The Effect of Open Space on Residential Property Values in St. Paul, MN

Soren Tyler Anderson
Advisor: Sarah West

Macalester College
Department of Economics
1600 Grand Avenue
St. Paul, MN 55105

May 2001

Acknowledgements

There are many people who have contributed to the completion of this honors thesis, and I am grateful for their help. For meeting with me early on and providing helpful suggestions, I thank Steve Polasky and Steve Taff from the University of MN. Thanks to V. Kerry Smith of North Carolina State University and Noelwah Netusil of Reed College for their suggestions and working papers on open space. For providing help with a particular ArcView GIS function, I thank Hyun Kim of CEnREP. Also, thanks to Christopher G. Leggett of Industrial Economics, Inc. for his working paper on parks. For reviewing my thesis and providing helpful suggestions, I thank Karl Egge and Gary Krueger from the Macalester Department of Economics. A huge thanks to Brian Larson and Regional Multiple Listing Service of MN, Inc. for providing the home transaction data for this study. From the Macalester Geography Department, I would like to thank Carol Gersmehl and Ben Matuzas for graciously offering their time, teaching me GIS, and assisting with the GIS work on this project. Also from Mac Geography, thanks to Siri Eggebraten, Mary Reuling, and Liz Hajek for their help in the cartography lab. Finally, I would like to thank my advisor, Professor Sarah West. Her suggestions and insight were invaluable. Her constant encouragement kept me going during the spring semester. She is a wonderful person, and I cannot imagine having a more kind and supportive thesis advisor. She's the best!

Thank you, everyone!

The Effect of Open Space on Residential Property Values in St. Paul, MN

Abstract

In order to make informed policy decisions regarding the preservation of open space areas, communities need knowledge of individuals' preferences *for* and the value *of* open space amenities. There is no explicit market for these amenities, but the benefits of open space may be reflected in nearby home values. Using home transaction data from the St. Paul, MN area, I employ hedonic regression analysis to estimate the effect of proximity to open space (parks, golf courses, and cemeteries) on home sales price, controlling for home structural attributes, neighborhood characteristics, home location, and other amenities. Proximity measures were derived from regional land use data using geographic information systems (GIS) software. Overall, I find that proximity to parks and cemeteries has a negative effect on home value, while proximity to golf courses has a positive effect. As the first open space study to do so, however, I also compare how these spatial relationships differ within city and suburban sub-markets. Within the city of St. Paul, I find that proximity to parks has a *positive* effect on home value of $354 (0.25%) per 100 meters. In the suburbs, proximity to parks has a *negative* effect of $252 (0.18%) per 100 meters. These results confirm the importance of considering context when modeling complex spatial relationships in residential housing markets.

Key words: open space, hedonic, geographic information systems (GIS)

Table of Contents

I. Introduction

A. Motivation

On November 7, 2000—Election Day—voters in Washington County, MN voted on a countywide bond referendum. A "yes" vote would have authorized a property tax increase to help support the county's open space preservation program (Kaszuba, 2000). Although the referendum failed in an extremely close vote of 51,340 (51.57%) to 48,206 (48.43%), Washington County continues to identify and preserve open spaces through the purchase of development rights (conservation easements) from county landowners. Other communities have implemented or are considering similar measures to protect open areas. In order to make informed decisions regarding such policies, however, decision-makers need knowledge of individuals' preferences *for* and the value *of* open space amenities.

B. Summary

A number of studies have sought to estimate the value of various types of open space areas, including parks (Weicher & Zerbst, 1973; Leggett, 1999), urban greenbelts (Correll et al., 1978; Lee & Lineman, 1998), forests (Tyrvainen & Miettinen, 2000), golf courses (Do and Grudnitski, 1995), wetlands (Doss & Taff, 1996; Mahan et al., 2000) and open space in general (Breffle et al., 1998; Lutzenhiser & Netusil, 1999; Anderson, 2000; Bolitzer & Netusil, 2000; Smith et al., 2000). In this study, I estimate the amenity value of parks, golf courses, and cemeteries.

Open space areas may provide a number of benefits, including opportunities for recreation, privacy, a barrier to adjacent development, nice views, wildlife habitat, and protection of natural areas and native vegetation. Like many environmental amenities, however, there is no *explicit* market for these benefits. But the externality value of these amenities may be reflected *implicitly* in the value of nearby homes, as individuals bid for properties with desirable characteristics in competitive housing markets. Hedonic price analysis allows us to isolate and measure the magnitude of this externality, providing a partial valuation of the benefits of open space.

Hedonic theory (Rosen, 1974) defines a differentiated product, like a home, by its various characteristics, $Z = (z_1, z_2,..., z_n)$. In a competitive housing market, the interaction of suppliers and demanders leads to an equilibrium hedonic price function that relates the price of a home to its various attributes, $p = p(z_1, z_2,..., z_n)$. Differentiating the hedonic price function with respect to a particular attribute yields the *marginal implicit price* of that attribute. By regressing home value on a vector of housing characteristics to estimate the hedonic price function, marginal implicit prices can be retrieved empirically.

Using 1997 single-family home transaction data from the St. Paul, MN area,[1] I estimate the effect of proximity to parks, golf courses, and cemeteries on home value, while controlling for home structural attributes, neighborhood characteristics, home location, and other amenity characteristics. Home sale price and structural data come from Regional MLS of MN, Inc, a regional multiple listings service. Neighborhood data come from the 1990 U.S. Census. I derived accessibility measures from several data sources using GIS (geographic information systems) software. Finally, using GIS software, I derived measures of proximity to parks, golf courses, cemeteries, rivers, and lakes from regional land use data supplied by The Lawrence Group.

I first estimate the hedonic price function for the entire three-county study area. Overall, I find that proximity to parks has a negative effect on home value of $238 (0.17%) per 100 meters. For cemeteries, the effect is a negative $69 (0.05%) per 100 meters. And, proximity to golf courses has a positive effect of $33 (0.02%) per 100 meters. In a similar regression, I estimate that homes within 200 meters of a park sell for $4,700 (3.30%) less, homes within 200 meters of a cemetery sell for $12,300 (8.65%) less, and homes within 200 meters of a golf course sell at a $2,900 (2.04%) premium. Upon plotting the spatial error of this model's predicted values, however, I find a disturbing amount of autocorrelation. This suggests a spatial relationship that is uncontrolled-for in this first model.

In order to control for this spatial dependence, I employ a segmented-markets model and estimate two separate hedonic equations: one for homes in the city of St. Paul, and a second for homes in the suburbs (outside the city). As a result, spatial autocorrelation is noticeably reduced in this second model. In the city, I find that proximity to parks has a *positive* effect on home value of $354 (0.25%) per 100 meters. In the suburbs, proximity to parks has a *negative* effect of $252 (0.18%) per 100 meters. In a similar regression, I estimate that homes within 200 meters of a park are worth $2,650 (1.86%) *more* in the city and $3,480 (2.45%) *less* in the suburbs. These results demonstrate that there may be fundamental differences in the way open space amenities are valued and perceived, depending on whether a home is located in the city or in the suburbs.

[1] I use data from Dakota, Ramsey, and Washington County, MN.

C. Overview

The remainder of this paper is divided into eight sections. Section II defines open space and discusses its benefits. Section III discusses the valuation of non-market goods, the capitalization of amenity externalities, hedonic theory, and second-stage hedonic analysis. Section IV reviews important issues in residential hedonic literature, including specification of the hedonic model, functional form, segmented markets, and data issues. Section V discusses hedonic open space papers in more detail, paying special attention to the specification of the open space variable, open space data, and the empirical results of these studies. Section VI presents the empirical model, explanatory variables, and a discussion of expected signs. Section VII discusses the data used in this study. Section VIII presents the empirical analysis and results of this study. Finally, Section IX concludes with a summary and discussion of this study's contributions.

II. Open Space

A. Definition of Open Space

Defining open space can be a tricky endeavor. The phrase "open space" can mean a variety of different things, depending on the context of the discussion and the parties involved. For some, "open space" may simply mean "not houses," in which case urban parks, wetlands, farms, golf courses, baseball fields, and variety of other land uses would qualify as open space. Others may define "open space" as even less-developed types of land use, such as woodlands, forests, wildlife refuges, or nature conservancies.[2]

In dealing with these issues, the existing literature has adopted one of three approaches. The first and most common approach has been to examine the value of a very specific type of open space, such as urban parks (Weicher & Zerbst, 1973), golf courses (Do & Grudnitski, 1995), or urban greenbelts (Correll et al., 1978). Focusing on a specific type of open space, these studies render issues of definition and classification irrelevant. A second approach has been to define open space broadly by aggregating all open space types into one category (Anderson, 2000). The third and perhaps most useful approach has been to examine a broad range of open space areas, classify them by type, and then examine the value of open space based on these classifications (Lutzenhiser & Netusil, 1999; Smith, et al.,

[2] Discussion based on conversation with Dr. Steve Taff, Associate Professor and Extension Economist, University of MN Department of Applied Economics, October 2, 2000.

2000; Bolitzer & Netusil, 2000). I adopt the latter approach in this study by estimating the amenity value of public parks, golf courses, and cemeteries.[3]

B. Benefits of Open Space

In general, open space can be thought of as providing a set of particular benefits and amenities. This set includes (but is not limited to) opportunities for recreation, privacy, a barrier to adjacent development, nice views, habitat for wildlife, and the protection of natural areas and native vegetation. While open space areas share many of the same or similar attributes, however, the benefits provided by different types of open areas are not identical. For example, urban parks may provide countless recreational opportunities, but may fail to protect native vegetation. Similarly, the benefits conferred by distinct areas of a particular classification may also be diverse. For example, one urban park may have a baseball diamond and volleyball a net, while another may have a swing set and a merry-go-round.

The benefits of open space can be broken down into those that provide *use value* and those that provide *nonuse value*. Use value is derived from current use of the resource, such as use of the area for recreation, scenic views, privacy, or as a barrier to adjacent development. Nonuse value is derived from considering the possible future use of the area. Nonuse value may take one of several forms:

> Option value represents an individual's willingness to pay to maintain the option of utilizing a resource in the future. Existence value represents an individual's willingness to pay to ensure that some resource exists, which may be motivated by the desire to bequest the resource to future generations (Fausold & Lilieholm, p. 3, 1996).

That is, for instance, the pleasure derived from knowing that open space is being conserved (Breffle, et al., 1998).

III. Theory

A. Valuation of Non-market Goods

The benefits provided by open space give these areas amenity value. But like other environmental amenities, there is no *explicit* market for the benefits provided by open space. That is, for instance, one cannot simply open up Monday's edition of the *Wall Street Journal* and find the going rate on a marginal increase in open space amenities. There *is* a market for undeveloped land, and properties have market value as reflected in the real estate market.

---

[3] I also estimate the value of rivers and lakes, which may also generate certain open space benefits.

But this value does not necessarily reflect the value of undeveloped land *as open space*. For instance:

> In rural areas, where the highest and best use of land (i.e., most profitable use) is open space, one can examine market transactions. In urban or urbanizing regions, however, where the highest and best use (as determined by the market) has usually been development, the open space value of land must be separated from its development value (Fausold & Lilieholm, p. 4, 1996).

In most cases, therefore, real estate market price will not accurately measure the value of undeveloped land *as open space*. As a consequence, alternative methods are necessary to measure the value of non-market environmental amenities such as open space. There are several alternative techniques for valuing non-market goods. Two of these are *contingent valuation* and *hedonic regression analysis*.[4]

B. Contingent Valuation

Put simply, contingent valuation measures the value of an environmental amenity by using a series of survey questions. Specifically, these questions are designed to elicit information regarding an individual's willingness to pay for the effects of a particular policy (Breffle, et al., 1998). For example, in order to determine the willingness to pay for open space in a particular neighborhood, one might ask, "How much would you be willing to pay for the creation of a 20-acre green space on the southwest corner of Grand and Snelling Avenues?" In practice, the contingent valuation method can be quite complicated, since results often depend on survey design. There is also

> A potential for strategic bias, and a concern is that this will cause households to overstate their willingness to pay. Some respondents may overstate their willingness to pay, hoping to influence a [government body's policy] decision, if those respondents thought they would never have to pay (Breffle et al., p. 718, 1998).

An advantage of contingent valuation is that, by directly *asking* individuals about their willingness to pay, it is able to capture the *nonuse* value provided by a particular amenity. Contingent valuation can also be used to value environmental amenities when the

---

[4] There are other methods, as well. The *travel cost* method treats the cost of travel to an area as a sort of "entry fee." Thus, the value of a particular open space area (or other environmental amenity) will be reflected in individuals' willingness to pay to travel to the area. Other methods examine the benefits provided by a particular amenity and then measure the costs of providing those benefits by some alternative means. For instance, say a marshland provides the benefit of flood control. Part of its value is reflected, therefore, in the cost of providing flood control by some other means—such as the cost of building a dam.

data required for other techniques are not available.[5] Brookshire et al. (1982) estimate the willingness to pay for improvements in air quality using *both* hedonic and contingent valuation approaches. They find that contingent valuation survey estimates are bounded *above* by hedonic estimates and *below* by zero. Thus, although contingent valuation methods have the *potential* to capture both the use and nonuse values of open space, prior research indicates that contingent valuation survey estimates may actually *understate* value as compared to hedonic estimates. Given these results and the availability of the required data, I will use hedonic regression analysis in this study to estimate the value of open space amenities.

C. Capitalization of the Open Space Externality

In order to understand residential hedonic analysis, one must begin by thinking of open space as a residential amenity that benefits individuals residing adjacent to or near open space areas. Among the many benefits of open space are nice views, convenient access to recreational opportunities, a barrier to adjacent development, and increased privacy. Benefits like these accrue mainly to households residing adjacent to and near open space areas. These benefits are rivalrous. That is, when one individual chooses to reside near open space, she necessarily precludes someone else from residing in the same location and enjoying the same benefits. Because of this rivalry, we expect that prospective homebuyers will bid up the prices of homes near to open space in order to gain these benefits. Thus, externalities generated by proximity to open space will be reflected in nearby home values. Hedonic analysis allows us to isolate and measure this capitalized externality.

Unfortunately, hedonic analysis provides only a limited measure of the total economic benefits of open space. While open space confers *many* benefits, hedonic analysis measures only those that are *capitalized* in the value of a home. Certain benefits, such as reduced soil erosion, wildlife habitat, or improved water quality, which have a strong public good element, or nonuse benefits such as the pleasure derived of knowing that open space is being preserved, are unlikely to be reflected in home values (Lutzenhiser & Netusil, 1999; Breffle et al., 1998). Furthermore, a lack of spatial variation in open space amenities may also inhibit the capitalization these benefits. For example, if the externalities generated by open space amenities were identical for each home within a given housing market, this lack of variation would prevent the benefits from being identified in home values (Tyrvainen &

---

[5] These include home property value data or wage data, in the case of hedonic analysis.

Miettinen, 2000). In most cases, insufficient spatial variation is not an issue. The inability of hedonic analysis to measure the nonuse, public good benefits of open space is *always* an issue, however, and must be considered when interpreting hedonic estimates.

D. Hedonic Theory—Rosen's Model[6]

In his seminal article, Sherwin Rosen (1974) defines a differentiated product (such as a home or automobile), Z, by its various characteristics z_1, z_2, ..., z_n.

$$Z = (z_1, z_2, ..., z_n)$$

The hedonic price function, $p(Z)$, relates changes in the price of the differentiated product to changes in the quantities of its various attributes.

$$p = p(Z)$$

In a competitive market, consumers and producers of the good will take this price schedule as given when making their consumption and production decisions. The equilibrium *marginal implicit price* of an attribute is given by the partial derivative of the hedonic price function with respect to that attribute.

$$\frac{\partial p(Z)}{\partial (z_i)} = p(z_i)$$

Empirically, then, the marginal implicit prices of home characteristics may be retrieved by regressing home value on a vector of home characteristics.

This model depends on several assumptions that hold to varying degrees in actual housing markets. First, it is assumed that there exists a large number of differentiated products and sufficient variation such that that the range of product choices is essentially continuous. Second, it is generally assumed that housing bundles cannot be untied and repackaged linearly according to consumers' tastes and preferences. That is, for instance, two homes of 1000 square feet are not equivalent (in terms of the benefits they provide) to one home of 2000 square feet (Rosen, 1974).[7] Third, and finally, it is assumed that there are a large number of consumers and producers, and that these individuals are price takers.

---

[6] Note: this section also relies heavily on summaries of Rosen's model by Palmquist (1991) and Orford (1999).

[7] If product characteristics can be repackaged linearly in this way, then the hedonic price function will be linear. Generally, however, this type of repackaging is either impossible or highly costly, and the hedonic equation is nonlinear.

In Rosen's model (1974), the equilibrium hedonic price function is determined by the interaction of utility-maximizing consumers and profit-maximizing firms in the housing market. Consumers differ according to a vector of socioeconomic characteristics, α. Consumers gain satisfaction from the characteristics of the differentiated good, *Z*, and a composite good, *x*, and maximize utility

$$U(x, z_1, z_2, ..., z_n; \alpha)$$

subject to the budget constraint

$$y = x + p(z_1, z_2, ..., z_n)$$

where *y* is the consumer's income and the price of *x* equals 1. The first-order conditions for the consumer problem are satisfied by setting the marginal rate of substitution between one of the characteristics and the composite good equal to the marginal price of the characteristic.[8]

$$\frac{\frac{\partial U}{\partial z_i}}{\frac{\partial U}{\partial x}} = \frac{\partial p(Z)}{\partial (z_i)} = p(z_i)$$

A consumer's actions in the housing market can be represented by a bid function

$$\theta(Z, u, y; \alpha)$$

where θ is an individual's willingness to pay for a particular product with characteristics *Z*, given a certain level of income and constant utility. Modifying the consumer utility function slightly, the bid functions are defined implicitly by

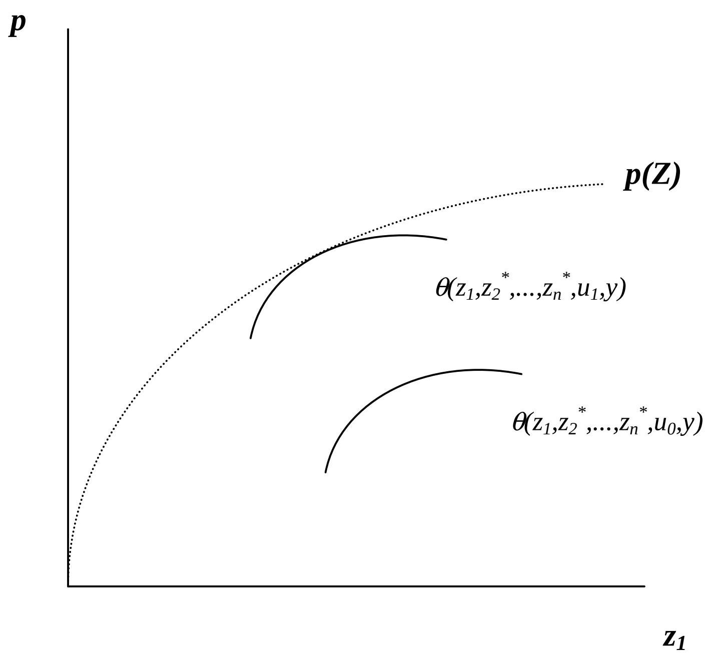
$$U(y - \theta, z_1, z_2, ..., z_n; \alpha) = u$$

and trace out a set of indifference surfaces relating changes in the z_i to corresponding changes in the total bid that hold consumers at the same utility level (see Figure 1).

