2[(I - \rho W)'(I - \rho W)]^{-1}$.

The Spatial Weight Matrix

In the aforementioned structures of spatial dependence, W, the spatial weight matrix, contains information that represents the pattern of dependence between observations. It is similar to a lag operator in time-series models, but again, within a spatial model, the lags are multidirectional and considerably more complicated. A particular element in W, w_{ij}, (the i^{th}, j^{th} element) represents the assumed dependence between the i^{th} and j^{th} observation. A nonzero element in each row i of W, defines j as a neighbor of i. However, because an observation cannot be a neighbor of itself, the diagonal elements of W are necessarily zero. If a neighbor of i is defined as any element j who shares a common border with i (a contiguity based weight), it is often the practice in devising W to standardize each row i so that it sums to one. Thus, each element in the standardized matrix will fall between 0 and 1 as imposed by $w^s_{ij} = w_{ij}/\sum_j w_{ij}$. This generalization is the most common assumption of spatial structure and is typically performed in order to make the spatial dependence easier to interpret and so that parameter estimates may be more readily compared between different models. However, it may be more appropriate to specify the spatial weight matrix based on a geographic relationship between observations, such as distance. In this case, a neighbor of i is defined only when a surrounding element j is within some critical distance of i, in which case, j would take the value of 1, and 0 otherwise. Such distance-based weight matrices are also often row-standardized. This latter case of imposing a distance-based weight matrix is most applicable to the microlevel situation of this investigation, in which households are scattered irregularly across the study area.

When estimating a hedonic model in the presence of a spatially correlated error term, the use of OLS will result in unbiased but inefficient parameter estimates. Additionally, estimates of standard errors will be biased leading one to make incorrect inferences. If instead the underlying spatial dependence is more of the structural nature, estimation using OLS will result in biased and inconsistent parameter estimates. Therefore, the typical approach to estimating models with spatial dependence relies on estimation with maximum likelihood (ML) techniques. However, generalized-moments (GM) have also been used to obtain consistent parameter estimates in micro-level data exhibiting spatial dependence (Bell and Bockstael, 2000).

As hedonic property value models describe the sale price of residential properties as a function of various characteristics that are also spatial in nature, relative locations are important for the estimation of land use spillovers. In order to obtain the most accurate results for application to policy, spatial dependence must be tested for and, if found, accounted for.

Chapter 6

RESULTS

The descriptive statistics of the hedonic variables used in this study appear in Table 6.1. Three models are estimated and are the subject of this chapter. The first results presented are those of the 'Continuous Distance Model', followed by the 'Discrete Distance Model', and the 'Public/Private Continuous Distance Model'. The diversity indices enter both the continuous distance model and the discrete distance model, in order to capture any effects that landscape pattern have on property values. The three models were estimated first with ordinary least squares (OLS), however, as the following discussion indicates, the OLS results are not appropriate for statistical inference.

Correcting for Heteroscedasticity

Given the cross-sectional nature of the data in this hedonic study, there is reason to believe that the error terms of the OLS estimators are non-spherical, or heteroscedastic. While OLS estimators remain unbiased under conditions of heteroscedasticity, they no longer have minimum variance. As a result, interval estimation and hypothesis testing can no longer be trusted. As Anselin (1988) notes, in the presence of spatial dependence the properties of several conventional tests for heteroscedasticity are no longer valid. Monte Carlo experiments reveal that the Glejser test is the most powerful, followed by the Breusch-Pagan test and the White test for identifying heteroscedasticity in the presence of spatial dependence (Anselin, 1988). Based on the relative performance of the three tests, the Glejser test is used in this thesis to test for heteroscedasticity.

Table 6.1: Descriptive Statistics of Hedonic Variables.

Variable	Mean	Std. Dev.	Minimum	Maximum
PRICE	214,651.25	186,034.22	18,066.85	2,024,793.39
Land				
ACRES	0.86	1.25	0.05	19.50
Structural				
AGE	44.37	35.04	0.00	313.00
INTSF	1,822.74	965.23	372.00	9,483.00
Neighborhood				
P_EDUC	0.59	0.13	0.38	0.82
P_AGE65	0.16	0.07	0.08	0.31
DIST_TSTOP	4,816.88	2,269.47	107.77	9,261.43
DIST_MJRD	426.09	421.78	1.28	1,990.62
DIST_COMRC	731.56	533.96	0.10	3,128.90
Environmental				
DIST_AG	2,099.00	1,341.74	0.10	6,067.08
DIST_CEM	2,480.36	1,487.71	12.85	7,271.38
DIST_CONS	351.44	263.84	0.10	1,268.89
DIST_GOLF	2,297.86	1,359.68	0.10	6,672.66
DIST_SPRT	623.48	453.98	0.09	2,804.60
DIST_GRM	3,714.60	1,920.22	28.34	8,198.88
AG_50	0.01	0.11	0.00	1.00
AG_100	0.01	0.10	0.00	1.00
AG_1000	0.26	0.44	0.00	1.00
AG_BASE	0.72	0.45	0.00	1.00
CEM_50	1.88E-03	0.04	0.00	1.00
CEM_100	3.13E-03	0.06	0.00	1.00
CEM_1000	0.16	0.37	0.00	1.00
CEM_BASE	0.83	0.37	0.00	1.00
CONS_50	0.08	0.27	0.00	1.00
CONS_100	0.07	0.26	0.00	1.00
CONS_1000	0.82	0.39	0.00	1.00
CONS_BASE	0.03	0.17	0.00	1.00
GOLF_50	0.01	0.08	0.00	1.00
GOLF_100	3.13E-03	0.06	0.00	1.00
GOLF_1000	0.16	0.36	0.00	1.00
GOLF_BASE	0.83	0.37	0.00	1.00
SPRT_50	0.03	0.17	0.00	1.00
SPRT_100	0.04	0.19	0.00	1.00
SPRT_1000	0.78	0.41	0.00	1.00
SPRT_BASE	0.15	0.36	0.00	1.00
GRM_50	3.13E-03	0.06	0.00	1.00
GRM_100	0.01	0.08	0.00	1.00
GRM_1000	0.08	0.27	0.00	1.00
GRM_BASE	0.91	0.28	0.00	1.00
DIVIND_SM	0.05	0.12	0.00	0.72
DIVIND_LG	0.38	0.18	0.01	0.99
DIST_PUBOS	283.64	210.19	0.09	1,191.36
DIST_PRIOS	1,232.92	977.88	0.10	4,049.73

The Glejser (1969) test checks for the presence of a systematic pattern in the variances of the errors of a particular model. First, an auxiliary equation is estimated for each model in which the absolute value of the OLS residuals is regressed on the explanatory variable(s) to which the heteroscedasticity in each model is thought to be related. The hypotheses of the test are:

H_0: the error variance is homoscedastic, vs.

H_1: the error variance is heteroscedastic.

The test statistic for each model is calculated as nR^2, where n is the sample size and R^2 is coefficient of determination from each auxiliary regression. Second, the test statistic is compared to the critical value of the Glejser test, distributed as $\chi^2_{(k-1)}$, where k is the number of explanatory variables in the auxiliary regression. The null hypothesis is rejected when the test statistic exceeds the critical value. The Glejser test was employed on the OLS residuals of all three models.[15] Each test statistic rejects the null hypothesis of homoscedasticity at the five percent level.[16]

Given the results of the heteroscedasticity tests, Weighted Least Squares (WLS) was used to generate weighted variables that would produce homoscedastic error terms. To create the weighted variables, the square of the residuals from each model were first regressed on each of the explanatory variables in the model independently to determine which variables were driving the heteroscedasticity (refer again to footnote 15). A

---

[15] Heteroscedasticity in the continuous distance model and the public/private continuous distance model was determined to be a function of AGE and ACRES. Heteroscedasticity in the discrete distance model was determined to be a function of AGE, ACRES, DIST_MJRD, AG_50, GOLF_100, SPRT_100, and GRM_1000.

[16] Continuous Distance (OLS): critical value ~ $\chi^2_{(0.05,2)} = 5.99$; test statistic = $nR^2 = 39.93$.

Discrete Distance (OLS): critical value ~ $\chi^2_{(0.05,7)} = 14.07$; test statistic = 81.93.

Public/Private Continuous Distance (OLS): critical value ~ $\chi^2_{(0.05,2)} = 5.99$; test statistic = 44.56.

variance function of each model was then estimated by regressing the square of the OLS residuals on the influential variables simultaneously, with the fitted value (v_i) from the variance regression used to calculate a weight (w_i) for each observation, where $w_i = 1/v_i$. Each variable was then transformed by multiplying each by the square root of w_i. The weighted model was estimated without an intercept and Glejser tests were again performed to test for heteroscedasticity. The test statistic for each model fails to reject the null hypothesis of homoscedasticity.[17]

However, the WLS estimates remain inefficient in the presence of spatial dependence. Recall that spatial dependence can appear in hedonic models in two forms, structural dependence in which the sale price of a particular residential property is influenced by the sale prices of other properties in the surrounding neighborhood, and/or spatial autocorrelation in which there is dependence across error terms as a result of the omission of variables from the hedonic function that follow a spatial pattern. To obtain efficient estimates, spatial dependence in each hedonic equation must also be corrected.

