9-1-2006

The Effect of Proximity to Commercial Uses on Residential Prices

John William Matthews

doi: https://doi.org/10.57709/1059068

THE EFFECT OF PROXIMITY TO COMMERCIAL USES ON RESIDENTIAL PRICES

A Dissertation
Presented to
The Academic Faculty

By

John William Matthews

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Public Policy

Georgia State University and the Georgia Institute of Technology

May, 2006

Approved by:

Dr. Geoffrey K. Turnbull, Advisor
Department of Economics
Georgia State University

Dr. Charles L. Jaret
Department of Sociology
Georgia State University

Dr. Arthur C. Nelson
Urban Affairs and Planning
Virginia Polytechnic and State University

Dr. Christine Roch
Department of Political Science and Urban Affairs
Georgia State University

Dr. David S. Sawicki
Department of City and Regional Planning
Georgia Institute of Technology

Date Approved: March 24, 2006

ACKNOWLEGEMENTS

I want to thank my committee for their time and interest. I especially want to thank my Chair, Dr. Geoffrey Turnbull. His guidance through this process has been a valuable capstone to a wonderful educational experience. I also want to acknowledge the contributions from Dr. David Sawicki and Dr. Arthur C. (Chris) Nelson, over many years, to my professional development. They brought me into the grove at *Academ*e in the first place.

The support and friendship I have received from Dr. David Sjoquist of the Andrew Young School's Fiscal Research Center, and his colleagues Dr. Lakshi Pandey and Dr. William (Joey) Smith, as well as that of Dr. Laura Wheeler, has been truly invaluable. Dr. Gregory Lewis, the Director of the Joint Program in Policy Studies has been a constantly available guide and great teacher.

Wives are always thanked for their patience and support. My wife Dot deserves a much bigger Thank You than most, I think. She had to go through this after I had finished a whole career was supposed to support her as a retired and available leisure companion. I'd also like to thank my children, the long grown Helen and Jacob, for not laughing at me, too much.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The subject of this dissertation is the effect of proximity to retail uses on the price of residential properties. To date, the literature does not contain much material directly addressing this topic. Urban economic theory is indefinite about the effect. On one hand, theory tells us both that prices should be enhanced by proximity due to convenience and because of reduced travel cost and be reduced by proximity because of disamenities associated with commercial development such as traffic congestion and noise.

Existing Studies

Existing studies have produced indefinite results. Some find a positive influence of commercial proximity on residential prices; others find a negative influence and a third group finds no effect at all. We review sixteen studies that include residential price effects as functions of proximity to non-residential development: five show a positive effect on residential price; two find a negative effect; five find no or indeterminate effects; one finds the effect varies with the relative strength of positive and negative factors; two find that effects vary with specific uses; and one finds that effect depends on design, maintenance and management of proximate non-residential uses, not necessarily the uses themselves. A final study finds the effect on residential price due to proximity to non-residential depend on the ratio of residential to non-residential uses in the neighborhood; where the ratio of residential to non-residential is high, increases in non-

residential uses increased residential prices. Most of these past studies do not treat the residential-commercial relationship directly. The relationship is often treated in very general terms, making no distinction in the size or type of commercial development, its design, age, or operating policies. Likewise, the neighborhood setting and design relationship between residential and commercial development is not included. No study addresses the profound differences between typical pre-war grid type neighborhood layouts and post-war urban design emphasizing isolated long curving non-connected streets terminating in cul-de-sacs. Similarly, there are obvious differences between neighborhood commercial areas such as that in Atlanta's Virginia-Highland neighborhood and the same city's Lenox Mall. Many researchers, in fact, include commercial development only as a control variable in studies focused on some other specific interest.

What are the different settings within which the residential/commercial interaction takes place? Do different settings influence the effect of commercial proximity on residential price? Does the design layout of a residential neighborhood influence that proximity to commercial uses has on residential prices? Does this effect vary with the layout of commercial areas, the type of businesses, operating hours, commercial property maintenance?

From a public policy perspective, some of these questions are more pliable and interesting than others. Design and size of commercial development, or neighborhood layout, for example, may be more easily and permanently regulated than operation and maintenance.

Study

This dissertation looks at the effect of proximity to commercial development on housing prices directly. The study includes controls for type of non-residential development and uses an innovative new technique to incorporate analysis of neighborhood layout as a specific independent variable. Incorporation of neighborhood layout is a unique contribution of this dissertation.

Looking at commercial development in greater detail will reduce some of the uncertainty found in existing literature about the effect of proximity to retail uses on residential prices and will identify circumstances in which effects are positive, those in which they are negative, and those in which there may be no effect.

Methodology.

There are many different methodological approaches to a study of this sort. They break into two basic types: 1) stated preference, for example questionnaires and attitude surveys, and 2) revealed preference methodologies, for example hedonic price modeling, an application of multiple regression techniques. Researchers have made the revealed preference approach, specifically hedonic price modeling, the method of choice. Hedonic modeling estimates the price impact of "unpriced" components of composite goods, such as the specific components of housing. This methodology is applied to two general categories of housing price analysis. The first attempts to predict or index housing price for mass appraisal of properties for tax assessment or construction of Consumer Price Index estimates of housing prices. The second attempts to estimate the effect of changes in a specific component of the "housing bundle" – adding an additional bathroom, for example – on housing price. More important for public policy, the hedonic approach can

also go well beyond price estimate of individual structural components of housing and estimate the price impact of neighborhood and environmental characteristics, such as the price impact of proximity to a landfill.

Hedonic methodology is widely used in housing price analysis, but it has both practical and methodological/theoretical problems. Practically, hedonic models used in past studies of housing prices have frequently suffered from under-specification, especially inabilities to adequately account for space and location, but also a paucity of information regarding specific details of individual properties. Most important has been inability to capture actual sales data as opposed to assessment values or owner estimates of value reported in the census. The recent advent of sales data driven computer based mass appraisal systems for property taxation is helping to meet this data problem.

Computerized tax assessment databases typically contain very detailed information about specific characteristics of each real property in a taxing jurisdiction. Because they are computer databases, they can be easily and cheaply copied for use in analysis. Additionally, the even more recent development of geographic information systems (GIS) has made possible sophisticated measurement of space and location variables when tax assessment databases are combined with geocoded tax parcel maps.

The idea of perception presents a second practical problem for hedonic housing price analysis. If, for example, an unattractive parking lot at a nearby shopping area cannot be seen or heard from a nearby house, does the parking lot affect the price of the house?

There are many methodological problems including such basic questions as specification of functional form of hedonic models and selection of independent

variables. Linear, semi-log and log-log forms are commonly found in hedonic studies. There is very little theoretical help for choosing one functional form over another. Many researchers prefer the very flexible Box-Cox approach; others reject this approach preferring to let theory guide specification. Selection of independent variables is another issue, especially in studies aimed at teasing out the effect of specific components. Problems of form and choice of variables lead to problems of multicollinearity and heteroskedasticity. Multicollinearity is generally not considered an important problem, or even one that can be easily addressed, but it can be a problem if it affects precision of estimates for specific individual variables.

Introduction of spatial relations is likely to introduce problems of spatial autocorrelation which in turn can increase the possibility of a Type I error; an incorrect rejection of a null hypothesis. Spatial autocorrelation is a concern because spatial relations between commercial and residential properties are at the heart of this dissertation.

Data

This dissertation takes advantage of a tax assessment database linked to a tax parcel GIS map from King County, Washington (Seattle). Tax assessor data are available for the entire property tax base in King County. Over 130 variables are listed for each residential property including not only the usual data such as size and number of rooms in the structure, but also scenic views, traffic noise, and other very detailed types of data. One of the most important types of data in the database is a record of sales prices and dates.

The GIS mapping system displays over 500,000 tax parcels, which can be linked to the tax assessment data. The topological mapping system that is the heart of GIS

allows calculation of the spatial relationship of data associated with each parcel to all other parcels. Additionally, there are over 19 major overlay maps, such as local taxing jurisdictions, parks, traffic controls, and so on, which add richness to hedonic analysis. The GIS maps can be linked to other spatially related census data, for example. In this way, structural characteristics of a specific house can be associated with individual neighborhood characteristics.

The study area is a swath of census tracts running from the Puget Sound water-front to the eastern urban growth boundary running north of Seattle's downtown eastward through portions of the cities of Kirkland and Redmond. There are 38 census tracts in the area encompassing 176 census block groups and approximately 58,700 tax parcels, residential and non-residential. Lake Washington divides the study area into two unequal parts. The two parts of the study area differ in their general design and general population characteristics. The portion of the study area west of the lake in Seattle developed in the era between WWI and WWII when automobiles were emerging in American culture, but had not come to dominate the culture. Houses in this sample have a median age of 72 years; gridiron street patterns predominate. The east side in Kirkland/Redmond is characterized by a curvilinear/cul-de-sac pattern. The median age of a house in this sample is 25 years.

These differences have important implications for the effect of proximity to retail properties on the price of houses.

Neighborhood Layout

Many methods of indexing neighborhood layout are being developed. Space syntax analysis, for example, is a new technology emerging from academic and analytic

approaches to the study of urban design. This study uses a less complex and more intuitive method –ratios of street segments and intersections - to quantitatively index neighborhood layout (design) and measures of integration. Inclusion of a design integration index as an independent variable permits analysis of the role of urban layout on the relationship between proximity to commercial uses on residential prices. Effects, both positive and negative, should be stronger in more integrated areas.

Intent

This dissertation advances the practice and application of hedonic analysis to issues of public policy. As residential development expands into new areas, commercial development follows as markets are created. Local elected officials and policy makers are frequently confronted with homeowners protesting that the presence of new commercial development, especially if it is close by, will run down their property values. This is the well-known NIMBY (Not In My Back Yard) syndrome. This NIMBY reaction occurs when regulatory permission is sought for development that ranges from "7-11" convenience stores to regional malls. Local homeowners make their claim of threatened property values arguing that "everybody knows" it is true. Local officials are confronted with difficult choices and little empirical information. These decision problems are exacerbated by the growing chorus for "new urbanism" and mixed use development in growing, low density, single use suburban areas.

Does proximity to commercial development adversely affect residential property prices? Even though residential property owners seem to be clear on the matter, existing professional and academic literature is unclear. This study provides empiric evidence of the effect of proximity to commercial development on residential property values.

Contributions

This dissertation makes several contributions to existing literature. 1) It uses a much larger and richer database than used in all but a very few studies. 2) The study incorporates GIS technology. Incorporation of GIS is not new, but is still unusual. Among other contributions, GIS is used to create precise location variables to control for spatial relations. 3) The study focuses on the effect of proximity to commercial uses and different types of commercial use and development on residential, as opposed to effects of generic non-residential uses. 4) The study explores, for the first time, the influence of neighborhood street patterns on the effect of proximity to commercial use on residential property values.

Past studies of the impact of proximity to commercial development on residential prices produce inconclusive results. This dissertation isolates factors that may confound past studies, explaining the contradictory results.

Home ownership is one of the most important, if not the most important, means available to U.S. families for accumulation of wealth. Community decisions that positively or negatively affect residential value should be taken seriously. This study helps inform individual decisions regarding investment in homes. It helps inform public decisions shaping comprehensive land use planning and land use regulation.

Hypotheses

There are two primary null hypotheses tested:

Hypothesis_0: Proximity of commercial development has no effect on prices of proximate residential properties.

Hypothesis$_1$: Proximity of commercial development has no effect on prices of proximate residential properties regardless of the layout of the neighborhood setting.

Results

In the older gridiron area in the Seattle portion of the study area, proximity to retail creates both a positive, or convenience, effect and negative, or spillover, price effect for residences; the effects play against one another. On the whole, the positive effect outweighs the negative effect, but up to about 250 feet, the negative effect of disamenities results in a net loss. Beyond a distance of around 250 feet, the effect is positive for almost another 1,000 feet. Neighborhood layout and density have a significant effect on the magnitude and reach of the travel and straight-line effects on price. As neighborhood layout becomes more integrated, the positive price effect of proximity increases.

In the eastern portion of the study area, the younger area featuring automobile oriented design, the positive effect of convenience to retail is not observed, even though the negative effect is approximately the same magnitude and reach as in the gridiron patterned area.

Both null hypotheses are rejected.

CHAPTER 1

INTRODUCTION

This dissertation examines the effect of proximity to retail uses on the price of residential properties. First, we show what the problem is and why it is important and outline the basic study question. Chapter 2 reviews the relevant literature. First there is literature that explores the price effect of proximity to retail and non-residential uses in general. There is also a review of literature relevant to methodology. Chapter 4 presents new and innovative technology, very valuable to this study, not available to past researchers. Chapter 5 presents the results of the study and Chapter 6 presents our conclusions.

Background

Local zoning hearings across the county resound with the arguments of homeowners objecting to proposed new nearby commercial development. For example:

- In Sylvan, North Carolina, homeowners appealed to the city council to reject a rezoning application for a small commercial project, claiming that "Any non-residential activity on the ... property would affect their quality of life, negatively impact the value of their property, alter the character of the neighborhood and materially diminish the property values within the surrounding areas…" (Hotaling & Hooper, 2001)

- In suburban Omaha, Nebraska, local residents opposed commercial development near Lake Hastings. "Speakers for the lake residents said the commercial development would hurt their property values and their quality of life...[a local real estate broker] also said owners of existing residential property would be damaged by the development. The presence of commercial property near a neighborhood does more to damage residences' desirability than does a busy street or road passing nearby, he said" (Raun, 1997).

Often, objections such as these, based on "conventional wisdom" and "common knowledge," but otherwise unsubstantiated, tend to carry the day, preventing integration of residential and commercial land uses.

The issue is not inconsequential; it is one of the key issues of new urbanism. New urbanism is a community design philosophy that favors the return to new home development mixed with multi-use buildings and housing clustered near commercial service areas. New urbanists claim substantial advantages for this approach to development, including lower over-all public infrastructure costs, and decreased automobile dependency.

There are those who argue that mixed development - commercial and residential together - may actually enhance residential property values by reducing the inconvenience of routine trips and reducing trip cost. In fact, the notion of residential rent gradients decreasing as distance from the CBD increases, found in basic urban economic theory (Muth, 1969), flows from the this idea of convenience. Proponents of new urbanism argue, for example, that "Location Efficient Development - residential and commercial

development located and designed to maximize accessibility" - minimizes total daily travel (Victoria Transport Policy Institute, 2003).

Long before new urbanists began advocating for mixed-use development, standard zoning ordinances and regulations throughout the nation recognized there could be both problems and benefits arising from adjacent development of residential and commercial uses. Consequently, measures such as buffers, walls, and setbacks are commonly required to protect bordering residential developments from commercial areas when they abut one another. This reflects another aspect of basic economic theory, the idea of "externalities," in this case, negative externalities (Miller, 1999). The buffers, walls, and set-backs are required to protect the abutting homes from lights and noise pollution assumed to be generated by non-residential uses and further assumed to reduce the desirability of nearby homes.

Those who object to mixing commercial and residential uses and those who advocate mixing uses both have a portion of theory backing their arguments. Convenience should positively affect residential property values, but negative externalities should drive down residential property values. When commercial and residential uses come together in a neighborhood, does the added convenience overcome the effect of any negative externalities with a net increase in residential values, or do the negative externalities trump convenience with a net loss of residential value?

Question

This dissertation explores theory and literature regarding price effects of mixing land uses, especially residential and commercial mixes. Further, this dissertation explores this relation within different neighborhood layouts. Following a review of theory and

literature, a study of residential property values in King County, Washington (Seattle) is presented. The basic question this study addresses is: Does proximity to neighborhood commercial development affect residential values? The second question is: does variation of neighborhood layout, ranging from grid to cul-de-sac, affect value of proximity?

CHAPTER 2

LITERATURE REVIEW

This literature review has three parts: (1) a general discussion of urban economic theory as it pertains to the relation of residential and commercial development, (2) a review of studies that directly or indirectly speak to the impact of proximity of commercial uses on residential property values, and (3) a review of studies of hedonic price modeling.

Urban Economic Theory

Two conflicting strains of urban economic theory bear on the question of the effect of proximity to commercial use on residential prices. Microeconomic theory applied to urban land holds that land values are determined by transportation costs. Generally, as distance to an "attractor" use (e.g. work or shopping) decreases, transport cost decreases and land cost increases. Consequently, residential properties located closer to retail use should, all else being equal, have a higher price than residential property farther away because travel cost to the retail use is lower.

A different strain of theory predicts prices will decrease with proximity because of disamenities associated with commercial development, such as traffic congestion and noise.

The notion of "externalities" is an important concept developed in welfare economics. An "externality" is a "consequence of an economic activity that spills over to

affect a third party" (Miller, 1999, p. 95). Thus, noise and traffic congestion generated by a shopping area are negative externalities that may adversely affect neighboring residences while convenient shopping may be a positive externality benefiting a residential area.

Zoning is a governmental police power that regulates land use to minimize external diseconomies "spilling" from one type of land use to another; usually from non-residential to residential uses (Mills, 1979). Zoning ordinances seek to minimize these externalities by first segregating land uses from one another and also by design controls imposed to minimize spill over of diseconomies (e.g. noise) from one zone to another. A thick planting of trees, for example, may be required between a shopping area and neighboring houses.

Review of Studies on Proximity: Residential to Commercial

Popularly, the nature and magnitude of the diseconomies associated with proximity of residential development to commercial uses are often seen to be quite large and threatening, as reflected in the news articles quoted earlier. Professionally and empirically, the extent of the diseconomies, either in terms of effect on price or in terms of geographic reach, is questionable. This section looks at both appraisal literature and empirical literature.

Appraisal Literature

Real property appraisal literature is not conclusive. For example, a large real property appraisal firm active in the northern mid-West states:

> "It is well established that the value of a new single-family residential property is lower when it is adjacent to commercially (as opposed to residentially) zoned land

and/or developments. However, it is not quite so clear what occurs when the residential property is already existing and the adjacent land use *changes.* This fundamental question is posed time and again, primarily by homeowners who challenge a proposed commercial re-zoning or development on the basis that the change would negatively impact the value of their adjacent residential properties." (Hosch & Koehlinger, 1997)

While these writers cite no general or consistent evidence for their claim of a "well established" relationship, a particular study finds, in a specific instance, that there is no negative impact on homeowners' property values when a neighborhood type retail project is built nearby:

> "In the town of Henniker, New Hampshire, a proposal to build a 9,800 square-foot pharmacy at the edge of a commercial zone met significant opposition from the community, especially abutting residential property owners. A study requested by the Henniker Planning Board to assess how the project might influence the values of adjacent residential properties ... demonstrated that in Henniker, commercial development does not have a measurable impact on abutting residential properties … No differentiation in values is apparent between properties with a commercial influence and those without. The minor differences in assessments reflect discrepancies in building sizes or amenities, such as a garage or a fireplace ... It is worth noting that the shopping center does not have any landscaping to serve as a buffer for the abutting residences." (Crafts, 1998, p. 8)

The study method in Henniker is instructive. The appraiser/author pairs similar houses from two groups - one near commercial and one not - based on their purchase

dates. Differences in paired sales values are attributed to the presence or absence of commercial influence and minor discrepancies in the configuration or amenities available in the homes.

Standard real property appraisal textbooks give no specific guidance regarding the impact of commercial development on residential values, but generally reflect the notion that there is a positive influence. "Residential locations, whether for homes or multiple units, are enhanced when they are ... well supplied with shopping..." (Kahn & Case, 1977) In the context of new residential growth on the fringe of metropolitan areas, Boykin and Ring write, "[al]though efforts are made to provide essential public conveniences for these outlying developments through construction of neighborhood shopping centers, ... there is often a considerable time lag ... [before] sufficient neighborhood shopping facilities ... become a reality." (Boykin & Ring, 1993) The implication is that nearby retail development is desirable.

Empirical Literature

There is no clear pattern of either positive or negative effects of proximity to non-residential uses on the price of housing. Formal studies look at the price effect of non-residential uses - or specifically retail and commercial uses – on residential prices. Almost exclusively these studies rely on hedonic price modeling developed through the cumulative work of Griliches (1961), Lancaster (1966), and Rosen (1974). As with less formal studies, and as might be expected from the conflicting points of theory, the results of these studies are generally inconclusive.

Crecine, Davis, and Jackson (1967) undertake to "ascertain the effects of certain externalities ... on the value of single family dwellings as this value is reflected in market

prices." The sales value of residential property is regressed on topography and accessibility to economic activities and amenities. The study uses Pittsburgh property sales data in the period 1956 to 1963. The researchers are unable to find any evidence of externalities - or negative influence on residential property values - from the presence of other types of land uses. Aside from speculating about effects from the model itself, they speculate that the findings may possibly result from (1) the possibility that negative externalities extend only "next door" or (2) that zoning may be effective. Finally, they consider that there may not be any externalities in the urban residential property market. In other words, no effect on residential property values is found associated with proximity to commercial development or wholesale, or light industry, or fourteen other types of non-residential land use because there is none.

Reviewing this and other studies, Mills (1979) writes that for "most nonresidential activities studied, the effects seem remarkably small. Coefficients are frequently insignificant and occasionally have the wrong sign. Even when significant, most effects are found to be small and decline rapidly with distance" (Mills, 1979, p. 521) Mills speculates on two reasons for these findings. First, as with Crecine, et al., he concedes that zoning might be effective. Second, he points out that commercial and industrial development does produce jobs and shopping. Proximity to these uses is valuable; "residential land values may even fall with distance from a nonresidential site; but that does not imply that there are no external diseconomies from the site, only that they are more than off-set by the advantages of proximity" (Mills, 1979, p. 521).

Michael MaRous (1996), a property appraiser, studies the impact of four very low-income family housing developments in four growing Chicago suburbs. "A

dampening effect of 3% to 5% on the market values of residential property adjacent to low-income and very low-income family housing was expected. However, there was no evidence of this. Instead the evidence showed market values consistent with property not adjacent to the low-income units, and values rising at rates consistent with the community as a whole" (MaRous, 1996, p. 32). Importantly, MaRous reports that factors that contributed to the success of the four projects include good community planning, good design and buffering of the sites, and good property management.

While MaRous' study is a market analysis, he does cite a study using statistical regression techniques. "Relationships Between Affordable Housing Developments and Neighboring Property Values," conducted by Paul Cummings of the Institute for Urban and Regional Development and John Landis of the University of California at Berkeley, reach similar conclusions:

> "Poorly designed, poorly maintained, and poorly managed projects can affect neighborhood property values--regardless of whether they are affordable or market-rate. Conversely, well-designed, well-managed, and well-maintained projects should not affect neighborhood property values, regardless of whether they are affordable or market-rate." (MaRous, 1996 p. 32)

In both MaRous' and Cummings and Landis' studies, even though they are undertaken with different types of analysis, the important finding is the same: effects on neighboring property values arise not necessarily from the general type of use, but from design, operation, and maintenance. In these studies, we see that negative externalities can be mitigated by appropriate design, maintenance, and management. While MaRous' study does not include commercial development, we can conclude that the same factors

could mitigate or diminish the effect of negative externalities coming from commercial development.

R. Gail Grass (1992) studies the effect of the advent of new heavy rail mass transit service on residential property values in Washington, D.C. Five areas within one-quarter mile of five transit stations are paired with five control (non-transit) areas. The study uses residential sales prices from 1970 (before the transit system was built) and 1980 (after the system was opened). A significant <u>positive</u> impact is found on residential property values related to proximity to transit stations (Grass, 1992).

The major finding in this report - the positive impact of transit - is of interest as it shows the positive influence of "convenience." The choice of one-quarter mile as the size of the impact area is also of great interest as the implication is that convenience extends for one-quarter mile in a pedestrian environment.

Bowes and Ihlanfeldt (2001) also study the effects of proximity to transit stations on residential property values. Along with Grass, they find a one-quarter mile distance from transit stations to be significant. They also find that positive effects extend further than do negative effects. "Large, positive, direct effects are found in high income neighborhoods between one-quarter and three miles of a station … [b]eyond one-quarter mile of a station, negative direct effects are generally restricted to low-income neighborhoods. Apparently, for middle- and high-income neighborhoods, the commuting cost savings provided by transit exceed any costs caused by negative externalities" (Bowes & Ihlanfeldt, 2001, p. 21). The measures of distance in this study are concentric rings drawn one-quarter, one-half, one mile, and so on from stations. The estimates of the reach of positive and negative effects are no more precise than these rings.

In a frequently cited paper, Grether and Mieskowski (1980) test effects of non-residential land uses on the prices of nearby dwellings in sixteen “market experiments” which use samples of home sales near a single nonresidential land use: industry, commercial, high-density dwellings, and highways. Regressions of physical characteristics of the dwelling, distance from the non-residential use, and the date of sale for each transaction on sales price show no systematic relationship between nonresidential land use and housing prices.

Li and Brown (1980) provide another oft-cited study that assesses the influences of “micro-neighborhood variables - aesthetic attributes, pollution levels, and proximity” to industries, thruways, and commercial establishments - on housing prices (Li & Brown, 1980, p. 125). The researchers are specifically not interested in large-scale geographic variables that are often researched in attempts to understand variation in real property prices across metropolitan areas. Instead, they are interested in the effect of proximity to

> “a corner grocery store, a neighborhood park, a school, a river, an ocean, or conservation land. Accessibility at this micro level is normally thought to increase the value of a house and has been an essential part of architects’ designs of new towns in the United States. At the same time, proximity to some of these non-residential uses can also be accompanied by external diseconomies such as congestion, noise, and air pollution that affect the value of residential property.” (Li & Brown, 1980, p. 126)

The study is conducted at the neighborhood level. Census tracts are used as neighborhoods. Li and Brown (1980) note two conflicting theories: first, convenience of proximity will enhance property value as distance increases. Second, protection from

negative externalities will enhance residential values as distance increases. Proximity to the various non-residential uses is measured in as a Euclidian distance in ten-meter increments and entered into the model in logarithmic form to model the authors' assumptions about the convenience of proximity. Distance is entered again in a negative exponential form to reflect assumptions about the effect of negative externalities. "Empirical findings suggest that proximity to certain non-residential land uses affects housing prices by having a positive value for accessibility and a negative value for external diseconomies (congestion, pollution, and unsightliness). Furthermore, visual quality and noise pollution have impacts on housing prices" (Li & Brown, 1980, p. 125). Their measure of distance is much finer grained than Bowes' and Ihlanfeldt's. Even so, they too find that positive effects have greater range than negative effects.

Nelson and McClesky (1990) use the Li and Brown model to examine the price effect of proximity to elevated transit stations in Atlanta. Noting that proximity to elevated transit stations can have a positive price effect arising from convenience and a negative effect due to exposure to traffic, noise, and other nuisance, they work to capture the interplay between the two effects. As with Li and Brown (1980), they use a single Euclidian measure of distance with logarithmic and negative exponential transformations to proxy convenience and negative externalities. Using 286 observations of home sales near elevated transit stations, they find a revealed price gradient that is positive. The implication is that the positive price effect of convenience outweighs the negative prices effects of disamenities arising from proximity to the elevated stations.

Frew and Judd (2003) also look at property prices at a micro or neighborhood level. They study apartments, not single family residential. For their research, zip codes

are the neighborhoods. Their results differed from Li and Brown. They found that an increase in level of commercial activity in the "neighborhood" (zip code), measured by total payroll in the zip code, is associated with a reduction in property value - but the reduction is not statistically significant. Distance was not accounted for in this study.

Mahan, Polasky and Adams (2000), find a negative relation between residential values and proximity to commercial uses. They use proximity to commercial uses as a control variable in a study investigating the relation of proximity to wetlands as an influence on residential prices. Their results show, all else held equal, that as distance increases between commercial and residential uses, residential price increases. They expect the opposite: based on the notion of convenience, they expect that residential value would increase with proximity to commercial uses. They reasoned that their result might reflect negative externalities such as congestion and noise overriding the positive effect of convenience.

The distance measure used by Mahan, et al. is taken in feet. The mean distance between observations of home sales and the variable of interest – wetlands – is over 3,580 feet, a little over two-thirds of a mile, with the maximum distance ranging out as far as 11, 930 feet, or 1.3 miles. This is much more fine-grained than a metropolitan or city-wide study, but is wider ranging than neighborhood oriented studies such as those of Grether and Miezkowski (1980) and Li and Brown (1980).

Another study, this one of multi-family housing in Brasilia, also shows mixed results. The specific focus in this study is the effect of proximity to sewage treatment plants on residential values. Proximity to specific types of commercial uses is used as a group of control variables. Dummy variables are included for drugstores, bakeries,

butcher shops, fruit and vegetable markets, bookstores, restaurants, gas stations, and bars. All are found to have negative price effects except fruit and vegetable markets and gas stations (Batalhone, Nogueira, & Mueller, 2002).

In Seattle, a study by Franklin and Waddell examines the influence of accessibility to different types of employment on single-family residential property values. The results show that access to commercial and university uses is positively associated with sales prices, while proximity to local schools and industries is negatively associated with sale prices (Franklin & Waddell, 2003).

Clearly, evidence concerning the impact of proximity to non-residential land-uses on residential property values is mixed. In theory, there are potentially both positive and negative effects; mixed results are an understandable outcome. In a study that may integrate findings from others, Cao and Cory (1981) seek to construct a theoretical model that can account for positive and negative effects and test their theory empirically, using hedonic pricing equations. Their theoretical model asserts that the effect of nearby non-residential uses on residential property values is *a priori* indeterminate. Outcomes depend on the relative strength of positive and negative external effects generated in any given setting. In this case, the setting is the proportion of non-residential to residential uses in a neighborhood. The test is conducted in Tucson, Arizona. Census tracts define neighborhoods. The results show that over ranges of low proportions of non-residential uses, increasing the amount of industrial, commercial, multi-family and public land-use activity in a neighborhood tends to increase surrounding residential property values. Their conclusion is that optimal mixes of land-use activities are possible and should be sought. They assert that separation of activities, as is common with contemporary zoning

practices, is not optimal. Findings such as this also bolster arguments for the integration of uses that is an important a part of new urbanism.

Conclusions from the Literature

These studies bring out several important points. First, they show inconclusive results. Of the sixteen articles reviewed, four show that proximity to commercial uses produces a positive impact on housing values, two show a negative impact, four find no effect, one says the effect is indeterminate, one asserts the effect depends on relative strength and weakness of the positive and negative effects, one finds the effect depends on the specific type of commercial use, and last, one concludes the effect is a function of design and property management, not necessarily general types of use.

Second, several writers – Mills (1979), Li and Brown (1980), Grass (1992), and Bowes and Ihlanfeldt (2001), for example – find that both positive and negative externalities on residential prices arising from proximity to non-residential uses tend to be small and tend to dissipate rapidly with distance. There are also some findings that negative effects dissipate more rapidly over distance than do positive effects. One implication is that the effect occurs at a neighborhood level over short distances. In fact, the studies by Grether and Mieszkowski (1980), Cao and Cory (1981), and Li and Brown (1980), as well as others, all use neighborhoods as a basic unit of study. But, even this is contradicted by Mahan, et al. (2000), finding effects reaching over much larger distances. Mahan et al. find that at these greater distances, the price effect of distance is positive; as distance between retail uses and residences increases, residential prices increase (Mahan et al., 2000). As a side observation, most of these studies use census tracts as

neighborhoods. Use of census tracts avoids some problems that may bias statistical analysis and assures that many socio-economic descriptors are available.

Third, even though many investigators use neighborhoods as their basic study area, the physical characteristics of neighborhoods are often not included. Cao and Cory (1981) produce a rare study that looks at physical and land use aspects of neighborhoods. Their findings regarding the effect of mixing non-residential and residential uses are significant and important. Beyond that, no one has included the neighborhood design relationship between residential and commercial development. Jo (1996), points out that the pre-war neighborhood design – typically highly interconnected grids – is profoundly different from post-war design – isolated, long, curving, non-connected streets terminating in cul-de-sacs. These design traits can affect both the positive value of convenience and negative diseconomies normally associated with proximity between the two types of uses. Highly integrated grid system neighborhoods may have high values associated with both influences similar, for example, to the situation described by Li and Brown (1980). Jo's (1996) example of a highly integrated neighborhood is Atlanta's Virginia-Highland neighborhood. On the other hand, the effect of intentional isolation inherent in cul-de-sac designs may be very effective in reducing the impact of disamenities. But, this type design may also impose lengthy travel distances, thus reducing the positive influence of convenience. Jo's (1996) example of a neighborhood with a non-integrated design is Atlanta's Ashford-Dunwody.

Lastly, MaRouse' (1996) and Cummings and Landis'(1993) findings regarding design, operation, and maintenance coupled with Li and Brown's (1980) conclusion concerning the significant effects of visual quality and noise present interesting possible

avenues of investigation. The implication is that it is the performance of a development, in terms of generating negative externalities, as well as the general type of use, that is relevant. Size, design, maintenance, and management can all affect this kind of performance. From a public policy perspective, design and size of commercial development may be more pliable factors than operation and maintenance. Neighborhood street layout may explain, at least in part, the balance of positive over negative and vice versa found in some settings and the complete absence of any effect at all found in other settings.

Review of Literature on Methods: Hedonic Modeling

Hedonic modeling is the method of choice for studying factors that influence the price of housing. Hedonic housing price modeling is a specialized multiple regression analysis. There are two applications of hedonic price modeling: the first deals with adjusting prices on the left-hand side that result from right-hand changes in quality. For example, hedonic models are used to improve the precision of housing price indexes; they are used by professional appraisers to make residential valuations; and, they are used in mass appraisals for tax purposes. The second category of hedonic applications relates to individual right-hand characteristics and coefficients: hedonic models can test the effect a change in one independent variable has on housing price*, ceteris paribus;* hedonic models test the standard urban (monocentric) model to see if housing prices vary with distance from the CBD (as predicted) or if multi-centric models are better descriptors; they measure effects of environmental quality on housing prices; and they study implied prices of individual housing characteristics, e.g. the value of a third bedroom (Hulten, 2003; Malpezzi, 2002).

This dissertation belongs to the second type; it seeks to find the effect of changes in a right hand variable on the price of residential property. The right hand variable of interest is the proximity of housing to retail uses, with the variable actually consisting of two measures, straight-line distance and travel distance between residential and commercial uses. Because structural and neighborhood characteristics tend to dominate housing price (Butler, 1982), teasing out implied values of other components, such as proximity to retail uses, often proves difficult. Additionally, application of hedonic modeling to seek influences of particular right-hand variables is sensitive to a number of statistical problems, such as omitted variable bias and multicollinearity.