---

[8] See Appendix I for justification of these optimizing conditions.

Figure 1: Consumer Optimization



In this context, utility is maximized when the marginal bid with respect to a given attribute is equal to the marginal implicit price in the housing market for that attribute.

$$\frac{\partial\theta(Z;u,y)}{\partial z_i} = \frac{\partial p(Z)}{\partial z_i}$$

Graphically, this optimization occurs where the two surfaces *p(Z)* and $\theta(Z;u,y)$ are tangent to each other (see Figure 1).

On the production side of the market, Rosen (1974) assumes that producers must choose both a particular version of a differentiated product, as well as the number of units to produce. Costs will vary from firm to firm by technology or factor prices β, and so the cost function can be represented by

$$c(M,Z;\beta)$$

where *M* is the number of units produced. Revenues will depend on the quantity sold and on the hedonic price schedule, so profits are simply revenues minus costs

$$\pi = M \cdot p(Z) - c(M,Z;\beta).$$

Solving the producer problem, the first-order conditions are satisfied by setting the marginal price for each characteristic equal to the marginal cost per unit of increasing the quantity of that characteristic.[9]

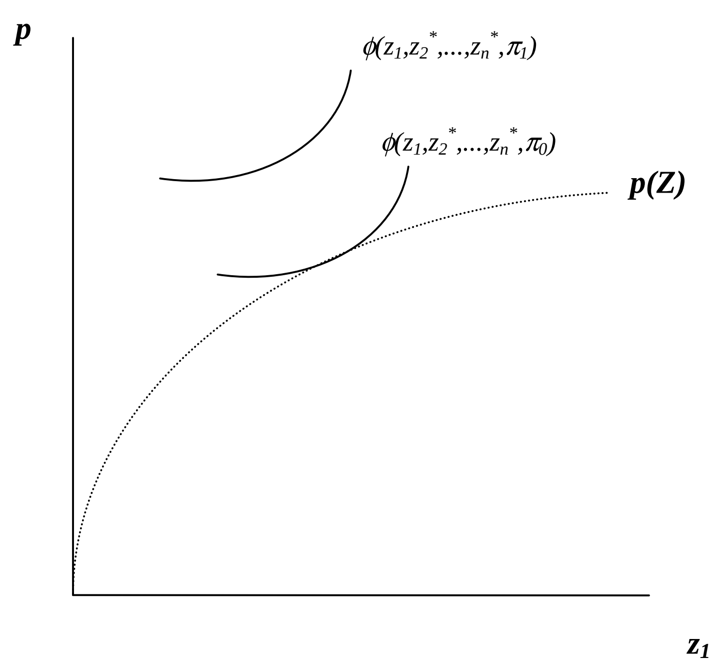
$$\frac{\partial p(Z)}{\partial(z_i)} = p(z_i) = \frac{\dfrac{\partial c}{\partial z_i}}{M}$$

Analogous to the consumer side, the behavior of firms can be represented by a set of offer functions

$$\phi(Z,\pi;\beta)$$

where ϕ represents the minimum price a producer can accept for a product with characteristics, Z, and still make profits π. Like the consumer side, these offer functions trace out a set of indifference surfaces. In this context, profits are maximized by setting the marginal offer with respect to a particular characteristic equal to the marginal price for that characteristic. Graphically, this relationship occurs when the surfaces $p(Z)$ and $\phi(Z;\pi)$ are tangent (see Figure 2).

Figure 2: Producer Optimization

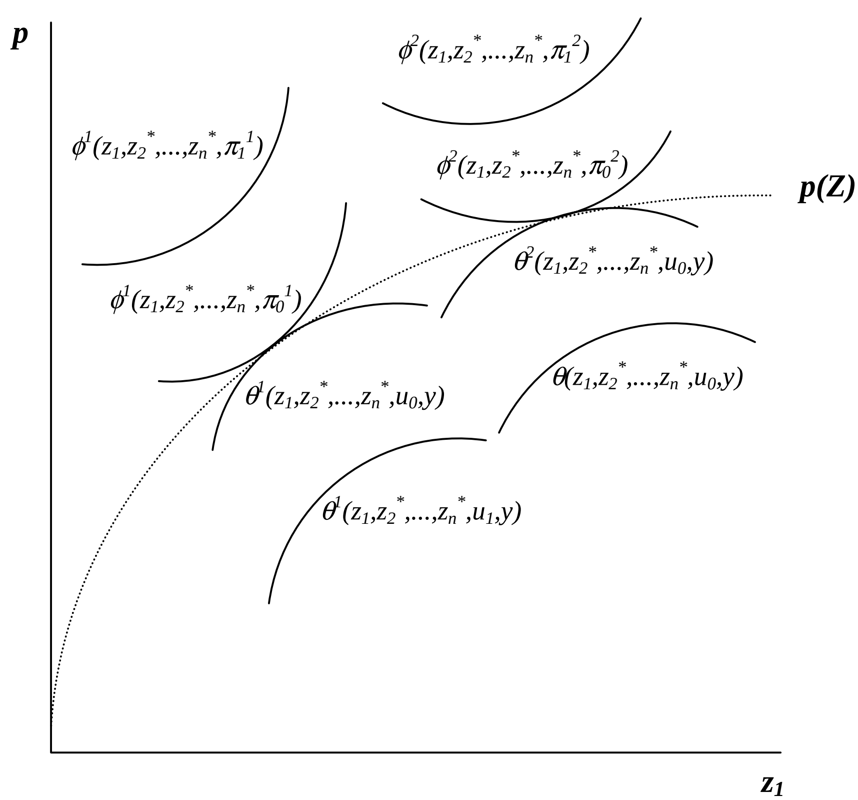


[9] See Appendix I for justification of these optimizing conditions.

While individual consumers and producers take the hedonic price schedule as given, it is the interaction of consumers and producers, represented by their respective bid and offer functions, that determines the hedonic price schedule in a competitive market. Consumers attempt to maximize utility by seeking out the lowest bid possible; producers attempt to maximize profits by seeking the highest possible offer. This interaction matches consumers with producers such that consumers cannot achieve a higher utility by choosing a product with different characteristics, nor can producers increase profits by altering the quantity or version of the product they produce.

This interaction is shown graphically in Figure 3. Each consumer's bid surface represents varying bids and quantities of a characteristic given a particular level of utility; each firm's offer surface represents varying offers and quantities of a characteristic given a particular profit level. The matching of consumers and producers produces a set of individual equilibria that trace out the hedonic price function, *p(Z)*.

Figure 3: Hedonic Equilibrium



E. Second Stage Hedonic

In Rosen's original article (1974), the hedonic property value method consists of two stages. The first stage, in which home values are regressed on housing attributes given some functional form, estimates the equilibrium hedonic price schedule. But this first stage does not necessarily contain information about individual behavior (i.e. demand functions). Rosen (1974) describes a second stage analysis in which the estimated hedonic price schedule is used to calculate the marginal implicit prices for each observation. Variation in these calculated marginal implicit prices is then used to identify the underlying bid and offer functions for each characteristic.

Researchers have employed various techniques to identify demand parameters from hedonic price equations in the market for homes (Blomquist & Worley, 1981; Palmquist, 1984), automobiles (Agarwal & Ratchford, 1980; Arguea & Hsiao, 1993), and other goods. Due to the many econometric difficulties associated with the second stage analysis,[10] however, most hedonic studies perform only the first stage. Palmquist (1992) shows that the first stage equation sufficiently measures total benefits in the case of localized externalities that affect a small geographic area and a small number of people. The first-stage is insufficient, however, when trying to measure the benefits of an amenity that affects a large geographic area and a larger number of homeowners. In such cases, the second stage hedonic analysis, which estimates the marginal willingness to pay function, is necessary (Freeman, 1993; Palmquist, 1991). Among open studies, only Mahan, Polasky, & Adams (2000) perform the second-stage analysis in their attempt to identify the marginal willingness to pay for proximity to wetlands. Like many other studies, they are plagued by econometric difficulties, and their attempt is unsuccessful.[11]

IV. Residential Hedonic Empirical Literature

A. Residential Studies and Environmental Goods

Hedonic regression analysis has a long history in the literature concerning the valuation of environmental goods. Hedonic techniques have been used to measure the externalities of nice views (Benson, et al., 1998; Gillard, 1981), air quality (Chattopadhyay, 1999; Palmquist, 1984), hazardous waste sites (Michaels & Smith, 1990), and countless other residential amenities. In addition to the issues discussed in the previous section, there are a

---

[10] See Palmquist (1991) and Freeman (1993) for full discussion of these difficulties.

number of other important issues regarding hedonic regression analysis that must be considered. Most of these issues relate to the econometric estimation of the hedonic price function.

B. Specification

Specification of the first-stage hedonic model itself can present unique problems. In theory, the price equation should include all housing characteristics that enter into the utility function of a household (Tyrvainen & Miettinen, 2000). Some studies apparently attempt to do this, including as many as 62 explanatory variables (Frech & Lafferty, 1984). Unfortunately, limited data and strong multicollinearity among explanatory variables do not make this a practical method. In the case of this study, however, my goal is to control for variation within the hedonic price function to generate unbiased estimates of the marginal implicit price for proximity to open space. Thus, the elimination of redundant variables is justified, as long as it does not lead to serious specification bias (Tyrvainen & Miettinen, 2000).

This, in fact, is what most studies do. For instance, Correll et al. (1978) eliminate various structural variables (bathrooms, bedrooms, dummy variables representing the presence of a basement and garage, and several structural design and construction dummy variables) due to problems with multicollinearity. Other studies do the same. In short there must exist some optimal specification of the hedonic price function that minimizes problems associated with hyper-specification (multicollinearity, lack of data) and under-specification (omitted variable bias). In the end, actual specification varies from study to study, based on data availability and the seriousness of multicollinearity among the independent variables.[12]

C. Functional Form

In general, the functional form of a hedonic price function cannot be determined by theory. Rather, it must be determined empirically from the given housing market data. Typical functional forms are linear (Bolitzer & Netusil, 2000; Frech & Lafferty, 1984; Doss & Taff, 1996; Correll et al., 1978; Weicher & Zerbst, 1973), semi-log (Bolitzer & Netusil,

---

[11] Their estimation of the willingness to pay function results in insignificant and unexpected signs for key demand parameters.

[12] It is unlikely that under-specification due to a lack of structural data will pose a severe threat. Most sources of housing data (assessors, realtors, etc.) include certain data for a reason—because it is known to affect home value.

2000; Smith et al., 2000; Frech & Lafferty, 1984; Leggett, 1999), double-log and inverse semi-log (Leggett, 1999). Others employ a Box-Cox flexible form (Lutzenhiser & Netusil, 1999; Tyrvainen & Miettinen, 2000).[13] In the end, choice of functional form is most often based on goodness-of-fit criterion.[14]

D. Segmented Markets

In addition to issues of specification and functional form, it is also important that hedonic specifications accurately model the structure of the housing market itself. Traditionally, most hedonic studies have assumed that housing attributes will have the same marginal implicit price across the entire study area, implying a single, homogenous housing market. In some cases, however, the complexity of housing market dynamics may actually lead to the formation of quasi-independent sub-markets. In such cases, implicit prices may vary by sub-market (Orford, 1999). Specifically, this occurs when *exogenous* factors constrain individual buyers and sellers to participation in certain segments of a larger market (Michaels & Smith, 1990). Examples of such exogenous factors may include geographic barriers or political boundaries, discrimination, or a lack of information (Palmquist, 1991). In the case of the latter, Michaels & Smith (1990) argue that real or perceived sub-markets may develop as a consequence of the amount and type of information available to participants in the housing market. In large housing markets, for instance, individuals may rely on housing market experts (realtors, developers, etc.) for housing market information, which may result in market segmentation.

After identifying sub-markets, researchers must model segmentation by estimating separate hedonic equations for each sub-market or by applying more advanced econometric techniques (see Orford, 2000). However, it is necessary to exercise caution. As Palmquist notes:

> If economists assume that there is a single market when it is actually segmented, their coefficients will be biased. On the other hand, if they assume that the markets are segmented when they are not, their estimates will be imprecise and they may have insufficient data in the segments (p. 89, 1990).

---

[13] In a Box-Cox, the dependent variable is transformed according to the equation $y^{(\lambda)} = (y^{\lambda}-1)/\lambda$ where the maximum likelihood for lambda λ is estimated according to the equation $y_i^{(\lambda)} = (y^{\lambda}-1)/\lambda = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + e_i$. This value is used to estimate the parameters of the model using OLS.

[14] In an influential paper, Cropper et al. (1988) test the performance of various functional forms in simulations of housing market equilibria in which the true marginal price paid for each attribute is known. They find that linear and quadratic functions of Box-Cox transformed variables perform best in estimating coefficient parameters when all attributes are observed. When certain variables are unobserved or measured with a proxy, then the linear Box-Cox function is the functional form of choice.

In addition to biased estimates, failing to adequately model market segmentation may result in spatial autocorrelation of error terms. The potential for spatial autocorrelation is magnified when aggregation of the data results in a significant amount of spatial heterogeneity, as is often the case in hedonic studies of large metropolitan areas. This violation of the ordinary least squares assumption of independently, identically distributed errors undermines the validity of OLS coefficient estimates (Orford, 1999).[15]

Identifying segmented markets can prove difficult, however. Many studies use dummy variables with the intent of capturing neighborhood or regional fixed effects (Do & Grudnitski, 1995; Doss & Taff; 1996; Bolitzer & Netusil, 2000). But differences in these effects do not necessarily imply segmented markets, as there are other omitted spatial variables. Nonetheless, some studies have attempted to use differences in regional dummies as a means to identifying segmented markets (Mahan, Adams & Polasky, 2000).[16] Some studies interact market segment dummies with attribute variables in order to capture differences in the marginal implicit prices of attributes across urban space (Lutzenhiser & Netusil, 1999). Finally, some studies use less numerical methods. In order to identify market segments in Boston, Michaels & Smith (1990) ask local realtors to distinguish areas of market segmentation based on their breadth of experience in that metropolitan area. Using these responses, Michaels & Smith identify sub-markets in the Boston area, and estimate separate hedonic equations accordingly.

In summary, it is unlikely that a single hedonic equation will accurately model spatial relationships in a large, complex housing market (Orford, 1999). In their study of the disamenity effects of hazardous waste sites in the Boston area, for instance, Michaels & Smith (1990) find that proximity estimates are unstable across sub-markets. In his multi-level study of the Cardiff housing market, [17] Orford (2000) finds that there is a clear inner city/suburban split in the implicit price of floor area, with floor space being more expensive in suburban areas than in the city. Furthermore, Orford (2000) finds corroborating evidence of a “stigma effect,” whereby home values in certain communities command an unexplained premium while home values in other communities are plagued by an unexplained discount.

[15] Orford also notes that spatial autocorrelation may arise as a consequence of uncontrolled for nonlinear relationships, by the omission of relevant variables, or when spatial data are incorrectly modeled.

[16] Mahan, Polasky & Adams attempt to use market segmentation as a means to identifying demand parameters in their second-stage hedonic estimation of the willingness to pay for the benefits of urban wetlands. After identifying sub-markets, they estimate separate hedonic equations for each segment and compute marginal implicit prices for their second stage analysis. See Section IV, E: Second Stage Hedonic.

[17] Cardiff is the capital of Wales.

This and other evidence suggest that many traditional hedonic studies do not fully capture the spatial element in their models. As Orford states:

> The main implication is that the contextual nature of the supply and demand of housing attributes needs to be explicitly incorporated into the hedonic specification if the spatial dynamics of the housing market are to be successfully modeled (p. 1669, 2000).

E. Data

Five types of data are generally needed in order to perform the first stage hedonic analysis. These are home value data, home characteristic data, neighborhood characteristic data, accessibility data, and amenity (including open space) data.[18] Researchers generally obtain home value and home characteristics data from the same source. These sources include county assessors offices, county property tax records, private companies that collect real estate data, and multiple listings services (realtors).

Actual home value data may come in one of several forms. These include census tract median value, listing value, assessed value, or transaction price. Which of these is superior is open to debate. As Doss and Taff (1996) discuss:

> The hedonic equation seeks to track the "true price"—that dollar value agreed upon by willing buyers and sellers, each with full information and no coercion. This price is never directly observable...Even the [reported transaction price], which is frequently assumed to be equal to true price, necessarily contains some error: reported transaction prices are often not adjusted for time or terms, they are subject to recording error, and they may be intentionally misreported. Consequently, any observed price—whether the transactions price or the assessor's estimate—is only a proxy for the unobservable true price (p. 123).

As noted by Mahan, Polasky and Adams (2000), however, "actual sales prices of individual properties are preferred to other forms of data on property values such as assessed, appraised, or census tract estimates because sales come closest to reflecting equilibrium prices" (p. 103). Still, when dealing with home transaction data it is vitally important to make sure that all home value observations represent "arms-length" transactions between utility-maximizing buyers and profit-maximizing sellers. A transaction in which a seller sells a home to a relative for $1, for instance, would *not* constitute an arms-length transaction, and sales price would *not* reflect market value in this case.

Home value data, especially transaction price data, may also be susceptible to housing market price level changes, changes in preferences for residential goods, and changes in the relative supply and demand for residential goods. Collecting home value data from a narrow window of time can minimize this problem. Data collected for wider time ranges are

[18] For the second stage analysis, additional household characteristics (income, age, etc) are generally needed.

susceptible to price level and other changes. Such fluctuations should be controlled for when aggregating data in this way (Palmquist, 1991). This could be difficult however, and may require more than a simple adjustment of home sale price by the CPI or some index of housing value. As Frech and Lafferty (1984) discuss, different housing characteristics are likely to be affected differently by price level changes. Nonetheless, aggregating data from different time periods to obtain enough observations is sometimes necessary. Hence, temporal stability of parameters must be studied when aggregating data in this way.

As discussed earlier, data for home value and structural characteristics usually come from the same source. As a result, the structural data available are generally limited to those variables provided by the particular data source. For example, if the county assessor does not note the number of bathrooms for each home sale transaction, it is unlikely that this data could be retrieved from some other source. I expect, however, that data are included because they influence home value. In essence, all available variables have been compiled because they are, by definition, relevant data.

Neighborhood characteristics data often come from the U.S. Census and/or other government data sources (state, local and regional governments, school districts, etc. There are two potential problems associated with neighborhood characteristic data: aggregation and relevance. First, by using census tract data, for instance, we are making the implicit assumption that a neighborhood is defined by the tract in which it is located. This may or may not be the case. Aggregated by census tract (or township, or school district), such data may or may not reflect the actual characteristics of individual neighborhoods. Second, neighborhood effects are assumed to remain constant during the relevant time period (Frech & Lafferty, 1984). Often, census tract data come from a different year than the year in which a home transaction occurred. Depending on the stability of census tract demographic characteristics over time (10 year intervals), this may or may not affect the relevance of the neighborhood characteristic data.[19]

Accessibility data, such as proximity to the central business district (CBD), proximity to the nearest shopping center, or city quadrant often come from geographic information systems (GIS) data or other map-based data. Certain characteristics, such as distance to CBD, which is intended to proxy for overall accessibility to employment, recreational, and shopping opportunities, may or may not be an effective proxy.