Identifying the Appropriate Structure of Spatial Dependence

Both structural dependence and spatial autocorrelation are likely to be present in hedonic property value models. It is common practice to test for both types of spatial dependence; however, it is also common that only one of the two forms is accounted for in estimation. First, the presence of spatial dependence in general is examined via the Moran's I test statistic. The Moran's I statistic represents the slope of the regression line

[17] Continuous Distance (WLS): critical value ~ $\chi^2_{(0.05,2)} = 5.99$; test statistic = $nR^2 = 4.15$.
Discrete Distance (WLS): critical value ~ $\chi^2_{(0.05,7)} = 14.07$; test statistic = 5.91.
Public/Private Continuous Distance (WLS): critical value ~ $\chi^2_{(0.05,2)} = 5.99$; test statistic = 5.11.

that results from regressing a standardized version of the dependent variable (*Y*) on a spatially lagged version of the dependent variable (*WY*) (Anselin, 2005). As a result, the Moran's I statistic depends on the assumed structure of spatial dependence (the spatial weight matrix *W*). A highly significant statistic indicates the presence of spatial dependence. In order to determine which form of spatial dependence is most apparent, Lagrange Multiplier (LM) test statistics must be examined.

Four LM test statistics are utilized for identifying the most dominant form of spatial dependence: the LM-Lag, the Robust LM-Lag, the LM-Error, and the Robust LM-Error (Anselin, 2005). The LM-Lag and Robust LM-Lag test statistics correspond to the spatial lag model, while the LM-Error and Robust LM-Error statistics correspond to the spatial error model. All of the test statistics are distributed as χ^2 with one degree of freedom. In order to determine the particular form of spatial dependence to model, the test statistics are considered in a specific sequence. The decision sequence is presented in Figure 6.1.

Figure 6.1: Decision Criteria for Modeling Spatial Dependence.

Run OLS Regression and Examine the LM-Error and LM-Lag Diagnostics:

- A. Neither the LM-Error nor the LM-Lag statistic is significant
    1. Proceed with OLS results
- B. One is significant
    1. Run a Spatial Error Model
    2. Run a Spatial Lag Model
- C. Both the LM-Error and the LM-Lag statistics are significant
    1. Examine the Robust LM-Error and Robust LM-Lag Diagnostics
        - i. One is significant
            - a. Run a Spatial Error Model
            - b. Run a Spatial Lag Model
        - ii. Both the Robust LM-Error and the Robust LM-Lag are significant
            - a. Run the model associated with the statistic of greatest significance

First, the model is estimated with OLS and the standard versions of the ML statistics are considered. If neither statistic is significant, such that both fail to reject the null hypothesis of no spatial dependence, the analysis can proceed with OLS. Rejection of the null hypothesis for the LM-Error statistic suggests that the spatial error model is most appropriate, while rejection of the LM-Lag statistic suggest that the spatial lag model is most appropriate. If both standard statistics reject the null hypothesis, the robust versions are considered. Identification of the most appropriate model using the robust statistics follows the same procedure as the standard LM statistics. If both the Robust LM-Lag and Robust LM-Error statistics reject the null hypothesis, then the relative significance of the test statistics are considered. The robust statistic of greatest significance identifies the particular form of spatial dependence that should be accounted for in the econometric model. If both robust statistics are highly significant, the statistic of the largest magnitude is the form of spatial dependence that should be modeled (Anselin and Rey, 1991). For all of the models estimated in this thesis, LM test statistics identified spatial autocorrelation as the dominant form of spatial dependence. This suggests there is dependence across error terms as a result of the omission of variables from the hedonic function that follow a spatial pattern. Specific test statistics will be presented with the results of each model.

While testing can reveal whether spatial dependence of a particular form is present in a hedonic property value model, there is little quantitative guidance for determining the most appropriate structure of spatial dependence, as represented by the spatial weight matrix, *W*. In order to shed light on the most appropriate specification of *W* for the study area under investigation, five alternative spatial weight matrices were considered. All five structures of *W* are distance-based, row standardized matrices, such that a neighbor of *i* is defined only when a surrounding element *j* is within some critical distance of *i*, in which case, *j* would take the value of 1, and 0 otherwise. For instance, $w_{ij} = 1$ if $d_{ij} \leq c$, 0 otherwise, where *c* is the cutoff distance beyond which no spatial dependence is assumed to exist. The five structure of *W* considered here are based on five cutoff distances: 200 meters, 400m, 600m, 800m, and 981m.[18] The approach taken here to use a cutoff distance, as opposed to a distance-decay structure of *W*, has two motivations. First, it seems appropriate that the dependence between observations will become insignificant at some critical distance, and second, assuming a cutoff distance for spatial dependence produces a sparse weight matrix that simplifies *W* and improves the probability of obtaining ML estimates (Bell and Bockstael, 2000).

Each structure of *W* was estimated in separate spatial error models using maximum likelihood. As Anselin (2005) notes, three classic tests can be utilized to compare the null model (the OLS specification assuming no spatial dependence) to each alternative spatial error model to investigate the most appropriate specification of *W*. The three tests are the Likelihood Ratio test (LR), the Wald test (W), and the Lagrange Multiplier test (LM).[19] In finite samples the three test statistics should follow the ordering: W > LR > LM. Failure of the estimated error models to meet this condition suggests that misspecification may invalidate the asymptotic properties of the maximum likelihood estimates (Anselin, 2005). Specific test statistics for W, LR, and LM will be presented with the results of each model.

---

[18] 981 meters is the minimum distance such that every observation in the data has at least one neighbor.
[19] The Wald test statistic is equal to the square of the asymptotic t-value (the z-value) of the parameter for the spatial weight matrix, while the LM-Error statistic is based on OLS residuals.

Results of the Continuous Distance Model

The continuous distance model includes the six continuous distance variables that measure the linear distance in meters from each property to the closest open space of each type.[20] In order to capture any effects that landscape pattern have on property values, the open space diversity indices also enter the continuous distance model. The continuous distance model takes the form:

$$\ln(P_i) = \beta_0 + \beta L_i + \sum\beta_k S_{ki} + \sum\beta_l N_{li} + \sum\beta_m DIVIND_{mi} + \sum\beta_n DIST\_OS_{ni} + \varepsilon_i$$

where $\ln(P_i)$ is the natural logarithm of the sale price of property i, L_i is a land characteristic (acreage) of property i, S_{ki} is a vector of structural characteristics of property i, N_{li} is a vector of neighborhood characteristics of property i, $DIVIND_{mi}$ is a vector of the measures of open space diversity surrounding property i, $DIST\_OS_{ni}$ is a vector of the measures of distance to the closest open space of each type from property i, and ε_i is the observation specific error term.

Of the five structures of W considered for the continuous distance model, only estimation with the 200 meter spatial weight matrix satisfied the inequality W > LR > LM, thereby suggesting that estimation with the 200 meter spatial weight matrix is the most appropriate specification.[21] The LM test statistics identified the spatial error model as the appropriate remedial measure for spatial dependence. Based on the 200 meter weight matrix, the Moran's I statistic is equal to 3.44 (p-value = 0.0006), and the LM-Error statistic is equal to 9.28 (p-value = 0.0023).

[20] An exception is DIST_GRM that measures the linear distance to Great Meadows, either the Concord or Sudbury division depending on which is closest.
[21] W = 9.94; LR = 9.53; LM = 9.28.

The results of the continuous distance model are presented in Table 6.2. The estimates were generated via maximum likelihood (ML) in a spatial error model using the weighted variables. As such, the ML estimates have unbiased standard errors and are appropriate for statistical inference. The marginal implicit price associated with each variable (calculated as the first partial derivative of the hedonic price function with respect to the variable of interest times the mean sale price) is also presented in Table 6.2. The OLS and WLS estimates are presented in the Appendix.

The land and structural characteristics have the expected signs. The neighborhood variables P_EDUC and P_AGE65 are positive indicating that the prices of homes in a neighborhood increase as the percentage of residents in that neighborhood with at least some college education increases, and the percentage of residents in that neighborhood over the age of 65 increases, respectfully. Additionally, the negative sign on the coefficient for DIST_TSTOP indicates that property values decrease with increased distance to commuter rail stations, while the positive sign for DIST_MJRD indicates that property values increase with increased distance to a major road. This suggests that the commuter rail is an amenity for residential location while major roads are not. It seems that residential location proximate to commuter rail stations has a premium associated with greater accessibility to locations such as employment and shopping centers. However, for residential location near major roads, the associated traffic or noise appears to offset any positive effects resulting from improved accessibility. Finally, distance to the closest commercial land use does not have a significant effect on the sale price of residential properties in the study area.[22]

[22] This result is likely a product of the fact that commercial land use is extremely heterogeneous, ranging from high-density traditional downtowns to low-density shopping malls.

Table 6.2: Continuous Distance Model.