The review of the literature of hedonic models that follows relies heavily on three existing reviews: 1) a review published by Batemen, Day, Lake, and Lovett in "The Effect of Road Traffic on Residential Property Values: A Literature Review" (Batemen, Day, Lake, & Lovett, 2001), 2) another review by Stephan Sheppard published as "Chapter 41: Hedonic Analysis of Housing Markets" in the *Handbook of Regional and Urban Economics, Volume 3, Applied Urban Economics* (Sheppard, 1999), and 3) a third by Stephan Malpezzi, "Hedonic Price Modeling: A Selective and Applied Review," found as a working paper at the University of Wisconsin, Madison, CULER website (Malpezzi, 2002).

Housing as a Complex Good

Economists have divided the world into simple goods and complex goods. A simple good can be thought of as a commodity. It has only one aspect and one accepted price. A simple good can be priced according to marginal utility or a consumer's marginal willingness to pay.

A complex good, on the other hand, is made up of not one, but many features. A complex good can be thought of as a bundle of variable attributes. Rosen (1974), for example, would characterize housing as a good that is differentiated by the amounts of the various characteristics it contains; the number of square feet of livable area, number of bathrooms, quality of neighborhood schools, distance to work, and so on. Think of complex goods as shopping carts filled with different bundles of "stuff."

Vectors of the Housing Model

The hedonic technique holds that any real property can be described by its constituent characteristics. Writers have generalized these characteristics somewhat differently, but are largely in agreement. Freeman (1993), for example, uses the vectors of environmental amenities, structural characteristics, and neighborhood while Batemen, et al. (2001) use these general vectors, but include a fourth - access - perhaps reflecting a different theoretical understanding of the basis of value.

A house is typically thought of as containing several kinds of stuff – vectors - in its bundle. Commonly found vectors include (Batemen et al., 2001; Li & Brown, 1980; Mahan et al., 2000; Sheppard, 1999):

- Structural characteristics such as the number of rooms, the floor area, whether there is a garage, and so on. Lot size is generally included in this group. Tax assessors' databases are common sources of this data.
- Neighborhood characteristics including race, income, family size, school quality and other similar characteristics. Butler (1982) cautions that social characteristics that reflect or proxy housing demand, such as income, should be left out of hedonic models. Data predominately comes from census data, but some comes

from other sources. School quality, for example, is commonly proxied with standardized test scores (Butler, 1982)

- Accessibility, a characteristic generally including distance or travel time to the CBD as well as access to employment centers, shopping areas, and cultural and recreational opportunities (Butler, 1982). Modern geographic information systems can calculate several measures of distance and accessibility.
- Environmental characteristics commonly thought of as including air pollution, noise, and local traffic congestion as well as amenities such as a good view (Batemen et al., 2001). Access to these measures is more problematic than the others. Some are carried in tax assessors' databases while others are in more specialized sources or may not be available at all.

None of these components – or many others that could be included in this list - of a house are bought and sold as individual goods or services. Consequently, there is no market price for them as individual or separable goods and services. They do have values that are included in the total value or price of the housing bundle. They are thought of as cumulating in a house's market price.

Hedonic price estimating techniques are based on the idea that while structural/ neighborhood/accessibility/environmental goods and services are not directly traded or priced individually, they are traded indirectly as part of a complex good. For example, if the task were to find the value of "peace and quiet" as a feature of a residential unit, it would not be possible to find "peace and quiet" traded as an individual commodity. One way to arrive at this price is to find houses that are in all respects the same except the noise level. Any price difference, therefore, could be called the price of peace and quite

(Batemen et al., 2001). As it is generally not possible to find houses that are the same except for noise – or any other single characteristic – multiple regressions with sizeable samples are used to estimate the prices of the individual components. This is the essence of hedonic price estimating.

Alternate Approaches

Hedonic price modeling is not the only method used for deriving non-market prices of components of complex goods and services. Some other methods are used in analysis of environmental projects and policy and can be used, sometimes, to value environmental components of complex goods. A brief listing of these methods as presented by Batemen, et al. (2001) includes:

- Analysis of Opportunity Costs: One of the costs of a reforestation project, for example, would be the value of foregone crops on agricultural land turned back to the wild.
- Cost of Alternatives: The cost of preserving a beautiful view from highway construction could be the cost of an alternate transportation solution less the cost of the original plan.
- Shadow Project Costs: Cost of providing an equal project somewhere else.
- Contingent Valuation (CV) Method. Asking people directly to value a gain in a specified non-market good – e.g. how much more would you pay for this house and one more unit of peace and quiet?
- Contingent Ranking (CR) and Shared Preference (SR) Methods. One-on-one interviews in which respondents are asked to make choices between goods.

None of these techniques are as suitable to the task as hedonic price modeling. Many of these methods simply cannot isolate, or cannot isolate with precision, individual components of a complex good or service. Hedonic price estimation does hold the possibility of teasing out individual prices of even very abstract goods and services, such as the market value of an attractive vista, buried within the housing bundle. Hedonic modeling estimates actual marginal willingness to pay rather than individual subjective valuations.

Formal Hedonic Housing Price Definition

In general, any house can be described mathematically by the vector

$$z = (z_1, z_2, \ldots z_n)$$

where z_i (i = 1 to n) is the level or amount of any one of many characteristics, structure, neighborhood, etc., describing a property (Batemen et al., 2001; Sheppard, 1999).

The price of a house, therefore, is determined by its characteristics:

$$P = \alpha + \beta_1 S + \beta_2 N + \beta_3 E + \beta_4 A + \mu$$

where:

S = vector of structural characteristics

N = vector of neighborhood characteristics

E = vector of environmental characteristics

A = vector of accessibility characteristics

α is the y (price) intercept

βs are parameters to be estimated

μ is the error term (Batemen et al., 2001; Cao & Cory, 1981; Li & Brown, 1980)

The price of a property is a function of the vectors of values describing its characteristics. The quality of any one characteristic may be isolated to show how it affects the price of the house given that all other characteristics are held constant.

$$P = \alpha + \beta_1 S + \beta_2 N + \beta_3 E + \beta_4 A + \beta_5 VI + \mu$$

With P, S, N, E, A, α , and μ are defined as above and VI is the variable of interest. β_5 is the value attached to the variable of interest, holding all other variables constant.

Sheppard (1999) lays out the basic hedonic theory applied to house pricing. There is an assumption of a large variety of bundles and consumers are constrained only by income and price.

The theoretical reasoning starts by assuming that consumers derive utility from consumption of two goods: 1) a housing commodity with several vectors of characteristics plus 2) the consumption of a composite good which includes all other goods and services:

$Z(z_S. z_A, z_N, z_E)$ plus the composite good Y (all other goods and services)

There is a fixed income: M. The Price function is P(Z)

- ά = observed and unobserved parameters which are preferences
- ν = ν (Z, Y, ά) = utility
- household consumption is a function of preferences and income: f(ά, M)
- A household's bid rent is a function of the housing bundle, income, utility, and preferences:

  β(Z, M, ν, ά)

- It follows that
  - $\nu = \nu\,(Z, M\text{-}\,\beta, \acute{\alpha})$ -> $M\text{-}\,\beta = Y$, and
  - $\text{Max}_{Z,Y\,\nu}\,(Z, Y, \acute{\alpha})$ subject to $M >= P(Z) + Y$
- $\delta\beta/\delta z_i = \nu_i / \nu_Y = P_i$: -- P_i is the hedonic price of characteristic i.

  P(Z) is the hedonic price function

The implication is that the hedonic price (or implicit price) for a particular characteristic or group of characteristics can be observed or estimated (Sheppard, 1999, p. 1601).

Problems with Hedonic Housing Price Modeling

Hedonic housing price estimation is a relatively simple idea that, as applied over the years, is much more complex in execution. There are extensive practical, statistical, and econometric complications that affect hedonic estimation of housing price functions. Among the practical problems are data availability and accuracy, definition of market, presence of market gaps, the assumption of market equilibrium, and questions of the perception of characteristics. Statistical issues include specification of functional form, selection and inclusion of variables and multicollinearity, heteroskedasticity and spatial autocorrelation. Econometric complications revolve around estimation of implicit demand, as well as price, for individual components of composite goods.

<u>Practical Problems</u>

The first group of hedonic issues arises from general unavailability of detailed housing data, such as sales price, specific structural characteristics and so on, for individual dwellings.

Incomplete data is a severe problem for early studies and continues to be

somewhat of a problem today. Perhaps the most serious of these problems are among studies using something other than actual sales prices of individual properties as dependent variables. Many early studies use "price" from the census. This is not a real price, but rather owner estimates of selling price, aggregated at block levels. This aggregated estimate reduces not only accuracy of prices, but also reduces ability to control for specific location characteristics (Freeman, 1993). Other studies use tax assessment data. While assessments do provide data on a site-specific basis, they introduce assessors' error. The data is only an individual assessor's best guess at the price a property might fetch if it were to sell. Because different assessors are assigned to different areas, systematic error from assessment area to assessment area may be the result. The best data are records of actual sales; sales price is the preferred dependent variable (Batemen et al., 2001; Mahan et al., 2000). Nevertheless, Malpezzi cautions that sales data may have a selection bias, "several papers … have tested the presence of such bias. Test statistics often reject the null, but … most studies have found the magnitude of the bias to be modest" (Malpezzi, 2002, p. 18). Potential bias can be tested for and corrected with application of the Heckman procedure, but this approach is recommended only for large samples (Kennedy, 1998).

Similar problems of data availability arise with explanatory variables. For example, Lake, et al. (2000) in their study of GIS application to hedonic modeling, not only have to estimate lot size and interior area of houses from GIS plat maps, which is otherwise unavailable, but have to proceed with no interior structural information. Likewise, spatial information, the relation of properties to one another in two dimensions and distance, often is not included. Measures of spatial relationships are particularly

important independent variables. It is easy to argue that housing prices will vary with location. We have all heard that the three most important influences on real property value are location, location, and location. "Given this observation, it is surprising how many hedonic models lack … a variable that explicitly identifies the location of the structure" ... (Sheppard, 1999, p. 1616).

Accessibility is a troublesome characteristic to measure. Earlier studies used Euclidian or straight-line distance from point to point. Contemporary computer and data technology now promise to address this issue more completely. Contemporary tax assessor mass appraisal databases contain actual sales data and dates as well as very detailed interior and exterior information. Emerging GIS mapping systems that include tax parcel maps allow parcel-by-parcel mapping of the specific structure and parcel information from mass appraisal databases. Additionally, these systems provide abilities to calculate spatial relations of individual parcels and the comparison of their attributes in space. Actual travel distances can be calculated along streets.

Computer assisted mass appraisal (CAMA) systems and GIS technology present the potential of including literally hundreds of right-hand variables in hedonic housing price models. CAMA systems, in fact, use rich databases and forms of hedonic estimations to “predict” sales prices of houses for tax assessment. The inclusion of many variables from CAMA databases can provide full model specification needed to meet the potential of omitted variable bias. However, inclusion of many variables introduces potential problems of multicollinearity.

Market identification, the ability to identify the market or markets from which data comes, is important because of the possibility of introducing error in an analysis. It

is possible to make one of two errors in defining the size of a market. First, the market areas tested can be set too large. In this case, data will actually be collected and analyzed from more than one true market. This type error may create serious bias in any resulting hedonic price estimates. Second, the definition may be too small and data may be collected and used from only a small portion of an entire market (Butler, 1982). This type error may lead to greater variance of the estimates (imprecision), but not bias. If there are questions of the proper definition of a market area, statistical tests of market segmentation are available (Batemen et al., 2001).

A market may not be able to supply houses that meet all demands. Batemen, et al, (2001) give an example in Boston of a wealthy neighborhood in an area of high air pollution. The area also enjoys good access to downtown and its cultural amenities. Clearly, the choice of wealthy families to live in this area requires a compromise and acceptance of pollution. There are not unpolluted areas with equal access to downtown amenities. In general, presence of market gaps, unmet demands such as that described here by Batemen, et al, cause larger variance and greater imprecision in hedonic modeling. Additionally, if gaps are not randomly distributed, but rather characterize a particular sub-market, they can introduce bias into hedonic price estimates (Batemen et al., 2001; also see Freeman, 1993, p. 385).

A basic assumption of hedonic modeling is that the market is in equilibrium and is efficient. Freeman (1999) lists three assumptions that, if not met, introduce error to hedonic models. First, households have perfect information. Without perfect information, the price that families pay for different characteristics and bundles of characteristics will vary from sale to sale. Hedonic price functions will be poorly defined,

will contain large variances, and will yield imprecise price estimates. In the face of imperfect information, hedonic pricing models will yield functions containing increasing variance (Batemen et al., 2001; Sheppard, 1999). A second assumption is that transaction costs are zero; that there are no expenses on top of actual housing cost. If there are additional costs (closing costs, moving expenses, and so on) and if they are too high, households will choose to stay in a house that does not have their most preferred bundle. The market will not be efficient. True preferences and prices for bundles and housing characteristics will be poorly defined and price estimates will have large variances (Batemen et al., 2001; Sheppard, 1999). The third assumption is that housing prices and prices of individual characteristics adjust instantaneously to changes in supply and/or demand. If there is a shock to the market and data is gathered during the period of adjustment while real prices are unsettled, greater variance in hedonic price estimates will result (Batemen et al., 2001).

It is unlikely that housing markets will be able to strictly meet all three of these criteria. Thus, it can be expected that prices paid for housing sometimes will be higher and sometimes lower than would be the case in a state of perfect efficiency. On average, the highs and lows should balance each other out and larger variances are likely, but such inefficiency will not bias estimation results (Batemen et al., 2001).

Perception of characteristics included in hedonic pricing models is a third practical problem. There is an assumption that when an individual characteristic is included in an hedonic pricing model, that characteristic is actually perceived by buyers and included in their valuation decisions. Bateman et al. ask if air pollution is included in a model, but the only significant pollution in the study area is a colorless, odorless gas,

will buyers know that pollution is present and consider it when bidding (Batemen et al., 2001). More to the point of this dissertation, if an unattractive parking lot at a shopping area cannot be seen from a given house, does it affect that house's price? If disamenities are not perceived, how can they have an effect on price?

<u>Statistical Issues</u>

Statistical issues include specification of omitted variable bias, use of spatial data, multicollinearity, and definition of functional form.

Omitted variable bias may affect hedonic pricing estimates. Hedonic models commonly include a large number of explanatory variables: structural attributes, variables describing accessibility, measures of the neighborhood, and the characteristics of the environment. As discussed earlier, structural characteristics are relatively easy to identify and settle on. The questions of what constitutes accessibility, what are relevant neighborhood characteristics, and what are pertinent environmental characteristics are all problematic. Omitting such variables leads to biased estimation parameters.

Including accurate measurements of all relevant explanatory variables is very important. Leaving out or mis-measuring explanatory variables can lead to bias in estimation of the parameters. The entire effect of an omitted variable may not be captured in the error term, but subsumed in other closely related variables (Batemen et al., 2001; Rogers, 2000; Woolridge, 2000). For example, if noise pollution and visual pollution both arise from a highway, and both diminish value of nearby property, leaving visual pollution out of the model will probably increase both the value of the noise pollution parameter as well as the error term – the effect of visual pollution being captured in both - and possibly lead to "wrong" conclusions about the strength of the

effect of noise pollution (Batemen et al., 2001).* Error terms will be correlated with explanatory variables and parameter estimated will be biased and inconsistent.

Butler raises the question of which right hand variables to include. In principal, he states, all should be included. In practice, not all are known or available and those that are available may have substantial measurement error. Additionally, inclusion of all possible right hand variables becomes unmanageably large and introduces substantial problems of multicollinearity. Consequently, Butler argues that any hedonic estimate will be mis-specified. Butler tested the biasing effects of omitted variables by using both a "fully specified" and a "restricted" model. The restricted model contained only structural characteristics while the fully specified model added non-structural characteristics. If misspecification is truly an issue, the models should differ significantly. Comparing the standard errors of the estimates, he finds little difference in the coefficients and estimates produced by the two. "The nonstructural variables … have little impact on the accuracy of the regression… biases [that] arise from excluding other [non-structural] housing characteristics have little practical impact" (Butler, 1982, p. 105). Butler's empirical work suggests that the practical impact of omitted variable bias may be small with little impact on the explanatory and predictive powers of the estimate. But, as discussed earlier, a reduced model can introduce a great deal of error and, potentially, large

---

* In the case of the Batemen, et al, being able to separate the effects on price of noise and visual pollution has important practical value. Their study was undertaken to develop a method of making monetary compensation estimates according to the Land Compensation (Scotland) Act of 1973. Under this law, payments are made for diminishment of value due to noise pollution, but not visual pollution (Batemen et al., 2001). An inability to separate the effect of the two factors can lead to serious monetary consequences.

variances in the estimates of individual variables of interest. Omitted variable bias has greater consequences in efforts to explain the influence of a single variable of interest than it does on efforts to predict an overall house price.

Spatial or location data are cited several times as an important but commonly omitted independent variable. Sheppard calls this omission a persistent and potentially serious error (Sheppard, 1999). Inclusion of independent variables accounting for location is important to hedonic housing price estimation. Techniques for developing, analyzing, and correctly applying such variables is relatively new. A review of the issue and statistical techniques is included here.

Sheppard's observation regarding space is worth repeating: "It is easy to argue that land price will vary with the location and the variance in price produces variance in type and intensity of land use … Given this observation, it is surprising how many hedonic models lack … a variable that explicitly identifies the location of the structure" (Sheppard, 1999, p. 1616). Wiltshaw reinforces this observation about the importance of location, writing, "Geographical location is a fundamental characteristic of property. Any methodology concerned with assessing valuation accuracy must incorporate location explicitly into its diagnostic procedures" (Wiltshaw, 1996, p. 275). Anselin (1988) illustrates the effect of including or omitting spatial variables with a study comparing two models of the effect of household income and house values on neighborhood crime incidents in 49 neighborhoods in Columbus, Ohio. One model includes the effect of spatial relations; the other does not. Comparing the results of the two, he finds that about 10 percent of the variation in crime rates is explained by spatial relations (reported in LeSage, 1998).

In the past, generating location information was at best, difficult and laborious. Recently, GIS has developed as an important new tool for capturing spatial dimensions and deriving other important, previously difficult to obtain, variables. GIS provides techniques for adding spatial detail as well as precision and different ways to measure accessibility. For example, in addition to straight-line distances from point-to-point, use of GIS street networks can provide estimates of travel distances along road networks and travel times to use as accessibility measures. As another example, GIS allows overlay of census data on parcel plat maps of individual properties so that neighborhood variables from the census can be associated with specific properties (Clapp, Rodriguez, & Thrall, 1997; Des Rosiers, Thériault, & Kestens, 2003; Geohegan, Lisa A, & Bockstael, 1997).

Inclusion of location information in a hedonic model is not without complications. Luc Anselin is generally credited with producing the first major text dealing with spatial econometrics (Anselin, 1988). Much of the material included here is from LeSage as well as Anselin (LeSage, 1998). There are two new problems when location is included in an econometric model: (1) spatially related heterogeneity, and (2) spatial dependency between observations.

Spatial heterogeneity refers to variation in relationships over space. Different types of development, activity, behavior, and so on are found at different places. As an econometric issue, spatial heterogeneity is likely to be found with models that use datasets that include dissimilar areas, such as different housing markets. Both Anselin (1988) and Butler (1982) point out that an effect of this dissimilarity is that functional forms and parameters vary with location. This observation relates to the assertions of Butler (1982), Sheppard (1999), and others that there is little theoretical guidance to

choosing functional forms for hedonic models; different forms are more appropriate for different areas as well as different data specifications.

Addressing spatial patterning of observations requires a specification for variation over space. Coding data to reflect location in a census tract or block, a political jurisdiction, rural or urban places, and so on is a way of including possibilities of spatial variation in a model. LeSage points out that the specification of location variables should be parsimonious; only a few parameters can be used for modeling spatial variation. The example given by LeSage is the simple (and parsimonious) division of places into ones that are urban and ones that are rural. Many questions surround development of parsimonious specifications, including the sensitivity of a specification to spatial variation, consistency with the data, and comparisons of competing specifications. For example, if data is classified as belonging to either an urban or rural region, we must ask: (1) is the classification consistent with the data, or is there reason to believe there should be more than the two groups, (2) are the estimates biased if the classification system is inconsistent with the data, and so on (LeSage, 1998).

The second complication introduced along with spatial variables is spatial dependence. Spatial dependence, or spatial autocorrelation, means that an observation taken at one location is dependent on other observations at other locations. Anselin cites Tobler's first law of geography to illustrate the meaning of spatial dependence, "everything is related to everything else, but near things are more related than distant things" (Anselin, 1988, p. 8). Spatial dependence comes about for at least two reasons. First, a series of observations associated with a particular area, e.g. a zip code or a census tract, might be affected by some form of common measurement error (LeSage, 1998).

Wiltshaw, discussed above, provides a good example when he writes about use of tax assessment data as a proxy for price (Wiltshaw, 1996).

Another reason to expect spatial dependence in the data is that the spatial component of an economic activity may be truly important and truly present. One example of this is the Anselin study of crime in Columbus, Ohio, reported above. A second example comes from LeSage who calculates the distance to the CBD for each of 35,000 housing sales over a five-year period in Lucas County, Ohio. On average, the homes are found to be progressively younger as distance from the CBD increases. The pattern indicates distinctive heteroskedastic patterns and spatial dependence that needs to be taken into account in any modeling of economic activity in this area (LeSage, 1998). Spatial heterogeneity creates difficulties for econometric techniques that do not account for spatial variation in the relationships being modeled.

Spatial dependence, or spatial autocorrelation, is similar to time series autocorrelation regression except the errors are related over space not time; they are multi-dimensional, not one-dimensional. When faced with spatial autocorrelation, "[p]roxial observations should have closely related error terms and the strength of this relationship should diminish as distance between the observations increases. The error terms should become independent when the distance separating the observations becomes sufficiently large" (Dubin, 1988, p. 466).

Spatial heterogeneity and spatial dependence are often related problems in hedonic models. A number of methods estimate relationships varying over space. Likewise, a great number of methods also make corrections for spatial autocorrelation (Anselin, 1988; Can, 1990; Carter & Haloupek, 2000; Dubin, 1988; LeSage, 1998).

As is the case with multicollinearity, spatial autocorrelation affects efficiency of estimation, but does not bias results. As with multicollinearity, spatial autocorrelation can be addressed with a fully specified model. "It may be possible to reduce spatial dependence through the choice of explanatory variables…[methods to correct for spatial autoregression] should only be used if spatial dependence remains after all relevant variables have been included" (Dubin, 1988, p. 466). Batemen, et al. echo this caveat, writing that if all characteristics of structure, neighborhood, accessibility, and environment are included as explanatory variables in an hedonic estimation, then similarities of selling prices of neighboring houses will be accounted for. When this is not the case, there is some resulting correlation between selling prices of neighboring houses not explained in estimated hedonic equations. Spatial autocorrelation generally does not lead to bias, but it does lead to greater variance and less precision in the coefficients. On the other hand, if the hedonic model can account for some similarity in prices due to spatial correlation, if it can remove some noise, the variables will become more precise and it is possible to make clearer inferences. To Batemen, Day, Lake and Lovett (Batemen et al., 2001) dealing with the loss of precision due to spatial autoregression is not as high a priority as accurate identification of functional form and working through multicollinearity.

There are techniques that incorporate the effects of spatial autocorrelation into regression analysis. The most widely used is spatial autoregression (SAR). This technique requires creation of a "contiguity matrix" relating the spatial location (i.e. the point) of each observation to every proximate observations (see, for example, Anselin, 1988). The technique is complex, can be time consuming, and it may not always be

needed, as spatial autocorrelation may not be present. Batemen, Day, Lake and Lovett advise always using diagnostic tests to see if there is a problem before proceeding with this expensive method (Batemen et al., 2001).

Multicollinearity, already briefly mentioned several times, is a problem that commonly arises with hedonic modeling of housing prices. Generally, multicollinearity is ignored: it usually has little effect on a model's overall predictive power and there is little that can be done about it (Woolridge, 2000). But, if the objective is to understand how individual independent variables impact housing price, multicollinearity can be a substantial problem.

The term "multicollinearity" refers to the presence of a high degree of linear correlation between independent – explanatory – variables. In a housing model, for example, the number of rooms may be highly correlated with floor area. Both may be correlated with price. If multicollinearity is present, "the coefficients cannot be estimated with great precision or accuracy" (Gujarati, 1995, p. 322). The larger standard errors lead to wider confidence intervals. In the room/floor area example, the effect of the number of rooms on price will become "muddled" with the effect of floor area. If floor area were the variable of interest, the resulting lack of precision would be a problem. Gujarati explains that where there is high multicollinearity, many coefficients individually have statistically insignificant *t* tests, yet the R^2 for the entire model may be quite high, over 0.9, for example, and the *F* test may be significant. The presence of a high degree of multicollinearity, while not affecting the development of "best fit" can hide the significance of individual variables. Therefore, if the object of a study is to find the hidden price of a specific attribute, the presence of a high degree of multicollinearity may

be a problem if the variables(s) of interest are correlated with other independent variables. Because the objective of this dissertation is to estimate the impact of specific independent variables, the research should be sensitive to possible problems of multicollinearity.

Problems of multicollinearity can occur in any of the vectors of independent variables in the standard hedonic housing model. In the vector of structural characteristics, as mentioned, the number of rooms in a house can be strongly correlated with floor area. The problem also occurs frequently with environmental variables. For example, as Batemen et al (2001) explain, both noise and air pollution arise from a highway and higher levels of traffic result in greater levels of both. If, for some reason, these are both variables of interest, the problem becomes one of teasing the two apart to find their separate influence on property value. This is a difficult task.

Others have made these same points. Woolridge (2000) states that the only truly effective approach to address problems of multicollinearity is to add more explanatory variables to the equation. Likewise, Sheppard (1999) advises that the best approach is a larger and richer data source. Nevertheless, unless a truly fully specified model is developed, adding more variables may exacerbate the problem rather than alleviate it.

> "Another important point is that a high degree of correlation between independent variables can be irrelevant to how well we can estimate other parameters in the model. For example, consider a model with three independent variables:
>
> $$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \mu$$
>
> where x_2 and x_3 are highly correlated. Then Var(βhat$_2$) and Var(βhat$_3$) may be large. But the amount of correlation between x_2 and x_3 has no

direct effect on Var(βhat$_1$)… If β_1 is the parameter of interest, we do not really care about the amount of correlation between x_2 and x_3 …[This] observation is important because economists often include many controls to isolate the causal effect of a particular variable … But high correlations among these variables do not make it more difficult to determine the effects of [a variable of interest]" (Woolridge, 2000, p. 97).

Consequently, as discussed above, the presence of a high degree of multicollinearity may have the effect of "hiding" the price effect of an individual variable in a combined effect of numerous correlated independent variables. But, a high degree of multicollinearity may not be a problem if the objective of a study is to find the "hidden price" of a specific attribute and that attribute is not highly correlated with other variables in the equation or if the degree of correlation can be reduced.

There is no easy solution to the problem of multicollinearity. Gujarati lists over seven approaches; none are especially satisfactory. One of the simplest, and most tempting, is to drop one of the collinear variables. But, this may lead to specification, or omitted variable bias (Gujarati, 1995).

Batemen, et al. (2001) report that it is sometimes possible to overcome multicollinearity problems with more accurate measurement of variables, and/or developing full specification of independent variables. Remembering the problem of trying to isolate the effects of both air pollution and noise arising from highways, inclusion of trees and the location of trees, steep banks and their location, and other spatial features which dissipate noise may help separate estimation of the effects of noise and air pollution.

Another approach to address multicollinearity used with some frequency is to combine highly correlated variables into one index through use of "principal components" (Des Rosiers et al., 2003; Kain & Quigley, 1970; Lake, Lovett, Bateman, & Langford, 1998; Orford, 2002). Principal components analysis is used as a data reduction method. A regression line representing the best summary of the linear relationship is fitted among collinear variables and the variables themselves are discarded. The new line is a new variable that captures most of the essence of the original multiple items. This line is a principal component variable. Because it has little correlation to other independent variables, it can be used to substitute for two or more highly correlated variables related to variables of interest. This addresses imprecise estimates of variables of interest. Use of principal components introduces problems of interpreting variables folded into a principal component, but if they are variables of little interest, the problem is not great.

Functional form definition for hedonic models is another common challenge. The basic approach to hedonic price estimations is comparatively straightforward and simple. The quantity (or quality) of each characteristic is regressed against the house price. The method controls for the other characteristics in the model (all else is held constant) and implies a straight-line relationship between the dependent variable (price) and the independent or explanatory variables (Batemen et al., 2001).

On its face, this notion of the hedonic relationship is over-simplified. There are reasons to believe that relationships between explanatory variables and price are not straight-line and constant. Freeman (1993) shows that the relationship between housing characteristics is generally non-linear as substitution of characteristics is generally not

feasible: "two living rooms with six-foot ceilings are not equal to one living room with a twelve foot ceiling" (Freeman, 1993, p. 371). Jones (1988) has shown that in equilibrium hedonic price functions should be convex, not straight-line. People cannot trade the individual characteristics of complex goods directly; they must be traded as they exist, as bundled goods. So, even in a Pareto non-optimal situation, it may not be possible to find mutually advantageous trades of the complex goods, even though characteristics could be traded advantageously if they could be separated. Therefore, in equilibrium, the marginal rates of substitution for two parties must be different and prices cannot be linear.

Another reason for putting aside a strict assumption of linear relationships lies in the notion of diminishing marginal utility. This idea leads to the belief, for example, that the three-thousandth square foot of floor area in a house is worth considerably less to most consumers than the six-hundredth square foot. The relationship of floor area to price, then, might best be stated as a logarithmic or quadratic, not linear, relationship (Batemen et al., 2001). As another example, both positive and negative environmental effects often play off against each other creating complicated non-linear relations with housing price. Li and Brown (1980) use a complex form to model the positive effects of proximity to commercial areas as they play off against the negative influences of commercial disamenities such as congestion and noise. These authors assume that positive price associated with access declines with distance (which they represented with a logarithmic function), and that negative price influence associated with diseconomies also declines with distance (which they modeled with an exponential function). They

further assume that the negative influences decline more rapidly than positives influences. The net effect is a downward sloping convex (from below) rent gradient.

The functional form of a model describes the relationship between the dependent variable and the explanatory variable(s). An incorrect functional form leads to misspecification bias. In hedonic modeling, researchers typically use one of four specifications: (1) a linear specification in which the dependent and independent variables are all left untransformed, (2) a "semi-log" form in which the dependent variable is transformed to a logarithmic form, but the independent variable(s) is left in linear form, (3) a "log-linear" form with a linear form dependent variable and the independent variable(s) transformed to a logarithmic form, and (4) the "log-log" form with both the dependent and independent variable(s) are transformed to logarithmic form.

Which functional form is correct? There is no theory on which to base selection of one of these specifications (Batemen et al., 2001; Butler, 1982; Malpezzi, 2002; Sheppard, 1999). "Theory offers little guidance about the form … so researchers have tended to regard the choice of functional form as an empirical question" (Butler, 1982, p. 97). Early investigators tended to rely on linear or logarithmic forms. Malpezzi (2002) argues five points that recommend semi-log models: 1) they allow value added to vary with size and quantity, 2) ease of interpretation, 3) they tend to mitigate heteroskedasticity, 4) they are computationally simple, and 5) it is possible to build flexibility into the right hand side using dummy variables. But, Sheppard (1999) reports that beginning in the 1980s, the field began to see the use of flexible Box-Cox modeling. Box-Cox transformations have the objective of finding exponential values for a regressor, a combination of regressors, and/or the dependent variable that have the effect of

minimizing total error in the residuals. Maximum likelihood methods are always used to find Box-Cox transformations.

$y(\lambda) = ((y\hat{}\lambda) - 1) /.$ for λ not equal to zero

$y(\lambda)=\log(y)$ for $\lambda=0$ (Box & Cox, 1964)

Thus, the data, in a way, find the functional form that is their own "best fit."

But the issue is not settled. First, Butler finds comparison of various forms shows little reason for choosing one over another: many functional forms could be equally good (Butler, 1982). Second, Cropper, et al. (1988) compared different Box-Cox specifications. They evaluated the output not in terms of best fit (the usual way), but rather in terms of estimating "true marginal bids." Contrary to Butler, they did find one form to be superior. If all attributes were included and there was no measurement error, they found the linear Box-Cox performed best and the quadratic Box-Cox performed worst.

In the face of all this, Sheppard (1999) has concluded that selection of functional form is determined by the data and the objective of the research. It is common to use flexible functional forms, such as Box-Cox. Importantly for this dissertation, he cautions that if the objective of the research is to determine implicit price of an attribute - the variable of interest - and not predict the total price, a "best fit" (minimum squared error) estimation may not be as appropriate as a statistical procedure that yields low variance of the parameter estimates.

Econometric Issues

Econometric issues rise when the objective of estimation is not simply characteristics' prices, but rather demand, or even compensated demand, for characteristics. Hedonic

regressions are used to estimate price functions and study component prices of characteristics in composite goods. Ridker and Henning (1967) publish one of the first papers to attempt to estimate the residential price effects of air pollution using an hedonic price function. Freeman (1971) criticizes use of hedonic price functions independent of considerations of demand and supply.

> "This [hedonic price] equation only purports to explain the variation in mean property values among observations. The air pollution coefficient can be used to predict the difference in property values between two properties under *ceteris paribus* conditions, and these conditions must include no change in air quality over all other land in the system. But the regression equation cannot be used to predict the general pattern of property values or changes in the value of any given property when the pattern of air quality over the whole urban area has changed" (p. 415).
>
> "What is required is a model which can be solved to yield the pattern of land rents as a function of the pattern of air quality, among other things ... When such a model is developed and the appropriate expressions for regressing land values on air quality are deduced, we will find, I am sure, that the regressor coefficient for air quality contains both supply and demand elements" (p. 416).

A two-step theory of hedonics that goes beyond basic hedonic price estimation and ties the price estimates to basic demand-supply theory was presented by Rosen (1974). “The hedonic equation is determined by the bids that consumers are willing to make for different bundles of characteristics and the offers of those bundles by suppliers”

(Palmquist, 1984 p. 395). Hedonic price estimation becomes the first stage in a two-stage least squares procedure developed by Rosen. This theoretical approach has gained wide use when the effort is to estimate individual demands in a supply and demand framework (Palmquist, 1984).