[19] Carol Gersmehl of the Macalester College Department of Geography suggests that census tract demographics are relatively stable over periods of 10 years.

Land use data are usually obtained from the same sources as accessibility data. Like neighborhood characteristic data, land use and other amenity data may not exactly match the home sales data. For instance, let us assume that the land use data in a particular study come from the year 2000; they contain a golf course that was constructed 1998. Now, let's assume the home transaction data come from 1997. The combination of these data sets would imply that some homes sold near a golf course. Interpreting the golf course as affecting home value would be incorrect, however, since the golf course did not exist at the time of the home transaction! Smith et al. (2000) argue that because there are irreversibilities (or high costs) in changing from one land use to another current land use is a good proxy for recent historical land use. This is not true in all cases, however, especially in rapidly growing areas (Smith et al., 2000).

V. Hedonic Open Space Papers

A. Studies

Several studies have explicitly examined the effect of open space on home values. Others have studied the effects of a particular "open space" land use amenity. Mahan et al. (2000) and Doss & Taff (1996) study the effects of various wetland types on residential home values. Weicher & Zerbst (1973) and Leggett (1999) measure the externality effect of urban parks on property values. Correll et al. (1978) and Lee & Lineman (1998) estimate the effect of proximity to urban greenbelts on home values in Boulder, CO and Seoul, Korea, respectively. Tyrvainen & Miettinen (2000) examine how proximity to urban forests affects home value in Salo, Finland. Frech & Lafferty (1984) study the determinants of home value before and after the California Coastal Commission's decision to protect coastal open space areas. Finally, Lutzenhiser & Netusil (1999), Bolitzer & Netusil (2000), Anderson (2000), and Smith et al. (2000) all explicitly study the effect of various open space types on residential home values.

B. Specification of the Open Space Variable

Hedonic regression analysis requires that all home characteristics be converted into numerical quantities or well-defined qualitative attributes (measured empirically by dummy variables). Translating a general definition of open space like "vacant land, parks, and wetlands" into a usable empirical definition has proven to be a perplexing matter. That is, once we have defined what types of land use constitute open space, we must next decide how to measure the effects of these land uses on home values. This issue is critical. As Smith et al. (2000) state, "decisions about how to measure the effects of open space on private residential lots are significant considerations in determining the diverse conclusions drawn from hedonic property value models" (p.3).

The St. Paul (Doss & Taff, 1996) and Portland (Mahan, Polasky, & Adams, 2000) wetland studies provide an illustrative example. Although both studies examine the effect of proximity to urban wetlands on home values, and although their models are relatively similar, the Portland authors state that "it is difficult to compare the results across studies" (p.110). Even interpreting the results within a single study can sometimes be difficult. For example, many studies include a measure of "distance to nearest open space." As the Portland authors put it, the problem with this type of specification

> Is that [nearest wetland of] a given wetland type may be several wetlands distant from a residence. For example, the closest forested area wetland may be the tenth wetland from a house and would have little impact on its value (p. 110).

Smith et al. also discuss the difficulties with specification and interpretation of the open space variable, particularly minimum distance specifications. They state:

> Measuring the distance to the closest...open space implies we can define a set of homogeneous land uses that convey approximately the same service to individuals. When we consider only one [open space area], we must be prepared to argue that home buyers know about it and that measuring proximity to the park is an indirect means to create a proxy measure for a pattern of land use that is consistent in the neighborhoods that surround it. Often this is not the case...Land uses can be very diverse (p. 16).

In the end, Smith et al. suggest using either a more restrictive (disaggregated) definition of open space or a distance measure that takes into account the multiple opportunities to for access to a class of land uses within the study area.

Table I summarizes the empirical specifications used by previous open space (and wetland) studies. Note the variety of specifications and the frequent use of the "distance to nearest" type of variable. These incomprehensive types of specifications do not measure the impact of other close (but not closest) open space areas. Furthermore, the results of previous open space studies indicate that the size, distance, and type of nearby open space are *all* relevant in determining its impact on residential home values.[20] Thus, in order to measure these effects, we would like an empirical specification of open space that simultaneously

---

[20] See Table II.

measures all three of these attributes, without neglecting other nearby land uses and their effect on the value of a home.

TABLE I

Open Space Specifications

Authors / Year	Location	Variables in Specification
Smith, Poulos, & Kim (2000)	Raleigh, NC	Distances to nearest type of each open space (3 types) Adjacent to open space of each type (3 types)
Mahan, Polasky, & Adams #1 (2000)	Portland, OR	Acreage of nearest wetland Distance to nearest wetland Nearest wetland type dummy (6 types) Distance to nearest park
Mahan, Polasky, & Adams #2	Portland, OR	Acreage of nearest wetland Distance to nearest wetland of each type (6 types) Distance to nearest park
Lutzenhiser & Netusil #1 (2000)	Portland, OR	Open space type within 1500 feet dummy (5 types) Acreage of open space area (from above) Acreage of open space area squared (from above)
Lutzenhiser & Netusil #2	Portland, OR	Open space type within distance range dummies (5 types) Acreage of open space area (from above) Acreage of open space area squared (from above)
Tyrvainen & Miettinen #1 (2000)	Finland	Distance to nearest forested area View of forest dummy
Tyrvainen & Miettinen #2	Finland	Forested area within distance range dummies
Bolitzer & Netusil #1 (2000)	Portland, OR	Open space within 1500 feet dummy Open space size
Bolitzer & Netusil #2	Portland, OR	Open space type within 1500 feet dummy (4 types) Open space size
Bolitzer & Netusil #3	Portland, OR	Open space within distance range dummies
Anderson (2000)	Mpls, MN	Adjacent to open space
Leggett #1 (1999)	Wash, D.C.	Distance to large regional park Neighborhood park within distance range dummies
Leggett #2	Wash, D.C.	Size of nearest park
Lee & Lineman (1998)	Seoul	Distance to greenbelt Outside of greenbelt ring Size of greenbelt
Doss & Taff (1996)	St. Paul, MN	Distance to nearest wetland of each type (4 types) Distance to nearest wetland of each type squared (4 types)
Do & Grudnitski (1995)	San Diego	Adjacent to golf course dummy
Correll, Lillydahl, & Singell (1978)	Boulder, CO	Walking distance to greenbelt
Weicher & Zerbst (1973)	Columbus, OH	Adjacency to park dummies

C. A Possible Alternative Specification

A more comprehensive specification of the open space variable is possible. For each home observation, we could measure the total acreage of each land use type, say neighborhood parks, regional "natural area" parks, vacant areas, and quasi-public open space (such as Macalester College's open areas, for instance) within a given radius of the home. We could then expand to larger and larger radii around the home, measuring the total acreage of each land use within each successive ring. These measures would then enter the hedonic specification (in addition to home structural characteristics, accessibility factors, neighborhood characteristics, and other amenities) as independent variables. Conceptually, we would have

Price = f (..., VAC500, VAC1000, VAC1500,..., PARK500, PARK1000, PARK1500)

where VAC500 measures the total vacant land acreage within 500 meters, VAC1000 measures the total vacant land acreage within the ring of 500 – 1000 meters, etc, and the PARK measures do the same for parks. This specification would allow for the simultaneous measurement of the effects of size (total acreage), distance (distance ranges), and type (different land use categories) of open space.

This specification poses potential problems. First, it is possible that the open space variables would be collinear. That is, for instance, the total park acreage within 500 meters would likely be correlated with the total park acreage within 500 – 1000 meters. There are ways of dealing with this potential problem. First, a sufficiently large number of observations could overpower the inflation in standard errors caused by collinearity. Second, increasing the distance range between radii could reduce the collinearity of the open space variables. Finally, estimating the model by running separate regressions with only one distance range variable at a time would eliminate the problem altogether.[21] A second problem with this superior specification is associated with GIS programs, land use data, and the feasibility of calculating this variable in GIS. This problem is discussed below.

D. GIS and Open Space

Earlier open space studies use traditional maps to measure the proximity of homes to amenities such as rivers, lakes, and various open space areas (Frech & Lafferty, 1984;

---

[21] Though interpreting the coefficients on these variables in the absence of the other open space variables could prove difficult.

Correll et al., 1978). More recent studies use GIS software to analyze data that relate homes geographically to open space and other amenities (Doss & Taff, 1996; Leggett, 1999; Lutzenhiser & Netusil, 1999; Smith et al., 2000; Bolitzer & Netusil, 2000; Mahan, Polasky & Adams, 2000).

At its most basic level, GIS data are simply a digital map—such as a map of parks within a particular county. GIS grows in complexity by tying additional data to the attributes of the map. For instance, in addition to displaying the parks visually, the GIS data may also include information on the size of each park. GIS data also grows in complexity by "layering" data. For example, data for the locations of homes that sold during May of 1997 could be displayed simultaneously with the park data. Now, complex geographic relationships between the two sets of data (home transaction observations and parks) could be studied. For instance, one could program GIS to calculate the distance from each home to nearest park. In GIS, such calculations require far less time than it would take to do the same calculations using an analog map, ruler, calculator, and "brute force."

Though powerful, GIS has its limitations. One limitation is that there are some functions that certain GIS programs are not able to do. Consider the "alternative specification" discussed above Section C. ArcView GIS software, for instance, has a function that will draw a "buffer" or ring around a map attribute. Having done this, it is possible to calculate the total "park area" within the buffer. Unfortunately, there is no function in ArcView that will perform this multi-step task repeatedly for a large group of map attributes. To do so manually would be highly inefficient. Implementation of the certain specifications is limited, therefore, by the capabilities of particular GIS software programs, and so I am unable to use the superior open space specification (as described in Section C) in this study.[22]

A second limitation arises as a result of GIS data. Like other data, GIS data never live up to the "ideal." They have to be collected, converted, cut, pasted, and entered into a computer just like any other type of data. This will inevitably lead to a certain amount of inaccuracy. Furthermore, the data that are actually collected may not correspond a researcher's concept of "open space," for instance. Instead of including all "open spaces," available land use data may only include parks, cemeteries, and golf courses (as is the case for the land use data in this study). Or, the data may not be of ideal form. GIS maps and data are comprised of points, lines, and polygons. For instance, data that contain only *points* representing the *location* of parks would give be sufficient for calculating the distance between parks and homes, but not park *areas*. Calculation of areas would require that parks be represented as *polygons*. Nevertheless, GIS has proven to be an invaluable tool in examining the spatial relationships between various geographic features.

E. Previous Results

It is extremely difficult to compare the results of open space studies, as nearly all studies specify the open space variable differently. These specifications include a mixing and matching of continuous and dummy variables that attempt to measure the effects of distance, adjacency, size and type of open space on home value. Because no two studies specify the open space variable/s identically, it is very difficult to interpret their relative results. Nonetheless, Table II attempts to break apart the specification of the open space variable into measures of distance, adjacency, and size.

The previous empirical literature has demonstrated that open space in general has a positive effect on nearby home values (Bolitzer & Netusil, 2000; Anderson, 2000). These results do vary, however, depending on the type of open space. All relevant studies demonstrate that golf courses have a positive effect on nearby home values (Smith et al., 2000; Bolitzer & Netusil, 2000; Lutzenhiser & Netusil, 1999; Do & Grudnitski, 1995). Proximity to cemeteries has also been shown to have a positive effect on home value (Bolitzer & Netusil, 2000; Lutzenhiser & Netusil, 1999).

The results regarding the effect of parks and other open areas on nearby home values are much less conclusive. Natural areas, forested areas, regional parks and greenbelts all seem to have a positive effect on home values (Tyrvainen & Miettinen, 2000; Leggett, 1999; Correll, et al., 1978). Neighborhood and urban parks are another story. Specifically, Leggett (1999) finds that proximity to neighborhood parks has a depressing effect on home values.[23] Smith et al. (2000) find that proximity and adjacency to public land (publicly owned land, open access parks, greenways, and private land with conservation easements) has a negative effect. Bolitzer & Netusil (2000) find that private parks negatively effect nearby home values. Weicher and Zerbst (1973)[24], Lutzenhiser & Netusil (1999), and Bolitzer & Netusil (2000) all find, however, that neighborhood parks have an overall positive effect on home

---

[22] Like other studies, I use a “distance to nearest” proxy.
[23] Leggett’s results (1999) are “preliminary” results only.

values. Finally, in one specification of their model, Mahan, Polasky, & Adams (2000) find that proximity to parks has a negative effect; in the second specification, the effect is positive.[25]

TABLE II
Open Space Estimates

Study	Variable	Percent Effect on Home Value
DISTANCE		*% Effect Per Meter of Proximity*
Tyrvainen & Miettinen (2000)	Proximity to nearest forest area	0.7*
Smith, Poulos, & Kim (2000)	Proximity to public land	-0.002*
Smith, Poulos, & Kim (2000)	Proximity to vacant lot	0.004*
Smith, Poulos, & Kim (2000)	Proximity to golf course	0.001*
Mahan, Polasky, & Adams (2000)	Proximity to nearest wetland	0.002*
Mahan, Polasky, & Adams (2000)	Proximity to nearest park	-5.6E-4, 3.6E-4
Bolitzer & Netusil (2000)	Open space within 30, 120, 210, 300, 400, & 450 meters (dummies)	5.2, 4.1*, 3.0*, 2.3*, 2.2*, & 1.5*
Leggett (1999)	Proximity to large regional parks	2.6E-5*
Leggett (1999)	Neighborhood park within 50, 100, 150, 200, & 250 meters (dummies)	-0.05*, -0.05*, -0.03*, -0.01, & -0.02
Doss & Taff (1996)	Proximity to nearest wetland	-0.01* – 0.03*
Correll, Lillydahl, & Singell (1978)	Proximity to greenbelt	0.03*
ADJACENCY		*% Effect of Adjacency*
Bolitzer & Netusil (2000)	Open space within 450 meters	3.2*
Anderson (2000)	Adjacency to open space	20.7*
Smith, Poulos, & Kim (2000)	Adjacent to public land	-3.4
Smith, Poulos, & Kim (2000)	Adjacent to vacant lot	1.1
Smith, Poulos, & Kim (2000)	Adjacent to golf course	6.4*
Bolitzer & Netusil (2000)	Public park within 450 meters	3.4*
Bolitzer & Netusil (2000)	Private park within 450 meters	-3.8
Bolitzer & Netusil (2000)	Cemetery within 450 meters	-0.008
Bolitzer & Netusil (2000)	Golf course within 450 meters	5.2*
Do & Grudnitski (1995)	Adjacency to golf course	7.6*
Weicher & Zerbst (1973)	Adjacency to neighborhood park	7.0 – 23*
SIZE		*% Effect Per Acre*
Bolitzer & Netusil (2000)	Open space size	0.05*
Mahan, Polasky, & Adams (2000)	Size of nearest wetland	5.0E-4*, 6.3E-4*
Lutzenhiser & Netusil (1999)	Size of nearest cemetery	1.50E-04
Lutzenhiser & Netusil (1999)	Size of nearest urban park	0.002*
Lutzenhiser & Netusil (1999)	Size of nearest natural area	6.90E-04*
Lutzenhiser & Netusil (1999)	Size of nearest golf course	1.20E-05*
Lutzenhiser & Netusil (1999)	Size of nearest specialty park	0.007*

* Significant at 95% level
Note: Many coefficient estimates did not come in the form of percent effect on home value. Coefficients on linear explanatory variables with logged a dependent variable can be interpreted directly in this way. Other estimates were converted to this form at mean price and characteristic levels. In the strictest sense, coefficients for different measures of open space (i.e. size and distance) from the same regression should not be interpreted separately. It is useful, however, to see how estimates for the effects of different attributes of open space compare across studies.

[24] Weicher and Zerbst (1978) find, however, that homes backing onto parks (not facing) or homes facing park areas with heavy recreational use and/or park buildings were worth less, on average, than other homes.
[25] Neither result is statistically significant.

VI. Empirical Model

A. Conceptual Model

As we have seen, home value can be conceptualized as a function of a home's structural characteristics, the characteristics of the surrounding neighborhood, its accessibility to important geographic locations, and environmental amenities.

Price = f(housing structure, neighborhood, accessibility, amenities)

Given this conceptual relationship, it is still necessary to define a set of specific explanatory variables within each of these categories. Table III presents the explanatory variables used in this study, their definitions, and expected relationship to the dependent variable.

TABLE III

Explanatory Variables, Expected Signs, and Definitions

Variable	Expected Sign	Definition
ACRE	Positive	Size of lot in acres
FSF	Positive	Finished square feet of floor space
BED	Positive	Number of bedrooms
BATH	Positive	Number of bathrooms
AGE	Negative	Age of home in 1997
FIRE	Positive	Equals 1 if home has fireplace; 0 otherwise
TAX	Negative	Property tax on home
RACE	Negative	Percent minority population
INC	Positive	Median household income
OWN	Positive	Percentage of population living in owner-occupied housing
PROP	Positive	Median home value
DENS	Ambiguous	Population density per square mile
CBD	Negative	Distance to nearest central business district, either St. Paul or Minneapolis
COM	Negative	Distance to nearest commercial area
JOB	Negative	Distance to nearest job center
HWY	Negative	Distance to nearest major highway
PARK	Negative	Distance to nearest park
GOLF	Negative	Distance to nearest golf course
CEM	Negative	Distance to nearest cemetery
RIVER	Negative	Distance to nearest major river
LAKE	Negative	Distance to nearest lake or pond

Note for Table IV: Neighborhood characteristics (RACE, INC, OWN, PROP, and DENS) are given by census tract block group in which home is located. Units for accessibility and amenity variables are given in meters. Note that for desirable amenities, we expect home value to *decrease* with increased distance *from* amenity. So, the larger the distance value, the less a home is worth. TAX variable measures property taxes on home during 1997.

B. Explanation of Expected Signs

Intuitively, the expected signs for most of the structural variables (ACRE, FSF, BED, BATH, and FIRE) all follow the logic that "more is better." That is, additional units of these characteristics will increase the value of a home. TAX represents an additional housing expenditure that should be capitalized (negatively) into the value of a home, so the expected sign on this variable is negative. With time, I expect that a home will deteriorate structurally, so the expected sign on AGE is negative. The justification for the expected signs on the neighborhood variables is slightly less intuitive. By measuring the relationship between home values and percent minority composition of neighborhoods, the RACE variable may reflect preferences for living in racially homogenous neighborhoods or the effects of racial discrimination in the housing market. Previous studies indicate that this variable is negatively related to home value (Palmquist, 1984; Gillard, 1981; Brookshire et al., 1982). Since higher median income may proxy for neighborhood quality factors like better schools and lower crime rates, the expected sign on INC is positive. A higher percent composition of neighborhood residents living in owner-occupied housing may reflect a more stable neighborhood community, so the expected sign on OWN is positive. Since a higher value of surrounding homes will likely have a positive externality effect on home value, the sign on PROP is expected to be positive. There are both benefits (i.e. close-knit community) and costs (i.e. more congestion) to higher population density, so the expected sign on DENS is ambiguous.

The expected signs on the accessibility variables are all fairly straightforward. Since proximity to the CBD, commercial areas, job centers, and major highways all represent increased accessibility to important geographic locations, the expected signs on these variables are negative.[26] Equivalently, *proximity* to these areas will have a *positive* effect on home value. These types of locations may also generate negative effects, however. Major highways, for example, create traffic, noise, and smog. For homes close to major highways, the net effect of *proximity* may be negative.