	Maximum Likelihood Results	
	Parameter Estimate (Standard Error)	Marginal Implicit Price
CONSTANT	**10.68568***** (0.10321)	-
AGE	**-0.00219***** (0.00034)	**-470.99*****
INTSF	**0.00025***** (0.00001)	**54.20*****
ACRES	**0.05533***** (0.01219)	**11,875.68*****
PEDUC	**1.84397***** (0.11592)	**395,808.93*****
PAGE65	**1.25640***** (0.21671)	**269,686.87*****
DIST_TSTOP	**-0.00003***** (0.00001)	**-6.21*****
DIST_MJRD	**0.00007**** (0.00003)	**15.14****
DIST_COMRC	**0.00001** (0.00003)	**2.42**
DIVIND_SM	**-0.08089** (0.08605)	**-17,362.83**
DIVIND_LG	**-0.21806**** (0.09401)	**-46,807.31****
DIST_AG	**0.00003**** (0.00001)	**5.94****
DIST_CEM	**0.00003***** (0.00001)	**5.51*****
DIST_CONS	**-0.00006** (0.00005)	**-13.87**
DIST_GOLF	**-0.00003***** (0.00001)	**-6.23*****
DIST_SPRT	**-0.00006**** (0.00003)	**-12.99****
DIST_GRM	**-0.00004***** (0.00001)	**-7.91*****
RHO	**0.09376***** (0.02974)	-
** significant at 5%; *** significant at 1%		

For the environmental focus variables, the large diversity index (DIVIND_LG) is negative indicating that property owners prefer less diversity of open space in the 1,000 meter neighborhoods around their homes. Five of the six coefficients for the continuous distance variables are significant at explaining the price of homes in the study area at the five percent level. The coefficients for distance to agricultural land (DIST_AG) and distance to cemeteries (DIST_CEM) are positive. The positive signs on the estimates indicate that an increase in a home's distance to agricultural land (cemeteries) results in an increase in the sale price of the home. Therefore, people prefer a residential location with greater distance to agricultural land and cemeteries. The coefficients for distance to golf courses (DIST_GLF), distance to sports fields (DIST_SPRT), and distance to Great Meadows (DIST_GRM) have negative signs. The negative signs on these estimates indicate that property owners prefer a residential location proximate to golf courses, sports fields, and the Great Meadows NWR – an increase in the distance to any of these open space types results in a decrease in house price. The coefficient for distance to natural parks/conservation land (DIST_CONS) is not significantly different from zero at the five percent level.

The fact that residential property values decrease with distance to certain open space types while property values increase with distance to other open space types reveals that not all open space is created equal. Specifically, Great Meadows is valued more highly than agricultural land, cemeteries, and natural parks/conservation land. Although, regression results alone do not indicate if Great Meadows is valued differently than golf courses, or sports fields. To examine if the amenity generating effect of Great Meadows is different than that of the other two open space types, likelihood ratio (LR) tests were performed between the maximum value of each restricted likelihood (under H_0) and the maximum value of the unrestricted likelihood (under H_1) using the ML parameter estimates.[23] The LR test is essentially identical to the Wald and Lagrange multiplier tests between the restricted and unrestricted sum of squared errors; however, the LR test statistic is distributed as $\chi^2_{(J)}$, where J is the number of restrictions under the null hypothesis. The null and alternative hypotheses of the first test are:

H_0: $\beta_{DIST\_GRM} - \beta_{DIST\_GOLF} = 0$, vs.

H_1: $\beta_{DIST\_GRM} - \beta_{DIST\_GOLF} \neq 0$

Given the single restriction under the null hypothesis and a five percent significance level, the critical value is distributed as $\chi^2_{(1,0.05)}$ and equal to 3.84. Consequently, H_0 is rejected when $\lambda_{LR} > \chi^2_c$. The LR statistic is 0.62 and fails to reject the null hypothesis that Great Meadows is valued no differently than golf courses at the five percent level. The null and alternative hypotheses of the second test are:

H_0: $\beta_{DIST\_GRM} - \beta_{DIST\_SPRT} = 0$, vs.

H_1: $\beta_{DIST\_GRM} - \beta_{DIST\_SPRT} \neq 0$

The LR statistic is 0.70 and also fails to reject the null hypothesis that Great Meadows is valued no differently than sports fields at the five percent level.[24]

The first objective of this thesis was to investigate if residential property owners value National Wildlife Refuges. With respect to this first objective, results suggest that property owners do value a residential location proximate to NWRs. The price effect of

---

[23] Formally, the LR test is based on the statistic: $\lambda_{LR} = 2[L(H_1) - L(H_0)]$.

[24] A joint test was also conducted of the form H_0: $\beta_{DIST\_GRM} - \beta_{DIST\_GOLF} = 0$ and $\beta_{DIST\_GRM} - \beta_{DIST\_SPRT} = 0$, vs. H_1: at least one restriction does not hold. Given two restrictions, the critical value is 5.99 and the test statistic is 3.54. The test fails to reject the null hypothesis at the five percent level of significance.

NWRs on surrounding properties has not previously been estimated and this research suggests that NWRs are an amenity to residential location.

The second objective of this research was to investigate if NWRs are valued differently than conservation land, agricultural land, sports fields, golf courses, and cemeteries. Results suggest that that property owners value NWRs more highly than agricultural land, cemeteries, and natural parks/conservation land. However, this research also suggests that property owners value NWRs no differently than they value sports fields and golf courses.

The result that NWRs are an environmental amenity is supported by the literature. Lutzenhiser and Netusil (2001) found that a residential location proximate to a natural area park has an associated price premium. Given that NWRs are naturally vegetated, it is not surprising that they too have associated price premiums on property values. The results here are also consistent with those of Mahan et al. (2000), which suggest that a wetland is an amenity to residential location. Great Meadows is almost entirely wetland, so the consistency of these results is also not surprising.

Comparing the implicit prices found here, in a relative sense, with those of other studies, Mahan et al. (2000) found that property owners value wetlands greater than urban parks. Further, Lutzenhiser and Netusil (2001) found that natural area parks have the largest average positive effect on a home's sale price, over urban parks, specialty parks, golf courses, and cemeteries. In this thesis, Great Meadows was not found to have a significantly different effect on surrounding properties than urban parks/sports fields or golf courses. However, this discrepancy may be insubstantial considering that neither

Mahan et al. (2000) nor Lutzenhiser and Netusil (2001) actually tested if the price effects were significantly different from one another.[25]

Results of the Discrete Distance Model

The continuous measures of distance to the closest open space of each type were used to create discrete measures of distance. These alternate measures of proximity are intended to identify any discontinuities in the hedonic price function that are not distinguishable when solely relying on continuous, linear measures of distance. The discrete distance model takes the form:

$$\ln(P_i) = \beta_0 + \beta L_i + \sum\beta_k S_{ki} + \sum\beta_l N_{li} + \sum\beta_m DIVIND_{mi}$$

$$\sum\beta_n ZONE1_{ni} + \sum\beta_o ZONE2_{oi} + \sum\beta_p ZONE3_{pi} + \varepsilon_i$$

where $ZONE1 = 1$ if the closest open space of type z exists within 50 meters of property i, 0 otherwise, $ZONE2 = 1$ if the closest open space of type z exists between 51 and 100 meters of property i, 0 otherwise, and $ZONE3 = 1$ if the closest open space of type z exists between 101 and 1,000 meters of property i, 0 otherwise.

The 200 meter spatial weight matrix was the only structure of the five that satisfied the inequality W > LR > LM.[26] Based on the 200 meter spatial weight matrix, the LM test statistic identified spatial autocorrelation as the most apparent form of spatial dependence. For the 200 meter weight matrix the Moran's I statistic is equal to 4.50 (p-value < 0.0001), and the LM-Error statistic is equal to 15.72 (p-value < 0.0001).

---

[25] While Mahan et al. (2000) never tested the null hypothesis that wetland is valued no differently than urban parks, the marginal implicit price of proximity to wetland was more than 13 times greater than that of proximity to urban parks (assuming a reduction in distance of 1,000 feet and an initial distance of one mile).

[26] W = 17.30; LR = 16.19; LM = 15.72.

The results of the discrete distance model appear in Table 6.3. As with the continuous distance model, the estimates were generated with ML in a spatial error model using the weighted variables. The OLS and WLS estimates are presented in the Appendix. The land, structural, and neighborhood variables have similar signs, magnitudes, and significance levels as the continuous distance model. An exception is the coefficient for distance to the closest major road (DIST_MJRD), which is insignificant. In the continuous distance model DIST_MJRD was positive at the one percent level.

In the discrete distance model, neither of the diversity indices is significantly different from zero, and only two of the open space zones are significant at the ten percent level. Specifically, the coefficient for GOLF_100 is negative at the ten percent level. The coefficient suggests that a property in which the closest golf course is located between 51 and 100 meters sells for $150,873 less than if the closest golf course was located at a distance greater than 1,000 meters. Additionally, the coefficient for SPRT_1000 is positive at the one percent level. The coefficient suggests that a property in which the closest sport field is located between 101 and 1000 meters sells for $17,145 more than if the closest sport field was located at a distance greater than 1,000 meters.

Table 6.3: Discrete Distance Model.