Many writers summarize and evaluate Rosen's contribution (for example Bartik, 1987; Follain & Jimenez, 1985; Palmquist, 1984). The following summary is taken principally from Bartik (1987). Rosen starts using a marginal bid function taken from estimated hedonic prices; in equilibrium, the marginal bid equals the marginal price. Further, in equilibrium, consumers' marginal bid equals suppliers' marginal offer:

$$\frac{\partial p}{\partial z_j}(Z_i) = W_{ij} = B_0 + B_1 Z_i + B_2 X_i + B_3 D_{oi} + e_{ij},$$

$$\frac{\partial p}{\partial z_j}(Z_i) = G_{ij} = A_0 + A_1 Z_i + A_a S_{oi} + u_{ij},$$

where

$(\partial p/\partial z_i)(Z_i)$ is the estimated hedonic marginal price of z_j,

W_{ij} is the marginal bid by consumer *i* for z_j,,

X_i is the consumer's expenditure for all other goods,

D_{oi} is a vector of observed consumer traits that can affect the marginal bid (e.g. income),

G_{ij} is the marginal offer by firm *i* for z_j,

S_{oi} is a vector of observed supplier traits that can affect the marginal bid (e.g. technology), and

e_{ij} and u_{ij} are error terms (Bartik, 1987).

Rosen argues that this system of equations can be solved with simultaneous two-stage least squares (2SLS) estimation methods, but many find difficulties with this approach. In addition to the data and statistical problems discussed already, the basic problem with this approach is the problem of simultaneity. Follain and Jimenez (1985) discuss two sources of simultaneity in this hedonic model. The first is correlation of error terms with independent variables in either equation. This problem arises if aggregate data are used, but not with micro-level data (Follain & Jimenez, 1985). The second arises from the non-linear nature of the price function, as discussed by Bartik (1987). Bartik argues that because of the assumption of a competitive market with both consumers and suppliers being price-takers, an individual consumer's decision cannot affect the hedonic price function. There is an estimation problem arising from consumers' ability to endogenously choose both quantities and marginal prices of Z_i due to the non-linearity of the hedonic price function. An unobserved consumer trait, taste, is in the error term and is correlated with the consumer's choice of Z and X. Therefore, estimation of the hedonic equations for demand and supply will be biased (Bartik, 1987).

The two-step approach also has an identification problem if aggregate data from a single market are used. Follain and Jinenez (1985) prescribe use of a Box-Cox functional form to address this issue. Freeman (1999), Palmquist (1984), and Whitehead (1999) all propose that calculating the demand function from multiple markets and tracing the locus of the individual demand functions creates a composite and unbiased estimate.

In addition to Rosen's two step approach, Follain and Jimenez (1985) describe several empirical approaches that use hedonic modeling: the simple hedonic approach, the bid-rent approach, and a discrete choice approach. None of the approaches is without

problems in addition to the statistical and data problems detailed earlier. The most important approach for this dissertation is the simple hedonic approach.

The simple hedonic approach uses only the first step; demand parameters are inferred directly from coefficients of estimated hedonic prices. The derivative of the regression with respect to a given characteristic is interpreted as the marginal willingness to pay for that particular characteristic. "In general, the hedonic equation will overstate the valuation of an additional unit of the characteristic … This difference can sometimes be important if one is doing a cost-benefit analysis" (Follain & Jimenez, 1985 p. 81). However, under some conditions or assumptions, the simple approach can be used to estimate demand parameters. If, for example, as many have contended, the supply of characteristics is perfectly inelastic at any location, ordinary least squares (OLS) may be used to estimate demand equations and the simple approach is adequate (Palmquist, 1984).*

Follain and Jimenez (1985) find that many papers – "too numerous to summarize" (p. 81) – use the simple hedonic approach to estimate market valuation of particular housing characteristics. Finally, Bateman, et al. (2001) conclude that the two-step approach is rarely used due to many econometric complications. They proceed to use the simple hedonic approach.

The bid-rent approach directly estimates bid-rent functions as opposed to demand equations. The method assumes that consumers that receive equal utility from consuming a particular characteristic can be identified and properly grouped. For this approach, the

---

* Folain and Jiminez caution that an assumption of an inelastic supply is an empirical question that needs to be examined when used. Hausman tests are used.

problem of identifying groups of individuals with "equal utility" is serious (Follain & Jimenez, 1985).

The discrete choice approach is the final approach discussed by Follain and Jimenez (1985). Discrete choice deals with characteristics that are not continuous: a municipal water system is available or it is not. Logit, probit, and other discrete choice models are used for estimation of probabilities that households of given characteristics will occupy housing of given characteristics. As the estimate necessarily produces probabilities and not estimates of price or willingness to pay, results are more difficult to interpret, especially for policy purposes. Additionally, the computational costs and restrictions are high and may be overly restrictive. Income, for example, always an important consideration in demand analysis, is more easily dealt with as a continuous variable, not discrete.

Follain and Jimenez (1985) conclude that choice of a procedure depends partly on the data as well as the objective of the research. Earlier we saw that the simple hedonic approach will often overstate the valuation of an additional unit of a characteristic. Any price estimate from a simple approach maybe upwardly biased and should be considered as upper bound, not a "median" estimate of demand (Follain & Jimenez, 1985; Whitehead, 1999). We also learned that this over valuation might not be important unless a cost-benefit analysis is the task. We also saw that if there is an assumption that the supply of characteristics is perfectly inelastic at any location, the simple hedonic approach may be adequate. Use of micro-level as opposed to aggregate data also is advantageous in that it leads to an assumption that price estimates in the simple approach are independent of error terms and analysis may go forward with direct price estimation.

Last, Freeman (1971) indicated in his criticism of Ridker and Henning that the simple hedonic approach is acceptable if the objective of the research is to estimate differences in property values ceteris paribus conditions and with no dynamic parameters. This review of the complexities of hedonic analysis has shown that biased estimates may always be a problem under any approach and that appropriate diagnostic tests should be employed.

Conclusions from hedonic literature

This review of hedonic literature applies to this dissertation in several ways. First, the purpose of the dissertation is to test differences in residential prices as they may arise from proximity to commercial development, estimation of demand is not an issue. Second, most data to be employed, and importantly the dependent variable and the variable of interest, are micro-level data. Correlation of error terms with independent variables will not be as great a potential problem as with aggregate data. Last, given the nature of the built environment in the observation area, it is reasonable to use an assumption that supply of housing characteristics is perfectly inelastic and the simple hedonic approach is adequate.

CHAPTER 3

NEW TECHNOLOGIES

Emerging new technologies, mentioned often in the foregoing discussions, speak to many of the issues raised above. Over the last two decades, computer technology has greatly expanded the capacity to keep, access, and use real property data. Geographic information systems (GIS) and computer assisted mass appraisal (CAMA) systems address issues of data availability and spatial analysis. Space syntax is among a number of new and still developing methods of analyzing design properties of urban layouts.

Computer Assisted Mass Appraisal

Contemporary CAMA databases typically hold the promise of providing the "rich" database many theorists call for full model specification (Sheppard, 1999; Woolridge, 2000). These possibilities are discussed extensively several places above. The King County, Washington, tax assessor's database, for example, contains records for approximately 591,000 tax parcels. The database has fields for over 150 structural and parcel characteristics plus a history of sales dates and prices. Among the characteristics fields included, in addition to the address and unique parcel identifier, are: number of stories, floor area, presence of a basement and level of basement finish, parking, environmental amenities (e.g. views) and disamenities (e.g. airport noise), current taxable value, age, and history of construction. Dating from 1992, records of over one million residential sales transactions are included in the database.

Geographic Information Systems

"Perhaps one of the most exciting means for extending hedonic modeling is making use of the spatial structure of the data, using the emerging technology of geographic information systems" (Malpezzi, 2002, p. 29). GIS provides an array of extremely powerful tools for storing and manipulating large amounts of information and the spatial relationships among that data. Using a GIS, virtually any kind of spatially related data can be placed on a digital map, then visualized, compared, measured, and analyzed. For example, the techniques described by Anselin (1988) for correcting spatial autocorrelation can be performed in a GIS. A great diversity of information can be mapped and analyzed ranging from population demographics, health statistics and epidemiology, utility and transportation networks, flood protection zones, crime patterns, historical sites, sales data, disaster areas, and much more (Davis, 2003). Much of this and other data are relevant to housing valuation. The potential of GIS for hedonic housing price modeling is only now becoming widely appreciated. Several inclusions in the literature reviewed above emphasized the importance of spatial relations to hedonic price modeling. Yet, only a few papers, and those of relatively recent vintage, make extensive use of GIS and its mapping and spatial analysis abilities. Not only is the technology relatively new, but the databases, such as the CAMA databases, needed to feed GIS input into hedonic housing studies are only now becoming easily accessible (Clapp et al., 1997).

A GIS that contains a topologically structured map of tax parcels can easily relate all the tax assessor detail from a CAMA database on all the individual parcels to each other and map the relation of all parcels. For example, a GIS map of King County's

parcel maps can be linked to the unique parcel identifier codes. Then assessment values, sales and sales values in the last two years, for example, can be mapped. Additional possible capabilities of a GIS, given the data, include relating each individual parcel to neighborhood attributes from the census and other sources, development of noise level estimates from roads and generation of other environmental characteristics, and measurement of access in sophisticated ways, e.g. estimated travel distances from point to point over a mapped transportation network. Thus, GIS and good tax assessor databases have the potential of dealing with two of the specific weaknesses of hedonic housing price modeling pointed out above: underspecification and lack of spatial information.

Measures of Neighborhood Layout

Several measures can develop interval values that index street patterns and neighborhood layouts. These techniques can be used to compare neighborhoods on several measures. Indices of integration – ease or difficulty of movement to all points in an area – can be created. Although there is little experience with these measures, an index may be useful as an independent variable describing neighborhood design in hedonic models.

Neighborhood settings and design are major environmental factors not included in studies of land use relationships. There is research that shows different types of neighborhood layouts are profoundly different from one another (Jo, 1996). Highly integrated grid system neighborhoods may have high values associated with both positive influences of proximity and negative influences of externalities. This would be a situation similar, for example, to that described by Li and Brown (Li & Brown, 1980). The effect of the intentional isolation of the cul-de-sac design may be very effective in

reducing the impact of disamenities. On the other hand, cul-de-sac design may also impose a real travel cost that far exceeds an apparent distance between residential and commercial uses, thus reducing the positive influence of convenience. In other words, in a neighborhood with a segregated type layout, there may be no proximity effect at all; neither the positive or the negative effects of proximity. It is possible that differences in neighborhood layout may explain, in part, why some past studies of the relation between housing prices and proximity to non-residential uses have found significant positive or negative relations and others have found no relationship at all. Assuming there is an effect, it is possible those studies resulting in significant relationships, either positive or negative, may use highly integrated neighborhoods while those with no relationship may be set in segregated neighborhoods.

Accessibility is a function of integration and integration, in turn, is a function of design (i.e. street layout) that is independent of land use and the relation of land uses (Peponis, Ross, Rashid, & Kim, 1996). Accessibility is one of the four vectors included in many of the general hedonic models used for housing price analysis. A neighborhood layout index can be used to introduce a new aspect of accessibility to the model. Street network configuration as it affects movement can be added to a model containing such measures as metric distance and network distance and even interacted with these measures.

Spatial integration is a function of accessibility and accessibility, in turn, is a function of design (i.e. street layout) that is independent of land use and the relation of land uses (Peponis et al., 1996). Accessibility is one of the four vectors included in many of the general hedonic models used for housing price analysis. Space syntax can be used

to introduce a new aspect of accessibility to the model. Street network configuration as it affects movement can be added to such measures as metric distance and network distance.

Researchers are exploring the use of several measures of street layout. Connectivity may include measures such as the ratios of street segments to intersections, street length per house, and the ratio of cul-de-sacs to connective streets in a neighborhood. These connectivity measures have been shown to have explanatory power related to housing value differences by neighborhood layout (Crane 2000; Song and Knaap 2003).

Space syntax analysis of street layout is the most esoteric of the methods in development. Referred to by its researchers as a research program as opposed to a theory, mathematically derived form graph theory, space syntax is based on the notion that movement in an urban space, all else being equal, is generated by the configuration of the space. This relation between layout and movement in fact underlies many other aspects of urban form: land use relations, for example the spatial relation of retail and residence, the spatial patters of crime, even patterns of building densities (Hillier, 1996; Hillier & Hanson, 1984).

A substantial number of studies show that spatial configuration correlates highly with observed movement of pedestrians and autos. Penn, for example, cites five studies showing high correlations between space syntax indices of spatial configuration and pedestrian movement and two studies showing high correlations with auto movement (Penn, 2003).

The basic tool of space syntax uses to begin analysis of street layout is an axial map. Axial maps describe – with interval numeric indices – urban street networks (Bafna, 2003). There are several measurements typically generated from axial maps. The measures that correlate best with movement, both pedestrian and auto, are measures of spatial integration (Penn, 2003). An axial map identifies the fewest and longest lines needed to cover all travel ways. "The integration value of a line is a function of the minimum number of other lines that must be used to reach all other parts of the system … Numerical integration values are ... relativized according to the mathematically possible range of integration, and then standardized so as to allow direct comparison between systems of different sizes" (Peponis et al., 1996, p. 7). More simply put, in an axial map, the number of distinct turns on a route is more crucial to access than actual route distance. Distance is counted as depth, and depth is measured as the number of turns along a path from one place to another, rather than the actual length of the trip. Thus, space syntax measures of neighborhood integration – directness of travel and access – can be developed.

Other measure of street layout, e.g. the ratio of street segments of intersections, are more intuitively understood than space syntax. See Appendix D for a more detailed discussion of measures of street layout.

CHAPTER 4

STUDY DESIGN CONSIDERATIONS AND METHODOLOGY

This study's design addresses some basic concerns: geographic focus, data, and methodology.

Geographic Focus

Hedonic housing price studies reviewed earlier focus on geographic areas that range from areas as large as several metropolitan areas down to individual neighborhoods. The choice of the type of area on which to focus is a function of the purpose of the study. Studies that examine the price effects of specific characteristics, especially neighborhood or environmental characteristics, focus on neighborhoods as study areas (Bowes & Ihlanfeldt, 2001; Cao & Cory, 1981; Grether & Mieszkowski, 1980; Kain & Quigley, 1970; Li & Brown, 1980). As Mills (1979) and others point out, these effects are small and fall off rapidly with distance. Further, virtually all of the studies that focus on neighborhoods use census tracts as stand-ins for neighborhoods (Bowes & Ihlanfeldt, 2001; Cao & Cory, 1981; Grether & Mieszkowski, 1980; Kain & Quigley, 1970; Li & Brown, 1980). Because this dissertation focuses on variables of interest that have relatively small effects, a small area focus - neighborhoods - is necessary. Census tracts are used as neighborhoods.

This dissertation uses data from King County, Washington (Seattle). The focus is at the neighborhood level. The study area is a swath of census tracts in King County, Washington running from the Puget Sound north of Seattle's downtown to the eastern urban growth boundary, east of the cities of Kirkland and Redmond. Lake Washington bisects the study area. The Seattle portion of the total sample, west of the lake, includes 28 census tracts and approximately 43,650 parcels. The Kirkland/Redmond portion with ten tracts and approximately 15,150 parcels is east of the lake. There are 19,085 observations on the west side and 6,740 on the east.

The portion of the study area west of Lake Washington in Seattle is older with smaller houses and lots developed on a gridiron pattern. The portion of the study area east of Lake Washington is the Kirkland/Redmond area, which is an edge city as defined by Garreau (1991), dominated by a curvilinear/cul-de-sac street pattern.

Data

This dissertation uses the King County, Washington (Seattle) tax assessor CAMA database coupled with the county's tax parcel level GIS mapping system as the principle source of data. While this data source is very complete, there are some data issues. The variables of interest in this dissertation are all access/proximity type variables, most of which are generated by GIS software.

An hedonic price study begins with gathering the selling price of a great many properties. For each property, there is need for structural, neighborhood, accessibility, and environmental characteristics.

The King County, Washington tax assessor's database is used for several reasons: 1) It is available in complete form on request from the King County Tax Assessor at

virtually no cost with delivery in less than a week. While many other tax assessors in the U.S. have large, reasonably complete databases, they are not so readily available. 2) The fact that the tax assessor advertises the database for sale on the internet attests to the confidence the assessor has in the accuracy and usability of the data. 3) The assessor's information includes records of residential sales back further than 1992. Since 1992, the database contains records of over one million sales. After culling (Mahan et al., 2000) this data for questionable records (the assessor codes for events such as sales to close relatives, and so on – another strength of this database), over 400,000 sales records are left to use. 4) The tax assessor's property assessment records "tie" to the King County GIS tax parcel map. The spatial relation between tax parcels - and between tax parcels and other features - can be estimated using the GIS map.

For the purpose of this dissertation, however, the King County database/GIS also has several weaknesses. The weaknesses and the steps to deal with them are:

- Weakness. Several writers have discussed the importance of perception to the impact of negative externalities (Batemen et al., 2001; Cummings & Landis, 1993; Espey & Lopez, 2000; Li & Brown, 1980; MaRous, 1996; Wilhelmsson, 2000). The tax assessor's data does not contain this information. The GIS contains only 100 foot contours; this is not fine grained enough to use for visual impact estimating. Screens and berms are not included in either the tax assessor's data or the GIS.

  Correction: Data on the visibility or non-visibility of commercial sites from residential properties is added to the database from field observations.

- Weakness. The smallest units in the GIS are tax parcels. To measure distances between parcels, whether straight-line distance or travel distance, the GIS system needs to work from specific points that reference each parcel.

  Correction: The GIS system is used to generate “centroids” (the calculated central point) of each parcel. Parcel centroids are used for all distance measurements in this dissertation.

- Traffic volume and traffic noise are two potentially important measures of negative externalities associated with proximity to commercial areas. The database does not record traffic volume (although there is a field for that data) and presents traffic noise as 0, 1 (low), 2 (moderate), and 3 (high).

  Correction: This dissertation uses only the traffic noise data. To a large degree, traffic volume will be captured in this variable. If these were specific variables of interest, traffic volume information would be found and added to the database. But, this is only a control variable; using only traffic noise suffices, given the cost of adding traffic volume data (Batemen et al., 2001).

- Weakness. School performance is an important control variable for estimating housing price (Li & Brown, 1980). Standardized test score data for each school is used (*School Guide*). There is a GIS overlay map with school locations, but there is no GIS map for school service areas. The lack of this map prevents associating school test scores with specific houses.

  Correction: A GIS overlay map of elementary schools services area is created using map images made available by both the Seattle School District and

the Lake Washington School District (*Lake Washington School District #414: Elementary School Boundary Map*, ; *Neighborhood Attendance Reference Areas*).

- Weakness. Sales data spans the period 1989 to 2003. There most certainly are substantial changes and fluctuations in market conditions effecting housing prices over this long period. Further, there may be influences other than inflation and effects may differ over the extent of the sample area.

  Correction. Rather than add a CPI inflation factor to the analysis, a non-linear trend, using squared and cubed terms, derived from the year of sale, controls for varying market trends.

Access/Proximity

Access and proximity measures are the principal variables of interest in this dissertation. The specific variables are 1) the Euclidian or straight-line distance, 2) the travel distance, and 3) the neighborhood index of integration. Euclidian and travel distances from residential to commercial uses in each neighborhood are measured on a parcel centroid to parcel centroid basis using the closet centroid possible. GIS spatial analysis capabilities are used to build these measures (Jenness, 2004a, , 2004b). All else being equal, the strength of positive effects should increase as travel distance decreases and the strength of negative effects should increase as Euclidian distance decreases. Measures of design integration are created for each neighborhood (census tract) using GIS.

Methodological Issues

The unsettled nature of hedonic housing price modeling discloses several methodological issues for this dissertation. These involve the nature of the variable of

interest, the functional form, the level of specification, and problems of collinearity, heterogeneity, heteroskedasticity, and spatial autoregression.

As this dissertation focuses on variables that seem to operate over small distances and have effects that are hypothesized to vary with neighborhood type, the analysis focuses at the neighborhood level – using several neighborhoods of various types – rather than on a city-wide or metro-wide level. This basic approach follows Grether and Mieszkowski (1980). Li and Brown's (1980) assumptions underlie the model used here, but there is an important difference. Their model postulates both a positive effect of convenience, measured as travel distance, and negative spillovers that travel over a direct line. The Li and Brown (1980) model assumes that the positive effect of convenience reaches further than that of negative externalities. The model they use reflects these expectations, using a logarithmic exponent for travel distance variables but an exponential exponent for the straight-line measures. This dissertation, using a much larger, more comprehensive, and more accurately measured database than that available to Li and Brown (1980) does not make the same assumptions, but rather tests these assumptions. This dissertation uses a simple OLS model with quadratic forms specified for both the travel distance and straight-line distances. If Li and Brown's (1980) assumptions hold, the quadratic specification for both variables in this OLS should produce a similar result.

Theory gives no good guidance on functional form. This analysis will include some complex quadratic and interacted variables and investigates areas that quite possibly are different markets. For generality and ease of interpretation, a simple linear form is used (Batemen et al., 2001).

Level of Specification

Review of the literature shows that reduced form models are appropriate for investigations that seek to develop high levels of explanatory and predictive power over entire markets (Butler, 1982). On the other hand, if the purpose is to explore the influence of a specific variable (or small range of variables) on housing price, use of fully specified models and sensitivity to omitted variable bias and multicollinearity is appropriate (Malpezzi, 2002). This dissertation is of the second type. Full specification is the preferred approach to developing the right-hand side of the model and is the approach used in this dissertation. The rich King County Tax Assessor's database, ability to generate spatial data with GIS, and the other accessible data facilitate development of a fully specified model.

Statistical Issues

Use of hedonic modeling for estimation of the value of components of housing prices raises many statistical issues: heterogeneity, omitted variable bias, collinearity, spatial autocorrelation, and others already discussed. Extensive diagnostics are used to test for these problems. Heteroskedasticity is present in the regressions; consequently robust estimators are used. As expected with a fully specified model, multi-collinearity is also present. As a result, some variables are removed from the model. Census tracts are highly collinear with measures of distance and direction to major centers. Census tract designations are removed (Bowen, Mikelbank, & Prestegaard, 2001). Likewise, census data itself, e.g. income, racial mix, etc. is also highly correlated with the distance and direction measures to major centers. In addition, inclusion of social characteristics such

as income, due to rationing effects and effect on bids for individual characteristics will produce bias (Butler, 1982). Census characteristics, though very important to an overall understanding of the setting, must be removed from the hedonic regression.

Model

The King County, Washington Tax Assessor's CAMA database contains a great deal of detail on real property. The database provides the dependent variable, residential sales price, and the sales date and detailed residential building descriptions used as major control variables. Additionally, the database contains detailed information on non-residential properties also used as control variables in this dissertation.

Sales Data

Sales price is the dependent variable. The Real Property Sale Record File contains data on over 1 million sales. There are 31 fields in the file. Of the 31 fields, those used in this dissertation are shown in Table 1.

Table 1: Variables from Real Property Sales Records

Field Name	Description	Comment
Major		Major and minor combine to form the parcel
Minor		code number, a unique identifier
Sale Price		Tax Assessor
Sale Date	MM/DD/YYYY	Tax Assessor: Transformed to *Trend*
Property Type	The type of property	Tax Assessor: Land only or land with improvement
Principal Use	Primary use of the property	Tax Assessor: Agriculture, residential, etc.
Historic Property	Special historic valuation	Tax Assessor: Used to cull non-market sales
Sale Warning	Coded entry	Tax Assessor: Used to cull non-market sales

From King County distribution disk

"Sale Price" is the dependent variable. Other information from the file classifies property as residential, culls non-market transactions, and "ages" the sales price.

Structure Data

The vector of data on each residential structure is used as control data. The Residential Building Description file contains records for over 427,000 properties and 43 fields of information for these properties. Residential structure data is:

Table 2: Variables Describing Residential Structures

Field Name	**Description**	**Comment**
Major	Major and minor combine to form the	Relates data to data in other files or maps
Minor	parcel code number, a unique identifier	contain a parcel code number
sqfttotliv	*Square feet: living area*	Tax Assessor
sqftliv	*Square feet: living area* ^{2}	Tax Assessor
bedrooms	*Number Bedrooms*	Tax Assessor
sqftxbedrooms	*Square feet: living area*Bedrooms*	Tax Assessor
bathrooms	*Number Bathrooms*	Tax Assessor
age	*Age*	Tax Assessor: from construction year
agesq	*Age* ^{2}	Tax Assessor: from age
condition	*Structural Condtion*	Tax Assessor: 1 to 5 - 5 is best
sqftlot	*Square feet: lot*	Tax Assessor
sqftlotsq	*Square feet: lot* ^{2}	Tax Assessor: from sqftlot

From King County distribution disk

In addition to the unique identifier, there are ten independent variables describing the structure and lot size of each house. Floor area is included as a quadratic term to test the notion of diminishing marginal utility of increasing floor area (Batemen et al., 2001). Lot size is included as a quadratic term based on the same reasoning. Bedrooms interacted with floor area is included to control for the possibility that increasing the number of bedrooms without increasing total floor area may decrease the value of the remainder of a house (Craig, Kohlhase, & Papell, 1991).

Neighborhood Data

Neighborhood data will be used to control for differences in neighborhood characteristics.

This dissertation uses census tracts as neighborhoods. Data describing neighborhoods is shown in Table 3.

What constitutes relevant neighborhood data is problematic (Batemen et al., 2001). Data that may reflect demand, i.e. social characteristics, is omitted from the hedonic regression (Butler, 1982). The neighborhood variables used in this dissertation are selected for a variety of reasons.

Table 3: Variables Describing Neighborhoods

Field Name	Description	Comment
itbs_read	*Iowa test: Reading*	Elementry school test score
density	*Density*	GIS Calculation: Census data
densitysq	*Density* ^{2}	GIS Calculation: Census data
nonresmix	*Proportion of nonresidential use*	GIS Calculation
nonresmixsq	*Proportion of nonresidential use* ^{2}	GIS Calculation
apt1_dis	*Distance apartment*	GIS Calculation
apt1_dissq	*Distance apartment* ^{2}	GIS Calculation
apt1_az	*Direction apartment*	GIS Calculation
cult1_dis	*Distance cultural use*	GIS Calculation
cult1_dissq	*Distance cultural use* ^{2}	GIS Calculation
cult1_az	*Direction cultural use*	GIS Calculation
govt1_dis	*Distance government use*	GIS Calculation
govt1_dissq	*Distance government use* ^{2}	GIS Calculation
govt1_az	*Direction government use*	GIS Calculation
hotel1_dis	*Distance hotel*	GIS Calculation
hotel1_dissq	*Distance hotel* ^{2}	GIS Calculation
hotle1_az	*Direction hotel*	GIS Calculation
off1_dis	*Distance office*	GIS Calculation
off1_dissq	*Distance office* ^{2}	GIS Calculation
off1_az	*Direction office*	GIS Calculation
hops1_dis	*Distance hospital*	GIS Calculation
hops1_dissq	*Distance hospital* ^{2}	GIS Calculation
hosp1_az	*Direction hospital*	GIS Calculation
ind1_dis	*Distance industry*	GIS Calculation
ind1_dissq	*Distance industry* ^{2}	GIS Calculation
ind1_az	*Direction industry*	GIS Calculation
sch1_dis	*Distance school*	GIS Calculation
sch1_dissq	*Distance school* ^{2}	GIS Calculation
sch1_az	*Direction school*	GIS Calculation
seg_tnodes	*Ratio: street segments to intersections*	GIS Calculation
seg_unodea	*Ratio: street segments to cul-de-sacs*	GIS Calculation

From King County distribution disk, *Seattle Times School Guide* , author's calculation

The "Iowa Test of Basic Skills: Reading" (*itbs_read*) operates as a measure of the quality of neighborhood elementary schools, an important consideration in many home purchase decisions and a variable often included in hedonic housing price analysis (Batemen et al., 2001; Clark & Herrin, 2000; Goodman & Thibodeau, 1995).

Residential d*ensity* (Butler, 1982; Li & Brown, 1980) is calculated dividing total land in a census tract devoted to housing, including apartments (taken from the GIS map), into total population from the census. *Density* is included as a quadratic to test the idea that the effect is non-linear.

The *proportion of nonresidential use* is also calculated from the GIS map; it is the acreage in nonresidential use as a proportion of total acreage. This variable is entered as a quadratic based on the findings of Cao and Cory (1981).

There is a series of variables describing the location of the closest nonresidential use, other than retail, relative to each house sale observation. These variables are used to control for the influence of these uses on house price and to differentiate their influence from the influence of retail uses. The classes of uses are apartments, cultural uses, government uses, hotels, offices, hospitals, industrial uses, and schools. The variables are the straight-line distance, distance squared, and the direction (azimuth). Distance and direction are generated using the NEAREST FEATURES extension to the ARCVIEW geographic information system software (Jenness, 2004a)

Neighborhood layout indices are generated for each neighborhood. This is a new variable, never before used in an hedonic housing price analysis. This variable is generated for each neighborhood (census tract) from the GIS map.

Environmental Data

The vector of environmental variables included in this dissertation include:

Table 4: Environmental Variables

Field Name	Description	Comment
noview	*No view*	Tax assessor, field observation
wfntlocati	*Water front location*	Tax assessor
trafficnoi	*Traffic noise*	Tax assessor - scale 1 to 3: 3 is worst
visibility	*Visible nonresidential use*	Field observation

From King County distribution disk, author's field observations

The environmental amenities of a view of mountain ranges or the downtown skyline or a waterfront location are so valuable in King County that they are entered as separate factors in the assessor's database. That, and informal discussions with residents indicate these are important control variables to include in this analysis. The tax assessor's database actually contains fields for ten different views, e.g. Mt Rainier, Cascades, Puget Sound, etc. with the quality of a view graded 0 to 4. For simplicity, this data is converted to a single dummy variable, there is a view or there is not.

Traffic noise also uses data taken from the Tax Assessor's database.

Visibility is a dummy variable indicating that nonresidential uses are or are not visible from a given house. One of the important tests included in this dissertation is a test of the price effects of proximity to retail uses. Whether or not visual pollution – measured with *visibility* – is par of that effect is an important consideration (Batemen et al., 2001; Patterson & Boyle, 2002) Field observations are used to generate this data. The variable is coded as "1" if a non-residential is visible from a given house, "0" otherwise.

Access Data.

As with the other non-structural vectors, the operating definition of "access" is not settled. Distance to the CBD is typically included as a primary measure of access and major control variable (Adair, McGreal, Smyth, Cooper, & Ryley, 2000; Bowes & Ihlanfeldt, 2001; Goodman & Thibodeau, 1995; Li & Brown, 1980; Rodriguez, Sirmans, & Parks, 1995) . But, the importance of the measure diminishes as cities become less monocentric. Consequently, some researchers are including measures of access to important places other than a CBD. This dissertation uses takes the approach of using several measures of general access: distance and direction from each house sale observation to downtown Seattle (the actual point is the Bank of America Tower), distance and direction to downtown Bellevue (the actual point is the Microsoft headquarters building), and distance and direction to the nearest expressway on-ramp. Distance and direction are generated using the NEAREST FEATURES extension to the ARCVIEW geographic information system software (Jenness, 2004a)

Table 5: Access Variables

Field Name	Description	Comment
dis_bofa	*Distance Seattle CBD*	GIS Calculation
dis_bofasq	*Distance Seattle CBD* ^{2}	GIS Calculation
az_bofa	*Direction Seattle CBD*	GIS Calculation
dis_mic	*Distance Bellevue CBD*	GIS Calculation
dis_micsq	*Distance Bellevue CBD* ^{2}	GIS Calculation
az_mic	*Direction Bellevue CBD*	GIS Calculation
dis_xway	*Distance expressway*	GIS Calculation
dis_xwaysq	*Distance expressway CBD* ^{2}	GIS Calculation
az_xway	*Direction expressway CBD*	GIS Calculation

Compiled by author

Variables of Interest. The two primary access variables are "proximity" and "distance." Proximity is the straight-line distance from the centroid of a residential parcel to the centroid of the nearest retail property. Distance is the travel distance, measured on street centerlines, from the centroid of a residential parcel to the centroid of the nearest retail property.

Table 6: Variables of Interest

Field Name	Description	Comment
r_net1	*Travel distance to retail*	GIS Calculation
r_net1sq	*Travel distance to retail* ^{2}	GIS Calculation
areu1_dis	*Straight distance to retail*	GIS Calculation
areu1_dissq	*Straight distance to retail* ^{2}	GIS Calculation
areu1_az	*Direction to retail*	GIS Calculation
average	*Average distance between retail*	GIS Calculation
avesq	*Average distance between retail* ^{2}	GIS Calculation

Compiled by author

The variables measuring the average distance between retail is a measure of the clustering of retail uses around the nearest retail use. It is a secondary measure of convenience and an additional control. If additional retail sites are easily accessible for the first, value of convenience is enhanced.

Both distance measures are entered as quadratic terms, based on the assumptions of Li and Brown, but the forms are left to act for themselves. In addition, the travel distance variable is interacted with the space syntax variable to test the idea that street layout affects access and convenience. In the sane vein, the straight-line distance measure is interacted with density to test the idea the effects of negative spillovers are exacerbated by increasing density.

Model

The model hedonic to be estimated is:

$P = \beta_0 + \beta_1 S + \beta_2 N + \beta_3 E + \beta_4 A + \beta_5 \text{ Proximity} + \beta_6 \text{Distance} + ų$

Where:

S is the vector of structural characteristics

N is the vector of neighborhood characteristics

E is the vector of environmental characteristics

A is the vector of accessibility characteristics

Proximity and Distance are the variables of interest

β_0 is the y (price) intercept

ų is the error term

Software

Several major pieces of software are used in the analysis:

- ARCVIEW is the GIS program used. It links databases on geographic bases, measure “proximity,” “distance,” and the other access variables. It is also used to create location data needed for spatial analysis.
- AXWOMAN is a spatial syntax extension to ARCVIEW that is used to create neighborhood layout indices.
- SAS is a database management and statistical analysis program. It is used to manipulate the very large databases called for in this dissertation.
- STATA is a statistical analysis program. It is used to estimate regression equations.