Like the accessibility variables, the expected signs on the amenity variables are all negative.[27] Equivalently, *proximity* to these areas will have a *positive* effect on home value. Again, however, particular open space areas may generate negative externalities. Parks, for

---

[26] That is, as distance *from* these locations increases, we expect home values to decrease.
[27] That is, as distance *from* these amenities increases, we expect home values to decrease.

instance, may generate noise and congestion. For homes nearer to a park, the net effect of proximity may actually be negative. These issues must be examined empirically.

VII. Data

A. Home Transaction and Structural Data

Home sale transactions and structural characteristic data were provided by Regional Multiple Listing Service of MN, Inc. These data represent single-family residential property transactions in the seven-county, Twin Cities metropolitan area during 1997. These data include home sales price, the address of each home, and a variety of structural characteristics for approximately 30,000 observations. As these data were provided by a multiple listing service (MLS), all transactions represent "arms-length" market transactions,[28] and home value data are given by sales price. Thus, as opposed to other measures of home value, such as listings price or assessed value, these data represent actual market transactions. Since these data come from an entire 12-month period, it is possible that they are biased by fluctuations in housing market supply and demand. Table IV shows the average selling price of homes in each month of 1997 and monthly price trends (page 32) for homes of different sizes (as measured by the FSF variable).

TABLE IV

Mean Home Sale Price by Month Sold

Month of Home Sale	Mean Price	Standard Deviation
JANUARY	142568.00	95007.62
FEBRUARY	131538.00	63687.82
MARCH	135053.00	68221.26
APRIL	137782.20	76075.59
MAY	140361.40	70640.24
JUNE	144804.70	81691.59
JULY	144573.40	82351.03
AUGUST	148955.40	74655.56
SEPTEMBER	143308.20	83295.49
OCTOBER	141214.60	80382.28
NOVEMBER	142406.10	79886.39
DECEMBER	143547.60	73499.68

[28] That is, transactions between utility-maximizing consumers and profit-maximizing sellers, where the buyer seeks the lowest possible price, and the seller seeks the highest possible price. Multiple listing services collect data primarily for homes transacted through real-estate brokers. Since it is unlikely that individuals *not* engaged in arms-length transactions would employ the use of a real-estate agent, it is unlikely that MLS data would include such a transaction. See Section IV, part E for more on "arms-length" transactions.

Home Sale Monthly Price Trends

Size of Home in FSF	Monthly Trend	t-statistic
ALL SIZES	702.11	2.82**
UNDER 750	1583.02	1.27*
750-1000	-548.39	-1.82
1000-1250	326.54	1.18
1250-1500	706.30	2.80**
1500-1750	987.34	3.97**
1750-2000	338.19	1.26
2000-2500	1102.02	3.65**
ABOVE 2500	670.99	0.72

**Significant at 99% level
*Significant at 90% level
Note: Sales price was regressed on month index to find monthly trend in home sales price in each of the categories. Constant terms have been omitted.

The results in Table IV indicate that there is some variation in home prices during 1997. Average sales price during the early months of the year seems slightly lower than at the end of the year. Indeed, there is an overall upward trend in home sales price of approximately $700 per month. Breaking this down by home size (as measured by the total finished square-footage of the home), this upward trend in home sales price is relatively consistent homes of all sizes.[29] So, while the upward trend may create some bias in overall home values, the bias should be fairly consistent across all home types. This means that hedonic coefficient estimates should be *comparatively* unbiased.

Most of the structural characteristic data are free of major problems. One possible exception is the ACRE variable. The home transaction data included two types of lot size data: lot dimensions and lot acreage. The lot dimensions data were incomplete and inconsistent, making them unusable for most of the sample. While the acreage data were more complete, nearly all observations had zero acreage. Since all non-zero values were of one acre or above, it seems this variable was probably reserved to record the acreage of large lots only. Thus, the acreage data will not fully measure variation in lot size, especially for properties with smaller lots (since no acreage data was recorded for these properties).

[29] With the exception of homes of 750 – 1000 square feet. These experience a negative downward trend of approximately $550 per month.

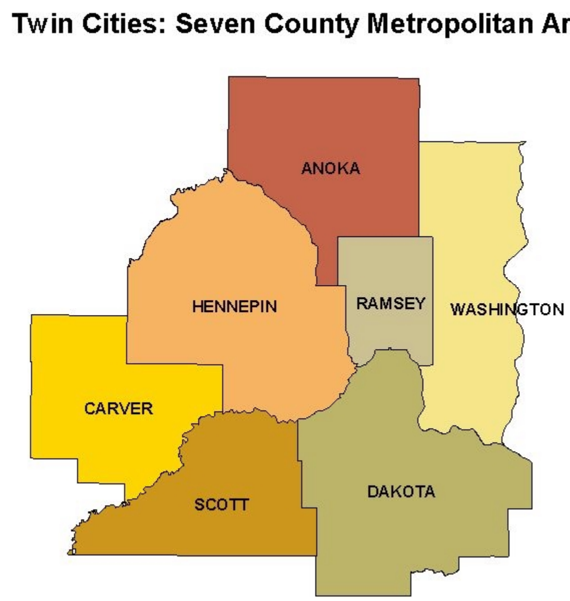


Figure 4

B. The Geo-coding Process

Using ArcView GIS software and a Twin Cities metropolitan area street address database[30], I geo-coded (located geographically in GIS) home observations for Dakota, Ramsey, and Washington County (Figure 4).[31] Some observations had missing, incomplete, or inconsistent address information. As a result, a number of home transaction observations (approximately 5%) were not geo-coded or "unmatched." This resulted in approximately 10,000 (out of 10,574) geo-coded or "matched" observations for the three counties. The sales price and locations of the matched homes are shown in Figure 5 on page 34.

Table V (page 35) compares the homes that were "matched" with the homes that were "unmatched." On average, unmatched observations had higher sales price, were larger (as measured by FSF), had larger lots, and were newer. This makes sense. If the address data used to geo-code the homes were incomplete, it would likely be *more* incomplete in rapidly developing areas (i.e. the suburbs and rural areas) where the address data has not had time to "catch up" with the development. Homes in these areas generally have higher value, are larger, have larger lot sizes, and are newer (obviously). I conclude, therefore, that the sample of homes used in the analysis that follows will be slightly biased toward urban homes and older suburban homes.

---

[30] Street address database provided by The Lawrence Group.

[31] Geo-coding can be a time consuming process. It was for this reason that I geo-coded observations from only three counties. Since these three counties include inner city, suburban, and rural areas, however, they are a fairly representative cross-section of the Twin Cities metropolitan housing market.

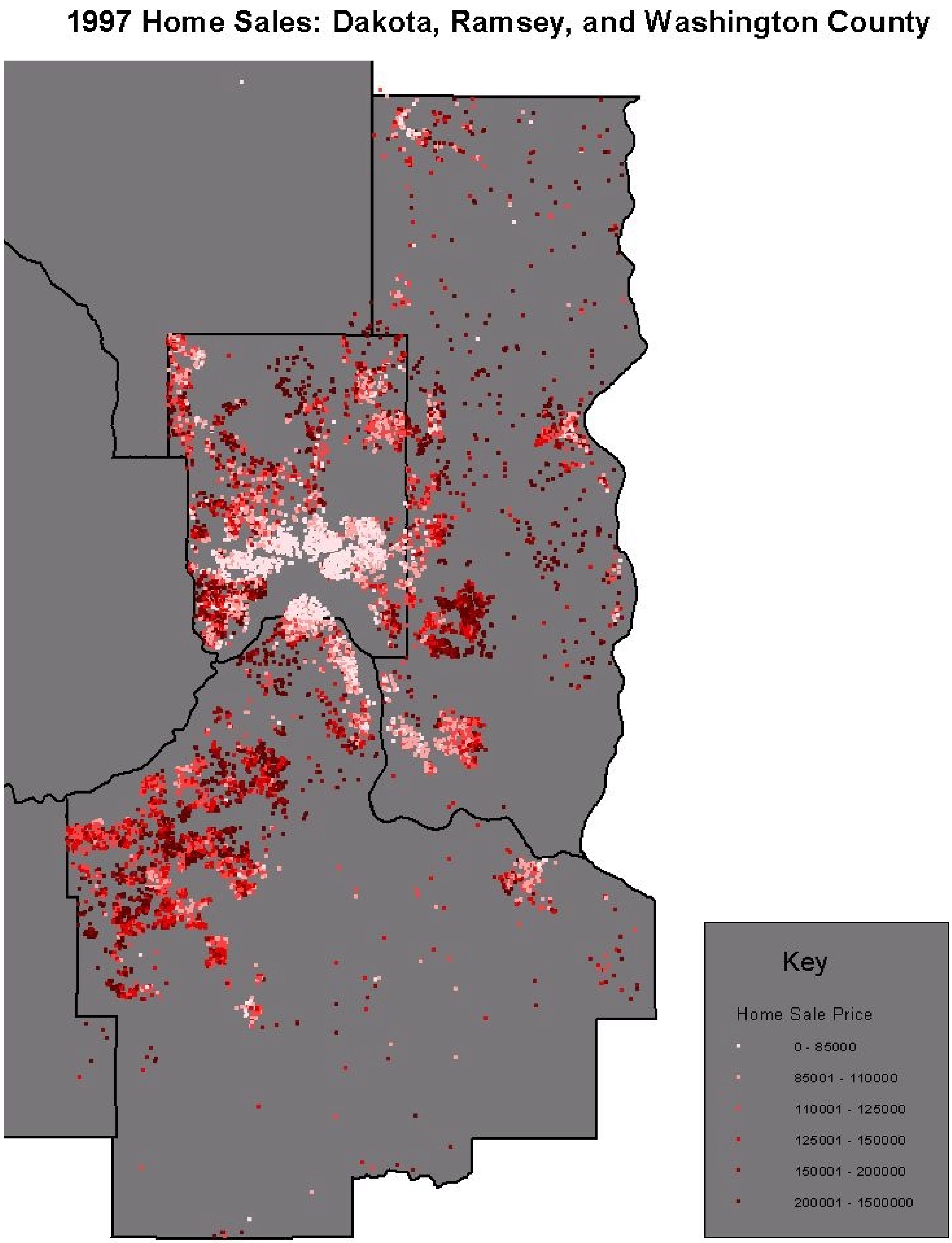


Figure 5: Home Sale Price and Location

TABLE V

Comparison of Matched and Unmatched Observations

	Matched (9990)		Unmatched (582)	
Variable	Mean	Standard Deviation	Mean	Standard Deviation
PRICE	141032.40	76985.40	161931.90	92450.60
FSF	1852.44	806.53	1940.51	877.31
ACR	0.32	3.67	0.58	2.68
AGE	32.92	34.26	20.58	38.25

C. Neighborhood Characteristics Data[32]

All neighborhood characteristic variables come from 1990 U.S. Census block group data for the Twin Cities metropolitan area. These data were spatially matched to the home transaction data using ArcView GIS. Aggregated by census tract block group, these data may fail to represent the actual characteristics of particular neighborhoods. Furthermore, these data are from 1990, while my home transaction data are from 1997. Block groups are relatively small in size, however, and census demographic data are fairly constant over time. Thus, it is likely that these data adequately reflect the relative characteristics of particular neighborhoods in 1997.

D. Accessibility Data[33]

Accessibility data come from a variety of sources, though all proximity *measures* were generated using ArcView GIS. I located the central business districts of Minneapolis and St. Paul using a GIS map of roads in the Twin Cities metropolitan area.[34] Data for commercial centers come from two sources. I located major retail centers using The Lawrence Group GIS data for regional landmarks, which included major shopping centers. I combined these data with the Metropolitan Council's 1997 Generalized Land Use data, which maps out areas of commercial land use (and other land uses) in the Twin Cities metropolitan area, to generate a GIS map of major shopping malls and commercial areas. I located major employment centers by combining Metropolitan Council commercial land use

[32] See Appendix III for neighborhood characteristics data maps.
[33] See Appendix III for accessibility locations data map.
[34] Road data is from The Lawrence Group street and address data.

data and MN Department of Economic Security Job Density data into a GIS map.[35] Finally, highway data come from a MN Department of Transportation GIS map of major highways. Using these four maps and the geo-coded locations of my home transaction data, I calculated the distance from each home to the nearest CBD, commercial center, employment center, and major highway in ArcView GIS.

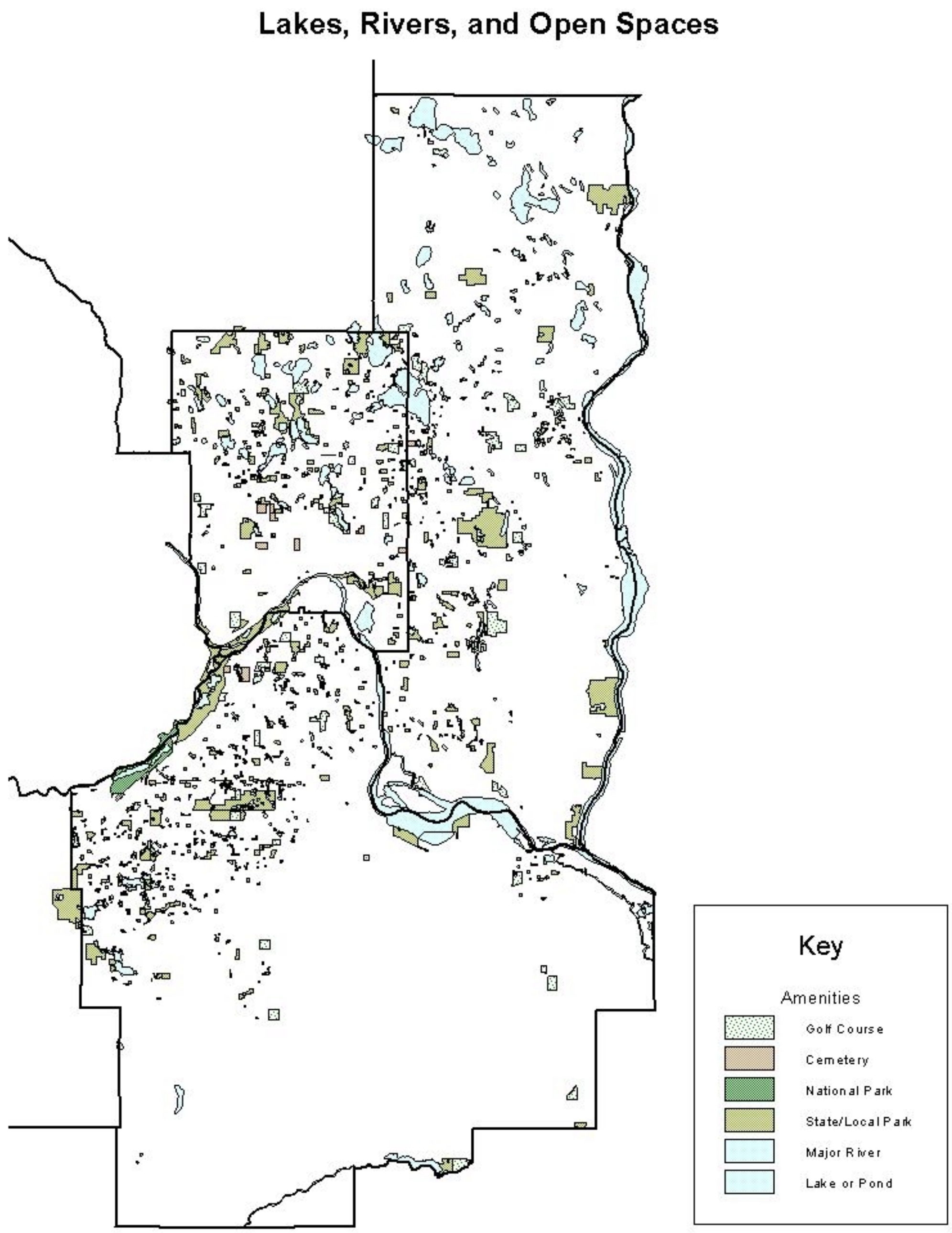


Figure 6: Amenity Data

[35] MN Department of Economic Security data are from 1990. I obtained them in map form from a January 1996 Metropolitan Council report entitled, “Growth Options for the Twin Cities Metropolitan Area.”

E. Amenity Data

My land use data for parks, golf courses, cemeteries, rivers, and lakes come from The Lawrence Group GIS data for regional landmarks.[36] I calculated the distance from each home observation to the nearest area of each type using ArcView GIS. It is possible that these data do not include other important open space amenity areas—like open areas provided by colleges and universities, public schools, etc. Nevertheless, these data *do* include parks, cemeteries, and golf courses, which all may provide open space benefits. The locations of these amenities can be seen in Figure 6 on page 36.

Summary statistics for the dependent and independent variables are presented in Table VI below.

TABLE VI

Descriptive Statistics of Variables

10574 Observations

Variable Name	Mean	Standard Deviation
PRICE	142162.90	78038.81
ACRE	0.33	3.63
FSF	1857.21	810.71
BED	3.27	0.95
BATH	2.15	0.88
AGE	32.25	34.60
FIRE	5306.00 observations	0.50
TAX	1837.45	1774.84
RACE	5.77	8.63
INC	42984.19	13606.53
OWN	82.75	17.46
PROP	97790.91	31349.19
DENS	3521.16	3121.17
CBD	15223.51	9594.79
HWY	1239.43	1214.02
COM	4646.68	3488.80
JOB	6833.52	5704.21
PARK	612.77	955.02
GOLF	2044.01	1467.49
CEM	5768.70	4881.44
RIVER	6082.82	4190.52
LAKE	1535.86	1497.24

Note: Variable Definitions are same as in Table III.

[36] The Lawrence Group data also include other regional landmarks that could be considered “open space” such as the State Fairgrounds. However, in order to generalize, it was necessary to classify open space areas by type. Parks, golf courses, and cemeteries we chosen for this reason. Data are from year 2000.

VIII. Empirical Results

A. Initial Results

Table VII shows simple correlation coefficients among the dependent and independent variables. Note the high correlation between FSF (finished square feet) and PRICE (sales price). This is not a surprising result, as home size is known to be a primary component of home value. Also notice the strong correlation between many of the explanatory variables. Again, this result is expected. Strong multicollinearity among the independent variables, however, may inflate the standard errors of their estimated coefficients.