	Maximum Likelihood Results	
	Parameter Estimate (Standard Error)	Marginal Implicit Price
CONSTANT	**10.19093***** (0.08353)	-
AGE	**-0.00210***** (0.00033)	**-451.84*****
INTSF	**0.00025***** (0.00001)	**52.97*****
ACRES	**0.05147***** (0.01177)	**11,047.99*****
PEDUC	**1.96190***** (0.10636)	**421,123.80*****
PAGE65	**1.60827***** (0.18733)	**345,215.91*****
DIST_TSTOP	**-0.00002***** (<0.00001)	**-4.95*****
DIST_MJRD	**0.00004** (0.00003)	**8.59**
DIST_COMRC	**0.00004** (0.00002)	**7.57**
DIVIND_SM	**0.14263** (0.18480)	**30,615.54**
DIVIND_LG	**-0.01764** (0.07924)	**-3,785.41**
AG_50	**-0.07997** (0.15117)	**-17,166.00**
AG_100	**0.00864** (0.09728)	**1,854.23**
AG_1000	**-0.00234** (0.02584)	**-502.27**
CEM_50	**0.02799** (0.30129)	**6,008.47**
CEM_100	**-0.15044** (0.17950)	**-32,291.35**
CEM_1000	**-0.01454** (0.02774)	**-3,121.79**
CONS_50	**-0.00290** (0.08369)	**-621.87**
CONS_100	**0.02990** (0.06840)	**6,418.05**
CONS_1000	**0.06730** (0.05526)	**14,446.98**
GOLF_50	**0.02782** (0.10407)	**5,971.92**
GOLF_100	**-0.70288*** (0.38264)	**-150,873.34***
GOLF_1000	**-0.00452** (0.03013)	**-970.85**
SPRT_50	**0.06387** (0.07848)	**13,709.15**
SPRT_100	**0.01237** (0.07559)	**2,654.51**
SPRT_1000	**0.07988***** (0.03097)	**17,145.34*****
GRM_50	**0.08450** (0.14335)	**18,138.24**
GRM_100	**-0.04855** (0.09958)	**-10,420.63**
GRM_1000	**0.03370** (0.05062)	**7,233.25**
RHO	**0.12199***** (0.02933)	-
* significant at 10%; ** significant at 5%; *** significant at 1%		

The lack of significant coefficients in this model might seem surprising at first given that five of the coefficients for open space type were significantly different from zero in the continuous distance model. However, examining frequencies of the number of observations in each open space zone provides evidence of a weakness of this model for use in the current study area.[27] As indicated in Table 6.4, very few properties have specific types of open space in certain zones. For instance, there are only three properties in the study area that have the closest cemetery located within 50 meters. Similarly, there are only five properties that have the closest cemetery located between 51 and 100 meters. Notice that for GOLF_100, which is significant at the ten percent level, there are only five properties that have the closest golf course located between 51 and 100 meters. This is a very small number of observations for such a large parameter estimate for this variable. Given the frequency of this variable and others, the results of the discrete distance model should be used with caution.

Table 6.4: Frequency of Sales per Open Space Zone.

Zone	Properties with OS in Zone	Without OS in Zone
AG_50	21	1,576
AG_100	15	1,582
AG_1000	416	1,181
CEM_50	3	1,594
CEM_100	5	1,592
CEM_1000	261	1,336
CONS_50	128	1,469
CONS_100	112	1,485
CONS_1000	1,307	290
GOLF_50	10	1,587
GOLF_100	5	1,592
GOLF_1000	252	1,345
SPRT_50	45	1,552
SPRT_100	29	1,538
SPRT_1000	1,248	349
GRM_50	5	1,592
GRM_100	10	1,587
GRM_1000	124	1,473

[27] Alternative specifications of zones, such as those in Table 3.1, were also created and estimated in the hedonic model. However, none of the specifications drastically improved the estimation results.

The third objective of this research was to investigate if the hedonic price function exhibits discontinuities with respect to distance to each open space type using discrete measures of distance. Given the nature of the variables in the discrete distance model, this investigation is inconclusive. Again, the spatial distribution of properties in the study area is a limiting factor for estimation of the discrete distance model, and the results cannot be used to verify the price-distance relationships identified by the continuous distance model. This situation is not likely to hold for other study areas, depending on the spatial location of property sales relative to the environmental variable(s). For instance, Tyrvainen and Miettinen (2000) estimated hedonic price functions using both continuous and discrete measures of distance to forested areas and found consistent price effects across the alternate measures of proximity. However, the researchers used a different delineation of open space zones than those used here (see Table 3.1).

Both the continuous and discrete distance models included as environmental variables the 100 meter and 1,000 meter diversity indices. The fifth objective of this research was to investigate if property owners value open space diversity. In the continuous distance model the coefficient for the 100 meter diversity index was insignificant while the coefficient for the 1,000 meter diversity index was negative, suggesting that property owners prefer less open space diversity in the 1,000 meter neighborhoods around their homes. For the discrete distance model neither open space diversity coefficient was significantly different from zero. These results suggest that open space diversity in the 100 meter neighborhoods around homes is not important in explaining residential property values. However, the effect of open space diversity on property values in the 1,000 meter neighborhoods around homes is uncertain – the negative effect of open space diversity is not consistent across the alternate measures of distance.

While this is the first known application of an open space diversity index in a hedonic model, at least two studies have estimated values for land use diversity. Geoghegan et al. (1997) found that property values increase with increased land use diversity at the 1,000 meter scale, while Acharya and Bennett (2001) found that property values decrease with increased land use diversity at the 0.25 and 1.0 mile scale. These mixed results suggest that measures of spatial pattern may be highly sensitive to the unique landscape characteristics of each study area.

Results of the Public/Private Continuous Distance Model

In order to examine if the public or private accessibility of open space affects the price of surrounding properties, variables were created that represent the continuous distance from each property to the closest publicly and closest privately accessible open space. The public/private continuous distance model takes the following form:

$$\ln(P_i) = \beta_0 + \beta L_i + \sum\beta_k S_{ki} + \sum\beta_l N_{li} + \beta DIST\_PUBOS_i + \beta DIST\_PRIOS_i + \varepsilon_i$$

where $DIST\_PUBOS_i$ is a measure of distance to the closest publicly accessible open space from property i, and $DIST\_PRIOS_i$ is a measure of distance to the closest privately accessible open space from property i.

The 200 meter spatial weight matrix was the only structure of the five considered that satisfied the inequality W > LR > LM, thereby suggesting that estimation with the 200 meter W is the most appropriate specification.[28] Further, the LM test statistics identified the spatial error model as the appropriate remedial measure for spatial

[28] W = 14.36; LR = 13.79; LM = 13.63.

dependence. For the 200 meter weight matrix the Moran's I statistic is equal to 3.94 (p-value < 0.0001), and the LM-Error statistic is equal to 13.63 (p-value = 0.0002).

The ML results and corresponding marginal implicit prices are presented in Table 6.5. The OLS and WLS estimates are presented in the Appendix. The land, structural, and neighborhood variables have the appropriate signs and are of similar magnitudes and significance levels to those of previous models. For the environmental variables, the coefficient for distance to the nearest publicly accessible open space (DIST_PUBOS) is insignificantly different from zero. However, the coefficient for distance to the nearest privately accessible open space (DIST_PRIOS) is negative at the five percent level indicating that as the distance to privately accessible open space from a home decreases, the price of the home increases.

With respect to the fourth objective of this thesis, which sought to investigate if privately and publicly accessible open space is valued by residential property owners, it seems the answer is yes. The marginal implicit price of residential location relative to privately accessible open space indicates that for each meter closer a home is located to privately accessible open space the price of the home increases by $7.42. While the marginal implicit price of residential location relative to publicly accessible open space is of a larger magnitude, it is not significantly different from zero.

Table 6.5: Public/Private Continuous Distance Model.

	Maximum Likelihood Results	
	Parameter Estimate (Standard Error)	Marginal Implicit Price
CONSTANT	**10.54308*** ** (0.088842)	-
AGE	**-0.00231*** ** (0.00034)	**-495.671*** **
INTSF	**0.00025*** ** (0.00001)	**54.39*** **
ACRES	**0.05325*** ** (0.01198)	**11,430.79*** **
PEDUC	**1.75121*** ** (0.11942)	**375,899.40*** **
PAGE65	**1.57931*** ** (0.18124)	**339,001.30*** **
DIST_TSTOP	**-0.00002*** ** (<0.00001)	**-5.20*** **
DIST_MJRD	**0.00003** (0.00003)	**6.60**
DIST_COMRC	**<0.00001** (0.00002)	**0.98**
DIST_PUBOS	**-0.00006** (0.00005)	**-12.88**
DIST_PRIOS	**-0.00003** ** (0.00001)	**-7.42** **
RHO	**0.11187*** ** (0.02949)	-
** significant at 5%; *** significant at 1%		

In a related study, Bolitzer and Netusil (2000) found that a residential location within 1,500 feet of a public park has an associated price premium, while location within 1,500 feet of a private park has no significant effect on property values. The results of Bolitzer and Netusil (2000) are inconsistent with the results found here. However, the authors incorporated public vs. private open space based on ownership, whereas this research uses public vs. private accessibility. Additionally, Bolitzer and Netusil (2000) used measures of proximity that were bounded at 1,500 feet (457 meters), while the measures used in this thesis are not bounded. In the current study the mean distance to the closest publicly accessible open space is 284 meters while the mean distance to the closest privately accessible open space is 1,233 meters (see Table 6.1). Further, the variables used by Bolitzer and Netusil (2000) excluded cemeteries and golf courses, whereas the variables used here include all open space types. This difference is likely driving the apparent discrepancy. Consider the results of the continuous distance model presented earlier. Golf courses had a positive effect (negative distance coefficient) on property values and cemeteries had a negative effect (positive distance coefficient) on property values. As depicted in Table 6.6, almost all of the golf courses in the study area are privately accessible, whereas all of the cemeteries are publicly accessible open space. Differences aside, both studies suggest that the characteristics that make one tract of open space public while another is private are important in explaining valuation.