CHAPTER 5

RESULTS

This chapter presents results of the study. A description of the study area and great difference that exist between its two distinct parts highlight the setting for the study. The demographics are quite different as are housing types and the physical layout of the development. The hedonic analysis looks not only at the study area as a whole, but also divides the area into four sub-parts - a western and eastern part and within them, observations that are within and not within walking distance of retail sites. Specific results for the price effect of residential proximity to retail are presented for all four areas. Finally, hedonic analysis is used to examine effects of neighborhood layout on the price effects.

Study Area

The study area is a swath of census tracts running from the Puget Sound water-front to the eastern urban growth boundary (See Figures 1A and 1B). The area runs north of Seattle's downtown and the University of Washington and through portions of the cities of Kirkland and Redmond as well as Seattle. There are 38 census tracts in the area encompassing 176 census block groups and approximately 58,700 tax parcels, residential and non-residential. Lake Washington divides the study area into two unequal parts. The Seattle portion is west of the lake. This portion includes 28 census tracts, 133 block

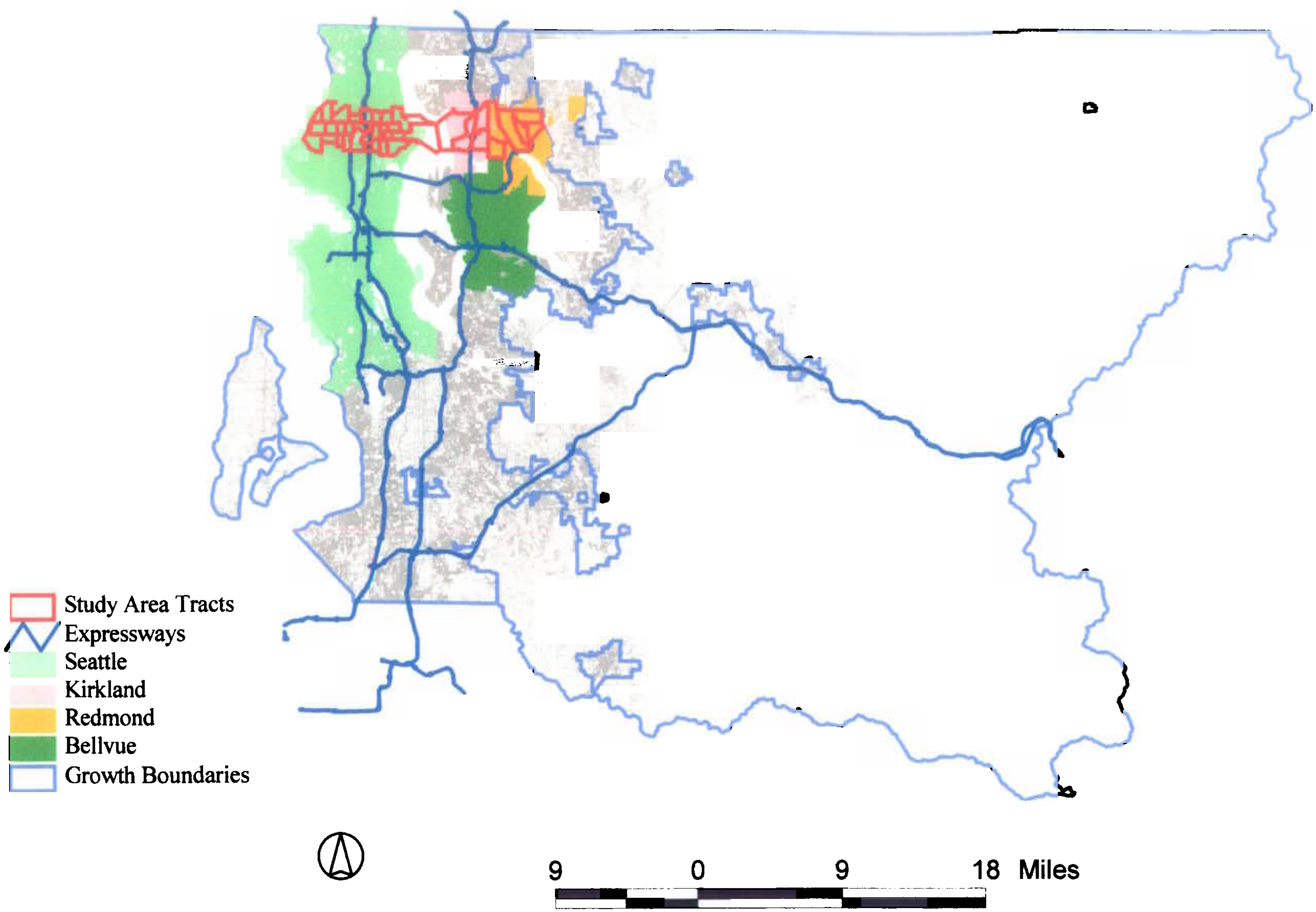

Figure 1A:
Study Area Within King County

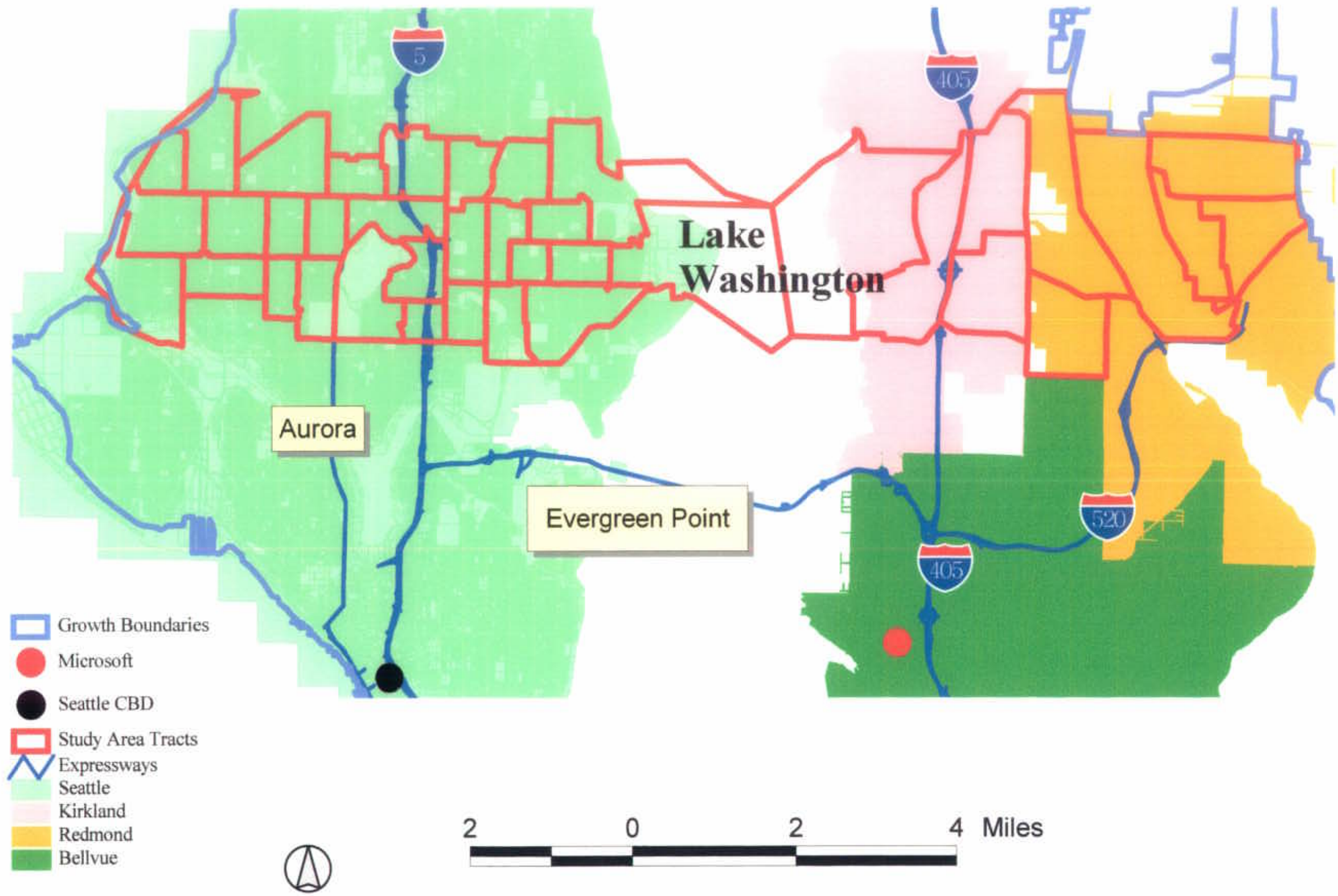

Figure 1B:
Detail of Study Area

groups, and approximately 43,650 parcels. The Kirkland/Redmond portion with 10 tracts, 43 block groups, and approximately 15,150 parcels is east of the lake. There are 19,085 observations on the west side and 6,740 on the east. The straight-line distance between the two sides is about 2.2 miles over water, but the travel between the two, requiring use the Evergreen Point Bridge well to the south, is a distance over eight miles.

The two parts of the study area differ in their general design, design implications, and general population characteristics. These differences have important implications for the effect of proximity to retail properties on the price of houses.

General Design of the Study Area

The portion of the study area west of Lake Washington in Seattle developed in the era between WWI and WWII when automobiles were emerging in American culture, but did not dominate the landscape. The median age of a house in this sample is 72 years. The older development on the Seattle side of the lake is predominantly a gridiron pattern while the east side in Kirkland/Redmond is characterized by a curvilinear/cul-de-sac pattern.

The difference between the two areas is illustrated with the "Edge City" concept developed and defined by Joel Garreau in *Edge City: Life on the NewFrontier* (Garreau 1991). Bellevue, the home of Microsoft, is a prototypical edge city and is listed as such by Garreau. Kirkland and Redmond are contiguous to Bellevue and can be viewed as edge city extensions of Bellevue. Seattle is a traditional core city.

An edge city is a new city, not just a subdivision or small suburb, which has developed on the edge of an existing metropolitan area. As defined by Garreau, an edge city is substantial, meeting five criteria: 1) over 5 million square feet of office space, 2)

Table 7: Cities East-of-the Lake - Edge City Criteria

Office Space and Period Built

Period	Bellevue Square Feet	Redmond Square Feet	Kirkland Square Feet
1900-1909	14,213	6,474	22,688
1910-1919		9,960	4,070
1920-1929		8,512	7,292
1930-1939	11,940	1,694	7,645
1940-1949	29,173	15,466	8,022
1950-1959	203,082	22,676	22,164
1960-1969	377,860	204,170	113,279
1970-1979	2,071,097	318,792	230,812
1980-1989	7,646,040	1,843,216	1,486,274
1990-1999	1,369,419	7,173,853	1,231,502
2000-2003	1,647,067	2,779,243	462,028
Current Office Inventory	13,369,891	12,384,056	3,595,776
Current Retail Space	4,025,030	1,016,363	828,326
Jobs/Bedroom Ratio	2.63	3.97	2.50

Source: King County Tax Assessor Data
Table 1: Statistics bt Economic Sector
http://www.census.gov/epod/www/g97aff.htm

over 600,000 square feet of retail space, 3) more jobs than bedrooms, 4) not a developed place 30 years ago, and 5) recognized as a distinct place (Garreau 1991). Typically, edge cities are relatively low density, automobile oriented and do not facilitate pedestrian travel or mass transit. Along with Bellevue (not a city in this study area), Redmond also meets all the criteria and Kirkland fails narrowly on only one measure. Table 7 shows how Bellevue and the two cities in the east section of the study area measure against the edge city criteria.

Table 7 shows all three cities meet the retail space criteria (600,000 square feet), there are more jobs than bedrooms in all three cities, and that well over 90% of the office space in all three have been developed in the past 30 years (92.9% in Bellevue, 97.7% in Redmond, and 94.8% in Kirkland). While the median age of a house in the Seattle

sample, the median age of a house in the Kirkland/Redmond sample is only 25 years. Kirkland does fall short on total office space. Figure 2 shows the cumulative percentage of residential growth from 1900 to 2002 in both samples. Notice that most of the housing in the Seattle sample was built before growth began to accelerate in the Kirkland/Redmond area.

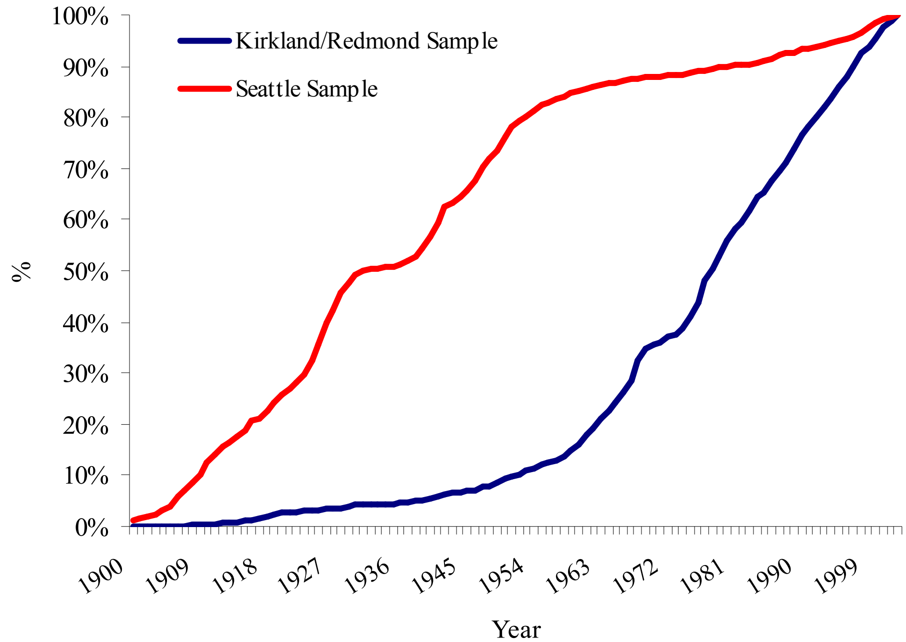

Figure 2:
Cumulative % of Housing Built: 1900-2003

Figures 3A and 3B, maps of street and retail patterns in the west and east sides of the study area, graphically display important differences implicit in the core city/edge city distinction. The western side of the study area, a portion of the Seattle core city, is characterized by a grid layout, short distances between intersections, and integrated land uses with small commercial properties "sprinkled" throughout the area lining many streets. The edge city side of the study area has fewer interconnected streets; there are

curvilinear and cul-de-sac streets in residential areas. The non-residential development tends to be segregated in to large commercial areas.

Design Implications

The implications for the two parts of the study area of being or not being an edge city are apparent in several measures. Edge cities are characterized by transportation dominated by automobiles. Table 8 compares travel mode commute time for the two sides of Lake Washington in the study area.

Eighty-eight percent of the workers on the east side, the edge city, travel to work

Table 8: Trip to Work Comparison in West and East Areas

					Comparison	
	West %	**Std Dev**	**East %**	**Std Dev**	**t**	**P>t**
Mode						
Automobile	71.51%	0.0800	88.06%	0.0277	9.7213	0.0000
Single Occupant	60.43%	0.0693	77.20%	0.0278	10.8829	0.0000
Car Pool	11.08%	0.0302	10.86%	0.0350	-0.1802	0.8596
Public Transportation	15.88%	0.0434	4.51%	0.0174	-11.7950	0.0000
Bicycle	2.82%	0.0188	0.44%	0.0047	-6.3284	0.0000
Walk	3.56%	0.0446	1.85%	0.0129	-1.8836	0.0674
Work at Home	5.47%	0.0172	4.45%	0.0198	-1.4477	0.1700
Travel Time to Work						
<5 minutes	1.28%	0.0078	2.45%	0.0167	2.1380	0.0573
5 - 10 minutes	5.98%	0.0206	10.53%	0.0236	5.4165	0.0001
10-15 minutes	11.11%	0.0195	16.55%	0.0392	4.2069	0.0016
15-20 minutes	16.04%	0.0340	18.95%	0.0261	2.7971	0.0110
20-25 minutes	17.83%	0.0204	15.42%	0.0221	-3.0318	0.0000
25-30 minutes	8.05%	0.0153	6.39%	0.0157	-2.8940	0.0109
30-35 minutes	15.80%	0.0177	12.12%	0.0200	-5.1581	0.0001
35-40 minutes	3.12%	0.0131	2.34%	0.0098	-1.9752	0.0615
40-45 minutes	4.28%	0.0105	2.57%	0.0088	-5.0218	0.0001
45-60 minutes	6.79%	0.0183	4.32%	0.0146	-4.3190	0.0003
60-90 minutes	3.19%	0.0150	2.90%	0.0132	-0.5291	0.5613
>90 minutes	1.18%	0.0071	0.99%	0.0078	-0.6891	0.5016

Source: U.S.Census 2000, Summary File 3 (SF3), CT P23 Journey to Work 2000, CT P31 Travel Time to Work 2000
http://factfinder.census.gov/servlet/QTSubjectShowTablesServlet?_ts=128811154750

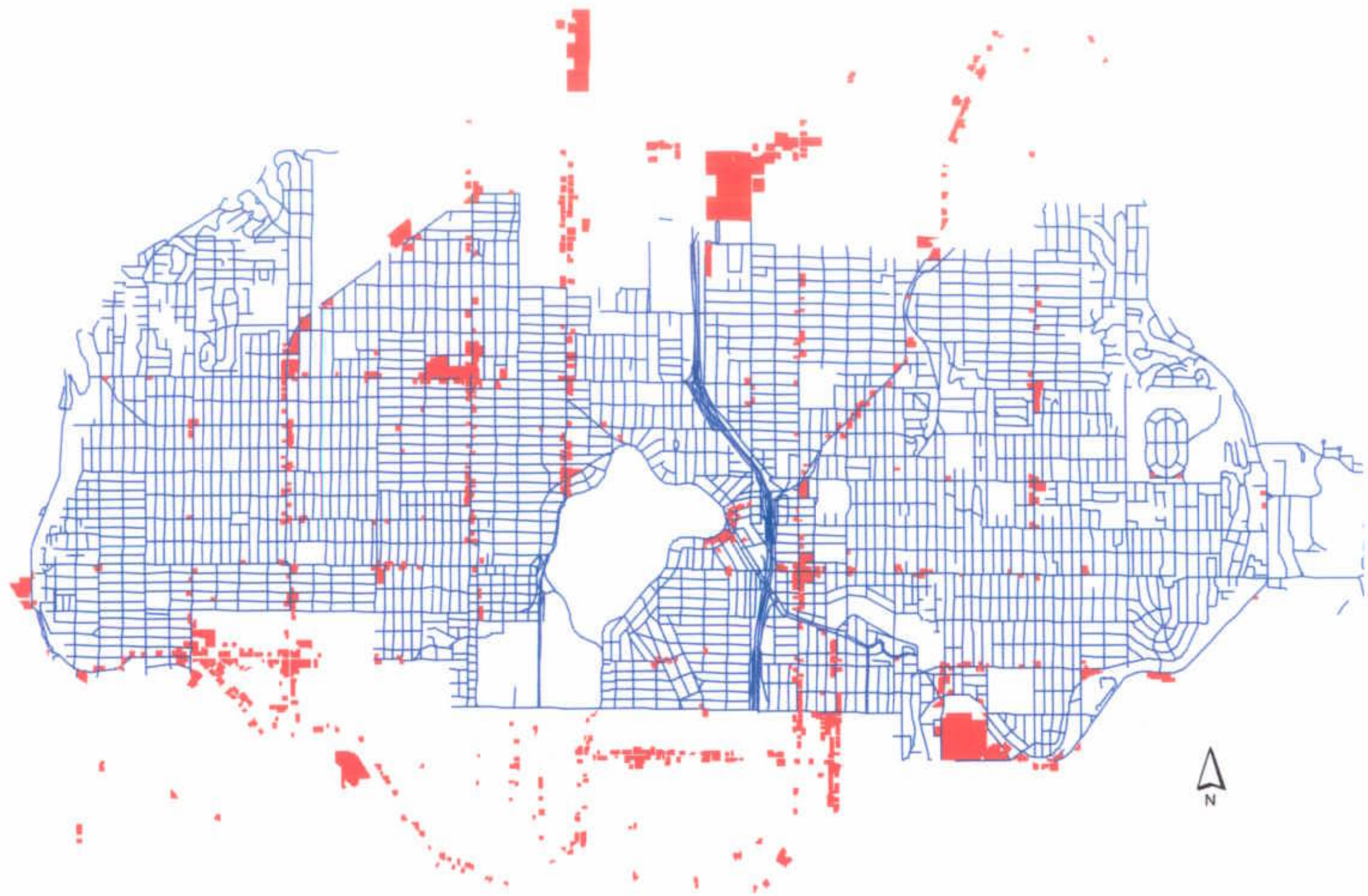

Figure 3A: Retail Patterns West of Lake

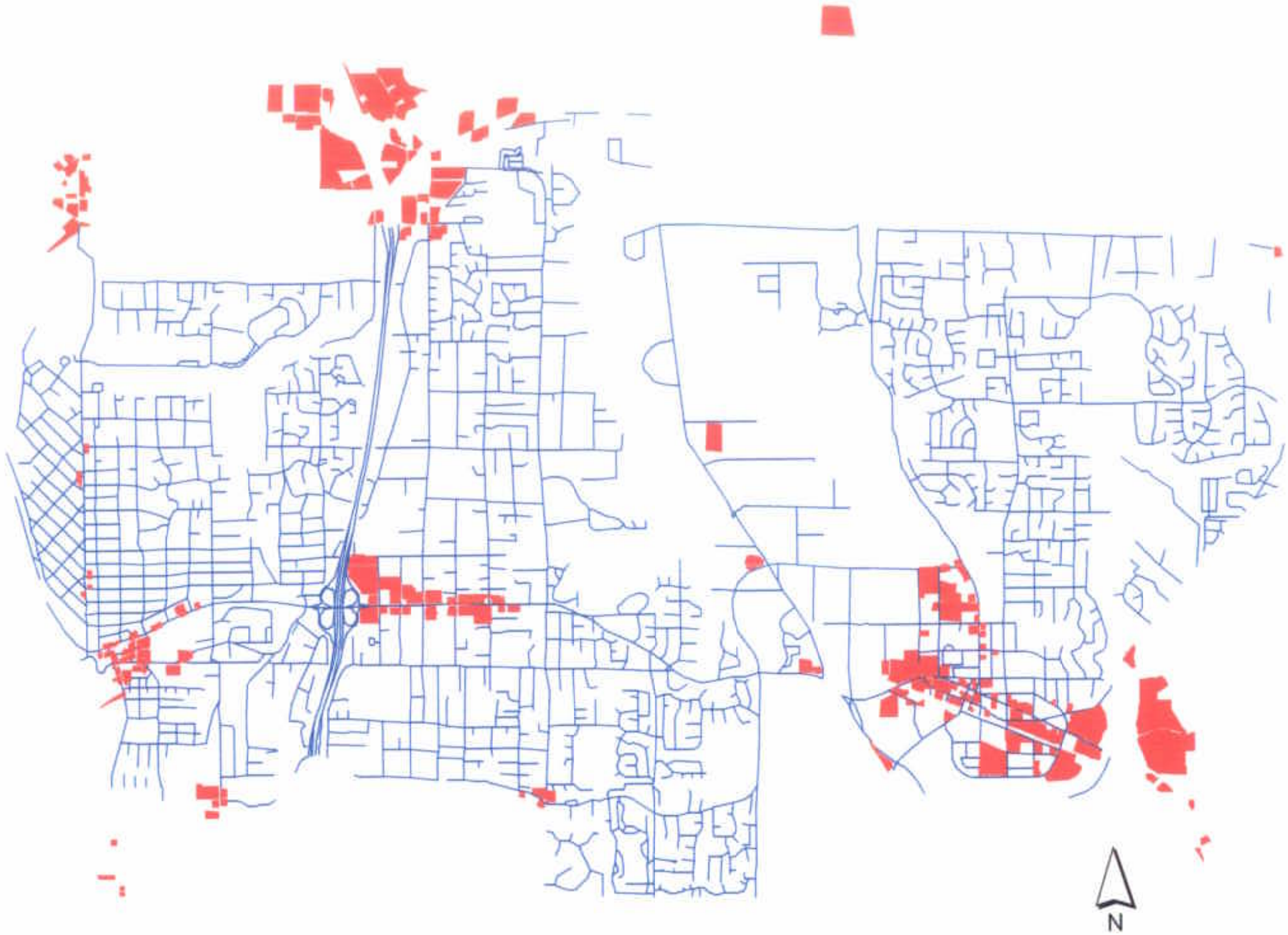

Figure 3B: Retail Patterns East of Lake

by car – over 77 percent in a single occupant car – while only 71.5 percent of the west side workers travel by auto. On the other hand, almost 16 percent on the west side use public transportation while almost no one on the east side does. Travel time to work tends to be less for east side residents. A higher percentage of east side commuters have shorter travel times to work, e.g. 10 to 25 minutes, whereas west side commuters have greater proportions in longer commute time brackets, e.g. 40 to 60 minutes. Both the travel mode and travel time data show the effects of the stronger automobile orientation of the east side.

Table 9: Comparison of Selected Variables in the West and East Areas

	West Sample Size	Mean	Std.Err	Std Dev	East Sample Size	Mean	Std.Err	Std.Dev	Comparison t	P>t
Saleprice	19,085	239,544.10	827.70	114,345.60	6,740	266,475.00	1,788.27	146,812.20	13.67	0.000
Square feet: house	19,085	1,603.45	4.75	655.88	6,740	1,991.80	8.61	706.49	39.52	0.000
Number Bedrooms	19,085	3.03	0.01	1.09	6,740	3.35	0.01	0.79	25.22	0.000
Number bathrooms	19,085	1.75	0.01	0.86	6,740	2.43	0.01	0.83	56.99	0.000
Square feet: Lot	19,085	5,102.96	16.06	2,218.49	6,740	10,189.12	106.21	8,719.62	47.35	0.000
Distance: nearest apartment	19,085	587.27	3.96	546.39	6,740	1,410.21	12.01	985.77	56.10	0.000
Distance: nearest cultural/entertainment	19,085	851.45	3.95	546.30	6,740	1,454.40	11.58	950.37	49.29	0.000
Distance: nearest government facility	19,085	2,741.76	9.12	1,260.27	6,740	4,129.55	26.60	2,183.50	49.36	0.000
Distance: nearest hotel	19,085	4,645.05	15.81	2,184.54	6,740	6,748.41	32.08	2,633.96	58.80	0.000
Distance: nearest office	19,085	894.85	5.05	698.11	6,740	2,111.34	22.26	1,827.61	53.29	0.000
Distance: nearest hospital	19,085	4,555.04	11.86	1,639.11	6,740	7,460.14	29.70	2,438.68	90.82	0.000
Distance: nearest industry	19,085	581.96	3.02	417.55	6,740	798.26	6.64	545.28	29.64	0.000
Distance: nearest elementary school	19,085	1,122.44	4.80	662.53	6,740	1,452.78	11.23	921.84	27.05	0.000
Straight distance: retail	19,085	890.04	4.46	616.51	6,740	2,366.69	14.16	1,162.52	99.46	0.000
Street distance: retail	19,085	1,231.16	6.51	899.68	6,740	3,610.46	22.06	1,811.32	103.43	0.000
Average distance: nearest 4 retail	19,085	791.23	4.45	615.23	6,740	2,044.25	24.18	1,984.93	50.96	0.000
Residential Density	19,085	24.91	0.06	7.75	6,740	14.08	0.04	3.13	-160.00	0.000
Non-Residential Mix	19,085	13.93	0.06	8.11	6,740	12.76	0.07	5.71	-12.87	0.000
Space syntax "integation"	19,085	1.61	0.00	0.60	6,740	0.81	0.00	0.37	-130.00	0.000
Street Segments/Intersections	19,085	1.84	0.00	0.65	6,740	1.13	0.00	0.12	-150.00	0.000
Street Segments/Cul-de-Sacs	19,085	113.65	0.93	128.01	6,740	4.79	0.03	2.66	-120.00	0.000

Source: Derived from King County Tax Data and King County GIS map

Significant and important differences between the two sides are also found in other distances and house and lot size. Larger houses and lots are hallmarks of suburban development, as compared to older core cities. Table 9 compares the west and east areas in terms of distances from residences to non-residential uses, house and lot sizes, residential density, and measures of neighborhood layout including the ratio of street

segments to intersections, the ratio of street segments to cul-de-sacs, and space syntax. Space syntax is a measure of street layout connectivity and integration based on the mathematics of graph theory presently used experimentally in the field of urban design (see Appendix D for a discussion of neighborhood layout measures).

There are significant differences on every comparison. Consistent with the maps in Figures 3A and 3B, distances between residential uses and non-residential sites are significantly and consistently shorter on the west side than on the east side. Of greatest interest here are the distances involving retail uses. On the Seattle side, the average straight-line distance between a house and store is 890 feet and on the Kirkland-Redmond side it is 2.367 feet. All else being equal, residences on the west side are more likely to be exposed to negative spillovers – noise, light pollution, etc. – than residences on the east simply because they are closer. On the east side only 4.2% of the residential observations are within a 500 foot radius of a retail site; on the west side the figure is 28.6%. For the two areas, the mean street or travel distances are 1,231 feet and 3,610 feet respectively. These are important distances to note: 1,231 feet is inside the maximum comfortable limit of about ¼ mile (1,400 feet) people will walk for shopping (Duany, Plater-Zybek, & Speck, 2000; Garreau, 1991). The average travel distance on the east side, 3,610 feet, is well beyond this comfortable walking distance.

Table 9 contains an entry called "Average distance: 4 nearest retail," a measure of retail clustering near residences. This measures the average distance from the retail site nearest a given residence to the four retail sites nearest that site. The average distance is much lower on the west where it is less than 1,400 feet, compared to the east where it is greater than 1,400 feet. These last two observations indicate that people on the west side

have a greater opportunity to walk to retail than people on the east. These factors are consistent with the characterization of edge cities being automobile oriented. Table 9 also shows that houses are smaller on the west side, with fewer bedrooms and bathrooms, and they sit on smaller lots. It follows that residential density is higher on the west side.

Measures of Neighborhood Layout

Looking over the maps in Figures 3A and 3B, it is apparent that the street and parcel configurations on the two sides of the study area differ. It is only recently that researchers have begun to develop methods measuring design and layout differences.

Space syntax is one such approach. Space syntax analysis is based on the notion that street layouts are lines that compose a graph with inherent patterns of connectivity. The patterns of connectivity may be simple and direct or complex and indirect, or anywhere in between. Generally, street intersections are analogous to nodes in graphs and streets themselves are analogous to "edges" in graph theory. Graph theory is used to analyze street layouts, with analysis leading to measures – or indexes – of ease of movement and access in the layout. Most interestingly, space syntax places a great deal of emphasis on the number of turns (including turns in curves, not necessarily intersections) needed in getting from one point to another, but does not include distance as a measure in any way (Bafna, 2003; Hillier, 1996, , 1999; Hillier & Hanson, 1984; Jo, 1996; Neiman, 2003; Peponis et al., 1996). Even though its predictive power in some empirical settings has been demonstrated (e.g. relation between space syntax measures and crime incidence in a neighborhood), space syntax is not well understood. The technique is ad hoc to the extent there is no underlying theory explaining why it does have the predictive power it seems to have (Penn 2003).

Other neighborhood design and layout measures have been developed to aid research on urban sprawl. Residential density and land use mix are included along with measures of connectivity of street and circulation systems (Crane, 2000; Song & Knapp, 2003b). Connectivity may include measures such as the ratios of street segments to intersections, street length per house, and the ratio of cul-de-sacs to connective streets in a neighborhood. Theoretically, the measures are related to spatial configuration and have greater intuitive appeal than space syntax. These connectivity measures have been shown to have explanatory power related to housing value differences by neighborhood layout (Crane 2000; Song and Knaap 2003). As with space syntax, distance is not included in these measures.

All these measures of neighborhood layout use the term integration. This term refers to how well streets in a neighborhood reach one another and also how well they reach the parts of the neighborhood. A street segment that is more easily accessed from more parts of the neighborhood is more integrated. Well integrated street segments are good commercial locations. In a cul-de-sac type layout, an arterial or backbone street is well integrated; the branching streets are not well integrated. In a grid system, all street segments are equally easy to reach by many different routes. The grid layout is more integrated.

Preliminary estimations show that space syntax measures differ between the east and west areas, but are not significant value determinants in these samples. On the other hand, measures based on the ratio of street segments to intersections and the ratio of street segments to cul-de-sacs proved more reliable in preliminary testing. Consequently, this study uses these measures rather than space syntax.

Both the space syntax comparisons and the node/segments ratios are shown in Table 9. Even though the space syntax measures are not used in the hedonic analysis that follows, it is still interesting to see that they differentiate the samples east and west of the lake. The east side has significantly lower scores on all three measure of street layout integration. As the space syntax index of integration increases, street layout connectivity increases (Jiang and Claramut 2002). The greater the ratio of street segments to intersections the greater the integration (Song and Knaap 2003) and the greater the ratio of street segments to cul-de-sacs the greater the integration. See Appendix D for further discussion of neighborhood layout.

General Population Characteristics

The 2000 Census population in the study areas west of the lake is 95,233 and east portion of the study area of the lake it is 30,757. Residential density is significantly higher in the west sample area at 24.91 persons per acre verses 14.08 per acre in the east sample (see Table 9). The Black population percentage west of the lake is significantly higher than east of the lake while the Asian population is a slightly higher percentage of the east of the lake population. The minority populations in either part of the study area are very small compared to the dominant majority population, which is 86.1% on the west and 84.4% on the east side. The school age population – the 5 to 17 year olds – is a much greater portion of the total population on the east side. The average family size is significantly larger on the east side of the lake. The proportion of female-headed households with children is significantly larger on the west side, but on both sides of the lake, female-headed households are a very small proportion of the total population.