TABLE VII

Simple Correlation Coefficients

	PRICE	ACRE	FSF	BED	BATH	AGE	FIRE	TAX	RACE	INC	OWN	PROP	DENS
PRICE	1.00												
ACRE	0.09	1.00											
FSF	0.76	0.05	1.00										
BED	0.42	0.01	0.55	1.00									
BATH	0.69	0.04	0.74	0.54	1.00								
AGE	-0.36	-0.01	-0.30	-0.22	-0.35	1.00							
FIRE	0.43	0.01	0.43	0.27	0.45	-0.24	1.00						
TAX	0.51	0.04	0.51	0.27	0.44	-0.13	0.28	1.00					
RACE	-0.20	-0.01	-0.11	-0.05	-0.12	0.24	-0.14	-0.09	1.00				
INC	0.55	0.05	0.47	0.29	0.47	-0.46	0.37	0.33	-0.34	1.00			
OWN	0.23	0.04	0.17	0.10	0.19	-0.33	0.15	0.09	-0.43	0.62	1.00		
PROP	0.62	0.05	0.53	0.31	0.51	-0.39	0.39	0.41	-0.22	0.83	0.38	1.00	
DENS	-0.40	-0.05	-0.33	-0.20	-0.35	0.59	-0.25	-0.16	0.40	-0.56	-0.42	-0.50	
CBD	0.26	0.05	0.22	0.17	0.23	-0.45	0.15	0.08	-0.35	0.36	0.34	0.27	-0.62
COM	0.02	0.09	0.00	-0.02	-0.02	-0.10	-0.07	-0.03	-0.11	0.04	0.14	-0.03	-0.21
JOB	0.06	0.09	0.00	-0.01	-0.01	-0.15	-0.06	-0.05	-0.17	0.07	0.19	-0.01	-0.31
HWY	0.16	0.02	0.12	0.12	0.16	-0.31	0.12	0.02	-0.11	0.30	0.32	0.21	-0.29
PARK	-0.05	0.12	-0.07	-0.08	-0.09	0.13	-0.09	-0.05	0.10	-0.13	-0.08	-0.13	0.12
GOLF	-0.03	0.08	-0.05	-0.02	-0.05	-0.03	-0.05	-0.08	0.07	-0.06	0.03	-0.12	-0.11
CEM	0.24	0.04	0.18	0.14	0.19	-0.40	0.12	0.09	-0.26	0.33	0.29	0.25	-0.49
RIVER	0.18	0.02	0.12	0.10	0.16	-0.42	0.09	0.03	-0.22	0.26	0.27	0.20	-0.48
LAKE	-0.11	0.00	-0.14	-0.08	-0.16	0.13	-0.10	-0.07	0.04	-0.15	-0.04	-0.18	0.12

	CBD	COM	JOB	HWY	PARK	GOLF	CEM	RIVER	LAKE
CBD	1.00								
COM	0.39	1.00							
JOB	0.64	0.66	1.00						
HWY	0.37	0.02	0.12	1.00					
PARK	0.03	0.36	0.31	0.15	1.00				
GOLF	0.24	0.25	0.28	0.20	0.37	1.00			
CEM	0.83	0.28	0.44	0.42	0.08	0.21	1.00		
RIVER	0.55	0.07	0.21	0.36	-0.04	0.10	0.65	1.00	
LAKE	0.08	0.25	0.23	0.10	0.36	0.23	0.23	-0.06	1.00

TABLE VIII

Regression Results: Model I
Dependent Variable is Ln(PRICE)

Variable	Coefficient	Standard Error	t-statistic	MIP
CONTANT	10.510010	0.063428	165.70*	
ACRE	0.003837	0.000630	6.10*	545.52
FSF	0.000164	0.000005	34.51*	23.34
BED	0.014156	0.003148	4.50*	2012.46
BATH	0.100800	0.004200	24.00*	14329.99
AGE	-0.001663	0.000090	-18.51*	-236.47
FIRE	0.102560	0.005334	19.23*	14934.56
TAX	0.000016	0.000002	10.13*	2.22
RACE	-0.006224	0.000325	-19.14*	-884.84
INC	0.000001	0.000000	1.76†	0.10
OWN	-0.001028	0.000188	-5.48*	-146.13
PROP	0.000003	0.000000	22.00*	0.46
DENS	-0.000006	0.000001	-4.92*	-0.84
LCBD	0.030379	0.006619	4.59*	0.28
LCOM	-0.021371	0.004066	-5.26*	-0.65
LJOB	0.013235	0.004805	2.76*	0.28
LHWY	0.016613	0.002222	7.48*	1.91
LPARK	0.010241	0.001879	5.45*	2.38
LGOLF	-0.004708	0.002539	-1.86†	-0.33
LCEM	0.028114	0.003321	8.47*	0.69
LRIVER	-0.019434	0.003267	-5.95*	-0.45
LLAKE	0.001817	0.002631	0.69	0.17
Number of Observations				9991
F-statistic				1484.74
Adjusted R-squared				0.76

* Significant at 99% level
‡ Significant at 95% level
† Significant at 90% level
Note: “MIP” refers to marginal implicit price; calculated at mean home value and attribute quantity.

Table VIII presents the results of the semi-log estimation, which outperformed other models in terms of best fit (highest adjusted R-squared). All the proximity variables in this model are logged (i.e. LPARK = Ln(PARK)),[37] and correspond to the variable definitions in Table III. The coefficient estimates for these logged variables measure the elasticity of home value with respect to the explanatory variables (i.e. the percent change in home value resulting from a percent change in the attribute level). The other coefficient estimates measure the percent change in sales price resulting from a *unit* change in the explanatory

[37] For some observations, value for proximity variable was zero. But Ln(0) is undefined. Thus, for observations where distance was less than one meter, I reset distance to one meter and *then* computed the logged value.

variable.[38] To ease the interpretation of these results, I translated all coefficient estimates into marginal implicit prices at mean home value and attribute levels.[39] I refer to these estimated marginal implicit prices in the discussion that follows.

All the coefficient estimates for the structural variables are statistically significant and, with the exception of TAX, have the expected sign. The marginal implicit price of an additional acre of land is approximately $550, as measured by the ACRE variable.[40] The coefficient on FSF implies that each additional square foot of finished floor space increases home value by $23. Each bedroom adds approximately $2010 to home value, as measured by BED; a bathroom adds $14,330, as measured by BATH. The coefficient on AGE implies that home value decreases by approximately $240 with each passing year. The coefficient on FIRE implies that a home with a fireplace is worth an additional $14,930. Finally, an additional dollar of yearly property tax adds $2.20 to the value of a home.[41]

The coefficient estimates for the neighborhood variables are also significant, and most have the expected signs. The coefficient on RACE suggests that homes in neighborhoods with a one-percent higher minority (non-white) population are worth $880 less, on average. Though significant at just the 90% level, the coefficient on INC implies that a $1,000 increase in neighborhood median income increases home value by approximately $100. Contrary to expectations, the estimate on OWN suggests that home value falls by approximately $150 with 1% increase in the neighborhood's population that lives in owner-occupied housing. A $1,000 increase in the neighborhood's median property value increases a home's sale price by approximately $460, as suggested by the coefficient on PROP. Finally, the coefficient on DENS implies that home value decreases by approximately $84 with an increase in population density of 100 persons per square mile.

---

[38] Except for FIRE. The estimated percent impact of a dummy variable with a logged dependent variable is given by $b' = 100*((e^{b - V(b)/2}) - 1)$ where b is the estimated coefficient and $V(b)$ is the estimated variance of b. This estimate is developed in Halvorsen, R. and Palmquist, P., "The Interpretation of Dummy Variables in Semilogarithmic Equations", American Economic Review, Vol. 70, 1980, pp. 474-475 and Kennedy, P., "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations", American Economic Review, Vol. 71, 1981, p. 801.

[39] With a logged dependent variable, the marginal implicit price of logged explanatory variables is given by $P*b/X$, where P is home value, b is the coefficient estimate, and X is the quantity of independent variable. The marginal implicit price of the linear continuous variables is given by $P*b$. The marginal implicit price of dummy variables is given by $P*b'$, where b' is the same as above.

[40] This estimate likely underestimates the marginal implicit price of an acre of land. This may be due to the problems with the ACRE variable, which are discussed in Section VII, part A.

[41] This result may seem counterintuitive since property taxes represent a housing expenditure that should be capitalized (negatively) in the value of a home. It may be, however, that TAX is picking up some positive effect that is correlated with high property taxes (i.e. good public schools in neighborhood, better provision of public goods in neighborhood, etc).

The coefficient estimates on the accessibility variables are all statistically significant and imply that proximity to the CBD (LCBD), job centers (LJOB), and major highways (LHWY) decreases home value by $280, $280, and $1910 dollars per kilometer, respectively. While it makes sense that proximity to highways may be a disamenity, the coefficient estimates on LCBD and LJOB are unexpected.[42] The coefficient estimate on LCOM implies that proximity to a shopping center or commercial area increases home value by $650 per kilometer.

Finally, I turn to the coefficient estimates measuring the effect of open space and other environmental amenities on home value. The coefficient on LPARK implies that proximity to a park decreases home value by approximately $240 per 100 meters.[43] Though less significant, the coefficient on LGOLF suggests that proximity to a golf course increases home values by $33 per 100. Proximity to cemeteries has a statistically significant negative impact of $70 per 100 meters, as measured by LCEM. Not surprisingly, proximity to rivers increases home value by $45 per 100 meters, as measured by LRIVER. Finally, though statistically insignificant, the coefficient on LLAKE implies that proximity to lakes has the unexpected effect of decreasing home value by $17 per 100 meters.

I also estimated a second version of Model I in which I replaced the environmental amenity proximity variables with “adjacent” dummy variables. For each observation, these dummies were set equal to one if the home was within 200 meters of the amenity (i.e. PARKADJ = 1 if PARK < 200). The results of this regression can be found in Appendix II. With few exceptions, this regression’s estimates for the structural, neighborhood, and accessibility coefficients are consistent with the estimates of the original Model I (Table VIII). The coefficient estimates on the amenity variables are all statistically significant. A park within 200 meters decreases home value by approximately $4,700. A golf course within 200 meters increases home value by $2,900. A cemetery within 200 meters decreases home value by $12,300. A river within 200 meters increases home value by $30,300, and a lake increases home value by $8,100. These estimates corroborate the results above.[44]

---

[42] Although, since many people hate to work, the coefficient on JOB may make sense, after all (ha ha).

[43] This does not necessarily mean that parks *cause* home values to decrease. In fact, the causal relationship may be reversed. As Leggett (1999) discusses, parks may be “*placed* in geographic areas that are less desirable for residential use...that there may be something inherently undesirable about the area” (p. 8) in which a park is located.

[44] And reassure me by showing that a lake has a positive effect on home value.

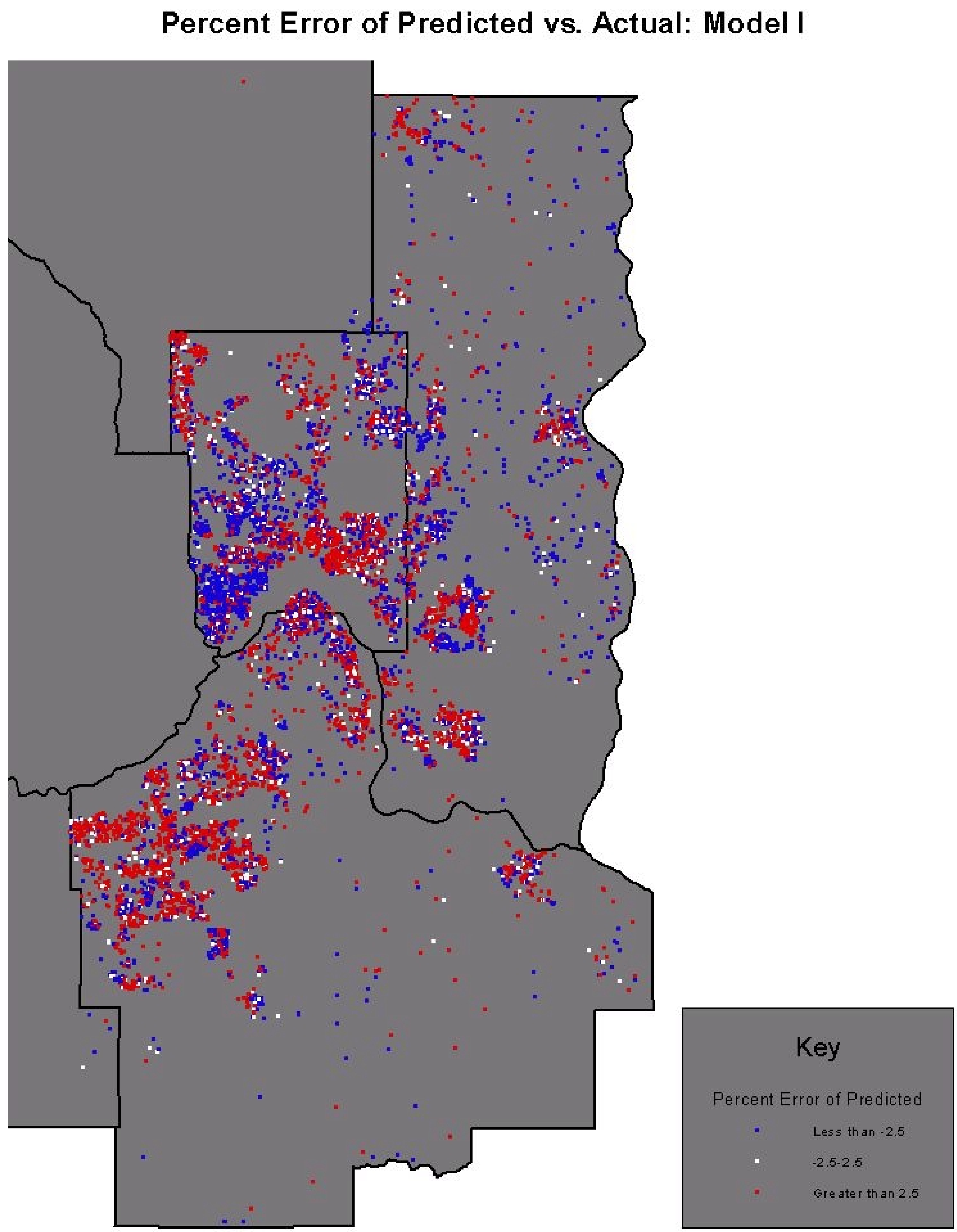


Figure 7

Figure 7 plots the percent error of home prices as predicted by Model I. Predictions between –2.5 and 2.5% error are represented by white dots. Blue dots represent cases in which the model under-predicted actual home value by more than 2.5%. Red dots represent cases in which the model over-predicted home value by more than 2.5%.[45]

Figure 7 shows that there is much spatial autocorrelation in the error of Model I's predicted values. For example, the model under-predicts home value in the southwest corner of Ramsey County (the area near and around Macalester College) and over-predicts home value in southeast portion of Ramsey County (near and around downtown St. Paul). Similarly, there are clumped areas of over-prediction and under-prediction throughout the three-county area. This spatial autocorrelation suggests that Model I does not adequately account for spatial factors in the metropolitan housing market. These factors could be market segmentation or possibly an omitted variable. Or, perhaps Model I overlooks "something special" about particular neighborhoods. The Macalester-Groveland neighborhood, for instance, is *perceived* and has a *reputation* for being a safe, friendly, desirable place to live. Summit Avenue, a beautiful tree-lined avenue with many beautiful and historic homes runs through the neighborhood. There are good schools. There is a strong and vibrant community. As a result, homebuyers are attracted to the area and bid up the price of homes. This "perception" may not be adequately addressed by Model I.[46]

B. Disaggregation of the Full Data Set

In order to control for spatial variation that may have been unaccounted for in Model I, I first disaggregated the full data set by county and estimated separate hedonic functions for each county separately. The results of this model (Model II) are given in Appendix II.[47] Then, assuming city/suburban market segmentation, I disaggregated the full sample into observations within the city of St. Paul and observations outside the city. Model III estimates separate hedonic price functions for these "city" and "suburban" sub-markets.[48] The results of Model III are presented in Table IX.[49]

---

[45] Percent error of prediction is given by *100*(P' – P)/P*, where *P* is actual home value and *P'* is predicted home value.

[46] The neighborhood variables (RACE, INC, OWN, PROP, & DENS) are intended to proxy for neighborhood effects. There is little variation within these variables in most of the sample however, as the majority of observations (75%) are located in the relatively homogeneous suburbs.

[47] Regression results and a map of percent errors of Model II are presented in Appendix II. These regression results and map may be interpreted the same as in Model I.

[48] By running regressions for each county separately.

[49] Coefficient estimates and marginal implicit prices may be interpreted the same as in Table VIII.

TABLE IX

Regression Results: Model III
Dependent Variable is Ln(PRICE)

	City of St. Paul				Suburbs			
Variable	Coefficient	SE	t-statistic	MIP	Coefficient	SE	t-statistic	MIP
CONS	9.245606	0.255901	36.13*		11.038990	0.07	158.63*	
ACRE	-0.000216	0.001329	-0.16	-30.74	0.004705	0.000638	7.38*	668.85
FSF	0.000086	0.000009	9.50*	12.15	0.000200	0.000005	39.33*	28.40
BED	0.083177	0.006893	12.07*	11824.64	-0.006947	0.003165	-2.20‡	-987.53
BATH	0.065593	0.008428	7.78*	9324.91	0.106645	0.004419	24.13*	15160.96
AGE	-0.001575	0.000163	-9.69*	-223.96	-0.001844	0.000099	-18.61*	-262.21
FIRE	0.124104	0.011510	10.78*	17860.98	0.068948	0.005406	12.75*	9736.57
TAX	0.000068	0.000006	11.85*	9.68	0.000006	0.000001	4.37*	0.89
RACE	-0.003547	0.000414	-8.57*	-504.24	-0.003887	0.000762	-5.10*	-552.52
INC	0.000001	0.000001	0.64	0.08	0.000001	0.000000	2.12‡	0.11
OWN	-0.000619	0.000373	-1.66†	-88.04	-0.000794	0.000205	-3.88*	-112.81
PROP	0.000004	0.000000	10.86*	0.56	0.000002	0.000000	12.18*	0.26
DENS	0.000001	0.000002	0.51	0.14	-0.000026	0.000002	-15.79*	-3.72
LCBD	0.198434	0.018992	10.45*	1.85	-0.021289	0.008406	-2.53‡	-0.20
LCOM	-0.101417	0.009771	-10.38*	-3.10	0.007191	0.004229	1.70†	0.22
LJOB	0.058097	0.015290	3.80*	1.21	0.022738	0.004968	4.58*	0.47
LHWY	0.033219	0.004528	7.34*	3.81	0.014290	0.002407	5.94*	1.64
LPARK	-0.015250	0.005733	-2.66*	-3.54	0.010878	0.001851	5.88*	2.52
LGOLF	-0.022696	0.005743	-3.95*	-1.58	-0.001256	0.002589	-0.49	-0.09
LCEM	0.027547	0.006185	4.45*	0.68	0.004940	0.004576	1.08	0.12
LRIVER	-0.032350	0.008987	-3.60*	-0.76	0.000360	0.003444	0.11	0.01
LLAKE	0.021934	0.009479	2.31‡	2.03	-0.020403	0.002557	-7.98*	-1.89
Number of Observations				2594				7397
F-statistic				353.34				987.64
Adjusted R-squared				0.74				0.74

* Significant at 99% level
‡ Significant at 95% level
† Significant at 90% level
Note: coefficient estimates must be interpreted within the context of their respective sub-market.

There are many notable differences in the coefficient estimates between the two market segments. First, the coefficient estimate on the ACRE variable is much larger and more significant in the suburbs than in the city of St. Paul.[50] This makes intuitive sense. Within the city, there is much less variation in lot size. In the suburbs, where lot sizes vary to a greater extent, it is not surprising that ACRE is a more powerful determinant of home value. For the same reason, home size (as measured by the FSF variable) is also more important in the suburbs, which is consistent with previous literature (Orford, 2000). Specifically, an additional foot of floor space is worth only $12 in the city; it is worth almost

$30 in the suburbs. Bedrooms increase home value by approximately $11,800 in the city, while decreasing home value by almost $1,000 in the suburbs.[51] Bathrooms are more important in the suburbs ($15,200 vs. 9,300), and fireplaces are more important in the city ($17,900 vs. $9,700). Taxes have a much larger effect on home value in the city than in the suburbs. In the city, an additional dollar of yearly tax increases home value by almost $10; in the suburbs the effect is only $0.90. This makes intuitive sense. In the city, there is a much larger variation in neighborhood type. If taxes somehow partially capture the effect of "neighborhood quality," then it is not surprising the effect of the TAX variable is stronger in the city of St. Paul.