Table 6.6: Distribution of Properties by Type of Nearest Public/Private Open Space.

	AG	CEM	CONS	GOLF	SPRT	GRM
Publicly Accessible	2.82%	1.19%	63.31%	0.94%	30.43%	1.31%
Privately Accessible	5.01%	0.00%	85.22%	4.76%	5.01%	0.00%

Open Space Categorization Results

It was also an objective of this thesis to investigate if the results produced by the interactive method of categorizing open space could be reproduced using an automatic method of categorizing open space at a fraction of the cost of time. Table 6.7 reveals the similarities and discrepancies between the two open space categorization methods.[29] The principal diagonal reflects the number of open spaces within each type that were assigned the same code in both methods. Reading across any given row identifies the number of open spaces that were assigned that particular open space type with the automatic method, and assigned that type, or one of the other types, with the interactive method. Similarly, reading down any given column identifies the number of open spaces that were assigned that particular open space type with the interactive method, and assigned that type, or one of the other types, with the automatic method. For example, there were nine open spaces that were labeled cemetery with both methods, one open space that was categorized as a cemetery with the automatic method, but was classified as an urban park/sports facility with the interactive method, and three open spaces that were categorized as a cemetery with the automatic method, but were classified as a natural park/conservation land with the interactive method. Overall, the alternative methods produced relatively similar open space data sets.

[29] For this comparative analysis Great Meadows is included within natural parks/conservation land because there were no discrepancies between alternative coding methods for NWR lands. Additionally, non-applicable open space is labeled as NA.

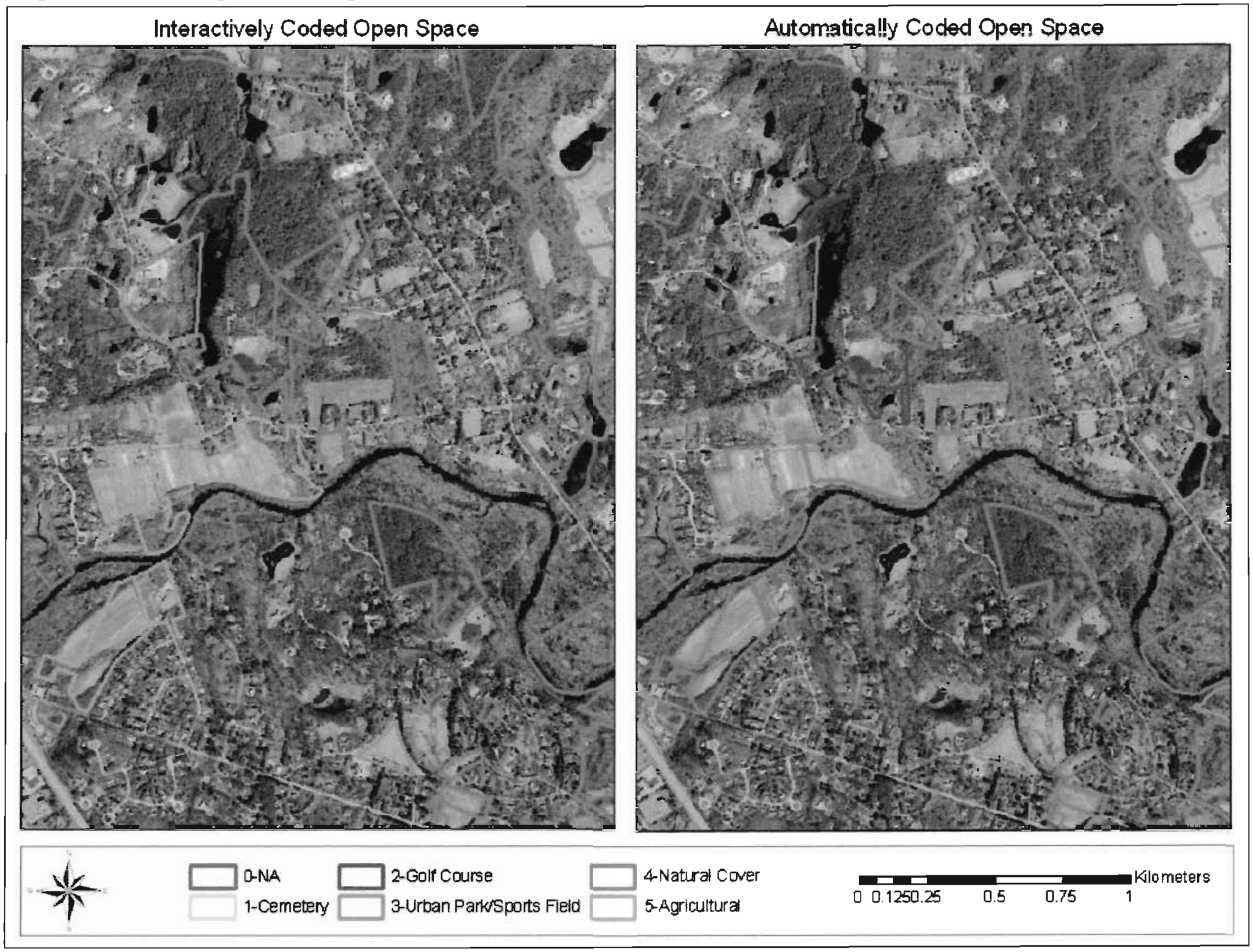
Table 6.7: Comparison of Categorization Methods.

		Interactively Coded Open Space						
		CEM	**GOLF**	**SPRT**	**CONS**	**AG**	**NA**	**% Correct**
Automatic. Coded	**CEM**	9	0	1	3	0	0	69%
	GOLF	0	6	4	5	1	3	32%
	SPRT	0	0	46	32	6	8	50%
	CONS	0	0	2	377	16	2	95%
	AG	0	0	0	6	10	1	59%
	NA	0	0	0	4	1	23	82%
	% Correct	100%	100%	87%	88%	29%	62%	

Nevertheless, there are noticeable differences between the alternative methods in the final data sets. For instance, Figure 6.2 depicts the differences between the alternative methods for a number of open spaces in an area in the town of Concord. Under the interactive method, in which aerial photographs were used to aid in interpretation and categorization, the assignment of open space type is very much in accord with the land cover and observable land use. Fields showing visible signs of cultivation are categorized as 'agricultural', open space that appears to be naturally vegetated as forest or scrub/shrub are categorized as 'natural cover', and open water is categorized as 'non-applicable'. Compare the same open space under the automatically coded data. Agricultural lands have been categorized as 'natural cover', an obviously naturally vegetated open space in the center of the map has been categorized as 'golf course', and a body of water has been categorized as an 'urban park/sports field'.[30]

[30] Some of this error in the automatic categorization method may have been avoided by using digital image processing techniques to perform a land cover classification of the aerial photos.

Figure 6.2: Example of Categorization Error.



The differences depicted in Figure 6.2 are the result of the heterogeneity present in the fields in the original open space data acquired from the Massachusetts Office of Geographic and Environmental Information. For instance, the agricultural open space in Figure 6.2 (in the Interactively Coded Open Space map) was assigned the primary purpose 'conservation', and level of protection 'protected'. However, these classifications were used almost exclusively for open space that is conservation land, while the majority of other agricultural open space in the data was assigned the primary purpose of 'agriculture'. Such discrepancies resulted from the plethora of volunteer organizations that participated with the interpretation and collection of the open space data. Therefore, upon development of the appropriate code for iteratively selecting individual open spaces, it was impossible to avoid the heterogeneity in the original GIS data in order to create delineations that would identically mirror that of the interactive process. It is important to note however, that the stark differences between the results of the alternative coding methods displayed in Figure 6.2 are not consistent throughout the study area. Nonetheless, the alternative methods of categorizing open space produced non-negligible differences in estimation results.

In order to compare the alternative categorization methods directly, two models were estimated, each based on continuous distance variables from a different set of open space data. Results based on each open space categorization method appear in Table 6.8. The results for the interactively categorized data are identical to those presented in Table 6.2 and were derived using the weighted data via maximum likelihood. The results for the automatically categorized data were also derived using transformed data via ML.[31]

---

[31] The error term of the automatically categorized data is a function of AGE, ACRES, and DIST_CONS.

Table 6.8: Results of the Two Open Space Categorization Methods.