Table 10: Statistical Comparison of Selected Population Variables in Study Area

	West		East		Comparison	
	mean	std	mean	std	t	P>t
% Black	1.79%	1.55%	1.23%	4.00%	-3.0011	0.0032
% Asian	6.66%	4.06%	8.14%	2.00%	2.0077	0.0486
% Age 5 or less	5.07%	1.45%	6.05%	18.00%	3.0806	0.0032
% age 5 to 17	10.92%	3.51%	15.44%	5.45%	5.1100	0.0000
% Age 50 to 64	13.98%	3.66%	15.65%	10.00%	2.3249	0.0233
% age 18 to 64	72.24%	7.78%	68.98%	27.00%	-2.4755	0.0156
Average household size	2.2033	0.2386	2.3779	0.28	2.4307	0.0187
Average family size	2.7385	0.1437	2.9065	0.03	3.7010	0.0005
% Married with children	16.55%	6.11%	23.91%	12.39%	3.7484	0.0005
% female head with children	3.56%	1.84%	4.60%	1.81%	3.2603	0.0017
% Vacant housing	3.07%	1.27%	4.05%	2.80%	2.2173	0.0314
Educational Attainment: Population 25 and Older						
Average years of schooling	14.93	0.5182082	14.45	0.3663013	-3.1624	0.0045
% no schooling	0.29%	0.33%	0.59%	0.58%	1.5744	0.1434
% to 4th grade	0.17%	0.34%	0.22%	0.22%	0.4581	0.6510
% 5th & 6th grades	0.24%	0.18%	0.22%	0.37%	-0.1094	0.9149
% 7th & 8th grades	0.77%	0.66%	0.64%	0.54%	-0.6039	0.5530
% 9th grade	0.45%	0.41%	0.53%	0.33%	0.6021	0.5541
% 10th grade	0.60%	0.50%	0.62%	0.33%	0.1359	0.8931
% 11th grade	0.77%	0.57%	0.80%	0.49%	0.1577	0.8764
% 12th grade, no diploma	1.39%	1.00%	1.72%	1.00%	0.9226	0.3701
% High school or GED	11.66%	3.89%	13.52%	3.62%	1.3718	0.1883
% some college< 1 year	4.73%	1.62%	6.59%	1.98%	2.6871	0.0182
% 1 year + college, no degree	14.00%	2.10%	17.07%	2.43%	3.5632	0.0031
% Associate Degree	5.72%	1.86%	7.34%	1.75%	2.4846	0.0240
% Bachelor's Degree	36.20%	4.30%	35.53%	5.49%	-0.3536	0.7293
% Master's Degree	14.04%	4.24%	10.60%	2.77%	-2.9221	0.0074
% Professional Degree	4.91%	2.13%	2.56%	0.90%	-4.8399	0.0000
% Doctorate Degree	4.06%	2.71%	1.46%	0.70%	-4.7279	0.0000
Median Household Income	$55,305	9,296.18	$65,364	6,523.40	3.7395	0.0011

Source: U.S.Census 2000, Summary File 3 (SF3), CT P6 Race, CT P10 Household Size,
P37 Sex by Educational Attainment Population 25 and Over, P53 Median Household Income
http://factfinder.census.gov/servlet/QTSubjectShowTablesServlet?_ts=128811154750

Educational attainment for the two populations is statistically undifferentiated except at the highest levels. There are significantly more persons with professional degrees and Ph.D.s on the west side (probably due to this part of the study area's proximity to the University of Washington). The east side, however, enjoys a significantly higher income; in fact, it is 18 per cent higher.

The development patterns differ as well. The average house size (1,972 square feet) and lot size (10,189 square feet) east of the lake are significantly larger than the average house (1,603 square feet) and lot (5,103 square feet) on the west side. The average number of bedrooms and bathrooms is also significantly higher on the east side

(see Table 9). Finally, as noted elsewhere, the average age of residences in the eastern area is much younger.

In summary, the two portions of the study area differ in the era and design style of development, mode of travel, size and price of housing, and in the general population characteristics of the people who live in the two different places.

Hedonic Analysis

This hedonic analysis studies the relationship between proximity to retail establishments and the price of housing. The analysis will examine these relations within the context of the very different neighborhood settings described above. The previous descriptions showed the significant differences in spatial configuration between the two sample areas. The hedonic analysis focuses on the effects of retail proximity on the price of housing and how differences in neighborhood configuration affect the relationship between retail proximity and the price of housing.

We expect that accessibility will have a positive effect on housing price while negative externalities will have a negative effect. The important results of the analysis include: 1) there are areas where proximity to retail sites has a significant effect on residential values and there are areas where the effect of proximity is insignificant, 2) in those areas where the effect is significant, the positive accessibility effect of proximity may outweigh the negative externality effect and the net effect is positive, 3) in those areas where there is no effect, the absence of effects appear to be due to highly segregated land uses, and 4) neighborhood design significantly influences the effect that travel distance (accessibility) and straight-line distance (negative externalities) have on housing price.

First, we will look at these relationships throughout the two parts of the study area. These areas are so different it can be argued they are different markets and should not be expected to behave similarly in this type analysis. Second, there will an analysis of only those observations within 1,400 feet of a retail site. One-thousand four-hundred feet or less is generally considered to be a walkable distance (Garreau 1991). Distances between residential and retail uses differ greatly in the two areas (see Table 9). Analysis will begin looking at the two areas on either side of Lake Washington as whole areas and then will proceed to use observations that are only within 1,400 feet of the nearest retail and then only those observations beyond 1,400 of the nearest retail.

Preliminary analysis contains evidence of heteroskedasticity in the regression analysis of all areas. Consequently, robust standard errors are used in all cases. Also, initial models include two sets of major spatial control variables: 1) variables locating each residence relative to the Seattle CBD and the Bellevue CBD and 2) census tracts. There is evidence of strong collinearity between these two sets that creates inconsistency in the signs and magnitudes of other variables. Because census tracts are used as neighborhoods in the analysis much data has been generated for and is attached to census tracts. There are several instances of collinearity involving use of census tracts as spatial location variables. Consequently, the Seattle CBD and Bellevue CBD variables are kept in the model and census tracts omitted.

Convenience and Negative Externalities

Convenience and negative externalities are the two unobserved variables believed to be important influences on housing price relative to proximity to retail establishments. If trips to the grocery store, for example, are inconvenient because of long travel distances,

the value of the inconvenience can lower housing price, just as locations that create long commutes to work can depress prices. Inconvenience could be a function of time or distance. Travel time may be a better proxy for convenience, but travel time information for these trips is not available. Travel distances, on the other hand, can be measured from GIS maps with relative ease. In neighborhood settings, loss of time due to congestion and other disruptions is generally not a problem and distance and time are substitutes. Therefore, travel distance to the closest retail establishment is a proxy for convenience/inconvenience and is a principal variable of interest.

Straight-line distance is a proxy for the influence of negative externalities and is the second principal variable of interest. Negative externalities are the potential annoyances that may arise from a retail site, for example noise, congestion, light and visual pollution, etc. Good measures of these influences are not available. The assumption is that the influence of negative externalities travel in a straight line, therefore, straight-line distances will reflect the influence of negative externalities. Of course, straight lines may be interrupted by trees, other buildings, changes in topography, and so on, and the influence of negative externalities may be stopped. A dummy variable measuring *Visibility* has been included to account for this possibility.

The two principal variables of interest are the straight-line distance and the on-street or travel distance from residential uses to retail. These distances are measured from the GIS map of land uses and streets. Squared terms of these variables are also included to pick up nonlinear effects. Secondary variables of interest describe neighborhood layout, density, and land use mix. All the other right hand variables usually found in hedonic price functions are included in the model to avoid omitted variable bias that

would affect these variables. For example age of each house has been included to account for potential deterioration, obsolescence, etc. Year of sale (*trend*), its square (*trendsq*), and cube (*trendcu*) are included to capture effects of inflation, cyclic property market trends, or other similar influences on each observation.

The variables squaring both the travel and straight-line distances are included on both the basis of theory and past analysis. Many past studies have shown that these type relations tend to be nonlinear (Li & Brown, 1980). Expectations are that housing price will diminish, all else being equal, with increasing travel distance to retail and that housing price will increase, all else being equal, as straight-line distance increases. Further, it is reasonable to expect to find limits to these effects, e.g., distances beyond which convenience or negative externalities do not extend. If there are nonlinear effects, these limits will be more apparent.

Additional Important Variables

In addition to the two principal variables of interest, travel distance and straight-line distance, several other variables are included to account for neighborhood characteristics that influence the relationship of residential and retail properties. These variables are discussed below.

The variable *Average* (and its square: *avesq*) is included to measure the clustering of commercial uses near residential properties. In the remainder of this discussion, this variable and it square, together, will be called the *clustering variable*. This variable is defined as the average distance from the closest retail site to the four retail sites closest to that retail site. These distances are also derived from the GIS map of parcels and streets in the study area. This is a broader measure of convenience supplementing convenience

proxies as travel distance from residence to retail. The shorter the average distance, the easier travel is between retail uses once the first retail use has been reached. The squared term is included to account for any declining effects.

The variable for *traffic noise* (*trafficnoi*) is taken from the tax assessor's data. *Traffic noise* may be a negative externality arising from commercial uses, but there may be commercial uses that do not generate traffic noise. Also, traffic noise may arise from major streets not directly related to commercial uses (Hughes & Sirmans, 1992). While this particular measure of traffic noise cannot be tied to retail sites, it is an important control variable. To the extent that traffic noise arises from a retail site and not from general traffic it should be picked up in the straight-line measure when this variable measuring background traffic is included in the model.

Visibility, as mentioned above, is also important variable of interest. This data comes from field observations. It is a dummy variable indicating that commercial properties are visible from individual residential properties. Clearly, visual pollution is not an issue if commercial properties are not visible. Other negative externalities may also be mitigated where there is no visibility.

Neighborhood Design Variables

Neighborhood design variables are new to residential hedonic price analysis. Song and Knaap (2003a; 2003b), for example, have only recently used this type analysis to assess the price effects of new urbanist neighborhood layout on residential prices.

In this dissertation, the neighborhood design analysis relies on several alterative measures including 1) a variable relating the number of street intersections to the number of street segments (*seg_tnodes*), 2) the ratio of total street segments to cul-de-sacs

(*seg_unodea*)*, 3) residential density (*density*), and 4) the ratio of residential to non-residential uses (*nonresmix*). All of these variables are calculated from the GIS maps and assessor databases linked to the GIS maps.

Interaction Variables

Last, there is a set of interaction variables linking the two principal variables of interest to neighborhood design variables. These are included to test the hypothesis that neighborhood design matters. First, travel distance is interacted with each of the measures of street integration, resulting in:

$$P = \beta_0 + \beta_1 TD + \beta_2 I + \beta_3 TD * I + \varepsilon$$

Where:

P = Sales Price

TD = Travel Distance

I = Integration (the ratio of street segments either intersections [*seg_tnodes*] or cul-de-sacs [*seg_unodea*],

Which may be interpreted using:

$$\frac{\delta P}{\delta TD} = \beta_1 + \beta_3 I$$

* Because several neighborhoods have no cul-de-sacs a value of 0.5 is added to the count of cul-de-sacs in every neighborhood. This is needed to avoid dividing by zero when constructing the ratio of segments to cul-de-sacs.

β_1, the coefficient on travel distance, is expected to be negative. Greater travel distance decreases convenience and price. If β_1 is negative and if:

- β_3 is positive then greater neighborhood integration reduces the marginal effect of travel distance on price.
- β_3 is negative then greater neighborhood integration increases the marginal effect of travel distance on price.

For example, if β_1 is –5.00, the price of a house will decrease $5.00 for every additional foot of travel distance to the nearest retail site, all else being equal. If the integration index is 2 and β_3 is +1, the price effect will be –3.00 (-5.00 + (1*2)) and the effect of travel distance reduced. On the other hand, if β_3 is –1, the price effect will be –7.00 (-5.00 + (-1*2). In this illustration, increased integration of street layouts enhances the effect of convenience, ceteris paribus.

The measure of straight-line distance is interacted with density (*densityXareu1*) and with the proportion of non-residential uses (*nonresXareu1*). As all of these measures potentially have negative effect on housing price, one may enhance the effect of the other.

$$P = \beta_0 = \beta_1 SD + \beta_2 D + \beta_3 SD * D + \varepsilon$$

Where:

P = Sales Price

SD = Straight-Line Distance

D = Density or proportion of non-residential uses

Which may be interpreted with the help of:

$$\frac{\delta P}{\delta SD} = \beta_1 + \beta_3 D$$

β_1, the coefficient on straight-line distance, is expected to be positive. As distance increases, negative effects (e.g. light and noise) decrease, and price is enhanced. If β_1 is positive and if:

- β_3 is positive greater density increases the price effect of distance.
- β_3 is negative greater density decreases the price effect of distance.

For example, if β_1 (the coefficient on straight-line distance) is +10.00, the price of a house will increase $10.00 for every additional foot of distance between it and the nearest retail site, all else being equal; the price is decreasing $10.00 for every foot closer. If density is five persons per acre and β_3 (the coefficient on the interaction) is +1, the price effect will be +15.00 (1* 5 [density] + 10.00 = 15.00) and the effect of distance on price is enhanced. Increased density would compound the effect of negative externalities. Alternatively, if β_3 is –1, the price effect will be diminished from $10.00 per foot to $5.00 and increased density would mitigate negative externalities.

Tables 1 through 6 present a full listing of all independent variables used in this analysis along with sources of the variables.

Pooled Samples

The first regression analyses are presented in Table 11 and 12. All observations on the west side of the lake are presented in Table 11 and those for the east are in Table 12. Sale price is regressed against a relatively full set of independent variables. The model used here is:

$$P = \beta_0 + \beta_1 S + \beta_2 E + \beta_3 V + \beta_4 O + \varepsilon$$

Where:

P = sales price

S = a vector of structural factors, such as size, lot area, number of bedrooms, and so on.

E = a vector of environmental factors such as view, distance to types of non-residential uses other than retail, location relative to downtown Seattle and the Bellevue employment center, etc

V = a vector of the variables of primary interest: travel distance and its square and straight-line distance and its square.

O = a vector of other variables of secondary interest including density, portion of non-residential uses, and so on.

The neighborhood design interactions are specifically not included at this point.

General Results

The two principal variables of interest are the straight-line distance and the on-street distance from residential uses to retail, most of the other right hand variables are included to avoid omitted variable bias that might affect these two variables. Tables 11 and 12 report the regression estimate using a simple linear form. The R^2s are in the range expected of hedonic price models, explaining about 73.3 per cent of the variation in housing sales prices in the west sample and about 71.6 percent in the east sample. In the west sample virtually every right hand variable has the expected sign and is statistically significant. In the east sample, not as many are significant. A Chow test of the equality of

the two samples (F = 39.80125, Prob F>= 0.0000) indicates the two samples should not be pooled.

Some of the more interesting variables in this estimate are:

The price effect of square footage of houses. Floor area enters the model as a quadratic term: both as a linear term (*sqfttotliv* or square feet of total living areas) and the square of the linear term (*sqftsq*) (Goodman and Thibodeau 1995). The squared term is significant only in the west sample. The linear term is expected to be positive and the squared term negative. Here in the west sample it is increasing at an increasing rate. A cubic function was tested in a supplemental regression and found to be not significant in the west sample, but significant in the east sample. Thus, in the east sample, price per square foot increases at an increasing rate for a while and then increases at a decreasing rate. In the west sample we are lead to the conclusion that there is a premium for larger floor area homes. Intuitively it is believable in this area. Houses were built in an earlier era and were typically smaller than more recently constructed residences. Demand for larger houses meeting modern tastes could create the premium indicated in the analysis.

*Bedrooms * Floor area*. This interactive term (*sqftxbedroom*) multiplies the number of bedrooms by the square footage of the total living area. While not significan in the west, the negative coefficient, found in both the east and the west, indicates that as the number of bedrooms increases against a constant floor area, prices decreases. This is consistent with past studies (Craig et al., 1991). The standard interpretation is that increasing bedrooms while holding floor area the same reduces size of other rooms, which negatively affects price.

Lot size and lot size squared. As with living area, lot size (*sqftlot*) often has a nonlinear relation to price (Li & Brown, 1980). In the west sample, somewhat consistent with the size of the living area, the price function of lot size and its square (*sqftlotsq*) is increasing at an increasing rate. The coefficient of the squared term is very small, indicating a small premium for large lots as well as large floor areas. In the east sample, the squared term is negative, indicating that the implicit per square foot price of residential property decreases with increasing lot size. Note the differences in the implicit per square foot prices of floor area and lots on the west and east samples. The implicit price of the built space in the west sample is significantly less than that in the east ($38.89 per square foot vs. $44.88 per square foot [t = 36.2341, P>|t| = 0.0000]). The older houses in the western section may suffer from a degree of functional obsolescence. The opposite is true of the land prices ($4.85 per square foot in the west sample vs. $1.76 per square foot in the east [t = -410.6686, P>|t| = 0.0000]). Much better access to the Seattle CBD from the western area is key to explaining this difference.

Condition. On the east side, the condition of a house is not a significant factor in its price but is significant on the west side. The average housing condition on the east is statistically significantly worse. On the Assessor's scale of 1up to 5, the average on the west is 3.47 while on the east it is 3.26 (t = -23.92, P>|t| = 0.0000).

View. *No view* is a dummy variable indicating a property does not enjoy a view of one or more of the natural features in the Seattle area or the downtown skyline. Data was gathered from field observations and confirmed in the Tax Assessor 's database. The large coefficient in both samples indicates there is a high premium for a view. This is believable given the spectacular views from only a few of houses in the study area.

Table 11: Regression of all Observations West of the Lake

Regression with robust stadard errors		
	Number of obs	19085
	F(64, 19020)	482.00
	Prob > F	0.000
	R-squared	0.7333
	Root MSE	59156

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-12245.97	1283.883	-9.54	0.000	-14762.5	-9729.448
trendsq	2639.946	181.4492	14.55	0.000	2284.29	2995.603
trendcu	-65.32847	7.603927	-8.59	0.000	-80.23284	-50.4241
sqfttotliv	38.89371	5.259855	7.39	0.000	28.58393	49.20349
sqftsq	0.0054131	0.0024255	2.23	0.026	0.0006588	0.0101673
bedrooms	3735.948	3928.74	0.95	0.342	-3964.731	11436.63
sqftxbedroom	-1.437656	2.289736	-0.63	0.530	-5.925743	3.05043
bathrooms	7893.387	1165.411	6.77	0.000	5609.077	10177.7
age	-837.4676	111.0313	-7.54	0.000	-1055.099	-619.8364
agesq	6.426758	0.8766124	7.33	0.000	4.70852	8.144996
condition	11230.25	816.4204	13.76	0.000	9629.991	12830.5
sqftlot	4.857408	0.6121582	7.93	0.000	3.657524	6.057293
sqftlotsq	0.0001133	0.0000277	4.09	0.000	0.000059	0.0001677
noviewd	-48442.1	2361.948	-20.51	0.000	-53071.73	-43812.47
wfntlocati	57483.17	13309.37	4.32	0.000	31395.63	83570.71
itbs_read	-435.4866	76.25331	-5.71	0.000	-584.9498	-286.0233
dis_bofa	-32.8828	5.181674	-6.35	0.000	-43.03934	-22.72626
dis_bofasq	-0.0004512	0.0000805	-5.61	0.000	-0.0006089	-0.0002935
az_bofa	8817.43	1794.793	4.91	0.000	5299.477	12335.38
dis_mic	-56.45038	11.12857	-5.07	0.000	-78.26337	-34.6374
dis_micsq	0.0008359	0.0001121	7.46	0.000	0.0006162	0.0010556
az_mic	41590.06	5120.359	8.12	0.000	31553.7	51626.41
dis_xway	-1.824695	0.6528485	-2.79	0.005	-3.104335	-0.5450537
dis_xwaysq	4.20E-06	0.0000759	0.06	0.956	-0.0001445	0.0001529
az_xway	-81.69544	12.3736	-6.6	0.000	-105.9488	-57.44209
apt1_dis	8.661034	3.191011	2.71	0.007	2.40637	14.9157
apt1_dissq	0.0062835	0.0017055	3.68	0.000	0.0029405	0.0096264
apt1_az	0.7564953	4.418846	0.17	0.864	-7.904834	9.417825
cult1_dis	14.25458	2.982271	4.78	0.000	8.409065	20.1001
cult1_dissq	-0.0065585	0.0015053	-4.36	0.000	-0.009509	-0.0036081
cult1_az	-0.8244553	4.257031	-0.19	0.846	-9.168613	7.519702
govt1_dis	-11.46246	1.528937	-7.5	0.000	-14.45931	-8.465605
govt1_dissq	0.0027363	0.0002801	9.77	0.000	0.0021872	0.0032853
govt1_az	-25.74987	5.040219	-5.11	0.000	-35.62915	-15.8706
hotel1_dis	3.968803	1.197527	3.31	0.001	1.621543	6.316062
hotel1_dissq	-0.0000614	0.0001311	-0.47	0.639	-0.0003183	0.0001955
hotle1_az	22.89071	6.191672	3.7	0.000	10.75448	35.02694
off1_dis	3.617194	3.522736	1.03	0.305	-3.287681	10.52207
off1_dissq	-0.0021355	0.0013117	-1.63	0.104	-0.0047066	0.0004356
off1_az	-1.858714	5.395973	-0.34	0.731	-12.4353	8.717872
hops1_dis	7.467621	1.565577	4.77	0.000	4.398951	10.53629
hops1_dissq	-0.0009923	0.0001774	-5.59	0.000	-0.0013399	-0.0006446
hosp1_az	-26.91787	4.820133	-5.58	0.000	-36.36576	-17.46998
ind1_dis	17.0183	5.136983	3.31	0.001	6.949361	27.08724
ind1_dissq	-0.0067888	0.002982	-2.28	0.023	-0.0126338	-0.0009439
ind1_az	-0.279431	5.244005	-0.05	0.958	-10.55815	9.999284
sch1_dis	-5.258741	3.054309	-1.72	0.085	-11.24546	0.7279757
sch1_dissq	0.0034456	0.0011725	2.94	0.003	0.0011474	0.0057437
sch1_az	8.973383	4.503226	1.99	0.046	0.1466613	17.80011
r_net1	-16.75717	3.891787	-4.31	0.000	-24.38542	-9.128924
r_net1sq	0.0051108	0.0007024	7.28	0.000	0.003734	0.0064875
areu1_dis	5.033572	5.94402	0.85	0.397	-6.617234	16.68438
areu1_dissq	-0.0036625	0.0019126	-1.91	0.056	-0.0074113	0.0000863
areu1_az	14.45764	5.01515	2.88	0.004	4.627498	24.28778
average	-7.324813	3.085767	-2.37	0.018	-13.37319	-1.276436
avesq	0.0029922	0.0014043	2.13	0.033	0.0002395	0.0057448
trafficnoi	-14390.08	744.2301	-19.34	0.000	-15848.84	-12931.32
visibility	-7512.378	1444.895	-5.2	0.000	-10344.5	-4680.255
tnodes_seg	22866.17	4253.967	5.38	0.000	14528.02	31204.33
unodes_seg	-177671.6	25830.49	-6.88	0.000	-228301.7	-127041.6
density	-4332.014	705.428	-6.14	0.000	-5714.716	-2949.313
densitysq	71.15417	10.39624	6.84	0.000	50.77663	91.53171
nonresmix	444.5034	301.7404	1.47	0.141	-146.9346	1035.941
nonresmixsq	-4.120385	6.747437	-0.61	0.541	-17.34596	9.105189
cons	-3902289	821107.1	-4.75	0.000	-5511732	-2292847

Table 12: Regression of all Observations East of the Lake

Regression with robust stadard errors						
				Number of obs		6740
				F(64, 19020)		200.02
				Prob > F		0.000
				R-squared		0.7163
				Root MSE		78570

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-8309.059	2762.236	-3.01	0.003	-13723.92	-2894.194
trendsq	2006.847	415.1898	4.83	0.000	1192.943	2820.752
trendcu	-33.08205	18.17253	-1.82	0.069	-68.70602	2.541913
sqfttotliv	44.87953	25.36079	1.77	0.077	-4.835711	94.59477
sqftsq	0.0123357	0.0079664	1.55	0.122	-0.003281	0.0279524
bedrooms	15171.16	9340.051	1.62	0.104	-3138.326	33480.64
sqftxbedroom	-9.44647	4.678118	-2.02	0.043	-18.61708	-0.2758642
bathrooms	1556.925	3060.161	0.51	0.611	-4441.968	7555.818
age	-3545.694	286.858	-12.36	0.000	-4108.027	-2983.36
agesq	31.49007	3.006639	10.47	0.000	25.5961	37.38405
condition	1873.04	2490.897	0.75	0.452	-3009.913	6755.993
sqftlot	1.763203	0.2923044	6.03	0.000	1.190193	2.336213
sqftlotsq	-6.57E-06	1.76E-06	-3.73	0.000	-0.00001	-3.11E-06
noviewd	-33514.2	6696.874	-5	0.000	-46642.21	-20386.18
wfntlocati	66472.53	11715.65	5.67	0.000	43506.11	89438.95
itbs_read	855.5314	224.0669	3.82	0.000	416.2887	1294.774
dis_bofa	-63.02387	42.13581	-1.5	0.135	-145.6235	19.57577
dis_bofasq	0.000394	0.0004469	0.88	0.378	-0.0004821	0.0012701
az_bofa	1408.987	15393.96	0.09	0.927	-28768.09	31586.06
dis_mic	-18.89355	18.99103	-0.99	0.320	-56.12204	18.33495
dis_micsq	0.0006099	0.0001977	3.09	0.002	0.0002225	0.0009974
az_mic	3680.226	957.6021	3.84	0.000	1803.02	5557.432
dis_xway	3.07E+00	4.098286	0.75	0.453	-4.960601	11.1073
dis_xwaysq	0.0005329	0.0003969	1.34	0.179	-0.000245	0.0013109
az_xway	85.00916	30.22026	2.81	0.005	25.7678	144.2505
apt1_dis	-20.26596	4.156983	-4.88	0.000	-28.41497	-12.11694
apt1_dissq	0.0028282	0.000985	2.87	0.004	0.0008973	0.0047591
apt1_az	11.26145	12.53411	0.9	0.369	-13.30942	35.83231
cult1_dis	0.979098	5.761914	0.17	0.865	-10.31609	12.27429
cult1_dissq	0.0024861	0.0019166	1.3	0.195	-0.0012711	0.0062433
cult1_az	32.96383	12.62059	2.61	0.009	8.223438	57.70422
govt1_dis	0.9431599	3.656703	0.26	0.796	-6.225146	8.111466
govt1_dissq	0.0007291	0.0003636	2.01	0.045	0.0000163	0.0014418
govt1_az	19.11531	18.28457	1.05	0.296	-16.72829	54.95891
hotel1_dis	-3.415989	5.557444	-0.61	0.539	-14.31035	7.478376
hotel1_dissq	-0.0000434	0.00039	-0.11	0.911	-0.0008078	0.0007211
hotle1_az	42.7651	19.68502	2.17	0.030	4.17618	81.35403
off1_dis	4.992128	3.73066	1.34	0.181	-2.321157	12.30541
off1_dissq	-0.0015939	0.0006831	-2.33	0.020	-0.002933	-0.0002548
off1_az	-13.03153	13.44772	-0.97	0.333	-39.39336	13.33029
hops1_dis	21.34044	4.397089	4.85	0.000	12.72074	29.96014
hops1_dissq	-0.0011813	0.0002971	-3.98	0.000	-0.0017638	-0.0005989
hosp1_az	-19.78787	18.83728	-1.05	0.294	-56.71495	17.13922
ind1_dis	18.70703	7.302412	2.56	0.010	4.391967	33.02209
ind1_dissq	-0.0137365	0.0029084	-4.72	0.000	-0.0194379	-0.0080352
ind1_az	40.71885	9.982056	4.08	0.000	21.15083	60.28687
sch1_dis	-25.72446	7.292025	-3.53	0.000	-40.01915	-11.42976
sch1_dissq	0.008703	0.0021873	3.98	0.000	0.0044152	0.0129909
sch1_az	-1.912528	13.63818	-0.14	0.888	-28.64772	24.82267
r_net1	15.5182	4.551614	3.41	0.001	6.595579	24.44081
r_net1sq	-0.0007461	0.0004588	-1.63	0.104	-0.0016456	0.0001534
areu1_dis	6.301124	7.389891	0.85	0.394	-8.185423	20.78767
areu1_dissq	-0.003333	0.0012496	-2.67	0.008	-0.0057825	-0.0008835
areu1_az	-21.40526	13.44011	-1.59	0.111	-47.75217	4.941643
average	-15.12134	4.550055	-3.32	0.001	-24.0409	-6.201775
avesq	0.0019277	0.0007164	2.69	0.007	0.0005232	0.0033321
trafficnoi	-6847.463	2164.616	-3.16	0.002	-11090.8	-2604.124
visibility	10432.67	11077.85	0.94	0.346	-11283.45	32148.79
tnodes_seg	124336.8	86652.33	1.43	0.151	-45529.48	294203
unodes_seg	194650.9	88005.61	2.21	0.027	22131.79	367170
density	-14539.24	5109.142	-2.85	0.004	-24554.79	-4523.692
densitysq	543.6976	164.2274	3.31	0.001	221.7595	865.6358
nonresmix	-2778.063	2512.482	-1.11	0.269	-7703.33	2147.203
nonresmixsq	51.80016	69.00132	0.75	0.453	-83.46446	187.0648
_cons	1277279	4874951	0.26	0.793	-8279182	1.08E+07

School Quality. ITBS reading score is the score of the reading component Iowa Test of Basic Skills for the elementary school serving each house. It is used as an indicator of school quality. Common wisdom would hold that this variable would be positive and significant. On the west side it is negative and significant. The same counter-intuitive finding appears in an earlier study in King County (Franklin & Waddell, 2003). This result is likely heavily influenced by School Board policy. The assignment policy of the Seattle Board of Education does not guarantee that children will attend the nearest school. Consequently, we would expect that school quality would not have the effect on housing price it would have otherwise, but the significant negative sign is puzzling.

In the east sample, the ITBS reading score is a positive and significant factor in hosing price, as expected. The two cities on the east side are part of the Lake Washington School district, not Seattle School District. Here, students do attend schools closest to their homes.

Distance to Non-residential Uses. There are a number of variables that measure the distance and distance squared from residences to various non-retail commercial uses (and apartments and elementary schools). These variables are included to account for price effects of proximity to these different types of uses and, more importantly, to avoid these effects being confounded with effects of retail proximity. Also, a measure of the direction or azimuth (e.g. *apt1_az*) is included to control for location differences. For convenience, Table 13 repeats the mean distances from residences to non-residential uses from Table 9 as well as the housing price coefficients from Tables 11 and 12. Note how

Table 13: Comparison of Price Effect of Distance to Non-Retail Uses

	West			East		
Use	**Mean Distance**	**Price Coefficient**	**Comment**	**Mean Distance**	**Price Coefficient**	**Comment**
Apartment	587.27	8.66	non-linear	1410.21	-20.26	non-linear
Culture/Entertainm	851.45	14.25	non-linear	1454.4	0.98	not significan
Government	2741.76	-11.46	non-linear	4129.55	0.94	not significan
Hotel	4645.05	3.97	linear	6748.41	-3.42	not significan
Office	894.85	3.62	non-linear	2111.34	4.99	non-linear
Hospital	4555.04	7.47	non-linear	7460.14	21.34	non-linear
Industry	581.96	17.02	non-linear	798.26	18.71	non-linear
School	1122.44	-5.26	non-linear	1452.78	-25.74	non-linear

sharply different the distance coefficients are for the two sample areas in terms of magnitude, sign, and significance. It is likely that they arise for complex reasons including different types of land use controls, automobile friendly infrastructure, as well as the predominant economic forces in the different eras in which the areas developed.

Mean straight-line distance from the retail site nearest a given residence to the four retail sites nearest that site (*average*) is a measure of retail clustering and convenience of traveling to more than one store. In both the west and the east it and its squared term are significant. Close clustering of nearby stores positively influences house prices. As distances between retail sites increase, residential prices decline, but at a decreasing rate. Note that the size of the coefficient for the average variable in the east side sample is twice as large as that for the west sample. The result is differing curves with a price effect going to zero at 1,254 feet on the west and 3,787 on the east. Again, the west side mean distance is a walkable distance and that on the east is not. As discussed above, because of different emphases on automobiles, convenience may have different meanings in the two areas.

Traffic noise (*trafficnoi*) is drawn from the Tax Assessor's database. It is measured by an index, assigned by assessors, that ranges from 0 to 3. As reported in

Tables 11 and 12, this variable is significant, negative, and has a relatively large price effect in both the east and the west samples. As discussed earlier, this variable is associated with general traffic and not necessarily traffic at or generated by retail areas.

Visibility is a dummy variable indicating visibility of retail and other commercial uses from individual residences. Data was gathered from a field survey. As with traffic noise the variable is significant, negative, and has a relatively large effect on price, but only in the west sample; the variable does not have a significant effect in the east sample. The absence of an effect in the east is probably due to the very large distances between residences and retail sites found there (See Table 9). The average straight-line distance to the nearest retail use on the east side is 2,366.7 feet while on the west it is only 890 feet. As distance increases, for example, a view of a retail site would not dominate the views from a residence at it would if it is close. Other unobserved negative externalities that might be associated with visibility over short distances, such as noise, would not travel over such long distances.

Residential density (*density*) is derived by dividing the population of a neighborhood by the acreage devoted to residential (single- and multi-family) use. From Table 9, mean residential density is 24.9 people per acre in the west and 14.1 in the east. From Tables 11 and 12, the density and density squared terms are all significant in the hedonic price regression with signs indicating that price decreases nonlinearly with increasing density. For the coefficients in Table 11, the west, the turning point is at 29.7471 persons per acre and for the east the turning point is much lower at about 14.955 persons per acre. The shape of the density-price relationship is the same in both sample areas, but the magnitudes differ.

Non-residential mix (*nonresmix*) is a measure of the proportion of total land in a neighborhood (census tract) used non-residentially. The values of this variable are calculated from the GIS map of the land use information in the assessor's database. Cao and Cory (1981) find that an increasing proportion of non-residential uses in a neighborhood have a positive effect on residential prices, but only up to a point. Beyond that point, the effect is negative. In Table 11 (west sample), this variable is positive and , its square is negative, neither is significant. Together, they are jointly significant (F = 6.52, Prob F> = 0.0015). These results from a reduced model provide weak evidence of consistency with Cao and Cory's (1981) findings. More complete models discussed later provide greater consistency.

The coefficient for *non-residential mix* in the east sample is not significant, the square is also not significant; but they are jointly significant at the .05 level. The sign of the primary coefficient is negative and the sign on the square is positive, indicating that as the proportion of non-residential use increases, residential prices decrease to a point and then increase. This pattern is not consistent with Cao and Cory (1981). The distances between residential and non-residential uses are much greater on the east than the west and the proportion of non-residential uses is significantly less; that the effect of the proportion of non-residential uses on housing price varies between the two is understandable.