Among the neighborhood variables, RACE and INC have similar effects in both the city and the suburbs. In neighborhoods with a 1% higher minority (non-white) population, home values are approximately $500 less on average in both the city and the suburbs. A $1,000 increase in median neighborhood income increases home value in the city and suburbs by $80 and $110, respectively. There are, however, notable differences among the neighborhood variables. Neighborhood median property value has a much stronger effect on sale price in the city than in the suburbs. Given a $1,000 increase in neighborhood median home value, sales price increases by $560 in the city and by only $260 in the suburbs. This makes intuitive sense, because there is much more variation in home value in the city. Density has a small positive (though insignificant) effect on home values in the city. In the suburbs, the effect is significant and negative.

Again, there are many notable differences in the accessibility variables. Proximity to the CBD has a negative effect on home values in the city ($185 per 100 meters), while the effect is positive in the suburbs ($20 per 100 meters). This makes intuitive sense. In the city, homes are already close enough to the CBD to take full advantage of its benefits, but proximity magnifies the negative effects (increased noise, congestion, etc). In the suburbs, there are fewer negative side effects of being closer to the CBD, so proximity has a positive effect on home values. In the city, proximity to commercial areas has a positive effect on home values ($10 per 100 meters), while the effect is negative in the suburbs ($20 per 100 meters). Proximity to job centers and highways has a negative effect on home values in both the city and suburbs, though in both cases the effect is stronger in the city. Job centers have a

---

[50] As before, I refer here to the estimated marginal implicit prices.

[51] The unexpected sign may be a result of multicollinearity with the other structural variables. Professor Gary Krueger of Macalester College also notes that there is little variation in the number of bathrooms in the suburbs, particularly Washington County.

negative effect of $380 per 100 meters in the city, and $50 per 100 meters in the suburbs. Highways have a negative effect of $700 per 100 meters in the city, and $160 per 100 meters in the suburbs. In the case of highways, this makes intuitive sense, since city homes are already "saturated" with access to highways. In this context, it is logical that proximity to highways represents a stronger disamenity.

I now turn to a discussion of the amenity variable coefficients, where there are also some notable differences between the city and suburbs. Proximity to parks has a positive effect of $266 per 100 meters in the city and a negative effect of $252 per 100 meters in the suburbs. Proximity to golf courses has a positive effect of $395 per 100 meters in the city and $49 per 100 meters suburbs, though the effect is insignificant in the suburbs.[52] Proximity to cemeteries has a negative effect of $68 per 100 meters in the city and $12 per 100 meters in the suburbs. Again, the effect is insignificant in the suburbs.[53] Proximity to rivers has a positive effect of $360 per 100 meters in the city and virtually no effect on home value in the suburbs. Finally, the effect of proximity to a lake is negative $203 per 100 meters in the city and positive $189 in the suburbs.[54]

As before, I also estimate a second version of Model III in which I replace the environmental amenity proximity variables with "adjacent" dummy variables. The results of this regression can be found in Appendix II. For the most part, the estimates for this regression's structural, neighborhood, and accessibility coefficients are consistent with the estimates of the original Model III (Table IX). The estimated coefficient estimates for the amenity variables are also consistent with the results above. In the city, a park within 200 meters increases home value by $2,650; in the suburbs, a park within 200 meters decreases home value by $3,480. A golf course within 200 meters has a positive effect of $3,420 in the city and $3,560 in the suburbs. A cemetery within 200 meters has a negative effect of $5,610 in the city and $11,580 in the suburbs. A river within 200 meters has a positive effect of approximately $10,000 in both the city and suburbs. Finally, a lake within 200 meters has a negative effect of $5,300 in the city and a positive effect of $10,000 in the suburbs.

Figure 8 on page 47 plots the percent error of home value as predicted by Model III. As before, predictions between –2.5 and 2.5% error are represented by white dots. Blue dots represent cases in which Model III under-predicted actual home value by more than 2.5%.

---

[52] The relatively smaller number of golf courses outside the city may explain this insignificance.
[53] As above, the relatively smaller number of cemeteries outside the city may explain this insignificance.

Red dots represent cases in which the model over-predicted home value by more than 2.5%. The city of St. Paul is outlined in white.

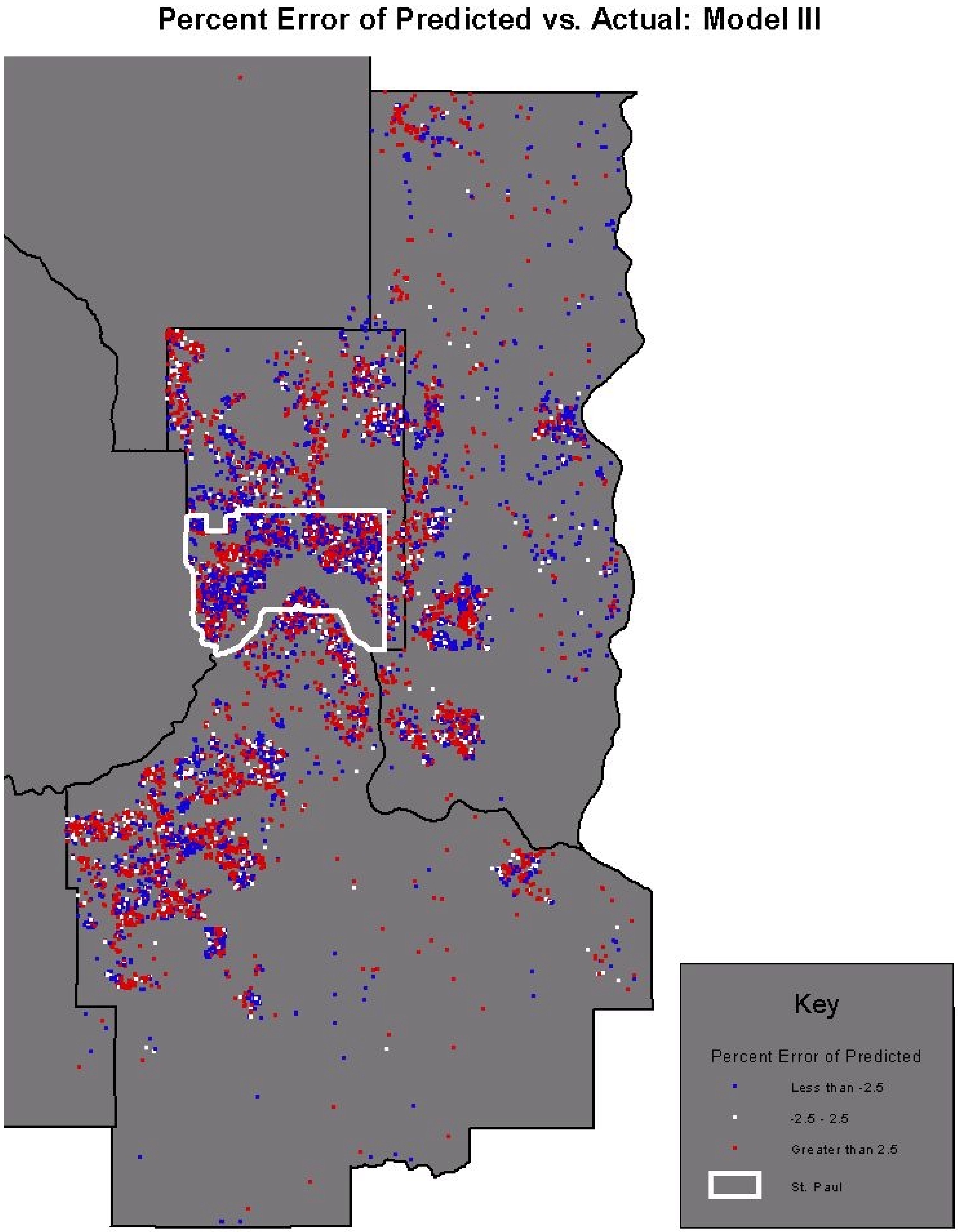


Figure 8

[54] Compared to the suburbs, there are very few lakes and ponds in the city of St. Paul. The homes around these particular areas are not of high value. Though the reason for these homes' low value may be something other than proximity to a lake, the proximity measure is picking up this negative effect.

Comparing this map to its counterpart from Model I (see Figure 7), it is clear that there is less spatial autocorrelation in the segmented model. This improvement is especially evident within the city of St. Paul itself. Model I consistently under-predicts home values on the west side of St. Paul and over-predicts home values on the east side. In Model III, this pattern of autocorrelation is less distinct. Outside the city of St. Paul, it is difficult to see any improvement. This is not surprising. There are significantly more observations outside the city of St. Paul than within, so the suburbs are less likely to be affected by the segmentation. Thus, Model III is probably superior for predicting home values within the city of St. Paul. Model III may also be superior for predicting suburban home values, though only slightly so.

There are two possible explanations for the improvement demonstrated by Model III. First, it may be that the ability of the neighborhood variables to control for spatial relationships is hindered in the full sample by the relative homogeneity of the suburbs. Overall, neighborhood characteristics are of little importance in Model I, since the suburban observations dominate the full sample. But this leads to spatial autocorrelation in the city of St. Paul due to its relative spatial heterogeneity. In the segmented model, the neighborhood characteristics adequately control for spatial heterogeneity, as they are unhindered by the homogeneity of the suburbs. A second explanation is that there actually *is* city/suburban segmentation in the housing market.[55] The differences in marginal implicit prices between the city and suburban sub-market regressions in Model III support this hypothesis.

C. Comparison and Discussion of Results

The estimated effects of open space on home value in this study are fairly consistent with prior research (see Table II). My estimates for the effect of adjacency to a golf course range from 2.1 – 2.5%.[56] These estimates are slightly lower than the 5.2 – 7.6% range set forth by the previous literature. I estimate that proximity to golf courses increases home value by 6.3E-5 – 0.001% per meter. This is within the 0.001 – 0.03% range of previous estimates. I estimate that adjacency to cemeteries decreases home by 3.95 – 8.64%. This is effect is stronger than the previous estimate of a negative 0.008%.[57] In the case of parks, I

[55] The existence of city/suburban market segmentation does not imply that there is not a more complex pattern of market segmentation, which may in fact be the case.
[56] I convert my coefficient estimates to "percent effects" in order to compare to previous studies. Estimate ranges come from Model I and Model III.
[57] No previous studies estimate the effect of proximity to a cemetery. But, at a distance of 500 meters, my proximity estimates imply an *adjacency* range of negative 0.24% to negative 042%, which is a stronger effect than the previous adjacency estimate.

estimate the effect of adjacency to range between negative 3.3% and positive 1.86% per meter. This is consistent with prior estimates, which range from negative 3.8% to positive 23%. My estimate for the effect of proximity to parks is also consistent with previous estimates. I estimate the effect to be approximately negative 00018% to positive 0.0025% per meter. Previous estimates range from negative 0.002 to positive 0.03%.[58]

My results are interesting, however, because they show that open space is valued differently in the city than in the suburbs. This difference is most distinct in the case of parks, where proximity has a *positive* effect in the city and *negative* effect in the suburbs. Previous researchers have argued that the disamenity effects of proximity to urban parks, such as noise, congestion, and other negative externalities, may outweigh the benefits (Weicher & Zerbst, 1973; Leggett, 1999). This would explain the negative effect in the suburbs. Why, then, is the effect positive in the city? There are several possible explanations. In the city, where there is much more heterogeneity and density, the concept of "neighborhood" may be of increased significance. Proximity to a park, therefore, may serve as a proxy for two effects. First, the straightforward amenity (or disamenity) associated with parks in general. But also, proximity to parks may act as a "good neighborhood" signal to prospective homebuyers, giving them the *perception* that the neighborhood is a desirable place to live.[59] Another plausible explanation is that city parks are *fundamentally different* than parks in the suburbs. The causes of this difference in valuation need to be examined in greater detail.

IX. Conclusions

A. Limits and Qualifications

In this paper, I employ hedonic regression analysis to estimate the effect of proximity to parks, golf courses, and cemeteries on residential home values in the St. Paul, MN area. Unfortunately, there are several limitations that must be considered in tandem with these estimates. First, these estimates are only as reliable as the data from which they are derived. In addition to other issues,[60] my open space data only include public parks, golf courses, and

---

[58] Note that it is harder to compare estimates of the effect of parks, since there is significant inconsistency across studies of what constitutes a "park."

[59] Michaels and Smith (1990) use a similar argument in distinguishing between the types of disamenity effects associated with hazardous waste sites.

[60] Other limiting data issues include: temporally separated home transaction (1997), neighborhood characteristic (1990), location (1990 – 1997), and amenity (2000) data; aggregation of neighborhood data by census tract; "unmatched" observations during geo-coding process; problems with ACRE variable; monthly home value trends. See Section VII.

cemeteries. As a result, my land use data fail to account for other important open space areas (like the green space along Summit Avenue, the State Fairgrounds, the open areas at Macalester College, etc.). Furthermore, because my specification classifies open space areas by type, it does not consider relevant differences between particular open space areas in the same category. For instance, it may be that some parks generate positive externalities, while others generate negative externalities.[61] My model does not consider these factors. Finally, my measure of proximity to open space areas is not ideal. Measuring only the "distance to nearest," this proxy does not measure other important features of the area (i.e. open space size) and may fail to capture the externality effects of other nearby open space amenities.

In addition to these data and specification issues, my estimation of the hedonic price function results in noticeable spatial autocorrelation, which may indicate biased coefficient estimates. This spatial autocorrelation is not as severe after the implementation of market segmentation in Model III. While this and other evidence support the segmented markets hypothesis,[62] however, I have no *a priori* reason to believe that the city of St. Paul and its suburbs operate as distinct sub-markets. Furthermore, some spatial autocorrelation persists even in Model III, despite the reduction. This autocorrelation may arise as the result of an omitted variable, or the failure of my model to capture some other complex spatial relationship.

Finally, my hedonic estimates represent only a partial measure of the total economic benefits (or costs) of open space preservation. Specifically, the estimates of this study yield important information regarding individuals' residential preferences and the externalities of open space amenities in a residential setting. They do not, however, measure many public good benefits, such as reduced soil erosion, wildlife habitat, or improved water quality, since these benefits are unlikely to be reflected in the value of a home.

B. Future Research

While this study presents useful information regarding the value *of* and preferences *for* open space amenities in St. Paul, MN, there is still room for further research. The hypothesis of city/suburban market segmentation and the reasons for different coefficient

---

[61] A city-suburban spatial pattern in such effects could explain the difference in the marginal implicit price of parks between the city and suburban sub-markets.

[62] Such as the severe spatial autocorrelation in Model I, the reduced amount of spatial autocorrelation in Model III, and the varying coefficient estimates between the city and suburban sub-markets.

estimates between city and suburban sub-markets need to be examined in greater detail.[63] The use of GIS as a means to understanding complex spatial relationships in urban housing markets needs to be better incorporated with housing market and hedonic theory.[64] Finally, issues regarding the ideal specification of the open space variable and the technical capabilities of GIS need to be reconciled. The calculation of an open space variable that simultaneously incorporates the effects of size, proximity, and type of open space would be a significant step in this direction.

C. Contributions

In spite of these limitations, this study makes important contributions. While there are similar studies in other cities, this is the first that explicitly examines the effect of open space on home values in the St. Paul metropolitan area.[65] Like Smith et al. (2000), Bolitzer & Netusil (2000), and Lutzenhiser & Netusil (2000), I examine and compare the different effects of several types of open space (including parks, golf courses, and cemeteries) on home value. Breaking from most previous open space literature, however, I examine the effect of open space on home values across a relatively large geographic area that includes three counties. There are disadvantages associated with examining such a large area—namely, the possibility of uncontrolled-for spatial relationships. By examining a large area, however, I am able to study the effect of geographic context (city vs. suburb) on spatial relationships in a metropolitan housing market.[66]

On a more technical note, this study is the first open space study I know of that uses GIS software to examine the spatial autocorrelation of errors in predicted home value.[67] This study confirms the usefulness of the technique (which is a fairly easy task in ArcView GIS).[68] By visually examining the spatial autocorrelation of the model's predicted values, I am able to quickly distinguish the model's inadequacy in accounting for spatial relationships.

---

[63] This difference may be the consequence of segmented markets. Or, differences in types of park amenities may be correlated with the city/suburban split.

[64] Orford (1999) provides a huge first step in this direction.

[65] With the notable exception of "The effect of open space on single-family, residential home property values" Anderson (2000), which won the Spring 2000 MN Economics Association award for best term paper. An extremely small sample size (approximately 65 observations) and atrociously poor home value and amenity data weaken the results in this paper, however. Doss and Taff (1996) study the effects of *wetlands* on home value in St. Paul, not *open space* per se.

[66] See Orford (1999) or Orford (2000) for a more detailed treatment of this issue.

[67] Waddell (1993) is the only residential hedonic study I know of that does this, plotting the "percent residuals" of the model spatially on a GIS map of the study area.

It would be interesting to see if these problems exist for other residential hedonic and open space studies, as well.

Finally, this is the first open space study that examines the different effects of open space on home value according to city/suburban market segmentation. My results demonstrate that there may be fundamental differences in the way open space amenities are valued and perceived, depending on whether a home is located in the city or in the suburbs. Namely, proximity to parks has a positive effect on home values in the city ($354 or 0.25% per 100 meters) and a negative effect in the suburbs ($254 or 0.18% per 100 meters). Although the reasons for these differences need to be examined in greater detail, they nevertheless demonstrate the importance of considering context when modeling complex spatial relationships in residential housing markets.

D. Concluding Remarks

As urban areas continue to grow, develop, and expand, issues concerning the preservation of open space areas will continue to demand attention. As this happens, information regarding the valuation of open space amenities will become more and more relevant. Though there is significant room for future research, this study has provided important preliminary evidence regarding the value *of* and preferences *for* open space amenities in the St. Paul metropolitan area.

---

[68] This simply involved estimating the hedonic price function in STATA (a statistical software program), calculating the percent errors and pasting them into a spreadsheet, and then joining this spreadsheet to the geo-coded home location data in ArcView.

Bibliography

Agarwal, M. K. and B. T. Ratchford. (1980). "Estimating Demand Functions for Product Characteristics: The Case of Automobiles." *Journal of Consumer Research*, Vol. 7: 249 – 262.

Anderson, S. T. (2000). "The effect of open space on single-family, residential home property values." *Unpublished paper*. Macalester College.

Arguea, N. M. and C. Hsiao. (1993). "Econometric issues of estimating hedonic price functions." *Journal of Econometrics*, Vol. 56: 243 – 267.

Benson, E. D., J. L. Hansen, A. L. Schwartz, Jr., and G. T. Smersh. (1998). "Pricing Residential Amenities: The Value of a View." *Journal of Real Estate Finance and Economics*, Vol. 16 (1): 55 – 73.

Blomquist, G. and L. Worley. (1981). "Hedonic Prices, Demands for Urban Housing Amenities, and Benefit Estimates." *Journal of Urban Economics*, Vol. 9: 212 – 221.