	Interactively Categorized		Automatically Categorized	
	Parameter Estimate (Standard Error)	Marginal Implicit Price	Parameter Estimate (Standard Error)	Marginal Implicit Price
CONSTANT	**10.68568***** (0.10321)	-	**10.78404***** (0.11746)	-
AGE	**-0.00219***** (0.00034)	**-470.99*****	**-0.00204***** (0.00033)	**-437.66*****
INTSF	**0.00025***** (0.00001)	**54.20*****	**0.00026***** (0.00001)	**54.91*****
ACRES	**0.05533***** (0.01219)	**11,875.68*****	**0.05648***** (0.01215)	**12,123.78*****
PEDUC	**1.84397***** (0.11592)	**395,808.93*****	**1.61134***** (0.12353)	**345,875.96*****
PAGE65	**1.25640***** (0.21671)	**269,686.87*****	**1.35515***** (0.20197)	**290,884.09*****
DIST_TSTOP	**-0.00003***** (0.00001)	**-6.21*****	**-0.00003***** (0.00001)	**-5.93*****
DIST_MJRD	**0.00007**** (0.00003)	**15.14****	**0.00007**** (0.00003)	**14.46****
DIST_COMRC	**0.00001** (0.00003)	**2.42**	**0.00002** (0.00003)	**3.56**
DIVIND_SM	**-0.08089** (0.08605)	**-17,362.83**	**-0.05291** (0.08178)	**-11,357.28**
DIVIND_LG	**-0.21806**** (0.09401)	**-46,807.31****	**-0.16671*** (0.08900)	**-35,784.32***
DIST_AG	**0.00003**** (0.00001)	**5.94****	**0.00005***** (0.00001)	**9.85*****
DIST_CEM	**0.00003***** (0.00001)	**5.51*****	**0.00001** (0.00001)	**2.15**
DIST_CONS	**-0.00006** (0.00005)	**-13.87**	**-0.00007** (0.00005)	**-15.33**
DIST_GOLF	**-0.00003***** (0.00001)	**-6.23*****	**-0.00003**** (0.00001)	**-6.43****
DIST_SPRT	**-0.00006**** (0.00003)	**-12.99****	**-0.00004** (0.00003)	**-7.68**
DIST_GRM	**-0.00004***** (0.00001)	**-7.91*****	**-0.00005***** (0.00001)	**-10.05*****
RHO	**0.09376***** (0.02974)	-	**0.10947***** (0.02952)	-
* significant at 10%; ** significant at 5%; *** significant at 1%				

Comparing the results for the environmental focus variables, DIVIND_LG is negative at the five percent level with the interactively categorized data and negative at the ten percent level with the automatically categorized data. Turning to the continuous distance variables, with the interactive data DIST_AG is positive at the five percent level, while it is positive at the one percent level with the automatically coded data. DIST_CEM is positive at the one percent level with the interactively coded data, but not significantly different from zero with the automatically coded data. Across both data sets, DIST_CONS is not significantly different from zero. DIST_GOLF is negative at the one percent level with the interactive data and negative at the five percent level with the automatic data. With the interactive data DIST_SPRT is negative at the five percent level, but not significantly different from zero with the automatic data. Finally, DIST_GRM is negative at the one percent level with both the interactive and automatic data.

With respect to the second objective of this thesis, which sought to determine if property owners value NWRs differently than other types of open space, the differences between the two categorization methods are less severe than may first seem. Recall that there is no significant difference between the value property owners place on Great Meadows vs. golf courses vs. sports fields using the interactive data. With the automatic data Great Meadows is valued no differently than golf courses, but differently than the other open space types.[32] Therefore, the two sets of data produce similar, but not identical results. Further, if the estimates generated by this research are to guide public policy, then using one categorization method over the other will suggest targeting the

[32] A likelihood ratio test of the form H_0: $\beta_{DIST\_GRM} - \beta_{DIST\_GOLF} = 0$, vs. H_1: $\beta_{DIST\_GRM} - \beta_{DIST\_GOLF} \neq 0$ was also conducted with the automatic data. The critical value is 3.84 and the test statistic is 1.62. The test fails to reject the null hypothesis at the five percent level of significance.

preservation of different open space types in order to maximize the benefits of preserved land.

This comparative analysis was conducted to determine if automatically categorized open space would produce similar estimates to that of the interactively categorized open space, at a fraction of the cost of time. Results suggest that the answer is no, and that the type of open space GIS data matters. That is, an off-the-shelf analysis with secondary GIS data is likely to tell a different story than an analysis in which the data has been ground-truthed and referenced against other sources.

This research derived valuation estimates for the open space benefits accruing to residential property owners using a number of different open space variables. While not all of the results presented in this thesis are highly applicable to public policy or research methodology, this research has important implications.

Chapter 7

CONCLUSIONS

The results of this research have important implications, both in terms of policy and methodology. First of all, this research has revealed that NWRs are valued by residential property owners. Therefore, when local community leaders argue that NWRs reduce tax revenue, it is not necessarily true. In fact, as this research reveals, the presence of NWRs increases the value of nearby residential properties, which can actually lead to increased property tax revenues. Additionally, NWRs make payments to communities in lieu of taxes. This result will help the USFWS prioritize future land acquisitions for the NWR system. For instance, the benefits of preserving land as NWR can be maximized by acquiring land where the positive spillover effects of NWRs will be capitalized into surrounding residential properties.

This research also suggests that NWRs are valued more highly than agricultural land, cemeteries, and conservation land, but are not valued significantly different than sports fields and golf courses. This finding will allow planners and local officials to arrive at more informed decisions when evaluating the municipal commitment of land into preserved open space. Not only will the marginal amenity and disamenity values estimated by this research allow communities to target the preservation of certain types of open space, the values will also aid communities in determining if the benefits of preserving specific open space types outweigh the costs. In other words, if the municipal decision to preserve certain types of open space is made correctly, open space preservation can generate positive spillover effects on surrounding properties, thereby partially offsetting the costs of land acquisition or conversion.

In terms of methodological implications of this research, the results of both the continuous distance model and the public and private open space model suggest that people perceive the unique characteristics that make one type of open space different from other types of open space. Therefore, it is inappropriate to aggregate open space into single measures in order to estimate marginal values. A researcher can only obtain accurate and unbiased valuation estimates when a heterogeneous environmental variable is disaggregated into its component characteristics in a manner that is perceivable to individual agents in the market.

This research also showed that automatically categorized open space GIS data failed to replicate the results of data that had been meticulously categorized by the researcher. This result suggests that researchers cannot simply use secondary GIS data without question. Too often GIS users utilize data available on the Internet or from some other secondary source without taking the time to examine the quality or completeness of the data. As this research has revealed, failure to do so can result in biased estimates.

Limitations

Many of the results and conclusions in this thesis are based on measures of linear distance to the closest activity or service. Admittedly, distance to the closest open space of each type is a rather basic measure of relation between two activities. There are at least six other measures of proximity that geographers have proposed and that take into account various measures of distance and/or counts of the service or destination in some specified zone (Church and Marston, 2003).[33] However, it is measures of proximity to

[33] The other major measures of accessibility include: 1) counting; 2) total sums of distances; 3) gross interaction potential; 4) probabilistic choice; 5) net and maximum benefit; and 6) absolute.

the closest service or location that have been used as proxies for the effectiveness of the provision of public services (Hodgart, 1978). Indeed, if the goal of this research and related research is to inform public policy and to ultimately enhance local economies and environments, the measures used must be easily interpretable and readily applicable to land use policy. Therefore, the measures of proximity used in this research may be the most applicable to the issues that planners and local officials encounter on a routine basis. Further, in order to account for more complex spatial relationships between residential properties and open space, such as landscape pattern, the open space diversity indices were included in models where marginal values for open space [land use] type were estimated.

However, it is also important to note that the open space variables used in this research do not include any indicators of quality (besides land use type). Measuring proximity to the nearest open space of each type assumes that property owners are indifferent between additional characteristics that define the quality of different tracts of open space of a given type. As such, the estimates derived here could be subject to omitted variable bias.

Additionally, the marginal values estimated here are based only on the benefits property owners incur as a result of their residential location. Therefore, the benefits derived from these estimates are lower limits to the total value of a particular type of open space in the study area. Excluded from these estimates are the recreational benefits experienced by non-home owners and tourists visiting the area. Also excluded from the estimates are nonuse values such as existence values and bequest values.

It should also be mentioned that the marginal implicit prices estimated here are specific to the towns included in the study. Benefit transfers of the results should only be used if the characteristics of the policy site are very similar to the study area used in this research. This includes both the housing market and the presences of and relative quantities of open space types.

This research has statistical limitations as well. Realistically, the property sales used in this analysis are a sample of convenience and do not comprise a random sample, which is the case for all hedonic studies. As such, the results may be biased if the property sales included in the analysis vary systematically from those of the population. For instance, properties with certain unobservable characteristics may be less likely to be placed on the market, or less likely to sell once placed on the market, such that sale prices for these properties are observed less frequently than represented in the population. In this case, the properties are systematically not included in the hedonic data, which may induce bias in the parameter estimates.

Future Research

This thesis has presented the result that NWRs are an environmental amenity to residential location. Future research should test this result with other NWRs. Use of the hedonic method to examine the price effects of other NWRs will require that the study area be selected such that there are a sufficient number of residential property sales adjacent to the refuge(s).

Related future research could also expand on the findings of this study by incorporating alternate measures of open space in the hedonic equation. For instance, instead of distance to surrounding open space, residents may also (or instead) value the amount of open space surrounding their homes. This inquiry could be examined by using open space variables that measure the acreage of each open space type within a given neighborhood around each home (or alternatively the percentage of land that is in each of the six open space types). Comparing the results to those presented here would provide further understanding as to how property owners perceive and value open space. Further, it would be interesting to measure the effect of privately and publicly accessible open space based on acreage, rather than distance, in order to compare to the estimates found by Geoghegan (2002) when she examined developable (privately owned) and permanent (publicly owned) open space.

Future research could also apply the specification of open space zones proposed here to another study area in which there is a more appropriate spatial distribution of properties. This would allow for a more conclusive analysis of the suitability of the discrete distance zones for examining the effects of open space on property values.

REFERENCES

Allen, Ryan, Barry Bluestone, Bonnie Heudorfer, and Gretchen Weismann. 2002. *The Greater Boston Housing Report Card.* The Boston Foundation.