For both the west and east samples, supplemental regressions are tested with the non-residential proportion taken to increasingly greater powers. When taken up to the fifth power, all transformations of the non-residential proportion became significant at levels well beyond 0.000. In both samples, the pattern of signs is the same (positive,

negative, positive, negative, positive). Comparing the east sample and the west, the effect is not is as dissimilar as a simple quadratic function would indicate. The relationship between residential price and proportions of non-residential uses is very complex. Addressing this question is beyond the scope of this dissertation, but it is interesting and deserves further exploration.

Travel and Straight-Line Distance. Looking at Table 11, the west sample, street travel distance in feet from a residence to the nearest retail use is a proxy for convenience: the value of proximity or access is capitalized positively into residential value. In the west sample, the partial effect of both street travel distance and its square is significant. The negative coefficient of street travel distance indicates that as travel distance increases between houses and retail uses, ceteris paribus, the price falls. This result is consistent with expectations: convenience to retail is positively capitalized into housing prices. The positive coefficient on the squared term indicates a non-linear relationship with the price effect decreasing at a decreasing rate as distance increases.

The relevant coefficients from Table 11are:

$P = -16.75717TD + 0.0051108TD^2$

Where:

P = price

TD = Travel distance

The turning point for the curvilinear price function is at a distance 1,639.4 feet. In other words, the price effect of increasing the travel distance from a retail site to a house decreases at a decreasing rate to a point 1,639.4 feet from the retail site. Mathematically, at that point, the price effect would begin to increase. But, assuming that convenience

has no effect past some limiting distance, we will say the price does not change beyond this point, all else being equal. This distance is very similar to walking distance. The implication is that this convenience effect is not robust beyond walking distance.

Substituting the turning point back into

$P = -16.75717TD + 0.0051108TD^2$

We get

$P = -13{,}735.8$

At a travel distance of 1,639.4 feet from a retail site, the effect of convenience on a house's sale price is at its greatest at maximizes at -\$13,735.8. We must remember we are looking at the positive effects on price arising from convenience to a retail site. A reasonable interpretation of these effects is to say that a house at zero feet from a retail site gains \$13,735.8, all else held equal. As distance increases, the positive effect decays and price decreases nonlinearly until we reach a point, 1,639.4 feet from the retail site, where the curve reaches its turning point at zero and there is no longer a price effect from convenience and accessibility.

In the east sample (see Table 12) the coefficient of travel distance is significant but its squared term is not. The sign is positive indicting that as travel distance increases price increases as a constant function of distance. The coefficients on the travel distance and its square are consistent with expectations in the west sample, but not in the east sample.

Given the differences in the distances involved (much greater in the eastern than the western sample, see Table 9), the travel distance variable seems to be measuring two different things in the two samples. The mean distance of 1,231 feet between a house and

a retail site on the west side is a walkable distance while the mean distance of 3,610 on the east side is far beyond easily walkable distance. If walking trips to retail are generally not feasible, convenience takes on a different meaning. Convenience means saving time and frustration. When walking, distance directly translates to time and related physical exertion, the greater the distance the greater the time and exertion. In a car, physical exertion is not an issue, but congestion and frustration may be. A driver may find a longer distance more convenient if there is less congestion.

The straight-line distance in feet from a house to a retail site is the other principal variable of interest. For the west side (see Table 11) the straight-line distance variable is positive but not individually significant, its squared term is marginally significant at the 5% level and negative; their joint effect is significant at the 0.05 level (F=3.62, Prob F=0.0386). As residential distances increase in a straight-line from commercial sites residential price increases at a deceasing rate. This is consistent with the argument that negative externalities, which negatively effect residential prices, travel in a direct line from retail sites. On the east side (Table 12) the straight-line variable is not significant, but its square is significant and negative; they are jointly significant (F=18.35, Prob>F=0.0000). The value of the straight-line coefficients (5.03 and 6.3) and the squared terms (-0.0037 and –0.0033) are remarkably similar on both the west and east sides. They are both nonlinear with the price effect decreasing at an increasing rate.

The straight-line effect reaches its maximum in the west sample at 687.177 feet and 945.263 feet in the east sample. In the west sample the value is $1,729.48, in the east it is $2.978.11. Looking at the effect in the west sample, from Table 11, we have:

$$P = 5.033572SD - 0.0036625SD^2$$

Where:

P = price

SD = Straight-line distance

which maximizes at 687.177. The turning point for the nonlinear price function is at a distance of 687.177 feet. In other words, the price effect of increasing the straight-line distance from a retail site to a house increases at a decreasing rate to a point 687.177 feet from the retail site. Mathematically, at that point, the price effect would begin to decrease. Because we have assumed that negative externalities have no effect past a distance, we will say the price does not change beyond this point, all else being equal.

Substituting back into

$P = 5.033572SD - 0.0036625SD^2$

$P = 1729.48.$

At a straight-line distance of 687.2 feet from a retail site, the effect of negative externalities on a house's sale price maximizes at $1,729.48. The interpretation of this function is identical to that used above for travel distance. First, we must remember we are looking at the negative effects on price of exposure to noise, light, congestion, etc., of being close to retail site. The most reasonable interpretation of these effects is to say that houses at zero feet from a retail site suffer a loss in price of $1,729.48, all else held equal. As the *straight-line distance* increases, the negative effect decays and price increases in a curvilinear fashion until we reach a point where the curve reaches its turning point, after which there is no longer a price effect arising from exposure to negative externalities; this occurs at 687.2 feet in this case.

The same analysis applies to the east sample, except that Table 12 produces coefficients of 6.301 for *straight-line distance* and –0.003 for *straight-line distance squared*. These two coefficients produce a price effect of 2,978.11 at a distance of 945.26 feet. We interpret this as a house in the east sample zero feet from a retail site suffers a price loss of $2,978.11, ceteris paribus. This price effect decays nonlinearly with distance to a point 945.26 feet from the retail site where the effect is zero.

The main variables of interest are *travel distance* from residence to retail and its square and *straight-line distance* from residence to retail and its square. In the portion of the study area west of Lake Washington the interplay between the variables of interest is as expected. As travel distance on streets increases, the price of housing decreases, ceteris paribus. This reflects the notion that convenient access to retail increases as distance decreases and savings in time and effort are capitalized into the value of the house. We also expected price to increase as straight-line distance increases, reflecting a diminishing effect of negative retail site externalities over distance. In the west sample section, this price effect is in the direction expected and is statistically significant. In this part of the study area the average street distance between a residence and closest retail is 1,231 feet and the average straight-line distance is 890 feet. All of this is in an environment with relatively high residential density (a friendly situation for retail), with a greater portion of land devoted to non-residential uses (a convenience to residents), and a more integrated and easy to traverse general gridiron layout.

The east sample is designed, much more so than the western portion, to accommodate automobiles and to protect residential property from negative externalities arising from non-residential uses. As is clear from Figures 3A and 3B and Table 9, land

uses are much more highly segregated, residential densities are much lower, and the average distances between residences and retail sites are much greater measured either in travel distance (3,610 feet) or straight-line distance (2,366 feet). Here the travel distance variable (convenience) does not behave as it does in the west sample, but the straight-line distance variable (negative externalities) does behave as it does in the western sample. The street travel distance is significant, but its sign is positive, meaning that as distance increases, price increases. The distance here is so great that autos are probably required for short shopping trips. Convenience, as discussed earlier, may not be strictly related to travel distance. On the other hand, the straight-line measure is positive, its square is negative and significant – meaning that as distance increases price increases at a decreasing rate as a nonlinear function. The reach and dollar magnitude of the straight-line effect in the east sample is remarkably similar to that in the west sample. While convenience may be different, e.g. related to ease of access rather than travel distance, negative externalities, apparently are related to straight-line distance. In this environment, residential density is low – not a friendly situation for retail, but a situation mitigated by automobiles. Non-residential uses are not mixed in a fine-grained manner with residential uses. This is an inconvenience to residents, but one that also can be mitigated with automobiles. The street layout is not well integrated; all else held equal, travel distance compared to straight-line distance is greater in a less integrated layout.

Samples Split by Distance

For purposes of further hedonic analysis, the study area is divided not only into the two areas to the west and east of Lake Washington – the Seattle part and the Kirkland/Redmond part – but also is subdivided further into those observations within

1,400 feet (straight-line) of the nearest retail and those not within that distance. The proponents of New Urbanism argue, and we have seen, that walking distance is a significant consideration for this analysis. We create four areas for analysis. The areas within and beyond 1,400 feet of retail on either side of the lake are segregated from other housing only to the extent that retail sites are segregated

The models differ between the areas within 1,400 feet of retail and those beyond 1,400 feet. The models applied to areas within 1,400 feet of the nearest retail include the straight-line distance variables, analysis of the areas not within 1,400 feet do not. Li and Brown (1980) postulate, and the above analysis covering the total west and total east areas for this study finds, that the adverse price effects of any negative influences arising from retail activity decay relatively quickly over distance. The results discussed above suggest that they do not extend past 1,400 feet in our sample; in fact, the reach is far short of 1,400 feet. Consequently, those independent variables intended to capture the effects of negative externalities – straight-line distance itself, straight-line distance squared, and straight-line distance interacted with density - are not included in the new model analyzing the two areas beyond 1,400 feet.

Analysis of these divisions using two similar but different models yields additional insight beyond that gained by looking only at the pooled samples. We estimate these models for each of the four sub-samples.

The following discussion relies on two reduced models and a full model for each of the four sub-samples. The first regressions use a reduced model. This model is included for the sake of comparison, but will not be discussed at length. There are no neighborhood design variables – it omits traffic noise, visibility, density, and non-

residential mix. The interaction terms are also not included. This model includes independent variables describing the structure and its age; year of the recorded sale, its square and cube; condition; lot size; location; and relation to schools and various other classes of non-residential uses. Travel distance to the nearest retail (and square of travel distance), the retail clustering variable and its square are included. Straight-line distance to the nearest retail (and square of this distance), is included in analysis of the two areas within 1,400 feet of the nearest retail, but excluded in the analysis of the areas beyond 1,400 feet. Appendix A reports these reduced regression results. All have high R^2s, but as discussed in the pooled analysis, not all independent variables behave as expected.

The second regressions for each of the four analysis areas are included in Appendix B. Additional independent variables used to control for neighborhood design and environment are included in the second regression. These include *traffic noise (traffnoi)*, *visibility* of non-residential uses, the *ratio of street segments to intersections (seg_tnodes),* and the *ratio of street segments to cul-de-sacs (seg_unodea), density* (and its square), and the *proportion of non-residential (nonresmix)* uses (and its square) are included. *Straight-line distance* (and square of this distance) to the nearest retail, is included in analysis of the two areas within 1,400 feet of the nearest retail, but excluded in the analysis of the areas beyond 1,400 feet. The space syntax variable is highly correlated with the other two measures of the integration of street layout, although the other two are not highly correlated with each other. are highly correlated with one another, especially in the east sample. Consequently, space syntax variables are omitted form the current analysis. The interaction terms using distances and neighborhood variables are not included in this model. This set of variables captures important aspects

of the physical differences between places in the study area and confirm findings of Song and Gnaap (2003a; 2003b) that layout affects price.

The third set of regressions use full models. These include interactive terms between some of the neighborhood variables and the two principal variables of interest, travel and straight-line distance. This tests hypotheses that the effect of the travel and straight-line distances vary with neighborhood factors. Specifically, the interaction variables are 1) the *ratio of street segments to intersections interacted with travel distance* (*segtnodXrnet*), 2) the *ratio of cul-de-sacs to street segments interacted with travel distance* (*segunodaXrnet*), and, in the areas within 1,400 feet of the nearest retail, 3) *density interacted with straight-line distance* (*densityXareu1*) and 4) *non-residential mix interacted with straight-line distance (nonresXareu1)*. Straight-line distance (and the square of this distance) to the nearest retail, is included in analysis of the two areas within 1,400 feet of the nearest retail, but excluded in the analysis of the areas beyond 1,400 feet. The space syntax interaction, as is the case with the street layout measures, is highly correlated with the other two street layout measures, especially in the east sample. Therefore, the space syntax interaction is omitted form these models. This third set of regressions, the full model, is included in Appendix C.

The next section discusses the empirical results on an area-by- area basis. First is the Seattle side with the observations within 1,400 feet of the nearest retail followed by those beyond 1,400 feet. Then there is a discussion of analyses on the east side of the lake, taken in the same order. The discussion concentrates on the results of the model that includes the neighborhood development variables, Appendix B, and the full model –

the Appendix C regressions – that also includes the neighborhood/distance interaction variables. As might be expected, the results vary considerably among the areas.

There is no need to recite the meaning of most of the right-hand-variables again. Here, we concentrate on the variables of interest. The principal variables of interest are *travel distance* (*r_net1*) on the street and *straight-line distance* (*areu1_dis*) from retail to residences. To reiterate, the *travel distance* is meant to capture convenience, a positive price effect for housing and the *straight-line distance* is meant to capture negative spillovers from retail activity. Secondary variables of interest include the *retail clustering (average)* variable, *density*, and the effect of the *proportion of non-residential uses.* Additionally and very importantly, we include variables that express the effect of neighborhood street layout design: *seg_tnodes*, the ratio of segments to intersections and *seg_unodea* the ratio of segments to cul-de-sacs.

Space syntax is an interesting idea and has proven to be pragmatically useful, but to date has no recognized theoretical underpinnings. Its use is not included here, for reasons discussed earlier, and because we do not fully understand what we are measuring. Further research into and using this index could be interesting.

<u>West Sample Within 1,400 Feet</u>. This section of the analysis uses the Seattle side of the study area and includes only those observations within 1,400 feet of the nearest retail. As discussed extensively earlier, this part of the study area was developed primarily at a time when automobiles were not as dominant a travel mode as they have become, the street pattern is largely a grid system, houses and lots are smaller than they are in the other part of the study area, and retail uses are much more scattered among residential uses than they are concentrated together. Over 85 percent of the observations

in the west side of the study area are within 1,400 feet – walking distance – of the nearest retail.

Analysis in this sample – west of the lake and within 1,400 feet of a retail site - not only shows that, within this setting, the positive influence of reduced travel distance acts to off-set the depressing influence of negative externalities, (traveling over a straight-line) on the price of residences close to retail sites, but also that the magnitude of these effects is significantly affected by neighborhood layout and density.

Reduced model excluding interaction variables. Appendix B presents regression results from a more complete, but still reduced model. Variables measuring traffic noise, visibility, neighborhood layout, density, and the proportion of non-residential uses are added to the model.

In Appendix B, the table titled "West less than 1400 feet" presents the regression without neighborhood interaction variables. First we will discuss the secondary variables of interest followed by a detailed look at the primary variables of interest.

The retail clustering measure (*average*) and its squared term are jointly significant in this near west sample, although neither is individually significant. The signs, but not the P-value, indicate that increasing separation of retail uses negatively affects residential values in this setting. The negative effect decreases at a decreasing rate as distance between the retail uses increases.

Exposure to *traffic noise* negatively affects residential values. As discussed earlier, this noise factor may or may not arise because of the influence of retail uses, but its presence does control for traffic noise not arising from retail use.

In this setting, the west side of the lake for distances within 1,400 feet of a retail site, *visibility* negatively affects residential property values. Here where there is relatively dense development in a layout dominated by grid type street systems and using only observations within 1,400 feet of the closest retail use, being able to see commercial development has a significant negative effect on residential price separate from any other negative effects that may arise from proximity to retail. On the other hand, this also means that a potential negative price effect associated with proximity to retail in this setting can be mitigated if retail uses can somehow be visually shielded from residences.

There are two variables that measure connectivity of the neighborhood street systems in this reduced model; both have independent significant effects on residential price in this setting. Interestingly, by one measure, the ratio of street segments to intersections, price goes down as integration increases. By the other measure, the ratio of street segments to cul-de-sacs, price increases as integration increases. As the ratio of segments to intersections increases, street layout moves more toward a pure gridiron design and greater connectivity. Greater connectivity implies, but does not necessarily mean, greater direct access and greater choice of routes. Some access advantages of greater connectivity can be reduced if traffic calming or traffic channeling devices are introduced. Some of the Seattle neighborhood layouts do incorporate traffic calming devices. A higher ratio of all street segments to cul-de- sacs, a positive indicator of connectivity or fewer indirect routes in a neighborhood, has a statistically significant positive effect on housing price in the near west sample. (This does not imply, one way or the other, that houses on cul-de-sacs are not valued more or less highly. The measure pertains to entire neighborhoods, not individual streets.) We must conclude that while a pure gridiron layout is not valued here, neither are cul-de-sacs. But , as both are

statistically significant, we must also conclude that street layout is important in some way. This is not as odd as it may seem as there are many optional layouts, such as curvilinear layouts, off-set streets, etc. to gridirons and cul-de-sacs. There may be an optimal level of accessibility that is highly valued. This is an empirical question that can be explored.

Increasing neighborhood *density* in this reduced near west sample has a negative effect on housing price, all else being equal. This measure is jointly significant with the positively signed term *density squared*. The relation of density to prices is nonlinear. As density increases its negative effect on housing prices decreases at an increasing rate.

Recall that Cao and Cory (1981) find that an increasing proportion of non-residential use in a neighborhood has a positive effect on housing price, up to a point. Earlier, using pooled samples, we found only a weak effect of non-residential mix of uses on residential prices. But concentrating on residences in the west sample within 1,400 feet of retail sites, we do find effects similar to those the Cao and Cory (1981) find. Using a quadratic function, we find that values increases with increasing proportion of non-residential properties to a point, then decreases. An increasing proportion of non-residential uses (*nonresmix*) is positive and significant on house price. As the proportion of nonresidential uses increases, residential prices increase. The squared term has a negative sign, meaning that the price increase deceases. The squared term is individually not significant, but is jointly significant with the main term. The jointly significant variables show that as the non-residential proportion increases in the near west sample using this reduced model, housing price increases, but at a decreasing rate. The curve

derived from coefficients in the regression in Appendix B, for the west side less than 1,400 feet form retail sites

$$P = 1046.145\ (\textit{Non-Residential \%}) + -11.55593\ (\textit{Non-Residential \%})^2$$

attains a maximum at a relatively high non-residential proportion of 45.26%. The highest non-residential percentage in this setting is 43.62%, so the effect on housing price in this sample is always positive.

The principal variables of interest, the *travel distance* variables (*r_net1*) and the *straight-line distance* variables (*areu1_dis*) and their squares, are all significant and jointly significant and are of the expected sign in the Appendix B west sample, for houses within 1,400 straight-line feet of a retail site. Increasing travel distance negatively affects housing price. The street travel distance and its square are jointly significant (joint test, F = 3.92, Prob>F = 0.0199). Taking coefficients from the table, the equation for this function is:

$$P = TD + TD^2$$

$$P = -14.19011TD + 0.005778TD^2$$

Where:

P = sales price

TD = travel distance

The minima of this function occurs at $TD = -14.19/(2*(0.005)) = 1{,}227.94$. The net price effect at the minima is $P = (-14.19*1{,}227.94) + (0.005*1{,}227.94^2) = -\$8{,}712.32$. As we discussed at length earlier, we interpret this as a positive residential price effect of \$8,712.32 at zero feet from the closest retail, finally eroding to no price effect at a travel distance of 1,227.94 feet from the closest retail use.

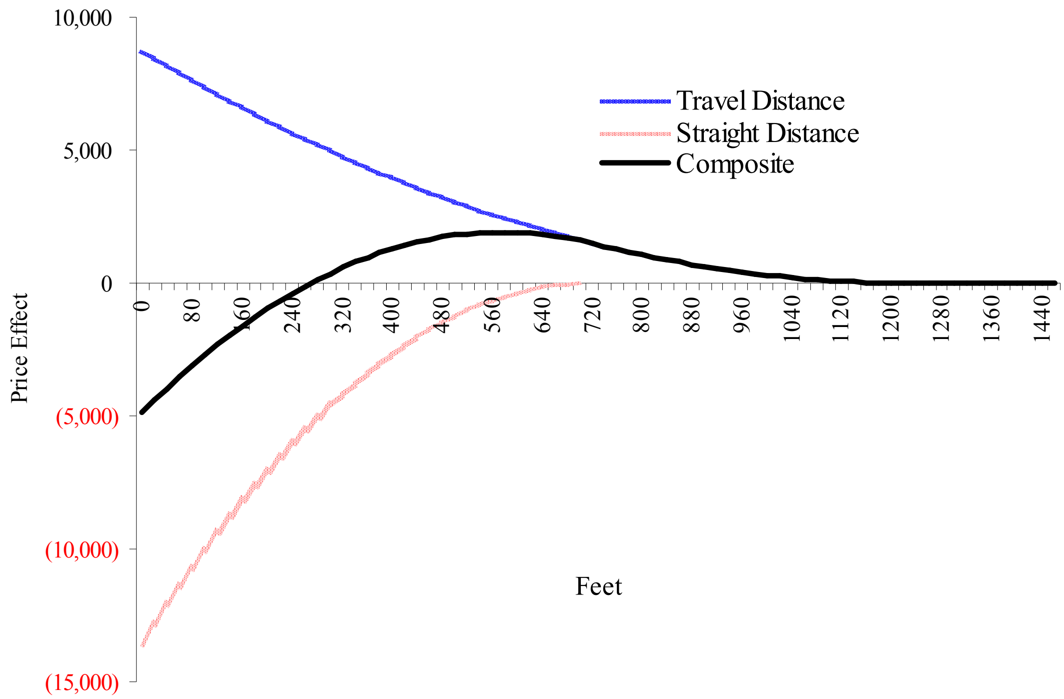

Figure 4:
Effect of Street and Straight Distances on the West Side Less Than 1,400 Feet

Increasing straight-line distance positively affects housing price. Both the straight-line distance and its square are significant and they are jointly significant (F = 13.66, Prob>F = 0.0000). From Appendix B, West Less than 1400 feet, the equation for this function is:

$$P = SD + SD^2$$

$$P = 37.75311SD - 0.0262474SD^2$$

Where:

P = sales price

SD = straight-line distance

The maxima of this function occurs at SD = 719.17. The net price effect at the maxima is P = $13,575.60. The interpretation is that the presence of negative

externalities lowers price \$13,575.60 at zero feet from the closest retail, finally eroding to no price at a straight-line distance of 719.17 feet from the closest retail use.

If the term indicating that a retail establishment is visible (*visibility*) from a given house is removed, the coefficient on the straight-line distance term increases from 37.75311 to 42.02337 ($t = -41.5835$, $P>|t| = 0.000$) indicating this term is picking up the negative externalities of any visual blight, as expected.

In this sample the travel distance and straight-line distance variables play against one another creating competing and offsetting effects on residential prices. Figure 4 shows this interplay using the coefficients of travel distance and its square and straight-line distance and its square. Figure 4 shows 1) the price effect of travel distance from a retail site by feet traveled, 2) the price effect of straight-line distance from a retail site, and 3) the composite effect of the two. The most interesting result of this plot is that the composite effect is positive beyond about 265 feet and the composite positive effect on value peeks at about 575.57 feet. The effect plays out at about 1,195 feet as calculated above. Six hundred feet, as discussed earlier, is often considered an optimum distance for the length of a shopping mall and 1,400 feet is considered a maximum walking distance. This composite curve is remarkably consistent with these distances.

The relationship between travel distance and straight-line distance in this setting is very similar to that derived by Li and Brown (1980) who found negative effects of externalities decaying over a shorter distance than the positive effects of convenience. The straight-line distance measures negative externalities. We find the shorter the straight-line distance, the more price is depressed. This effect is off set by the positive effect of convenience. Both effects vary at different rates over distance and the negative

externalities do not reach as far as the positive effect of convenience. The composite effect is that the negative externalities have only a very short reach and dominate the positive effect of convenience for only a very short distance. After that, the positive effect of convenience dominates over its reach, which in this type of older non-auto oriented neighborhood, play out over distances commonly considered as walking distance. The negative effect completely disappears at about 265 feet, but the positive effect of convenience extends to almost 1,230 feet. The maximum price effect is at approximately 575 feet. The differences between this analysis and that of Li and Brown (1980) are that there is no *a priori* assumption here that the composite curve will take the form it did and this analysis uses a simple linear form with quadratic variables rather than a form with forced exponential values for the travel and straight-line distances; a form intended to produce the this type composite curve. Further, as explained later, this analysis takes place in a neighborhood of a specific design type. Neighborhood design itself may play a role in the effect of proximity to retail on residential price.

Full model including the interaction variables: west sample within 1,400 feet of retail. The tables in Appendix C report the output of the analysis of the full model including the interactions of the variables of interest with neighborhood variables. Four interactions are included in the model: 1) the ratio of street segments to intersections interacted with travel distance (*segtnodXrnet*), 2) the ratio of street segments to cul-de-sacs interacted with travel distance (*segunodXrnet*), 3) residential density interacted with straight-line distance (*densityXareu1*), and 4) the proportion of non-residential use interacted with straight-lone distance (*nonresXareu1*).

West Side within 1,400 Feet. In the sample west of the lake, using only observations within 1,400 feet of the nearest retail site, three of the four interactions included in the model are statistically significant. These are the interaction between street segments to intersections and travel distance (*segtnodXrnet*), the interaction between density and straight-line distance (*densityXareu1*), and the interaction of the proportion of non-residential uses with straight-line distance. The interaction of the ratio of all street segments to cul-de-sacs and travel distance (*segunodXrnet*) is not significant, although the ratio without the interaction continues to indicate that a higher ratio of segments to cul-de-sacs positively affects residential prices. In the near west setting, neighborhood layout integration (measured by the ratio of street segments to intersections) influences the price effect of travel distance, density influences the price effect of straight-line distance, and the proportion of non-residential uses influences the price effect of straight-line distance.

Introduction of the interaction terms significantly reduces the value of the coefficients on the travel distance variable (*r_net1*) from –14.19011 to –4.823383 (t = -130, P>|t| = 0.0000) and straight-line distance variable from 37.75311 to 12.75976 (*areu1_dis*) (t = 234.83, P>|t| = 0.0000). The squared terms are also reduced. The coefficient on travel distance squared changes from 0.005778 to 0.00566 (t = 0.11511, P>|t| = 0.9084), which is an insignificant change, and the coefficient on straight-line distance squared changes from –0.0262474 to –0.0218542 (t = -73.7474, P>|t| = 0.0000), which is significant.

From the interaction term, the relationship between travel distance and neighborhood integration as measured by the ratio of street segments to intersections in the near west setting (Appendix C) is:

$P = - 4.823383TD - 4.407904SI$

Where:

P = Sales Price

TD = Travel Distance

SI = Street Segments/Intersections

In the full model west sample for observations less than 1,400 feet from the nearest retail site, as the travel distance from a retail site to a residence increases, the sales price of the house, all else held equal, decreases. We hold this to be due to a loss of convenience and the capitalized value of convenience. Further, as the ratio of street segments to intersections in a neighborhood's street layout increases the decrease in price becomes greater. For the sample west of Lake Washington the ratio of street segments to intersections ranges from 1.27 to 3.34 so the price effect on residences within 1,400 feet of the nearest retail site varies from -10.42 per foot of travel distance to -19.54 per foot of travel distance as the ratio of street segments to intersections increases. Notice that this range brackets the coefficient, -14.19011, on travel distance (*r_net1*) in the reduced model shown in Appendix B (no interactions) for the west sample less than 1,400 feet from the nearest retail site. As the ratio of street segments to intersections is a measure of connectivity, we see that the price effect of convenience is greater as connectivity increases. Intuitively this makes sense. One feature of greater street integration is more routes and routes that are more direct between various points. It is easy to think that more

choices of routes enhance convenience. For example, driver frustration may be reduced if one route is blocked but alternates are available.

A second significant interaction in this setting is the interaction between density and straight-line distance. We theorize that negative externalities from retail sites are compounded by negative externalities arising from increased residential density. The interaction of density with straight-line distance is positive and significant beyond the 0.01 level of confidence. As density increases, price effects of negative externalities arising from retail increase in this setting. In this part of the study area, neighborhood densities range from 13.56 persons per acre to 51.36 persons per acre. The price effect of straight-line distance from retail, including the effect of increasing density, ranges from $27.58 per foot to $68.90 per foot. Notice that this range brackets the coefficient of straight-line distance, 37.75311, in the reduced model discussed earlier and presented in Appendix B.

We calculate nonlinear functions for travel distance using the 20th centile, the median, and the 80th centile values for segments/intersections. Similar curves are plotted for the price effect of straight-line distance incorporating the partial effects of density's 20th centile, median, and 80th centile values. Figure 5 shows the composite curves resulting from combing these effects.[*]

The important point here is not the specific shape of these curves; an infinite number of curves could be drawn depending on values selected for density and street

---

[*] The interaction effect was tested at various values of street segments/cul-de-sacs using methodology described in Woolridge (2000) on pages 190-191. The coefficient for travel distance derived in these tests is used to produce the curves in Figurer 4 and similar charts of interaction effects that follow

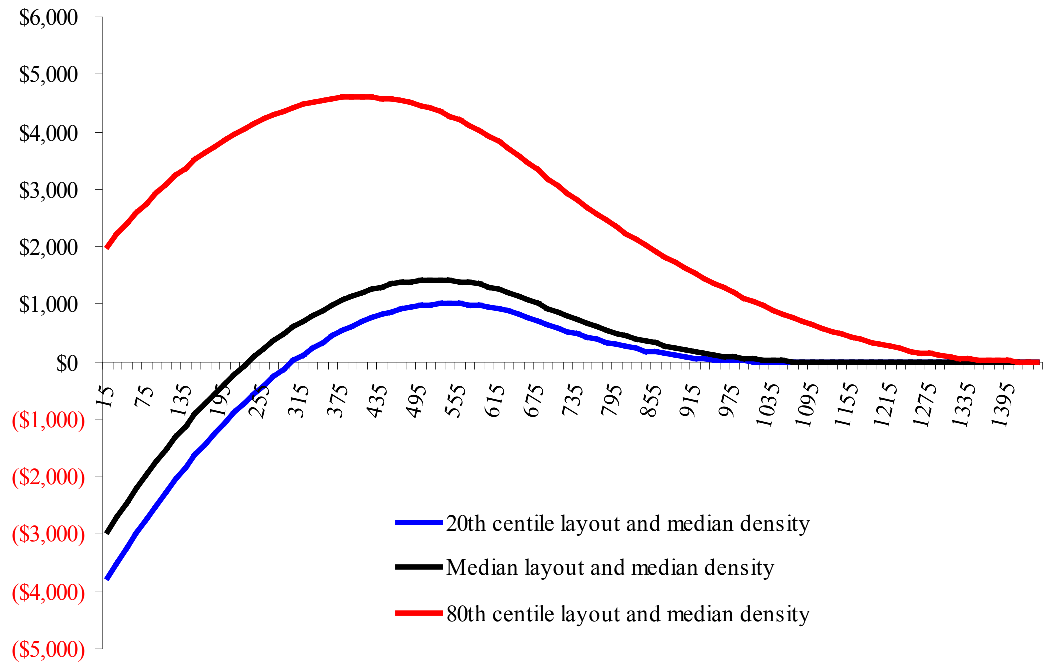

Figure 5:
Effect of Street Layout and Residential Density on the Price of Proximity to Retail in the West Sample < 1,400 feet

connectivity. The important point is that residential densities and different neighborhood street layouts do have a significant and discernable systematic influence on the price effect of residential proximity to retail in this setting.

The interaction between the proportion of non-residential uses and straight-line distance is significant at beyond the 0.01 level. Using

$$\frac{\delta P}{\delta SD} = \beta_1 + \beta_3 D$$

Where:

β_1 = the coefficient for straight-line distance

β_3 = the coefficient on the interaction term

D = Non-residential proportion (*nonresmix*)

In the west sample the non-residential proportion ranges from 0.38% to 43.62% of land use. But the nonresidential proportion of 43.62% is an outlier; at the 92nd percentile of the range of the nonresidential proportion of land use in the western sample, the proportion is 20.21%. From Appendix C, β_1 = 12.75976 and β_3 = -0.5930198. Looking at this interaction effect using non-residential land use mix proportions at the 20th centile, median, and 80th centile, we see:

P = 12.75976 + (-0.5930198*7.57) = 8.2706 at the 20th percentile

to

P = 12.75976 + (-0.5930198*19.23) = 1.356 at the 80th percentile

and

P = 12.75976 + (-0.5930198*13.3) = 4.8726 at the median

As the proportion of nonresidential uses in a neighborhood increases, the price effect of straight-line distance from a residence to the nearest retail site diminishes in magnitude, ceteris paribus. There are two points to be made about this result. First, the measure of the proportion of nonresidential uses includes not only retail, but also all other nonresidential uses such as schools, offices, hospitals, etc. The negative externalities arising from retail uses are mixed, and probably substantially diluted, in terms of a direct measurable impact, in this interacted variable. Second, while not providing specific information about the relation of retail proximity and residential price, it does provide further evidence that neighborhood design elements do make a difference in residential prices. In this case, for example, the diluted effect of retail externalities implies that the proportion of types of nonresidential uses within the overall residential/nonresidential

mix has measurable effects on residential prices. While not a subject for this dissertation, this finding does raise questions for future research

East Side Within 1,400 Feet. This section of the analysis looks at the Kirkland/Redland side of the study area and includes only those observations within 1,400 feet of the nearest retail. The cities in this sample area were developed more recently than the west sample and use design and street layouts that are more oriented to automobiles or, as in the case of long cul-de-sacs, made possible by automobiles. Not only is the street layout in the east sample more auto-oriented than that in the west sample area but the travel modal choices of the residents are much more oriented to auto use than that of residents in the western part of the study area. While over 82% of the observations in the west side of the study area are within 1,400 feet – walking distance – of the nearest retail, only 21% of the observations on the east side are within 1,400 feet of the nearest retail site.

In contrast to the results of the analysis of the full model in the west sample within 1,400 feet of a retail site, analysis of the east sample within 1,400 foot of a retail site does not show that proximity to retail has a strong influence on residential prices. In the reduced model, travel distance is significant at the 5% level, but not the 1% level. The squared travel distance term is not significant and travel distance and its square are jointly significant only at the 5% level. No other terms on the reduced model are significant except nonresidential mix and its square (*nonresmix* and *nonresmixsq*), which are individually and jointly significant close to the 1% level. The signs are not as expected; price decreases at a decreasing rate as the proportion of nonresidential uses increases.