Bolitzer, B. and N. R. Netusil. (2000). "The Impact of Open Spaces on Property Values in Portland, Oregon." *Journal of Environmental Management*, Vol. 59: 185 – 193.

Breffle, W. S., E. R. Morrey, and T. S. Lodder. (1998). "Using Contingent Valuation to Estimate a Neighborhood's Willingness to Pay for Undeveloped Urban Land." *Urban Studies*, Vol. 35 (4): 715 – 727.

Brookshire, D. S., M. A. Thayer, W. D. Schulze, and R.C. d'Arge. (1982). "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches." *The American Economic Review*, Vol. 72 (1): 165 – 177.

Chattopadhyay, S. (1999). "Estimating the Demand for Air Quality: New Evidence Based on the Chicago Housing Market." *Land Economics*, Vol. 75 (1): 22 – 38.

Correll, M. R., J.H Lillydahl and L. D. Singell. (1978). "The Effects of Greenbelts on Residential Property Values: Some Findings on the Political Economy of Open Space." *Land Economics*, Vol. 54 (2): 207 – 217.

Cropper, M. L., B. D. Leland, and K. E. McConnell. (1988). "On the choice of functional form for hedonic price functions." *The Review of Economics and Statistics*, Vol. 70: 668 – 675.

Do, A. Quang and G. Grudnitski. (1995). "Golf Courses and Residential House Prices:: An Empirical Examination." *Journal of Real Estate Finance and Economics*, Vol. 10: 261 – 270.

Doss, C.R. and S. J. Taff. (1996). "The influence of wetland type and wetland proximity on residential property values." *Journal of Agricultural and Resource Economics*, Vol. 21: 120 – 129.

Fausold, C. J. and R. J. Lilieholm. (1996). "The Economic Value of Open Space." *Land Lines*, Vol. 8 (5).

Frech III, H.E. and R. N. Lafferty. (1984). "The Effect of the California Coastal Commission on Housing Prices." *Journal of Urban Economics*, Vol. 16 (1): 105 – 123.

Freeman, A. M. III. (1993). *The Measurement of Environmental and Resource Values: Theory and Methods*. Johns Hopkins University Press, Baltimore, MD.

Gillard, Q. (1981). "The Effect of Environmental Amenities on House Values: The Example of a View Lot." *Professional Geographer*, Vol. 33 (2): 216 – 220.

Kaszuba, Mike. (2000). "A groundbreaking green space vote." *Star Tribune*. Monday, October 30.

Lee, C. M. and P. Lineman. (1998). "Dynamics of the Greenbelt Amenity Effect on Land Market – the Case of Seoul's Greenbelt." *Real Estate Economics*, Vol. 28 (1): 107 – 129.

Leggett, C. G. (1999). "The Effects of Neighborhood Parks on Residential Property Values." *Unpublished paper*, Department of Agricultural and Resource Economics, University of Maryland, August.

Lutzenhiser, M. and N. R. Netusil (1999). "The Effect of Open Space Type and Proximity on a Home's Sale Price: Portland, Oregon." *Unpublished paper*. Reed College.

Mahan, B., S. Polasky, and R. Adams. (2000). "Valuing Urban Wetlands: A Property Price Approach." *Land Economics*, 76 (1): 100 – 113.

Michaels, R. G. and K. V. Smith. (1990). "Market segmentation and valuing amenities with hedonic models: The case of a hazardous waste site." *Journal of Urban Economics*, Vol. 28: 223 – 242.

Orford, S. (1999). *Valuing the Built Environment : GIS and House Price Analysis.* (Ashgate Publishing Company). Brookfield, Vermont.

Orford, S. (2000). "Modeling Spatial Structures in Local Housing Market Dynamics: A Multilevel Perspective." *Urban Studies*, Vol. 37 (9): 1643 – 1671.

Palmquist, R. B. (1984). "Estimating the Demand for the Characteristics of Housing." *The Review of Economics and Statistics*, Vol. 66 (3): 394-404.

Palmquist, R. B. (1991). "Hedonic Methods" in *Measuring the Demand for Environmental Quality*, (J. B. Braden and C. D. Kolstad, Eds.). North-Holland, Amsterdam: 77 – 120.

Palmquist, R. B. (1992). "Valuing Localized Externalities." *Journal of Urban Economics*, Vol. 31: 59 – 68.

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

Smith, V. K., C. Poulos, and H. Kim. (2000). "Treating Open Space as an Urban Amenity." *Unpublished Paper*. Department of Agricultural and Resource Economics, North Carolina State University.

Tyrvainen, L. and A. Miettinen. (2000). "Property Prices and Urban Forest Amenities." *Journal of Environmental Economics and Management*, Vol. 39: 205 – 223.

Waddell, Paul. (1993). "Residential property values in a multinodal urban area: new evidence on the implicit price of location." *Journal of Real Estate and Finance Economics*, Vol. 7: 117 – 141.

Weicher, J. C. and R. H. Zerbst. (1973). "The Externalities of Neighborhood Parks: An Empirical Investigation." *Land Economics*, Vol. 49: 99 – 105.

Appendix I: Justification of Optimizing Conditions

The Consumer Problem

Consumers maximize utility

$$U(x, z_1, z_2, ..., z_n; \alpha) \tag{A.1}$$

Subject to the budget constraint

$$y = x + p(z_1, z_2, ..., z_n) \tag{A.2}$$

Combining (A.1) and (A.2), we set up the Lagrangian.

$$L(x, z_i, \lambda) = U(x, z_1, z_2, ..., z_n; \alpha) - \lambda \cdot [y - x - p(z_1, z_2, ..., z_n)] \tag{A.3}$$

First order conditions are given by differentiating (A.3) with respect to x, z_i, and λ.

$$\frac{\partial L}{\partial x} = \frac{\partial U}{\partial x} + \lambda = 0 \tag{A.4}$$

$$\frac{\partial L}{\partial z_i} = \frac{\partial U}{\partial z_i} + \lambda \cdot \frac{\partial p(Z)}{\partial z_i} = 0 \tag{A.5}$$

$$\frac{\partial L}{\partial \lambda} = y - x - p(z_1, z_2, ..., z_n) = 0 \tag{A.6}$$

Solving for -λ in (A.4) and (A.5) and equating, we have

$$\frac{\partial U}{\partial x} = \frac{\frac{\partial U}{\partial z_i}}{\frac{\partial p(Z)}{\partial z_i}} \tag{A.7}$$

Rearranging (A.7), we have

$$\frac{\frac{\partial U}{\partial z_i}}{\frac{\partial U}{\partial x}} = \frac{\partial p(Z)}{\partial z_i} \tag{A.8}$$

The Producer Problem

Consumers have revenue

$$M \cdot p(Z) \tag{A.9}$$

and costs

$$c(M, Z; \beta) \tag{A.10}$$

Combining (A.1) and (A.2), producers maximize profits

$$\pi(M, z_i, \delta) = M \cdot p(Z) - \delta \cdot c(M, Z; \beta) \tag{A.11}$$

First order conditions are given by differentiating (A.3) with respect to M, z_i, and δ.

$$\frac{\partial \pi}{\partial M} = p(Z) - \delta \cdot \frac{\partial c}{\partial z_i} = 0 \tag{A.12}$$

$$\frac{\partial \pi}{\partial z_i} = M \cdot \frac{\partial p(Z)}{\partial z_i} - \delta \cdot \frac{\partial c}{\partial z_i} = 0 \tag{A.13}$$

$$\frac{\partial \pi}{\partial \delta} = c(M, Z; \beta) = 0 \tag{A.14}$$

Solving for δ in (A.12) and (A.13) and equating,

$$\frac{p(Z)}{\frac{\partial c}{\partial M}} = \frac{M \cdot \frac{\partial p(Z)}{\partial z_i}}{\frac{\partial c}{\partial z_i}} \tag{A.15}$$

Rearranging (A.7), optimizing conditions are given by

$$\frac{\frac{\partial c}{\partial z_i}}{M} \cdot p(Z) = \frac{\partial p(Z)}{\partial z_i} \cdot \frac{\partial c}{\partial M} \tag{A.16}$$

Assuming that price is set equal to marginal cost,

$$p(Z) = \frac{\partial c}{\partial M} \tag{A.17}$$

Optimizing conditions are given by

$$\frac{\frac{\partial c}{\partial z_i}}{M} = \frac{\partial p(Z)}{\partial z_i} \tag{A.18}$$

Appendix II: Additional Regression Results

Regression Results: Model I with Adjacent Terms

Dependent Variable is Ln(PRICE)

Variable	Coefficient	t-statistic	MIP
CONSTANT	10.383260	188.18*	
ACRE	0.003942	6.27*	560.38
FSF	0.000163	34.26*	23.20
BED	0.015955	5.06*	2268.18
BATH	0.100136	23.83*	14235.65
AGE	-0.001619	-18.25*	-230.13
FIRE	0.103106	19.30*	15414.83
TAX	0.000015	9.84*	2.16
RACE	-0.006388	-19.92*	-908.11
INC	0.000001	2.68*	0.15
OWN	-0.001201	-6.44*	-170.79
PROP	0.000003	22.49*	0.46
DENS	-0.000003	-2.59*	-0.44
LCBD	0.048040	8.69*	0.45
LCOM	-0.022204	-5.49*	-0.68
LJOB	0.020958	4.45*	0.44
LHWY	0.015769	7.13*	1.81
PARKADJ	-0.031036	-5.81*	-4712.22
GOLFADJ	0.026746	2.07‡	2910.97
CEMADJ	-0.080046	-3.86*	-12288.79
RIVERADJ	0.214270	5.08*	30294.26
LAKEADJ	0.060362	6.34*	8128.93
Number of Observations			9991
F-statistic			1478.45
Adjusted R-squared			0.76

* Significant at 99% level
‡ Significant at 95% level
† Significant at 90% level
Note: MIP refers to marginal implicit price; calculated at mean home value and attribute quantity.

Regression Results: Model II

Dependent Variable is Ln(PRICE)

	Dakota County			Ramsey County			Washington County		
Variable	Coefficient	t-stat	MIP	Coefficient	t-stat	MIP	Coefficient	t-stat	MIP
CONS	11.025230	133.29*		10.246380	71.37*		9.991824	28.56*	
ACRE	0.001780	2.09‡	253.01	0.000637	0.51	90.61	0.005823	5.82*	827.81
FSF	0.000201	31.80*	28.59	0.000118	15.57*	16.75	0.000189	18.67*	26.87
BED	-0.007711	-2.13‡	-1096.25	0.047582	9.34*	6764.35	-0.003414	-0.43	-485.39
BATH	0.090484	16.68*	12863.45	0.086385	12.92*	12280.77	0.118419	12.58*	16834.77
AGE	-0.002806	-20.29*	-398.84	-0.001702	-12.59*	-242.00	-0.001391	-7.38*	-197.75
FIRE	0.062789	9.53*	8714.61	0.127384	14.49*	18605.19	0.069068	6.10*	9306.49
TAX	0.000000	-0.22	-0.06	0.000055	13.99*	7.86	0.000007	2.81*	0.94
RACE	0.001268	0.98	180.19	-0.005190	-13.48*	-737.84	-0.003836	-2.30‡	-545.29
INC	0.000000	0.24	0.02	-0.000004	-5.32*	-0.50	0.000003	4.03*	0.48
OWN	-0.000231	-0.71	-32.88	0.000334	1.11	47.47	-0.001853	-4.36*	-263.41
PROP	0.000002	8.49*	0.27	0.000004	14.92*	0.57	0.000002	5.07*	0.21
DENS	-0.000016	-7.47*	-2.29	0.000002	1.13	0.26	-0.000035	-8.67*	-4.93
LCBD	-0.041548	-3.75*	-0.39	0.162309	11.77*	1.52	0.222717	3.68*	2.08
LCOM	0.009109	1.34	0.28	-0.098947	-11.88*	-3.03	0.038221	4.37*	1.17
LJOB	0.009264	1.10	0.19	0.035073	3.42*	0.73	-0.095799	-2.91*	-1.99
LHWY	-0.000768	-0.23	-0.09	0.028946	7.71*	3.32	0.023810	4.69*	2.73
LPARK	0.012164	4.94*	2.82	-0.006499	-1.84†	-1.51	0.007409	2.04‡	1.72
LGOLF	-0.002576	-0.76	-0.18	-0.005820	-1.38	-0.40	-0.015573	-2.75*	-1.08
LCEM	0.037595	4.65*	0.93	0.017805	3.57*	0.44	-0.011685	-1.09	-0.29
LRIVER	0.014696	2.51‡	0.34	-0.067372	-9.11*	-1.57	-0.038250	-5.00*	-0.89
LLAKE	-0.018618	-5.56*	-1.72	-0.007838	-1.35	-0.73	-0.017650	-2.71*	-1.63
Number of Observations			3748			4107			2136
F-statistic			619.06			679.70			232.73
Adjusted R-squared			0.78			0.78			0.70

* Significant at 99% level
‡ Significant at 95% level
† Significant at 90% level

Percent Error of Predicted vs. Actual: Model II

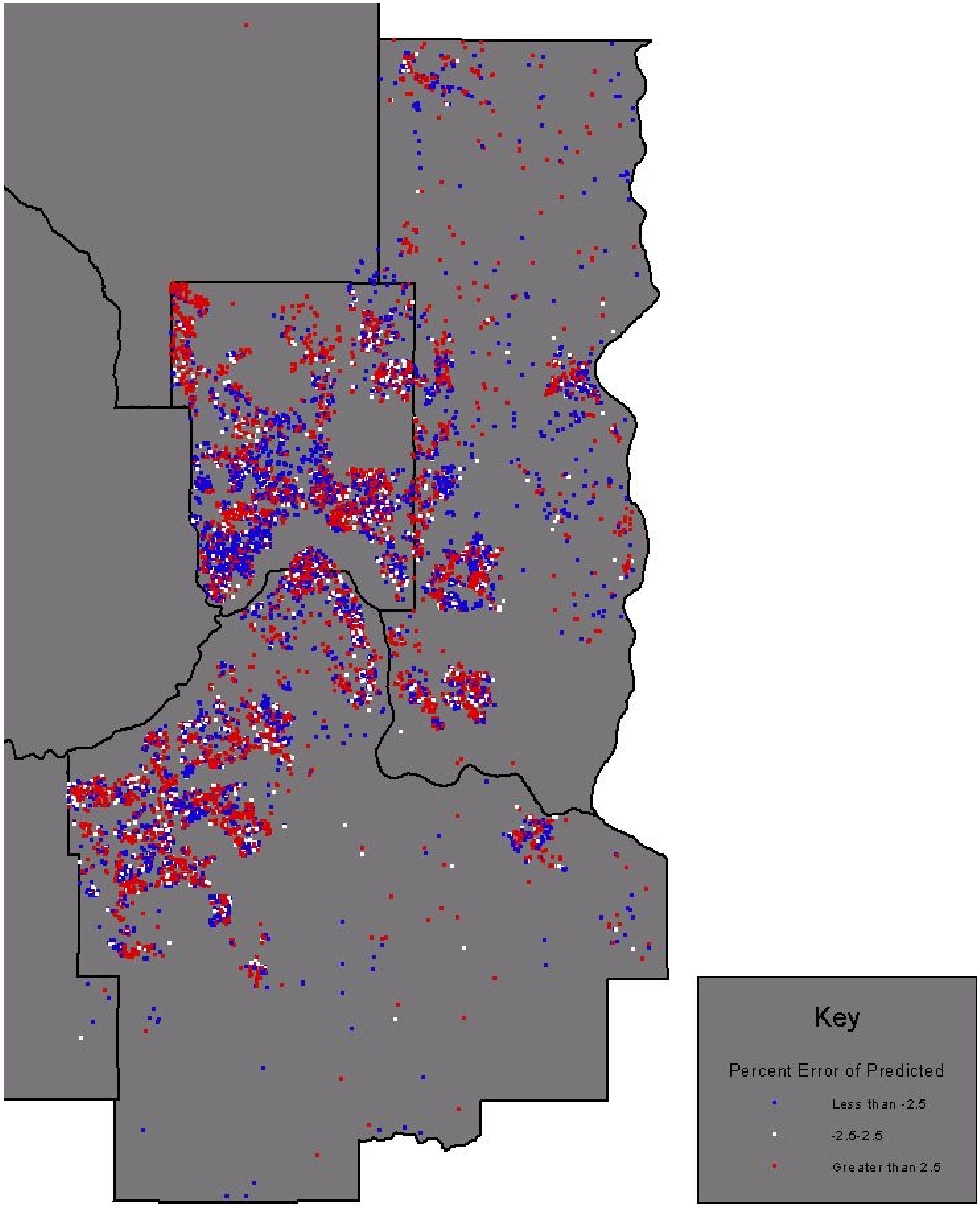


TABLE IX

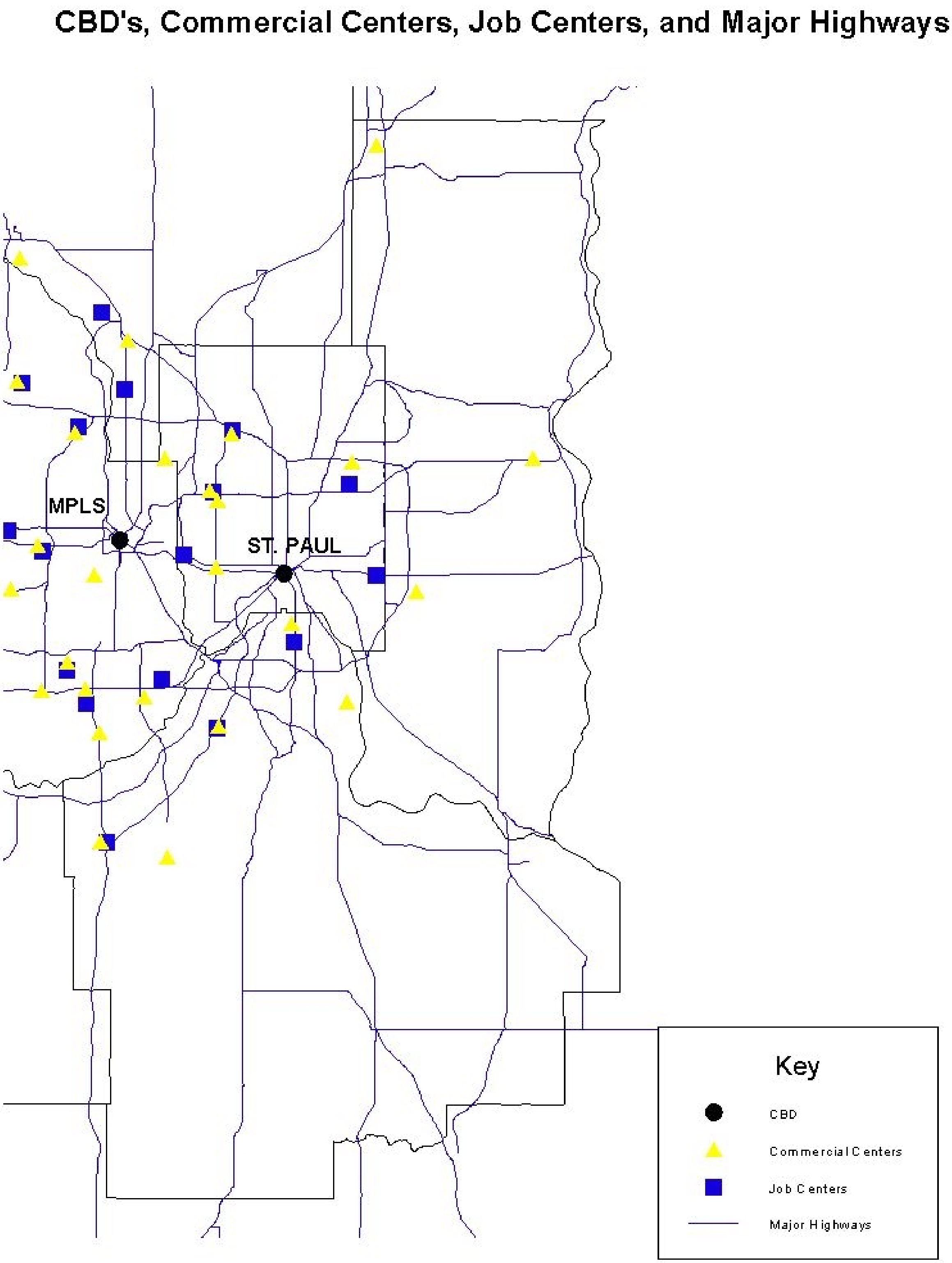
Regression Results: Model III with Adjacent Terms