Anselin, Luc. 2005. *Exploring Spatial Data with GeoDa: A Workbook.* Spatial Analysis Laboratory, Department of Geography: University of Illinois, Urbana-Champaign; Center for Spatially Integrated Social Science.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer Academic Publishers.

Anselin, Luc, and S. Rey. 1991. "Properties of Tests for Spatial Dependence in Linear Regression Models." *Geographical Analysis*, 23: 112-31.

Acharya, Gayatri, and Lynne L. Bennett. 2001. "Valuing Open Space and Land-Use Patterns in Urban Watersheds." *Journal of Real Estate Finance and Economics*, 22(2/3): 221-37.

Beasley, Steven D., William G. Workman, and Nancy A. Williams. 1986. "Estimating Amenity Values of Urban Fringe Farmland: A Contingent Valuation Approach: Note." *Growth and Change*, 17(4): 70-78.

Bell, Kathleen P., and Nancy E. Bockstael. 2000. "Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Microlevel Data." *The Review of Economics and Statistics*, 82(1): 72-82.

Beron, K. J., Y. Hanson, J. Murdoch, and M. Thayer. 2002. "Hedonic Price Functions and Spatial Dependence: Implications for the Demand for Urban Air Quality." In L. Anselin and R. Florax, *Advances in Spatial Econometrics.* Berlin: Springer-Verlag.

Bowes, David, and Keith Ihlanfeldt. 2001. "Identifying the Impacts of Rail Transit Stations on Residential Property Values." *Journal of Urban Economics*, 50: 1-25.

Breffle, William S., Edward R. Morey, and Tymon S. Lodder. 1998. "Using Contingent Valuation to Estimate a Neighborhood's Willingness to Pay to Preserve Undeveloped Land." *Urban Studies*, 35(4):715-27.

Brown, James N., and Harvey S. Rosen. 1982. "On the Estimation of Structural Hedonic Price Models." *Econometrica*, 50(3): 765-68.

Can, Ayse. 1990. "The Measurement of Neighborhood Dynamics in Urban House Prices." *Economic Geography*, 66: 254-72.

Cassel, Eric, and Robert Mendelsohn. 1985. "The Choice of Functional Forms for Hedonic Price Equations: Comment." *Journal of Urban Economics*, 18: 135-42.

Cheshire, Paul, and Stephen Sheppard. 1995. "On the Price of Land and the Value of Amenities." *Economica,* 62(246), 247-67.

Church, Richard L., and James R. Marston. 2003. "Measuring Accessibility for People with a Disability." *Geographical Analysis*, 35(1): 83-96.

Cliff, A., and J. K. Ord. 1973. *Spatial Autocorrelation*. London: Pion Publishing.

Correll, Mark R., Jane H. Lillydahl, and Larry D. Singell. 1978. "The Effects of Greenbelts on Residential Property Values: Some Findings on the Political Economy of Open Space." *Land Economics*, 54(2): 207-17.

Cropper, Maureen, Leland Deck, and Kenneth McConnell. 1988. "On the Choice of Functional Form for Hedonic Price Functions." *The Review of Economics and Statistics*, 70(4): 668-75.

Do, A. Quang, and Gary Grudnitski. 1995. "Golf Courses and Residential House Prices: An Empirical Investigation." *Journal of Real Estate Finance and Economics*, 10: 261-70.

Doss, Cheryl R., and Steven J. Taff. 1996. "The Influence of Wetland Type and Wetland Proximity on Residential Property Values." *Journal of Agricultural and Resource Economics*, 21(1): 120-29.

Dubin, Robin A. 1988. "Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms." *The Review of Economics and Statistics*, 70: 466-77.

Espey, Molly, and Kwame Owusu-Edusei. 2001. "Neighborhood Parks and Residential Property Values in Greenville, South Carolina." *Journal of Agricultural and Applied Economics*, 33(3): 487-92.

Freeman, A. Myrick, III. 1974. "On Estimating Air Pollution Control Benefits from Land Value Studies." *Journal of Environmental Economics and Management*, 1(1): 74-83.

Freeman, A. Myrick, III. 2003. *The Measurement of Environmental and Resource Values*. Washington D.C.: Resources for the Future.

Garrod, Guy, and Ken Willis. 1992. "The Environmental Economic Impact of Woodland: A Two Stage Hedonic Price Model of the Amenity Value of Forestry in Britain." *Applied Economics*, 24: 715-28.

Geoghegan, Jacqueline. 2002. "The Value of Open Spaces in Residential Land Use." *Land Use Policy*, 19: 91-98.

Geoghegan, Jacqueline, Lisa A. Wainger, and Nancy E. Bockstael. 1997. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS." *Ecological Economics,* 23: 251-64.

Geoghegan, Jacqueline, Lori Lynch, and Shawn Bucholtz. 2003. "Capitalization of Open Spaces into Housing Values and the Residential Property Tax Revenue Impacts of Agricultural Easement Programs." *Agricultural and Resource Economics Review*, 32(1): 33-45.

Gittell, Ross J., and Patricia M. Flynn. 1995. *"The Lowell High-Tech Success Story: What Went Wrong?"* *New England Economic Review*, March: 57-70.

Glejser, H. 1969. "A New Test for Heteroskedasticity." *Journal of the American Statistical Association*, 64: 316-23.

Graves, Phil, James C. Murdoch, Mark A. Thayer, and Don Waldman. 1988. "The Robustness of Hedonic Price Estimation: Urban Air Quality." *Land Economics*, 64(3): 220-33.

Griliches, Zvi. 1961. "Hedonic Prices for Automobiles: An Econometric Analysis of Quality Change." In *The Price Statistics of the Federal Government.* General Series No. 73. New York: Columbia University Press for the National Bureau of Economic Research.

Halvorsen, Robert, and Henry O. Pollakowski. 1981. "Choice of Functional Form for Hedonic Price Equations." *Journal of Urban Economics*, 10(1): 37-49.

Heckscher, August. 1977. *Open spaces: The Life of American Cities.* New York: Harper & Row.

Hodgart, R. L. 1978. "Optimizing Access to Public Services." *Progress in Human Geography*, 2: 17-48.

Howard, Ebenezer. 1902. *Garden Cities of Tomorrow.* London: Swann Sonnenschein.

Irwin, Elena G. 2002. "The Effects of Open Space on Residential Property Values." *Land Economics,* 78(4): 465-80.

Johnston, Robert J. 1998. *Aquidneck Island and Open Space: An Economic Perspective.* Technical Report: The Aquidneck Island Partnership; Coastal Resources Center, University of Rhode Island.

Johnston, Robert J., James J. Opaluch, Thomas A. Grigalunas, and Marisa J. Mazzotta. 2001. "Estimating Amenity Benefits of Coastal Farmland." *Growth and Change*, 32: 305-25.

Kim, Chong W., 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(1): 24-39.

Kluge, Kevin D., and Gregory K. White. 1999. *Economic Benefits of Wetlands: A Hedonic Study in Windham, Maine*. Miscellaneous Report 409: Maine Agricultural and Forest Experiment Station, University of Maine.

Knetsch, Jack L. 1962. "Land Values and Parks in Urban Fringe Areas." *Journal of Farm Economics*, 44(5): 1718-26.

Lupi, Frank Jr., Theodore Graham-Tomasi, and Steven J. Taff. 1991. *A Hedonic Approach to Urban Wetland Valuation*. Staff Paper P91-8. Institute for Agriculture, Forestry, and Home Economics: Department of Agricultural and Applied Economics. University of Minnesota.

Luttik, Joke. 2000. "The Value of Trees, Water, and Open Space as Reflected by House Prices in the Netherlands." *Landscape and Urban Planning*, 48:161-67.

Lutzenhisher, Margot, and Noelwah R. Netusil. 2001. "The Effect of Open Space on a Home's Sales Price." *Contemporary Economic Policy,* 19(3): 291-98.

Mahan, Brent L., Stephen Polasky, and Richard M. Adams. 2000. "Valuing Urban Wetlands: A Property Price Approach." *Land Economics*, 76(1): 100-13.

Michael, Holly J., Kevin J. Boyle, and Roy Bouchard. 2000. "Does the Measurement of Environmental Quality Affect Implicit Prices Estimated from Hedonic Models?" *Land Economics*, 76(2): 283-98.

More, Thomas A., Thomas Stevens, and P. Geoffrey Allen. 1988. "Valuation of Urban Parks." *Landscape and Urban Planning*, 15: 139-52.

Nelson, Arthur C., John Genereux, and Michelle Genereux. 1992. "The Price Effects of Landfills on House Values." *Land Economics*, 68(4): 359-65.

Netusil, Noelwah R. 2005. "The Effect of Environmental Zoning and Amenities on Property Values: Portland, Oregon." Land Economics, 81(2): 227-46.

O'Neil, R. V., J. R. Krummel, R. H. Gardner, G. Sugihara, B. Jackson, D. L. DeAngelis, B. T. Milne, M. G. Turner, B. Zygmunt, S. W. Christensen, V. H. Dale, and R. L. Graham. 1988. "Indices of Landscape Pattern." *Landscape Ecology*, 1(3): 153-62.

Palmquist, Raymond B. 1984. "Estimating the Demand for Characteristics of Housing. *Review of Economics and Statistics*, 64(3): 394-404.