This is contrary to the findings of both Cao and Cory (1981) and in the west sample in this dissertation.

Using the full model, including the interaction terms, with the east sample within 1,400 feet of a retail site, does not substantially change this outcome. With the full model, neither travel distance or its square nor straight-line distance or its square is individually or jointly significant. Two of the secondary variables examined as having possible nonlinear effects have an effect at a statistically significant level. Only one of the interaction effects is significant here. While not statistically significant, *straight-line distance*, the retail clustering variable, and *visibility* have unexpected signs.

The significant effects of variables of secondary interest are found with *density* and *nonresidential mix*. Neither density nor its square is individually significant in the reduced model or full model, they are jointly significant only in the full model. As density increases, price decreases to a point.

Nonresidential mix and it square are individually significant and jointly significant in the full model and nonresidential mix interacted with straight-line distance is significant. Here the signs are the opposite of what we expect. The price effect decreases through a nonresidential proportion of about 17 percent and increases thereafter. This is inconsistent with theory, the empirical findings of Cao and Cory (1981), and the results for the west sample. In the east sample less than 1,400 feet of a retail site this significant independent variable of interest produces a puzzling result.

What can explain the differences in the effect of proximity to retail uses and the proportion of non-residential between the two samples? A reasonable explanation lies in neighborhood layout and other land development factors. We have already seen that

Table 14: Layout and Distance Differences Between East and West Samples

		West	East	t	P>|t|
Street Segments/Intersections				-37.7539	0.0000
	Minimum	1.269231	0.9929578		
	Mean	1.821874	1.1987595		
	Maximum	3.341177	1.379822		
Street Segments/Cul-de-Sacs				-35.0009	0.0000
	Minimum	5.688889	2.553672		
	Mean	126.6481	6.402661		
	Maximum	412	10.44944		
Density				-58.6814	0.0000
	Minimum	13.56324	10.42336		
	Mean	25.87725	14.0787		
	Maximum	51.35686	27.86523		
Travel Distance to Nearest Retail				33.4779	0.0000
	Minimum	27	7.7		
	Mean	954.45	1461.3		
	Maximum	7371	6708		
Straight-Line Distance to Nearest Retail				19.3814	0.0000
	Minimum	17	56		
	Mean	692.56	873.3		
	Maximum	1399.9	1399.9		

these factors vary considerably between the two sample areas. The analysis in the west sample within 1,400 feet of a retail site demonstrates that neighborhood street layout and residential density can have significant influence on the price effects of travel (convenience) and straight-line (negative externalities) distances. In the west sample, we see that increasing straight-line distance does increase price, ceteris paribus, as the impact of negative externalities diminishes with distances. But, increasing travel distance reduces price, ceteris paribus, off-setting some or all of the straight-line effect, depending on specific distances. A principal objective of modern land use control, the kind of controls that have developed since the time the west area was being built, but certainly in place while the east area was built is to diminish or eliminate the negative effects on residential prices of proximity to retail and other non-residential uses (Mills

1979). In other words, an objective of more recent zoning practices is to minimize or eliminate the effect captured by our straight-line distance variable in our model. McMillen and McDonald (1989; 1991) have found evidence that considerations of externalities are an important influence on zoning decisions and that resulting land use patterns are inefficient in the sense that allowing some non-residential uses into more centralized locations would increase land values. Wallace (1986) investigated zoning patterns in King County, Washington, the site of the current study. She found that zone designations do not follow the market; that is they seem not to be made to maximize land value. In many cases King County zoning decreased parcel price below prices that could be expected in an unconstrained market. Coupled with increased use of autos and their ability to travel greater distances and carry more goods than possible for pedestrians, the differing approaches to land use and land use patterns have contributed to differences between the west sample within 1,400 feet of retail and the east sample within 1,400 feet of retail. Table 14, presents some differences for observations within 1,400 feet of a retail site.

The straight-line distance measure shows that even within the constraint of a maximum of 1,400 feet, separation of residences from retail sites is far greater in the east than the west. In fact, only 5 percent of the observations in the east are within 250 feet of a retail site; over 10 percent of the observations in the west are within 250 feet. Two hundred fifty feet is about the reach of negative effects in the west sample above. Also note that the mean travel distance in the east area exceeds 1,400 feet, which we interpret as the maximum walking distance. On the basis of these measures, it appears that the absence of any significant effects arising from proximity of residences to retail uses in the east sample within 1,400 feet of a retail site results from land use controls separating land

uses coupled with auto oriented design that increases distances between residential and other uses. The resulting segregation of land use types into exclusive concentrations, not only to mitigates the negative effects of proximity to retail, but also eliminate the positive effects of walking distance convenience. This is consistent with the conclusions of McMillen and McDonald (1989; 1991) and Wallace (1986). They find that that zoning decisions include consideration of externalities. They also conclude that zoning leads to inefficient allocation of land use in the sense that allowing some non-residential uses into more centralized locations would increase land values.

West Side Beyond 1,400 Feet. The analysis returns to the Seattle side of the study area, but now includes only those observations beyond 1,400 feet of the nearest retail; observations not within a standard walking distance to the nearest retail use. Only 15 percent of the observations in the west side of the study area are beyond 1,400 feet of the nearest retail.

Our earlier analysis of the pooled samples in this area, west sample beyond 1,400 feet from the nearest retail site, shows that increasing straight-line distance is not relevant - negative retail externalities do not reach to this distance. The model for observations beyond 1,400 feet from a retail sites drops the straight-line variable (*areu1_dis*), the square (*areu1_dissq*), and the interaction with density (*densityXareu1*).

The depressing effect of greater travel distances to retail on residential price persists. According the estimates presented in Appendix B (the sub-sample of observations less than 1,400 feet from a retail site) and Appendix C (the sub-sample of observations greater than 1,400 feet from a retail site), travel distance (*r_net1*) and its square (*r_net1sq*) are individually and jointly significant. In addition, in the full model,

the effect of the interaction between travel distance and neighborhood street layout, a significant factor in the west area within 1,400 feet of retail, is also significant in the sample beyond 1,400 feet of a retail site. But, in the west near sample, it is the interaction with the ratio of street segments to intersections that is significant. In contrast, for the sample beyond 1,400 feet, the interaction with the ratio of cul-de-sacs is significant. It could be that privacy is valued differently by those choosing to live in these different areas.

Coefficients for the travel distance are individually and jointly significant; housing prices diminish at a decreasing rate with increasing distance from retail. In the west close to retail, the price effect of negative externalities, the straight-line effect, diminishes to zero at between 570 and 750 feet from the nearest retail site. In the west sub-sample where closest observations are at 1,400 from retail, we cannot reasonably expect to find any significant price effect from the negative aspects of retail sites; the distance is too great. Preliminary analysis shows that coefficients for the straight-line distances are, as expected, individually and jointly not significant. There is no price effect on residential properties from negative retail site externalities in the west sample far from retail. For the final analysis these straight-line variables are dropped from the model for sub-areas beyond 1,400 feet from a retail site.

According to either the reduced or full model, retail site clusters have no effect on residential prices in this setting, 1,400 or more feet beyond the nearest retail site. This factor is significant in the west sample of observations within 1,400 feet of a retail site where residences are within walking distance of retail. The factor is not significant in the west sub-sample of observations 1,400 feet or more from retail where residences are not

within walking distance of retail. If people in this sample take a retail trip, because of the distances involved, they are likely to go in a car. Once in a car, proximity of retail sites to another is not nearly as important as it is to a pedestrian. People in the sample area close to retail are more likely to take a walking retail trip than are people in the sample far from retail. If shoppers close to retail do walk to the nearest retail, an easy walk among several retail sites is important.

As expected, *traffic noise* is a significant negative influence on residential prices in both the reduced and full models. Retail *visibility* does not have a significant effect on residential price using either model, probably because of the distances involved (only twenty-six of the 2,818 residences included in the far sample area can see retail sites).

In the west sample area beyond 1,400 feet of a retail site, residential density has no significant effect on residential property prices, all else being equal. Densities in the area are relatively low at 19.5 persons per acre. The range of densities is narrow, rising from 13.6 persons per acre to 40.8. (But the highest density is a bit of an outlier as the density at the 90^{th} percentile is 29.0.)

The proportion of non-residential uses similarly has no significant effect on residential price in this setting, west of the lake with observations beyond 1,400 feet of a retail site (in the full model, the squared term is significant, but the principal and squared terms are not jointly significant). Even though not significant, the coefficients do have the signs we expect based on Cao and Cory (1981) and our analysis in the west sample area within 1,400 feet of a retail site.

There is a significant interaction between a measure of neighborhood street connectivity and travel distance. In this case, the significant interaction is with the ratio

of cul-de-sacs to street segments (*segunodaXrnet*). The coefficient on the interaction term is 0.55. The range of the ratio of street segments to cul-de-sacs is 5.68 412 with a mean of 38.67. The lower the ratio, the greater the proportion of cul-de-sacs and the less connected a neighborhood layout.

The measure of travel distance is interacted with the ratio of street segments to cul-de-sacs. As discussed earlier:

$$P = \beta_0 + \beta_1 TD + \beta_2 I + \beta_3 TD * I + \varepsilon$$

Where:

P = Sales Price

TD = Travel Distance

I = Integration, the ratio of segments to cul-de-sacs (*segunodaXrnet*)

Which is interpreted using:

$$\frac{\delta P}{\delta TD} = \beta_1 + \beta_3 I$$

β_1, travel distance, is expected to be negative. As distance increases the positive effects of convenience decrease and price is lowered. A positive β_3 diminishes a negative price of distance on housing values. Here, for example, the coefficient on travel distance (*r_net1*) is –66.29133 and the coefficient on the interaction term is 0.5464004. In a neighborhood with the median ratio of street segments to cul-de-sacs, 54.00, the price effect of travel distance is:

$$-66.29133 + (0.5464004 * 54.00) = -36.7857084$$

But, in a neighborhood with a lower ratio (a higher proportion of cul-de-sacs), the 20th centile is 20.66667, the price effect of travel distance is:

$$-66.29133 + (0.5464001 * 20.66667) = - 54.99705$$

This is consistent with our expectation that a higher ratio of cul-de-sacs, a less integrated neighborhood street layout, makes shopping trips even less convenient.

The effect is pronounced. At the 70th percentile (ratio is 84.66 and the pattern unconnected) the negative price effect is –20.03 compared to –36.78 at the median (a more connected pattern) and –46.50 at the 20th centile (the most connected pattern). Clearly, less connected neighborhood street patterns compound the negative price effect on increasing travel distances to retail sites.

East Side Beyond 1,400 Feet. The final area studied is the Kirkland/Redland side and includes only those observations beyond 1,400 feet of the nearest retail. The automobile orientation of this area has been discussed at length. A little more than 78 percent of the all observations on the east side are beyond 1,400 feet – walking distance – of the nearest retail (compared with 15 percent for the west side). Average travel distance between a residence and the nearest retail site is over 4,200 feet, almost eight tenths of a mile (about a 40 minute walk at three miles per hour); it is only half that distance in the west side sub-area greater than 1,400 feet from a retail site. Travel distances are by far the greatest of any of the sub-samples studied. The average straight-line distance is greater than one-half mile. There is no effect from negative spillovers from retail uses; as with the west sample beyond 1,400 feet, the straight-line distance variables are dropped in the east model beyond 1,400 feet.

The measure of retail clustering is significant and is jointly significant with its squared term in both the reduced and full models for the east sub-sample for observations beyond 1,400 feet of a retail site. In the far west side sub-sample where retail trips are beyond walking distance, we argued that extended distance between retail uses is not significant; trips between retail would be made in a car, as a car is already in use. In this far east side sub-area, the clustering measure may be significant because the distances are so great (the mean in the far east sample is 2,275 feet -almost half a mile - compared to a mean of 1,165 feet in the far west sample) that inconvenience is a factor, even in a car.

In both the reduced and full models *traffic noise* has a significantly negative effect on residential price in both the far west sample and the far east sample, but *visibility*[*] of retail from residences does not affect residential price, nor does the proportion of *non-residential uses*. In analysis of areas beyond 1,400 feet of a retail site, *density* is a significant factor in the reduced model, but not the full model after interaction terms are introduced to the analysis.

As argued above, distances and public policy measures intended to minimize the negative effects of retail externalities may explain the lack of effect for any of these variables.

The variable for *travel distance* to the nearest retail is significant as is its squared term, indicating a nonlinear relationship. But the direction of the signs on the coefficients is not what we expect. The expectation is that the travel distance variable will be negative; price will diminish with distance. As distance increases, we expect price to

---

[*] Of the 5,258 observations on the east side at a distance greater than 1,400 feet from a retail site only 64 can see retail sites. In the majority of cases, these residences are on

decrease as the cost of travel is traded against willingness to pay for housing. (In the preliminary analysis, which included the variables for straight-line distance, travel distance has a negative sign but is insignificant. Straight-line distance has an unexpected negative sign, but is not significant.) This sign on *travel distance* is puzzling at first, but is consistent with the findings of Mahan et al (2000) who also found prices increasing with distance from retail sites at large distances. They speculate that the rise in value might reflect a general aversion on the part of some people to congestion, noise, etc. associated with more developed areas; the kind of areas that include retail sites.

Nelson supplies two related theories that can help explain the increase in price with increasing distance. First, he develops a theory of housing prices near urban growth boundaries (Nelson, 1986). In addition to the general incremental price increase expected throughout the rent gradient in a growth constrained area, an actual increase in housing price is expected at the boundary. The increase arises from proximity to open space amenities beyond the growth boundary. Thus, for a distance of about a mile (found in preliminary research) prices actually rise as a function of reduced distance to the growth boundary, not as a function of distance from developed areas and retail concentrations. In a second theoretical article, Nelson (1993) ideas about the effect on residential prices of a complex interaction between the influences of edge cities, the outer fringes of urban development, and CBDs. While prices fall with respect to distance from a CBD and from the edge of urban development, predicts they will increase with increasing distance from an edge city as some households value, as speculated by Mahon, et al, separation from the noise, congestion, and nuisances of edge cities. The area where we find prices increasing

---

high ground with extended vistas.

with increasing distance from retail, the area east of the lake more than 1,400 feet from retail, is also an area near both an urban growth boundary and an edge city. Perhaps these influences can be used to explain the unexpected rise of housing prices with respect to distance from retail in this part of the study area.

Interactions of travel distance with both indices of neighborhood connectivity are significant. The interaction with street segments related to cul-de-sacs is significant at the five percent level, but not at the one percent level. As is the case with the regressions in Appendix C for west and east observations less than 1,400 feet from retail, they have opposite signs.

The most likely explanation for the price curves not behaving as expected is that the wrong variable is being used to measure convenience. While travel distance works well in the settings on the west side sample, it may be that travel time is needed to reflect convenience on the east side. When walking, travel time and distance are highly correlated. This correlation does not necessarily hold automobile travel where network impedance – e.g. accidents, congestion, broken signals, street collapses, and so on – can disrupt the relationship between distance and time on both a temporary and long term basis.

CHAPTER 6

CONCLUSIONS

The two samples used in this study differ. They display significantly different social and development characteristics. One, the Seattle sample, has fewer children, more professionals, a larger portion of non-automobile commuters, longer trip to work times, older and smaller houses on smaller lots. In contrast, the east sample is an automobile oriented edge city. The Seattle side was developed at a time when there was more dependence on public transit and walking; the gridiron neighborhood street layout reflects the pedestrian orientation of the times. The edge city side has an auto oriented design featuring lower density development, less connectivity in the street patterns and lavish inclusion of cul-de-sacs in neighborhood layout. Land uses are integrated in the Seattle side, but highly segregated in the Kirkland/Redmond side.

In the older gridiron area, proximity to retail creates a positive price effect for residences; the further from retail, the lower the residential price, all else being equal. On the other hand, in the same setting, proximity to retail creates a negative price effect due to exposure to disamenities such as noise and congestion. The positive effect outweighs the negative effect. Up to about 250 feet, the negative effect of disamenities results in a net loss. Beyond a distance of around 250 feet, the effect is positive for almost another 1,000 feet. Whether or not retail sites are visible from residences significantly affects the

strength of negative disamenities. The less visible a retail site, the lower the effects on residential price. The implication for public policy is clear and is not new: measures should be developed to provide visual barriers between retail and residential sites. However, barriers should not be created that sacrifice access from residential sites to retail. Loss of access will reduce convenience and negatively affect residential value.

In the edge city portion of the study area, the positive effect of convenience to retail is not observed, even though the negative effect is approximately the same magnitude and reach as in the gridiron patterned area. The difference is that far fewer residences are in the negatively affected range. In this portion of the study area, reduction of the negative price effects of proximity to retail has been achieved at the cost of positive price effects of convenience. There seems to be a problem with analysis in the edge city area, especially with travel distance serving as a proxy for convenience. Analysis may be more useful if travel time data can be made available and incorporated.

Neighborhood layout and density have a significant effect on the magnitude and reach of the travel and straight-line effects on price. As neighborhood layout becomes more integrated, the positive price effect of proximity increases.

There is an important public policy implication. Modern land use controls have been oriented to mitigating the negative effects of disamenities on residential prices, as seen in the Kirkland/Redmond portion of the study area. While this policy has been successful to the extent residential development has been minimized within the reach of the disamenities, some of the positive benefit of being near retail seems to have been lost altogether as well. The policy design question revolves around whether or not the positive benefits of segregating uses outweigh the negatives. In pedestrian oriented

neighborhoods, as street patterns become more integrated, the probability that mixing land uses enhances residential property values increases and segregating land uses diminishes this positive effect. In automobile oriented neighborhoods, there are no significant residential price effects associated with proximity to retail uses, but increased neighborhood street connectivity itself has a negative influence on housing price.

The extremely different neighborhoods and development types in the two parts of the study area house extremely different populations. People may have selected these different types of areas because of preferences for convenience over privacy, vice versa, or for other reasons. In any event, it is apparent that there are markets and preferences for a great variety of development types. Further research to understand the convenience and privacy preferences of people choosing to live in these different areas could be very interesting.

This dissertation points to several avenues of future research. One set of questions encompasses the space syntax measures of neighborhood integration and connectivity. How does this measure differ from the simpler ratio indices used in theis dissertation? Do space syntax measures provide greater insight into the effect of neighborhood layout on residential price? How can these measures be used to positively inform public policy?

The effect of the mix non-residential uses has very uneven results in this dissertation. In different settings, this measure has been unrelated, positively related, and negatively related to residential price. The question of non-residential mix lies at the heart of zoning and land use control, so further exploration and greater understanding of this issue is warranted.

APPENDIX A: REDUCED MODEL

REGRESSION RESULTS

Appendix A: West less than 1400 feet

Regression with robust standard errors

Number of obs	=	16266
F(56, 16209)	=	465.10
Prob > F	=	0.0000
R-squared	=	0.7105
Root MSE	=	56356

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-12929.15	1330.112	-9.72	0.000	-15536.32	-10321.99
trendsq	2688.634	188.3814	14.27	0.000	2319.386	3057.882
trendcu	-66.9837	7.917247	-8.46	0.000	-82.50238	-51.46502
sqfttotliv	42.87393	6.27603	6.83	0.000	30.57222	55.17564
sqftsq	.0033097	.003106	1.07	0.287	-.0027784	.0093977
bedrooms	2320.541	3999.371	0.58	0.562	-5518.668	10159.75
sqftxbedroom	-.3134401	2.341504	-0.13	0.894	-4.903047	4.276166
bathrooms	7330.354	1196.074	6.13	0.000	4985.918	9674.79
age	-797.6224	115.3801	-6.91	0.000	-1023.78	-571.4647
agesq	6.034941	.9084725	6.64	0.000	4.254235	7.815648
condition	11441.27	846.5246	13.52	0.000	9781.993	13100.56
sqftlot	4.727663	.6509206	7.26	0.000	3.451786	6.003539
sqftlotsq	.0001184	.0000341	3.47	0.001	.0000516	.0001851
noviewd	-39173.1	2575.219	-15.21	0.000	-44220.81	-34125.38
wfntlocati	78221.7	20071.77	3.90	0.000	38878.81	117564.6
itbs_read	-572.4814	75.73707	-7.56	0.000	-720.9344	-424.0284
dis_bofa	-43.74119	5.305628	-8.24	0.000	-54.14081	-33.34157
dis_bofasq	-.0007399	.0000813	-9.11	0.000	-.0008992	-.0005806
az_bofa	14676.08	1829.465	8.02	0.000	11090.13	18262.04
dis_mic	-71.76938	11.47574	-6.25	0.000	-94.2631	-49.27566
dis_micsq	.0011552	.0001146	10.08	0.000	.0009305	.0013798
az_mic	61474.42	5211.33	11.80	0.000	51259.64	71689.2
dis_xway	1.56743	.6192576	2.53	0.011	.3536172	2.781244
dis_xwaysq	-.0001038	.000072	-1.44	0.149	-.0002449	.0000373
az_xway	-86.31673	13.93821	-6.19	0.000	-113.6372	-58.99631
apt1_dis	23.41816	4.658803	5.03	0.000	14.28639	32.54993
apt1_dissq	-.00692	.0035291	-1.96	0.050	-.0138374	-2.58e-06
apt1_az	.7449528	4.696311	0.16	0.874	-8.460335	9.950241
cult1_dis	1.754333	3.946462	0.44	0.657	-5.981168	9.489834
cult1_dissq	.0018334	.002244	0.82	0.414	-.002565	.0062318
cult1_az	3.990335	4.332841	0.92	0.357	-4.50251	12.48318
govt1_dis	-13.61697	1.557286	-8.74	0.000	-16.66943	-10.56452
govt1_dissq	.0034397	.0002857	12.04	0.000	.0028797	.0039997
govt1_az	-16.93751	5.234594	-3.24	0.001	-27.19789	-6.677126
hotel1_dis	7.529643	1.202501	6.26	0.000	5.172609	9.886677
hotel1_dissq	-.0007143	.0001403	-5.09	0.000	-.0009893	-.0004394
hotle1_az	44.78703	6.343192	7.06	0.000	32.35367	57.22039
off1_dis	-9.141017	4.901906	-1.86	0.062	-18.74929	.467259
off1_dissq	.0063278	.0022825	2.77	0.006	.0018539	.0108017
off1_az	-4.095424	5.365596	-0.76	0.445	-14.61258	6.421735
hops1_dis	8.980765	1.689486	5.32	0.000	5.669187	12.29234
hops1_dissq	-.0010588	.0001905	-5.56	0.000	-.0014322	-.0006855
hosp1_az	-36.49981	5.371682	-6.79	0.000	-47.0289	-25.97072
ind1_dis	18.45243	5.270078	3.50	0.000	8.122498	28.78237
ind1_dissq	-.0083423	.0032948	-2.53	0.011	-.0148005	-.0018841
ind1_az	-.4949463	5.172444	-0.10	0.924	-10.63351	9.643614
sch1_dis	-.4646443	3.062028	-0.15	0.879	-6.466557	5.537269
sch1_dissq	.0010596	.001227	0.86	0.388	-.0013455	.0034648
sch1_az	-.1553403	4.644827	-0.03	0.973	-9.259714	8.949034
r_net1	-2.470752	6.422431	-0.38	0.700	-15.05943	10.11792
r_net1sq	.0042635	.0021319	2.00	0.046	.0000848	.0084422
areu1_dis	40.63522	9.264488	4.39	0.000	22.4758	58.79463
areu1_dissq	-.0351209	.0051027	-6.88	0.000	-.0451227	-.025119
areu1_az	15.84151	5.152497	3.07	0.002	5.742044	25.94097
average	-7.809761	3.569435	-2.19	0.029	-14.80625	-.8132756
avesq	.0009386	.0017462	0.54	0.591	-.0024842	.0043613
_cons	-6689656	841781.3	-7.95	0.000	-8339641	-5039672

```
. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2, 16209) =    8.40
            Prob > F =    0.0002

. test areu1_dis areu1_dissq

 ( 1)  areu1_dis = 0
 ( 2)  areu1_dissq = 0

       F(  2, 16209) =   31.68
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2, 16209) =   14.98
            Prob > F =    0.0000
```

Appendix A: West greater than 1400 feet

Regression with robust standard errors

Number of obs = 2818
F(53, 2764) = 111.71
Prob > F = 0.0000
R-squared = 0.7670
Root MSE = 73119

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-13076.69	3825.57	-3.42	0.001	-20577.96	-5575.427
trendsq	2906.891	542.245	5.36	0.000	1843.645	3970.138
trendcu	-75.02234	22.53597	-3.33	0.001	-119.2114	-30.8333
sqfttotliv	29.63279	10.9941	2.70	0.007	8.075315	51.19026
sqftsq	.0052952	.002983	1.78	0.076	-.0005539	.0111442
bedrooms	1142.038	6037.691	0.19	0.850	-10696.8	12980.88
sqftxbedroom	-.7998521	3.144333	-0.25	0.799	-6.965331	5.365627
bathrooms	13196.12	3287.879	4.01	0.000	6749.168	19643.06
age	-2132.442	363.2964	-5.87	0.000	-2844.801	-1420.082
agesq	17.68335	3.006679	5.88	0.000	11.78778	23.57891
condition	11575.89	2618.627	4.42	0.000	6441.223	16710.55
sqftlot	3.580938	1.18994	3.01	0.003	1.247677	5.914199
sqftlotsq	.0001233	.0000341	3.62	0.000	.0000565	.0001901
noviewd	-57997.87	5102.573	-11.37	0.000	-68003.11	-47992.63
wfntlocati	48933.38	16470.5	2.97	0.003	16637.65	81229.1
itbs_read	1966.994	385.1461	5.11	0.000	1211.791	2722.197
dis_bofa	-142.0025	35.49798	-4.00	0.000	-211.6077	-72.39722
dis_bofasq	-.0001689	.0003136	-0.54	0.590	-.0007838	.000446
az_bofa	9528.567	7012.923	1.36	0.174	-4222.532	23279.67
dis_mic	-278.0377	60.93148	-4.56	0.000	-397.5135	-158.5619
dis_micsq	.0032452	.0007227	4.49	0.000	.0018282	.0046622
az_mic	119049.9	31421.47	3.79	0.000	57437.99	180661.9
dis_xway	-17.96768	3.564032	-5.04	0.000	-24.95612	-10.97925
dis_xwaysq	-.0002286	.0003036	-0.75	0.451	-.000824	.0003667
az_xway	-73.80866	36.87225	-2.00	0.045	-146.1086	-1.508727
apt1_dis	23.71401	9.368104	2.53	0.011	5.34482	42.0832
apt1_dissq	.0004546	.0035075	0.13	0.897	-.006423	.0073322
apt1_az	-49.38069	15.99549	-3.09	0.002	-80.74501	-18.01638
cult1_dis	25.97757	10.01137	2.59	0.010	6.347059	45.60808
cult1_dissq	-.0164077	.0037131	-4.42	0.000	-.0236884	-.0091271
cult1_az	.1328912	16.10694	0.01	0.993	-31.44995	31.71573
govt1_dis	-10.004	10.37311	-0.96	0.335	-30.34383	10.33584
govt1_dissq	.0021528	.0015807	1.36	0.173	-.0009467	.0052523
govt1_az	9.775465	28.35021	0.34	0.730	-45.81426	65.36519
hotel1_dis	-6.813349	5.06903	-1.34	0.179	-16.75282	3.126119
hotel1_dissq	.0000922	.0004311	0.21	0.831	-.000753	.0009375
hotle1_az	9.301087	27.79217	0.33	0.738	-45.19443	63.79661
off1_dis	41.08535	10.30817	3.99	0.000	20.87286	61.29785
off1_dissq	-.0137496	.0030361	-4.53	0.000	-.0197028	-.0077964
off1_az	5.313115	22.70511	0.23	0.815	-39.20758	49.83381
hops1_dis	-17.28892	6.250403	-2.77	0.006	-29.54485	-5.032989
hops1_dissq	.0009547	.0006486	1.47	0.141	-.0003171	.0022266
hosp1_az	-8.681328	21.51196	-0.40	0.687	-50.86247	33.49982
ind1_dis	6.085755	16.46373	0.37	0.712	-26.19669	38.3682
ind1_dissq	-.0019787	.0076525	-0.26	0.796	-.016984	.0130265
ind1_az	7.096318	20.19913	0.35	0.725	-32.51059	46.70323
sch1_dis	-8.392351	12.66569	-0.66	0.508	-33.22752	16.44282
sch1_dissq	.0063102	.0036046	1.75	0.080	-.0007577	.0133782
sch1_az	42.91735	23.41549	1.83	0.067	-2.996273	88.83098
r_net1	-19.61387	10.29809	-1.90	0.057	-39.8066	.5788638
r_net1sq	.004717	.0014333	3.29	0.001	.0019066	.0075275
average	-9.431519	8.599564	-1.10	0.273	-26.29374	7.430701
avesq	.0039222	.0033546	1.17	0.242	-.0026556	.0105
_cons	-4466812	3135865	-1.42	0.154	-1.06e+07	1682063

```
. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  2764) =   13.35
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  2764) =    0.68
            Prob > F =    0.5048
```

Appendix A: East less than 1400 feet

Regression with robust standard errors

Number of obs =	1482
F(56, 1425) =	48.13
Prob > F =	0.0000
R-squared =	0.7135
Root MSE =	1.0e+05

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-30572.75	8011.835	-3.82	0.000	-46289.01	-14856.49
trendsq	4886.543	1102.396	4.43	0.000	2724.049	7049.037
trendcu	-122.3213	44.77341	-2.73	0.006	-210.1502	-34.49246
sqfttotliv	34.27132	54.37685	0.63	0.529	-72.39595	140.9386
sqftsq	.0187771	.0157605	1.19	0.234	-.0121392	.0496933
bedrooms	17762.35	16846.62	1.05	0.292	-15284.48	50809.19
sqftxbedroom	-13.1136	8.710459	-1.51	0.132	-30.2003	3.973098
bathrooms	2168.531	7502.975	0.29	0.773	-12549.53	16886.59
age	-2763.74	600.5661	-4.60	0.000	-3941.829	-1585.652
agesq	25.38094	5.175365	4.90	0.000	15.22879	35.53309
condition	430.3184	4162.159	0.10	0.918	-7734.299	8594.935
sqftlot	1.439394	1.850471	0.78	0.437	-2.190544	5.069333
sqftlotsq	.0000184	.0000414	0.44	0.657	-.0000628	.0000996
noviewd	-22365.8	10655.84	-2.10	0.036	-43268.62	-1462.972
wfntlocati	91977.71	17006.49	5.41	0.000	58617.26	125338.2
itbs_read	2951.3	705.7699	4.18	0.000	1566.841	4335.76
dis_bofa	-747.6366	130.2508	-5.74	0.000	-1003.14	-492.1327
dis_bofasq	.0081525	.0014534	5.61	0.000	.0053016	.0110035
az_bofa	-208667.9	45715.35	-4.56	0.000	-298344.5	-118991.3
dis_mic	-392.0995	77.78289	-5.04	0.000	-544.6808	-239.5183
dis_micsq	.0032951	.0011893	2.77	0.006	.0009621	.005628
az_mic	-15718.4	3938.539	-3.99	0.000	-23444.35	-7992.441
dis_xway	-15.681	14.91285	-1.05	0.293	-44.9345	13.57249
dis_xwaysq	-.0020102	.0016886	-1.19	0.234	-.0053225	.0013022
az_xway	451.2827	150.4014	3.00	0.003	156.2509	746.3146
apt1_dis	-24.1252	22.68584	-1.06	0.288	-68.62643	20.37603
apt1_dissq	.0139627	.0106896	1.31	0.192	-.0070064	.0349318
apt1_az	24.37261	31.84414	0.77	0.444	-38.09381	86.83904
cult1_dis	22.63969	20.05749	1.13	0.259	-16.70568	61.98507
cult1_dissq	-.0013949	.0088322	-0.16	0.875	-.0187204	.0159307
cult1_az	9.157676	37.01389	0.25	0.805	-63.44989	81.76524
govt1_dis	33.39839	11.14544	3.00	0.003	11.53516	55.26163
govt1_dissq	-.0059616	.0015692	-3.80	0.000	-.0090399	-.0028833
govt1_az	-20.86551	38.84212	-0.54	0.591	-97.05938	55.32836
hotel1_dis	-4.61371	18.78577	-0.25	0.806	-41.46443	32.23701
hotel1_dissq	.0010844	.0010213	1.06	0.289	-.0009191	.0030879
hotle1_az	76.6516	38.68482	1.98	0.048	.7663039	152.5369
off1_dis	4.962159	17.44642	0.28	0.776	-29.26126	39.18558
off1_dissq	-.002327	.0034798	-0.67	0.504	-.0091532	.0044992
off1_az	33.69798	34.03673	0.99	0.322	-33.06949	100.4655
hops1_dis	-26.55154	21.03593	-1.26	0.207	-67.81626	14.71318
hops1_dissq	.0015632	.0011705	1.34	0.182	-.0007329	.0038593
hosp1_az	-7.706662	36.82041	-0.21	0.834	-79.93468	64.52136
ind1_dis	94.23233	31.46323	2.99	0.003	32.51311	155.9515
ind1_dissq	-.0490296	.0200504	-2.45	0.015	-.0883611	-.0096982
ind1_az	-25.10544	29.79307	-0.84	0.400	-83.54843	33.33754
sch1_dis	-53.9088	25.14084	-2.14	0.032	-103.2258	-4.59178
sch1_dissq	.0160994	.0066763	2.41	0.016	.0030029	.0291959
sch1_az	-16.397	42.7277	-0.38	0.701	-100.2129	67.41894
r_net1	27.83002	13.55039	2.05	0.040	1.249161	54.41087
r_net1sq	-.0023796	.0021397	-1.11	0.266	-.0065769	.0018178
areu1_dis	-53.22495	40.01833	-1.33	0.184	-131.7261	25.27622
areu1_dissq	.0081369	.0215781	0.38	0.706	-.0341913	.0504651
areu1_az	14.69474	30.43453	0.48	0.629	-45.00656	74.39603
average	25.02442	23.80713	1.05	0.293	-21.67635	71.7252
avesq	-.0022552	.0048121	-0.47	0.639	-.0116947	.0071844
_cons	7.42e+07	1.49e+07	4.98	0.000	4.50e+07	1.03e+08

```
. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  1425) =    7.11
            Prob > F =    0.0008

. test areu1_dis areu1_dissq

 ( 1)  areu1_dis = 0
 ( 2)  areu1_dissq = 0

       F(  2,  1425) =    2.65
            Prob > F =    0.0713

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  1425) =    1.13
            Prob > F =    0.3238
```