Dependent Variable is Ln(PRICE)

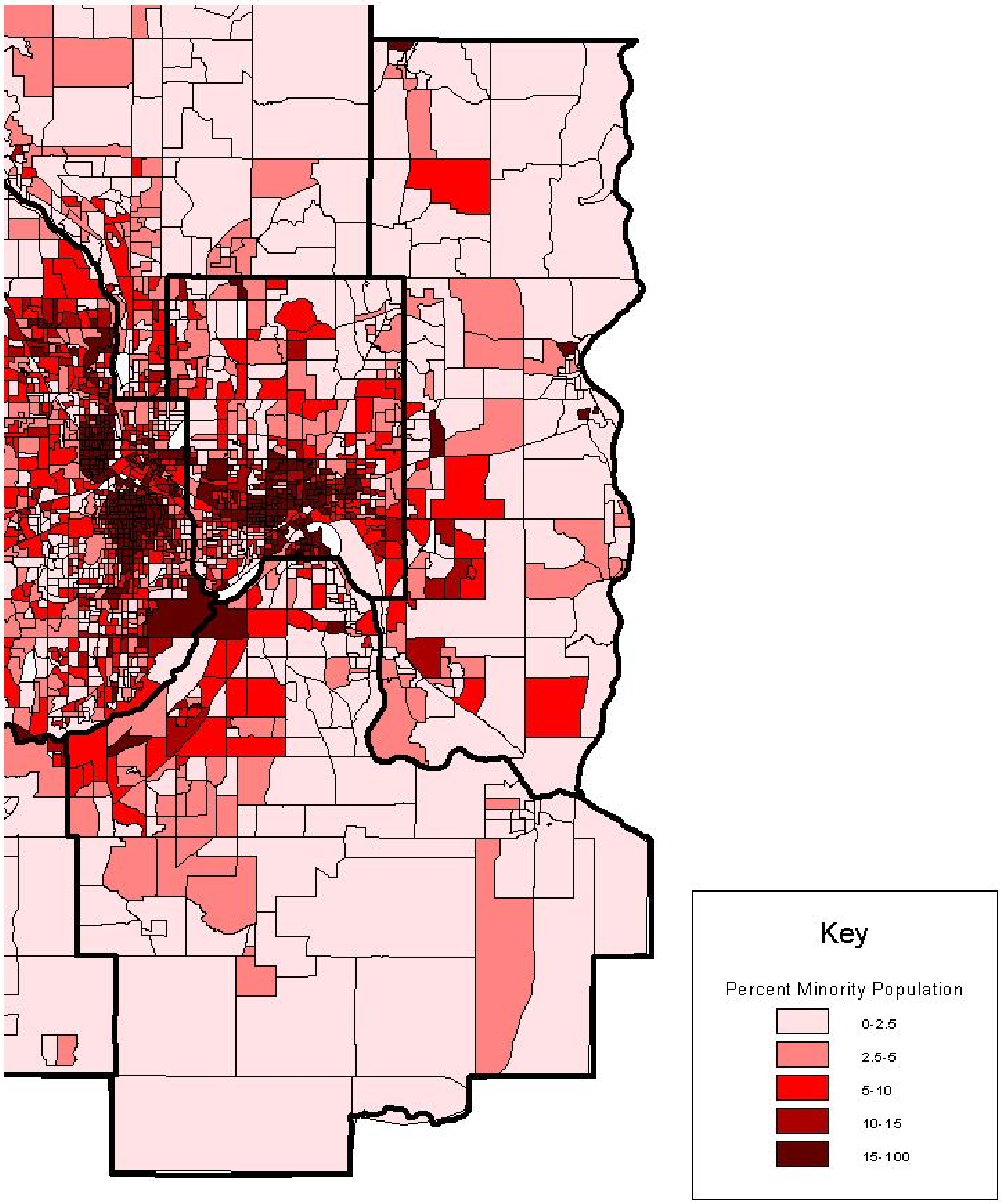
	City of St. Paul				Suburbs			
Variable	Coefficient	SE	t-statistic	MIP	Coefficient	SE	t-statistic	MIP
CONSTANT	9.115298	0.221543	41.15*		10.989090	0.058492	187.87*	
ACRE	-0.000287	0.001355	-0.21	-40.77	0.005005	0.000635	7.89*	711.54
FSF	0.000084	0.000009	9.17*	11.94	0.000199	0.000005	39.18*	28.25
BED	0.083297	0.007030	11.85*	11841.67	-0.005757	0.003169	-1.82†	-818.39
BATH	0.061562	0.008592	7.17*	8751.82	0.107372	0.004414	24.33*	15264.33
AGE	-0.001393	0.000164	-8.50*	-198.08	-0.001877	0.000097	-19.34*	-266.81
FIRE	0.133743	0.011696	11.44*	19395.87	0.068062	0.005407	12.59*	9601.78
TAX	0.000074	0.000006	12.57*	10.45	0.000006	0.000001	4.13*	0.84
RACE	-0.004112	0.000415	-9.90*	-584.53	-0.004257	0.000757	-5.62*	-605.12
INC	0.000002	0.000001	2.19‡	0.28	0.000001	0.000000	2.15‡	0.12
OWN	-0.001232	0.000373	-3.31*	-175.07	-0.000900	0.000203	-4.43*	-127.92
PROP	0.000004	0.000000	12.02*	0.62	0.000002	0.000000	13.12*	0.28
DENS	0.000003	0.000002	1.56	0.42	-0.000026	0.000002	-16.03*	-3.75
LCBD	0.209759	0.018024	11.64*	1.96	-0.015888	0.006347	-2.50‡	-0.15
LCOM	-0.105935	0.009298	-11.39*	-3.24	0.005281	0.004207	1.26	0.16
LJOB	0.048906	0.015084	3.24*	1.02	0.019203	0.004837	3.97*	0.40
LHWY	0.022725	0.004482	5.07*	2.61	0.014235	0.002378	5.99*	1.63
PARKADJ	0.027864	0.018816	1.48	2648.13	-0.022280	0.005003	-4.45*	-3479.70
GOLFADJ	0.037836	0.028133	1.35	3419.67	0.031257	0.013099	2.39‡	3556.20
CEMADJ	-0.027605	0.025323	-1.09	-5610.73	-0.067177	0.035648	-1.88†	-11584.66
RIVERADJ	0.106907	0.077127	1.39	10055.61	0.103518	0.045697	2.27‡	11943.51
LAKEADJ	-0.018437	0.039156	-0.47	-5302.86	0.076437	0.008756	8.73*	10622.20
Number of Observations				2594				7397
F-statistic				335.50				988.75
Adjusted R-squared				0.73				0.74

* Significant at 99% level
‡ Significant at 95% level
† Significant at 90% level

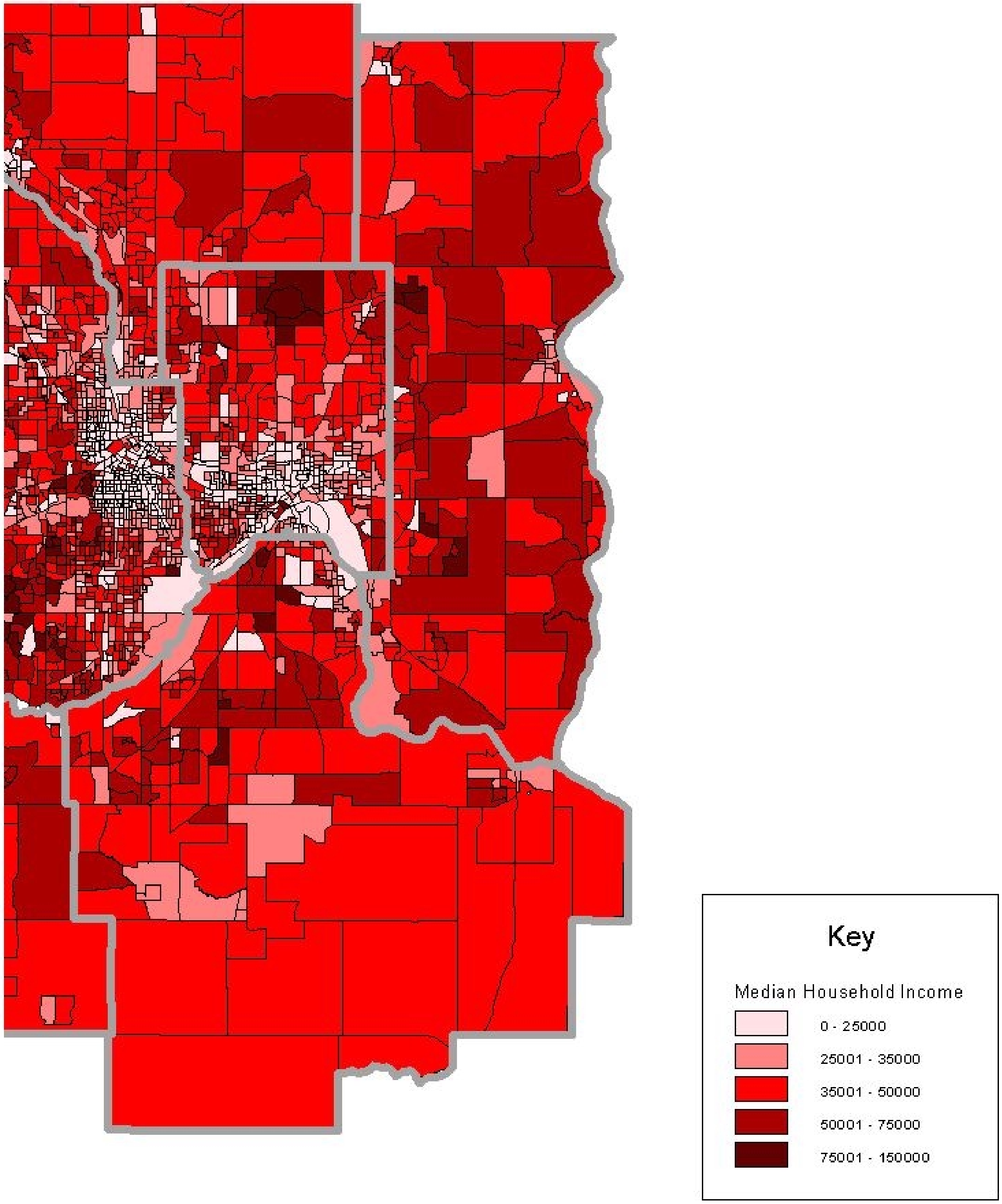
Appendix III: Neighborhood and Accessibility Data Maps



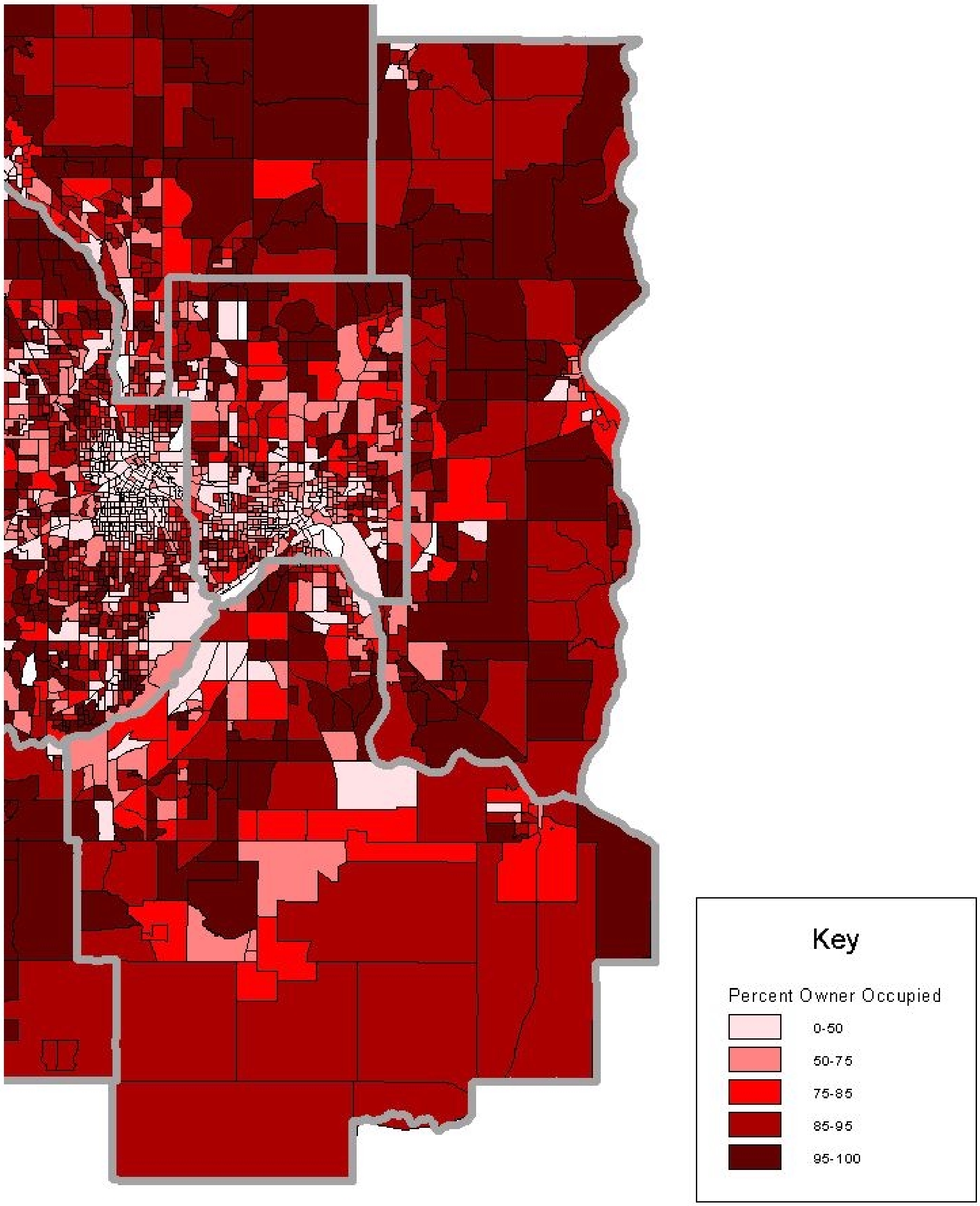
Percent Minority Population by Block Group



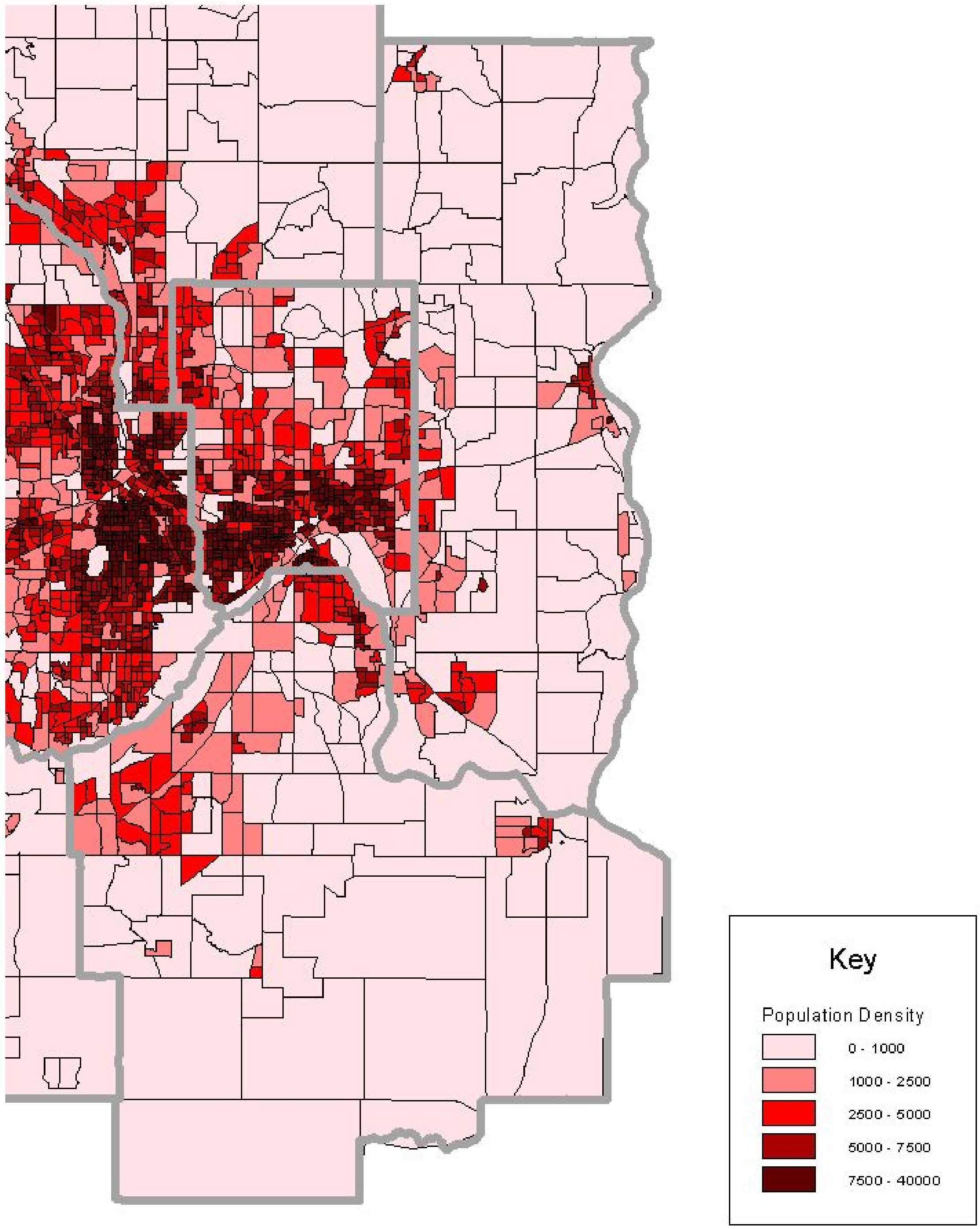
Median Household Income by Block Group



Percent Owner Occupied by Block Group



Population Per Square Mile by Block Group



Median Home Value by Block Group