Palmquist, Raymond B. 1991. "Hedonic Methods." In *Measuring the Demand for Environmental Quality*. J. B. Braden and C. D. Kolstad (Eds.). Amsterdam: North-Holland Publishers.

Powe, N. A., G. D. Garrod, C. F. Brunsdon, and K. G. Willis. 1997. "Using a Geographic Information System to Estimate an Hedonic Price Model of the Benefits of Woodland Access." *Forestry*, 70(2): 139-49.

Ridker, Ronald G., and John A. Henning. 1967. "The Determinants of Residential Property Values with Special Reference to Air Pollution." *Review of Economics and Statistics*, 49(2): 246-57.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Perfect Competition." *Journal of Political Economy*, 82(1): 34-55.

Shannon, C. E., and W. Weaver. 1962. *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Shultz, Steven D., and David A. King. 2001. "The Use of Census Data for Hedonic Price Estimates of Open-Space Amenities and Land Use." *Journal of Real Estate Finance and Economics*, 22(2/3): 239-52.

Smith, V. Kerry, C. Poulos, and H. Kim. 2002. "Treating Open Space as an Urban Amenity." *Resource and Energy Economics*, 24: 107-29.

Thorsnes, Paul. 2002. "The Value of a Suburban Forest Preserve: Estimates from Sales of Vacant Residential Building Lots." *Land Economics*, 78(3): 426-41.

Turner, Monica G. 1990. "Spatial and Temporal Analysis of Landscape Patterns." *Landscape Ecology*, 4(1): 21-30.

Tyrvainen, Liisa. 1997. "The Amenity Value of the Urban Forest: An Application of the Hedonic Pricing Method." Landscape and Urban Planning, 37: 211-22.

Tyrvainen, Liisa, and Antti Miettinen. 2000. "Property Prices and Urban Forest Amenities." *Journal of Environmental Economics and Management*, 39: 205-23.

U.S. Fish and Wildlife Service. 2000. *Eastern Massachusetts National Wildlife Refuge Complex* (brochure). Department of the Interior.

Waugh, F. V. 1928. "Quality Factors Influencing Vegetable Prices." *Journal of Farm Economics*, 10(2): 185-96.

Weicher, John C., and Robert H. Zerbst. 1973. "The Externalities of Neighborhood Parks: An Empirical Investigation." *Land Economics*, 49: 99-105.

APPENDIX

OLS and WLS Results

Table A.1: Continuous Distance Model - OLS & WLS Results.

	OLS	WLS
	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
CONSTANT	**10.60106***** (0.09756)	**10.67775***** (0.09695)
AGE	**-0.00157***** (0.00028)	**-0.00215***** (0.00034)
INTSF	**0.00027***** (0.00001)	**0.00026***** (0.00001)
ACRES	**0.04682***** (0.00893)	**0.05658***** (0.01215)
PEDUC	**1.88849***** (0.10955)	**1.83639***** (0.10872)
PAGE65	**1.28424***** (0.20711)	**1.25583***** (0.20245)
DIST_TSTOP	**-0.00003***** (0.00001)	**-0.00003***** (0.00001)
DIST_MJRD	**0.00009***** (0.00003)	**0.00007**** (0.00003)
DIST_COMRC	**0.000005** (0.00003)	**0.00001** (0.00003)
DIVIND_SM	**-0.05786** (0.08568)	**-0.08569** (0.08305)
DIVIND_LG	**-0.21241**** (0.09028)	**-0.22078**** (0.08793)
DIST_AG	**0.00003***** (0.00001)	**0.00003**** (0.00001)
DIST_CEM	**0.00003***** (0.00001)	**0.00003***** (0.00001)
DIST_CONS	**-0.00007** (0.00004)	**-0.00007** (0.00004)
DIST_GOLF	**-0.00003***** (0.00001)	**-0.00003***** (0.00001)
DIST_SPRT	**-0.00006**** (0.00003)	**-0.00006**** (0.00003)
DIST_GRM	**-0.00004***** (0.00001)	**-0.00004***** (0.00001)
** significant at 5%; *** significant at 1%		

Table A.2: Discrete Distance Model - OLS & WLS Results.

	OLS	WLS
	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
CONSTANT	**10.10198***** (0.08530)	**10.17248***** (0.07760)
AGE	**-0.00147***** (0.00028)	**-0.00202***** (0.00033)
INTSF	**0.00028***** (0.00001)	**0.00025***** (0.00001)
ACRES	**0.04744***** (0.00891)	**0.05237***** (0.01172)
PEDUC	**1.94865***** (0.10705)	**1.95004***** (0.09902)
PAGE65	**1.72572***** (0.18461)	**1.63186***** (0.17202)
DIST_TSTOP	**-0.00002***** (<0.00001)	**-0.00002***** (<0.00001)
DIST_MJRD	**0.00007***** (0.00003)	**0.00004** (0.00002)
DIST_COMRC	**-0.00001** (0.00002)	**0.00003** (0.00002)
DIVIND_SM	**0.20370** (0.22007)	**0.15490** (0.18034)
DIVIND_LG	**0.04258** (0.08467)	**-0.02737** (0.07354)
AG_50	**-0.12654** (0.10612)	**-0.08138** (0.14433)
AG_100	**-0.02568** (0.10822)	**0.00166** (0.09483)
AG_1000	**-0.01080** (0.02632)	**-0.00147** (0.02376)
CEM_50	**0.05396** (0.22372)	**0.03953** (0.28793)
CEM_100	**-0.15102** (0.17326)	**-0.15469** (0.17767)
CEM_1000	**-0.01314** (0.02717)	**-0.01240** (0.02566)
CONS_50	**-0.00811** (0.09503)	**-0.00822** (0.07995)
CONS_100	**0.03667** (0.07279)	**0.03308** (0.06494)
CONS_1000	**0.07750** (0.05640)	**0.07060** (0.05113)
GOLF_50	**-0.09438** (0.14113)	**0.02821** (0.10137)
GOLF_100	**-0.69668***** (0.17040)	**-0.67950*** (0.37143)
GOLF_1000	**-0.00472** (0.03184)	**-0.00197** (0.02783)
SPRT_50	**0.00925** (0.08954)	**0.06225** (0.07542)
SPRT_100	**0.00072** (0.06235)	**0.01946** (0.07364)
SPRT_1000	**0.06389***** (0.03098)	**0.08032***** (0.02893)
GRM_50	**0.09107** (0.16913)	**0.06585** (0.13489)
GRM_100	**-0.02148** (0.12307)	**-0.03574** (0.09529)
GRM_1000	**0.03340** (0.03850)	**0.03361** (0.04662)
* significant at 10%; ** significant at 5%; *** significant at 1%		

Table A.3: Public/Private Continuous Distance Model - OLS & WLS Results.

	OLS	WLS
	Parameter Estimate (Standard Error)	Parameter Estimate (Standard Error)
CONSTANT	**10.46247***** (0.08291)	**10.53828***** (0.08200)
AGE	**-0.00153***** (0.00028)	**-0.00224***** (0.00034)
INTSF	**0.00027***** (0.00001)	**0.00026***** (0.00001)
ACRES	**0.04813***** (0.00877)	**0.05465***** (0.01186)
PEDUC	**1.76687***** (0.11191)	**1.73327***** (0.11081)
PAGE65	**1.63155***** (0.17089)	**1.58393***** (0.16587)
DIST_TSTOP	**-0.00002***** (<0.00001)	**-0.00002***** (<0.00001)
DIST_MJRD	**0.00004** (0.00003)	**0.00003** (0.00003)
DIST_COMRC	**<0.00001** (0.00002)	**<0.00001** (0.00002)
DIST_PUBOS	**-0.00008*** (0.00005)	**-0.00006** (0.00004)
DIST_PRIOS	**-0.00004***** (0.00001)	**-0.00004***** (0.00001)
* significant at 10%; ** significant at 5%; *** significant at 1%		

BIOGRAPHY OF THE AUTHOR

Bradley C. Neumann was born in Ann Arbor, Michigan on June 29, 1980. He was raised in Petoskey, Michigan and graduated from Petoskey High School in 1998. He attended Northern Michigan University in Marquette, Michigan and graduated in 2003 earning Bachelor of Science degrees in Economics and Land Use Planning. While at Northern Michigan University, Brad worked as a research and teaching assistant in the Geography Department, a planning consultant with the city of Marquette, and as a Geographic Information Systems intern for Environmental Systems Research Institute, Inc. (ESRI).

In the summer of 2003 Brad entered the Resource Economics and Policy program at The University of Maine to obtain a Master of Science degree. He was awarded a National Science Foundation (NSF) GK-12 Fellowship with the *GK-12 Sensors!* Program at The University of Maine. As a NSF Fellow, Brad worked at Bangor High School teaching about spatial information technologies and stimulating students' interests and aspirations in science and technology. His work as a NSF Fellow is highlighted by the partnership he helped to create between Bangor High and the city of Bangor, Maine, in which students use GIS to develop spatial information for the city. Based on the quality of his graduate research, Brad was a finalist in the 2004 Dow, Griffee, and Clement Graduate Student Research Competition.

After finishing his degree, Brad plans to work as a consultant on land use related planning and/or economic issues. Brad is a candidate for the Master of Science degree in Resource Economics and Policy from The University of Maine in August, 2005.