Appendix A: East greater than 1400 feet

Regression with robust standard errors

Number of obs =	5258
F(53, 5204) =	233.44
Prob > F =	0.0000
R-squared =	0.7364
Root MSE =	68613

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-3665.555	2933.968	-1.25	0.212	-9417.365	2086.255
trendsq	1467.26	451.4764	3.25	0.001	582.1767	2352.343
trendcu	-20.86822	20.14075	-1.04	0.300	-60.35254	18.61611
sqfttotliv	60.29486	21.15471	2.85	0.004	18.82275	101.767
sqftsq	.0083083	.0077724	1.07	0.285	-.0069289	.0235454
bedrooms	16868.03	10396.97	1.62	0.105	-3514.395	37250.45
sqftxbedroom	-8.992341	5.102223	-1.76	0.078	-18.99484	1.010159
bathrooms	-19.42942	2777.114	-0.01	0.994	-5463.74	5424.881
age	-3673.442	312.6851	-11.75	0.000	-4286.436	-3060.448
agesq	31.79188	3.691746	8.61	0.000	24.55451	39.02926
condition	4933.43	2908.049	1.70	0.090	-767.5657	10634.43
sqftlot	1.424104	.2726476	5.22	0.000	.8895999	1.958607
sqftlotsq	-4.97e-06	1.59e-06	-3.12	0.002	-8.10e-06	-1.85e-06
noviewd	-51756.17	8183.674	-6.32	0.000	-67799.61	-35712.74
wfntlocati	62014.75	13903.59	4.46	0.000	34757.87	89271.63
itbs_read	1241.834	270.1798	4.60	0.000	712.1682	1771.5
dis_bofa	-26.44742	64.15473	-0.41	0.680	-152.2176	99.32279
dis_bofasq	.000025	.0006881	0.04	0.971	-.001324	.001374
az_bofa	10322.79	23348.45	0.44	0.658	-35449.97	56095.55
dis_mic	-10.20357	26.88961	-0.38	0.704	-62.9185	42.51136
dis_micsq	.0005966	.0002506	2.38	0.017	.0001054	.0010877
az_mic	4607.266	1272.637	3.62	0.000	2112.363	7102.169
dis_xway	6.445819	4.077064	1.58	0.114	-1.546938	14.43858
dis_xwaysq	.0002469	.0003668	0.67	0.501	-.0004722	.000966
az_xway	136.9762	30.20636	4.53	0.000	77.7591	196.1934
apt1_dis	-18.86289	5.492912	-3.43	0.001	-29.6313	-8.094473
apt1_dissq	.0015393	.0012666	1.22	0.224	-.0009438	.0040223
apt1_az	15.0214	15.03989	1.00	0.318	-14.46311	44.5059
cult1_dis	-4.085668	6.576269	-0.62	0.534	-16.97792	8.80658
cult1_dissq	.0024153	.0020666	1.17	0.243	-.001636	.0064667
cult1_az	13.80005	11.94245	1.16	0.248	-9.612165	37.21227
govt1_dis	5.299957	4.602165	1.15	0.250	-3.722219	14.32213
govt1_dissq	.0002433	.000431	0.56	0.572	-.0006016	.0010881
govt1_az	32.7484	22.46794	1.46	0.145	-11.2982	76.79499
hotel1_dis	-13.10105	8.407996	-1.56	0.119	-29.58426	3.382149
hotel1_dissq	.000506	.0005464	0.93	0.355	-.0005652	.0015771
hotle1_az	-8.806802	18.56305	-0.47	0.635	-45.19817	27.58457
off1_dis	-6.285167	3.633583	-1.73	0.084	-13.40852	.838182
off1_dissq	.0000779	.0007623	0.10	0.919	-.0014165	.0015722
off1_az	2.753881	15.69256	0.18	0.861	-28.01013	33.51789
hops1_dis	25.79958	6.363835	4.05	0.000	13.32379	38.27537
hops1_dissq	-.0011897	.0003712	-3.20	0.001	-.0019175	-.0004619
hosp1_az	-57.16847	21.79411	-2.62	0.009	-99.89408	-14.44286
ind1_dis	-9.162609	7.558141	-1.21	0.225	-23.97974	5.65452
ind1_dissq	-.0016749	.0029567	-0.57	0.571	-.0074712	.0041214
ind1_az	38.95078	10.56367	3.69	0.000	18.24156	59.66001
sch1_dis	-9.01746	6.469082	-1.39	0.163	-21.69958	3.664658
sch1_dissq	.00371	.0019472	1.91	0.057	-.0001073	.0075274
sch1_az	3.733429	13.59698	0.27	0.784	-22.92236	30.38922
r_net1	15.35486	5.240653	2.93	0.003	5.080983	25.62875
r_net1sq	-.0010773	.0005355	-2.01	0.044	-.0021271	-.0000274
average	-7.277625	3.966075	-1.83	0.067	-15.0528	.4975467
avesq	.0007415	.0006679	1.11	0.267	-.0005679	.0020508
_cons	-1998754	7485869	-0.27	0.789	-1.67e+07	1.27e+07

```
. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  5204) =   10.21
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  5204) =    4.84
            Prob > F =    0.0079
```

APPENDIX B: REGRESSION RESULTS

WITH NEIGHBORHOOD AND ENVIRONMENTAL VARIABLES

Appendix B: West less than 1400 feet

Regression with robust standard errors

Number of obs	=	16266
F(64, 16201)	=	424.84
Prob > F	=	0.0000
R-squared	=	0.7208
Root MSE	=	55359

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-12612.63	1314.134	-9.60	0.000	-15188.48	-10036.78
trendsq	2655.667	185.6956	14.30	0.000	2291.683	3019.651
trendcu	-65.98032	7.793988	-8.47	0.000	-81.2574	-50.70325
sqfttotliv	42.75185	6.262026	6.83	0.000	30.47759	55.02611
sqftsq	.003401	.0031638	1.07	0.282	-.0028004	.0096023
bedrooms	3043.922	4144.017	0.73	0.463	-5078.809	11166.65
sqftxbedroom	-.7288212	2.431346	-0.30	0.764	-5.494528	4.036885
bathrooms	6777.391	1187.581	5.71	0.000	4449.601	9105.181
age	-726.4394	115.7349	-6.28	0.000	-953.2926	-499.5862
agesq	5.128803	.9058484	5.66	0.000	3.353241	6.904366
condition	10761.6	833.7564	12.91	0.000	9127.344	12395.85
sqftlot	4.905612	.6653997	7.37	0.000	3.601355	6.209869
sqftlotsq	.0001191	.0000352	3.38	0.001	.00005	.0001881
noviewd	-41283.5	2594.652	-15.91	0.000	-46369.3	-36197.69
wfntlocati	77396.18	19875.22	3.89	0.000	38438.55	116353.8
itbs_read	-837.5874	81.82809	-10.24	0.000	-997.9794	-677.1953
dis_bofa	-38.70749	5.597181	-6.92	0.000	-49.67858	-27.73639
dis_bofasq	-.0005353	.0000845	-6.33	0.000	-.000701	-.0003696
az_bofa	11646.17	1915.808	6.08	0.000	7890.979	15401.37
dis_mic	-57.80291	11.50782	-5.02	0.000	-80.35951	-35.2463
dis_micsq	.0009191	.000116	7.92	0.000	.0006917	.0011465
az_mic	49412.65	5413.278	9.13	0.000	38802.03	60023.27
dis_xway	-1.111054	.6649806	-1.67	0.095	-2.41449	.1923811
dis_xwaysq	.0000321	.0000765	0.42	0.675	-.0001179	.0001821
az_xway	-100.3497	14.2466	-7.04	0.000	-128.2746	-72.42481
apt1_dis	26.46069	4.605212	5.75	0.000	17.43397	35.48742
apt1_dissq	-.0088098	.00349	-2.52	0.012	-.0156506	-.0019689
apt1_az	3.331887	4.630424	0.72	0.472	-5.744255	12.40803
cult1_dis	-2.560677	3.926395	-0.65	0.514	-10.25684	5.135489
cult1_dissq	.0030391	.0022117	1.37	0.169	-.001296	.0073743
cult1_az	-.583291	4.281706	-0.14	0.892	-8.975906	7.809324
govt1_dis	-10.6817	1.557689	-6.86	0.000	-13.73494	-7.628453
govt1_dissq	.0029169	.000285	10.23	0.000	.0023582	.0034756
govt1_az	-31.19587	5.336133	-5.85	0.000	-41.65528	-20.73646
hotel1_dis	6.633398	1.298075	5.11	0.000	4.089028	9.177768
hotel1_dissq	-.0003792	.0001448	-2.62	0.009	-.0006631	-.0000953
hotle1_az	29.56885	6.487475	4.56	0.000	16.85268	42.28502
off1_dis	-9.111109	5.052638	-1.80	0.071	-19.01484	.7926198
off1_dissq	.0058734	.0023396	2.51	0.012	.0012875	.0104593
off1_az	-4.561576	5.382607	-0.85	0.397	-15.11208	5.988928
hops1_dis	7.486326	1.785284	4.19	0.000	3.986973	10.98568
hops1_dissq	-.0009884	.0002	-4.94	0.000	-.0013805	-.0005964
hosp1_az	-36.34542	5.14756	-7.06	0.000	-46.43521	-26.25563
ind1_dis	20.29989	5.335371	3.80	0.000	9.841975	30.75781
ind1_dissq	-.0113221	.0033202	-3.41	0.001	-.01783	-.0048142
ind1_az	-4.066408	5.149511	-0.79	0.430	-14.16002	6.027201
sch1_dis	-2.380282	3.139782	-0.76	0.448	-8.534601	3.774037
sch1_dissq	.0016984	.0012421	1.37	0.172	-.0007363	.0041332
sch1_az	10.25573	4.633953	2.21	0.027	1.172675	19.33879
r_net1	-14.19011	6.424544	-2.21	0.027	-26.78293	-1.597299
r_net1sq	.005778	.0021297	2.71	0.007	.0016035	.0099525
areu1_dis	37.75311	9.333923	4.04	0.000	19.45759	56.04863
areu1_dissq	-.0262474	.0051558	-5.09	0.000	-.0363534	-.0161415
areu1_az	16.97924	5.030506	3.38	0.001	7.118897	26.83959
average	-3.373378	3.570519	-0.94	0.345	-10.37199	3.625234
avesq	-.0001604	.0017441	-0.09	0.927	-.0035791	.0032583
trafficnoi	-13563.26	735.6072	-18.44	0.000	-15005.13	-12121.39
visibility	-3826.505	1433.014	-2.67	0.008	-6635.371	-1017.639
seg_tnodes	-6676.773	1072.922	-6.22	0.000	-8779.819	-4573.727
seg_unodea	54.00085	5.580945	9.68	0.000	43.06159	64.94012

```
     density |  -2757.213   615.7688    -4.48   0.000    -3964.188   -1550.239
   densitysq |   46.70411   9.071067     5.15   0.000     28.92381     64.4844
   nonresmix |   1046.145    315.244     3.32   0.001     428.2323    1664.059
 nonresmixsq |  -11.55593   7.043578    -1.64   0.101    -25.36212    2.250265
       _cons |   -5172260   875036.9    -5.91   0.000     -6887429    -3457091
------------------------------------------------------------------------------

. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2, 16201) =    3.92
            Prob > F =    0.0199

. test areu1_dis areu1_dissq

 ( 1)  areu1_dis = 0
 ( 2)  areu1_dissq = 0

       F(  2, 16201) =   13.66
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2, 16201) =    5.80
            Prob > F =    0.0030

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2, 16201) =   18.11
            Prob > F =    0.0000

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2, 16201) =   20.36
            Prob > F =    0.0000
```

Appendix B: West greater than 1400 feet

Regression with robust standard errors

Number of obs = 2818
F(61, 2756) = 101.05
Prob > F = 0.0000
R-squared = 0.7710
Root MSE = 72585

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-13772.9	3816.775	-3.61	0.000	-21256.93	-6288.871
trendsq	3018.256	539.1603	5.60	0.000	1961.057	4075.455
trendcu	-79.58048	22.35615	-3.56	0.000	-123.417	-35.74399
sqfttotliv	29.39151	11.00647	2.67	0.008	7.80974	50.97328
sqftsq	.0048553	.0029953	1.62	0.105	-.001018	.0107285
bedrooms	-787.198	6111.302	-0.13	0.898	-12770.39	11196
sqftxbedroom	-.1314881	3.173302	-0.04	0.967	-6.353777	6.090801
bathrooms	13555.83	3291.645	4.12	0.000	7101.486	20010.17
age	-2227.571	363.6618	-6.13	0.000	-2940.648	-1514.494
agesq	18.30205	3.014287	6.07	0.000	12.39155	24.21254
condition	11584.3	2633.09	4.40	0.000	6421.274	16747.33
sqftlot	3.76891	1.20404	3.13	0.002	1.407998	6.129822
sqftlotsq	.000118	.0000347	3.40	0.001	.00005	.0001859
noviewd	-60132.25	5218.238	-11.52	0.000	-70364.3	-49900.19
wfntlocati	54798	16784.54	3.26	0.001	21886.46	87709.55
itbs_read	1728.13	431.6981	4.00	0.000	881.6459	2574.615
dis_bofa	-179.5413	40.96735	-4.38	0.000	-259.8711	-99.21151
dis_bofasq	.0012011	.0004644	2.59	0.010	.0002904	.0021117
az_bofa	2459.659	7590.22	0.32	0.746	-12423.43	17342.75
dis_mic	-231.5527	66.41677	-3.49	0.000	-361.7844	-101.321
dis_micsq	.0025388	.0007939	3.20	0.001	.0009821	.0040956
az_mic	81564.34	34479.68	2.37	0.018	13955.72	149173
dis_xway	-22.85043	4.004355	-5.71	0.000	-30.70227	-14.99859
dis_xwaysq	.0001273	.0003345	0.38	0.704	-.0005287	.0007832
az_xway	-78.90631	35.80668	-2.20	0.028	-149.117	-8.695671
apt1_dis	23.31013	9.606509	2.43	0.015	4.473451	42.14682
apt1_dissq	-.0012127	.0036553	-0.33	0.740	-.00838	.0059547
apt1_az	-38.92631	16.63103	-2.34	0.019	-71.53686	-6.315758
cult1_dis	11.64769	10.61525	1.10	0.273	-9.166959	32.46233
cult1_dissq	-.0120363	.0038265	-3.15	0.002	-.0195393	-.0045333
cult1_az	-28.31605	17.7651	-1.59	0.111	-63.15031	6.518211
govt1_dis	-17.24812	11.06807	-1.56	0.119	-38.95067	4.454436
govt1_dissq	.0017776	.0016266	1.09	0.275	-.0014118	.004967
govt1_az	47.75463	35.833	1.33	0.183	-22.50761	118.0169
hotel1_dis	-7.436645	6.798456	-1.09	0.274	-20.76723	5.893937
hotel1_dissq	.0002283	.0005909	0.39	0.699	-.0009303	.001387
hotle1_az	2.025723	28.77057	0.07	0.944	-54.38833	58.43978
off1_dis	35.63853	10.51694	3.39	0.001	15.01665	56.26041
off1_dissq	-.0120424	.00306	-3.94	0.000	-.0180426	-.0060422
off1_az	23.05004	22.68614	1.02	0.310	-21.4335	67.53359
hops1_dis	-6.642452	6.844652	-0.97	0.332	-20.06362	6.778713
hops1_dissq	.0007502	.000689	1.09	0.276	-.0006008	.0021012
hosp1_az	37.42193	26.8341	1.39	0.163	-15.19505	90.03892
ind1_dis	2.344985	16.9894	0.14	0.890	-30.96826	35.65823
ind1_dissq	-.0002645	.0078217	-0.03	0.973	-.0156015	.0150724
ind1_az	11.75505	21.55828	0.55	0.586	-30.51697	54.02707
sch1_dis	-5.984742	12.55603	-0.48	0.634	-30.60492	18.63544
sch1_dissq	.005131	.0035727	1.44	0.151	-.0018744	.0121365
sch1_az	41.24215	23.64258	1.74	0.081	-5.11681	87.6011
r_net1	-21.18459	10.30496	-2.06	0.040	-41.39082	-.9783721
r_net1sq	.004959	.0014555	3.41	0.001	.002105	.007813
average	-8.020705	8.655439	-0.93	0.354	-24.99251	8.951096
avesq	.0035637	.0033811	1.05	0.292	-.0030661	.0101934
trafficnoi	-14235.89	3069.986	-4.64	0.000	-20255.59	-8216.182
visibility	10540.46	14108.55	0.75	0.455	-17123.95	38204.87
seg_tnodes	9067.702	5763.167	1.57	0.116	-2232.86	20368.26
seg_unodea	103.525	46.03819	2.25	0.025	13.25217	193.7978
density	6778.281	4639.592	1.46	0.144	-2319.148	15875.71
densitysq	-132.8542	86.64135	-1.53	0.125	-302.7427	37.03436
nonresmix	914.8613	1327.459	0.69	0.491	-1688.053	3517.776

```
 nonresmixsq |  -42.60324   32.86074    -1.30   0.195    -107.0374    21.83093
       _cons |   194771.3    3431704     0.06   0.955     -6534200     6923742
------------------------------------------------------------------------------

. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  2756) =   12.28
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  2756) =    0.58
            Prob > F =    0.5575

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2,  2756) =    1.23
            Prob > F =    0.2924

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2,  2756) =    2.79
            Prob > F =    0.0614
```

Appendix B: East less than 1400 feet

Regression with robust standard errors

Number of obs = 1482
F(64, 1417) = 43.47
Prob > F = 0.0000
R-squared = 0.7157
Root MSE = 1.0e+05

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-30171.34	8090.528	-3.73	0.000	-46042.04	-14300.64
trendsq	4789.745	1113.601	4.30	0.000	2605.261	6974.23
trendcu	-116.7917	45.23622	-2.58	0.010	-205.5289	-28.05459
sqfttotliv	35.8728	54.72099	0.66	0.512	-71.47005	143.2156
sqftsq	.0182881	.0158445	1.15	0.249	-.0127931	.0493692
bedrooms	19106.09	16921.26	1.13	0.259	-14087.32	52299.5
sqftxbedroom	-13.37516	8.683639	-1.54	0.124	-30.40933	3.659014
bathrooms	1985.722	7504.52	0.26	0.791	-12735.44	16706.88
age	-2856.99	617.0861	-4.63	0.000	-4067.49	-1646.489
agesq	26.27547	5.314602	4.94	0.000	15.85014	36.70081
condition	367.8919	4166.911	0.09	0.930	-7806.086	8541.87
sqftlot	1.718796	1.882299	0.91	0.361	-1.973596	5.411189
sqftlotsq	.0000139	.0000413	0.34	0.736	-.0000671	.0000949
noviewd	-21122.63	10686.72	-1.98	0.048	-42086.12	-159.128
wfntlocati	91790.75	17141.09	5.36	0.000	58166.11	125415.4
itbs_read	2719.375	772.0428	3.52	0.000	1204.905	4233.844
dis_bofa	-759.1304	138.7239	-5.47	0.000	-1031.257	-487.004
dis_bofasq	.0081252	.0015425	5.27	0.000	.0050993	.011151
az_bofa	-181162.4	48717.83	-3.72	0.000	-276729.3	-85595.62
dis_mic	-352.6245	81.04426	-4.35	0.000	-511.6041	-193.6449
dis_micsq	.0031737	.001173	2.71	0.007	.0008726	.0054747
az_mic	-15743.49	4231.313	-3.72	0.000	-24043.8	-7443.181
dis_xway	-23.69989	15.63403	-1.52	0.130	-54.36822	6.968429
dis_xwaysq	-.0023339	.0017641	-1.32	0.186	-.0057945	.0011267
az_xway	141.7041	178.1845	0.80	0.427	-207.8296	491.2378
apt1_dis	-16.77568	25.7643	-0.65	0.515	-67.31596	33.76459
apt1_dissq	.0087136	.0137847	0.63	0.527	-.018327	.0357543
apt1_az	36.53441	32.06592	1.14	0.255	-26.36737	99.43618
cult1_dis	27.41483	20.39443	1.34	0.179	-12.59169	67.42134
cult1_dissq	-.0020867	.0090407	-0.23	0.817	-.0198213	.0156478
cult1_az	-22.06377	38.29826	-0.58	0.565	-97.19115	53.06361
govt1_dis	27.63131	11.50073	2.40	0.016	5.07102	50.19161
govt1_dissq	-.0058239	.0016096	-3.62	0.000	-.0089813	-.0026666
govt1_az	-25.05324	39.03569	-0.64	0.521	-101.6272	51.52071
hotel1_dis	15.2659	19.88555	0.77	0.443	-23.74238	54.27418
hotel1_dissq	-.0003352	.0011795	-0.28	0.776	-.0026489	.0019786
hotle1_az	70.29525	39.59621	1.78	0.076	-7.37824	147.9687
off1_dis	9.749704	19.13366	0.51	0.610	-27.78363	47.28304
off1_dissq	-.0018929	.0043243	-0.44	0.662	-.0103756	.0065898
off1_az	14.142	35.78004	0.40	0.693	-56.04554	84.32954
hops1_dis	-38.02657	23.37702	-1.63	0.104	-83.88385	7.830707
hops1_dissq	.000848	.001236	0.69	0.493	-.0015766	.0032726
hosp1_az	-8.09842	37.41614	-0.22	0.829	-81.49539	65.29855
ind1_dis	86.08108	33.75416	2.55	0.011	19.86759	152.2946
ind1_dissq	-.0410115	.0224291	-1.83	0.068	-.0850092	.0029862
ind1_az	-12.25088	30.43796	-0.40	0.687	-71.95918	47.45742
sch1_dis	-47.97438	25.05169	-1.92	0.056	-97.11676	1.168001
sch1_dissq	.0146738	.0067459	2.18	0.030	.0014409	.0279068
sch1_az	-23.68826	56.25197	-0.42	0.674	-134.0343	86.65782
r_net1	27.89182	14.03466	1.99	0.047	.3608686	55.42277
r_net1sq	-.0032311	.0022683	-1.42	0.155	-.0076808	.0012186
areu1_dis	-45.28029	41.52657	-1.09	0.276	-126.7405	36.17987
areu1_dissq	.0038731	.0221136	0.18	0.861	-.0395059	.0472521
areu1_az	12.14189	30.26399	0.40	0.688	-47.22515	71.50893
average	17.26283	25.24435	0.68	0.494	-32.25749	66.78316
avesq	.0016188	.0053676	0.30	0.763	-.0089104	.012148
trafficnoi	-7715.039	5016.586	-1.54	0.124	-17555.77	2125.694
visibility	10788.89	11168.44	0.97	0.334	-11119.56	32697.35
seg_tnodes	-439091.5	352746.6	-1.24	0.213	-1131053	252870.2
seg_unodea	2546.309	13914.23	0.18	0.855	-24748.4	29841.02

```
     density |   -4353.94   14104.98    -0.31   0.758    -32022.83    23314.95
   densitysq |  -91.96953   418.8205    -0.22   0.826    -913.5444    729.6054
   nonresmix |  -19859.62   8594.791    -2.31   0.021     -36719.5   -2999.738
 nonresmixsq |   558.4602   223.9126     2.49   0.013     119.2244     997.696
       _cons |   6.86e+07   1.58e+07     4.35   0.000     3.77e+07    9.95e+07
------------------------------------------------------------------------------

. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  1417) =    3.51
            Prob > F =    0.0303

. test areu1_dis areu1_dissq

 ( 1)  areu1_dis = 0
 ( 2)  areu1_dissq = 0

       F(  2,  1417) =    2.24
            Prob > F =    0.1065

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  1417) =    2.25
            Prob > F =    0.1063

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2,  1417) =    2.57
            Prob > F =    0.0772

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2,  1417) =    3.57
            Prob > F =    0.0285
```

Appendix B: East greater than 1400 feet

Regression with robust standard errors

Number of obs	= 5258
F(61, 5196)	= 207.49
Prob > F	= 0.0000
R-squared	= 0.7392
Root MSE	= 68302

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-3997.586	2894.483	-1.38	0.167	-9671.991	1676.819
trendsq	1487.927	448.9424	3.31	0.001	607.8116	2368.043
trendcu	-21.08333	20.09566	-1.05	0.294	-60.47928	18.31262
sqfttotliv	58.71207	21.16098	2.77	0.006	17.22765	100.1965
sqftsq	.0085031	.007866	1.08	0.280	-.0069176	.0239238
bedrooms	16356.8	10584.89	1.55	0.122	-4394.045	37107.64
sqftxbedroom	-8.823091	5.174493	-1.71	0.088	-18.96727	1.321092
bathrooms	22.91207	2806.795	0.01	0.993	-5479.586	5525.41
age	-3633.103	321.1023	-11.31	0.000	-4262.599	-3003.608
agesq	31.55954	3.790489	8.33	0.000	24.12859	38.99049
condition	4098.096	2969.15	1.38	0.168	-1722.687	9918.878
sqftlot	1.428371	.2749744	5.19	0.000	.8893057	1.967437
sqftlotsq	-5.29e-06	1.56e-06	-3.40	0.001	-8.34e-06	-2.25e-06
noviewd	-47228.34	8209.478	-5.75	0.000	-63322.37	-31134.31
wfntlocati	59460.88	13947.43	4.26	0.000	32118.06	86803.71
itbs_read	1064.289	304.8477	3.49	0.000	466.6595	1661.919
dis_bofa	-68.87364	72.34586	-0.95	0.341	-210.7019	72.95467
dis_bofasq	.0004011	.0007643	0.52	0.600	-.0010973	.0018994
az_bofa	2826.59	25389.98	0.11	0.911	-46948.45	52601.63
dis_mic	-27.39163	30.13848	-0.91	0.363	-86.47573	31.69247
dis_micsq	.000937	.0002649	3.54	0.000	.0004177	.0014564
az_mic	5702.95	1498.208	3.81	0.000	2765.831	8640.068
dis_xway	4.907505	4.063746	1.21	0.227	-3.059147	12.87416
dis_xwaysq	.0002165	.000359	0.60	0.546	-.0004873	.0009204
az_xway	87.40035	31.54262	2.77	0.006	25.56355	149.2372
apt1_dis	-17.89567	5.626187	-3.18	0.001	-28.92536	-6.865972
apt1_dissq	.0018763	.0012913	1.45	0.146	-.0006553	.0044079
apt1_az	20.23078	13.66264	1.48	0.139	-6.553749	47.0153
cult1_dis	-1.010204	6.620838	-0.15	0.879	-13.98983	11.96942
cult1_dissq	.0014458	.0021192	0.68	0.495	-.0027087	.0056003
cult1_az	13.46884	13.04932	1.03	0.302	-12.11332	39.051
govt1_dis	6.761411	4.689182	1.44	0.149	-2.431358	15.95418
govt1_dissq	-.0001182	.000457	-0.26	0.796	-.0010141	.0007776
govt1_az	51.53244	22.47444	2.29	0.022	7.473076	95.5918
hotel1_dis	-3.709237	8.127639	-0.46	0.648	-19.64283	12.22435
hotel1_dissq	.0000942	.0005527	0.17	0.865	-.0009894	.0011777
hotle1_az	6.390265	24.83196	0.26	0.797	-42.29082	55.07135
off1_dis	-6.60049	4.323314	-1.53	0.127	-15.076	1.875024
off1_dissq	.0001363	.0008356	0.16	0.870	-.0015019	.0017744
off1_az	6.988944	16.69257	0.42	0.675	-25.73551	39.7134
hops1_dis	22.9599	6.700922	3.43	0.001	9.823279	36.09653
hops1_dissq	-.0012676	.0004035	-3.14	0.002	-.0020585	-.0004766
hosp1_az	-35.50969	21.87533	-1.62	0.105	-78.39454	7.375168
ind1_dis	-7.745782	7.383291	-1.05	0.294	-22.22014	6.728575
ind1_dissq	-.0021704	.003	-0.72	0.469	-.0080516	.0037109
ind1_az	43.64477	10.37722	4.21	0.000	23.30106	63.98848
sch1_dis	-7.279354	6.675712	-1.09	0.276	-20.36656	5.807849
sch1_dissq	.0033445	.0019691	1.70	0.089	-.0005157	.0072047
sch1_az	-7.849239	12.89792	-0.61	0.543	-33.13458	17.4361
r_net1	20.25559	5.648257	3.59	0.000	9.182632	31.32855
r_net1sq	-.0015867	.0005732	-2.77	0.006	-.0027103	-.000463
average	-7.482007	4.159825	-1.80	0.072	-15.63701	.6730009
avesq	.0006809	.0006965	0.98	0.328	-.0006846	.0020463
trafficnoi	-7360.68	2280.511	-3.23	0.001	-11831.44	-2889.918
visibility	23331.73	24304.34	0.96	0.337	-24315	70978.45
seg_tnodes	-71394.56	117204.6	-0.61	0.542	-301164.9	158375.7
seg_unodea	-5831.858	5761.804	-1.01	0.312	-17127.42	5463.7
density	-12204.1	5702.472	-2.14	0.032	-23383.34	-1024.857
densitysq	378.1409	178.2221	2.12	0.034	28.75055	727.5313
nonresmix	-4299.047	3106.847	-1.38	0.166	-10389.77	1791.679

```
 nonresmixsq |   115.1591   84.80665     1.36   0.175    -51.09764    281.4158
       _cons |    1186669    8267520     0.14   0.886    -1.50e+07    1.74e+07
------------------------------------------------------------------------------

. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  5196) =   11.46
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  5196) =    6.53
            Prob > F =    0.0015

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2,  5196) =    2.29
            Prob > F =    0.1011

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2,  5196) =    0.96
            Prob > F =    0.3835
```

APPENDIX C: FULL MODEL

REGRESSION RESULTS

Appendix C: West less than 1400 feet

Regression with robust standard errors

Number of obs	=	16266
F(68, 16197)	=	401.58
Prob > F	=	0.0000
R-squared	=	0.7216
Root MSE	=	55286

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-12621.91	1314.335	-9.60	0.000	-15198.15	-10045.67
trendsq	2655.041	185.5763	14.31	0.000	2291.291	3018.791
trendcu	-65.87406	7.785069	-8.46	0.000	-81.13366	-50.61447
sqfttotliv	43.01357	6.248084	6.88	0.000	30.76664	55.26051
sqftsq	.0033607	.0031653	1.06	0.288	-.0028435	.009565
bedrooms	3219.131	4161.393	0.77	0.439	-4937.659	11375.92
sqftxbedroom	-.8554859	2.442051	-0.35	0.726	-5.642176	3.931204
bathrooms	6854.128	1186.957	5.77	0.000	4527.561	9180.696
age	-735.6377	115.7069	-6.36	0.000	-962.436	-508.8395
agesq	5.141939	.9054712	5.68	0.000	3.367115	6.916762
condition	10779.74	833.3351	12.94	0.000	9146.313	12413.17
sqftlot	5.014156	.6666141	7.52	0.000	3.707519	6.320793
sqftlotsq	.0001186	.0000352	3.37	0.001	.0000495	.0001876
noviewd	-41355.54	2596.349	-15.93	0.000	-46444.68	-36266.41
wfntlocati	77639.52	19892.7	3.90	0.000	38647.63	116631.4
itbs_read	-843.7878	81.99819	-10.29	0.000	-1004.513	-683.0623
dis_bofa	-36.91799	5.613159	-6.58	0.000	-47.9204	-25.91558
dis_bofasq	-.0005217	.0000844	-6.18	0.000	-.0006872	-.0003562
az_bofa	11139.98	1920.809	5.80	0.000	7374.981	14904.98
dis_mic	-55.97677	11.52455	-4.86	0.000	-78.56615	-33.38739
dis_micsq	.0008868	.0001162	7.63	0.000	.000659	.0011146
az_mic	47468.46	5436.434	8.73	0.000	36812.45	58124.47
dis_xway	-1.383874	.6713088	-2.06	0.039	-2.699714	-.0680348
dis_xwaysq	.0000405	.0000768	0.53	0.598	-.0001101	.000191
az_xway	-92.66658	14.33468	-6.46	0.000	-120.7641	-64.56902
apt1_dis	27.22832	4.617438	5.90	0.000	18.17763	36.27901
apt1_dissq	-.0087413	.0034983	-2.50	0.012	-.0155984	-.0018842
apt1_az	2.35461	4.638062	0.51	0.612	-6.736504	11.44572
cult1_dis	-3.629526	3.956271	-0.92	0.359	-11.38425	4.125201
cult1_dissq	.0032583	.0022503	1.45	0.148	-.0011526	.0076692
cult1_az	.0450496	4.308683	0.01	0.992	-8.400446	8.490545
govt1_dis	-10.80298	1.571696	-6.87	0.000	-13.88367	-7.72228
govt1_dissq	.0028887	.0002892	9.99	0.000	.0023219	.0034555
govt1_az	-31.62043	5.385436	-5.87	0.000	-42.17648	-21.06438
hotel1_dis	5.79006	1.315472	4.40	0.000	3.21159	8.36853
hotel1_dissq	-.0002902	.0001466	-1.98	0.048	-.0005775	-2.82e-06
hotle1_az	34.50163	6.551634	5.27	0.000	21.6597	47.34356
off1_dis	-10.13448	5.068211	-2.00	0.046	-20.06873	-.2002295
off1_dissq	.0058296	.0023477	2.48	0.013	.0012277	.0104314
off1_az	-4.589964	5.382373	-0.85	0.394	-15.14001	5.960082
hops1_dis	7.146047	1.805753	3.96	0.000	3.606571	10.68552
hops1_dissq	-.0009441	.0002018	-4.68	0.000	-.0013396	-.0005486
hosp1_az	-32.62297	5.201704	-6.27	0.000	-42.81888	-22.42705
ind1_dis	20.59829	5.431301	3.79	0.000	9.95234	31.24424
ind1_dissq	-.0112664	.0033887	-3.32	0.001	-.0179087	-.0046241
ind1_az	-4.682359	5.162133	-0.91	0.364	-14.80071	5.435991
sch1_dis	-1.802284	3.13263	-0.58	0.565	-7.942585	4.338017
sch1_dissq	.0015481	.0012387	1.25	0.211	-.0008797	.003976
sch1_az	10.96125	4.647424	2.36	0.018	1.851791	20.07072
r_net1	-4.823383	7.254046	-0.66	0.506	-19.04212	9.395349
r_net1sq	.00556	.0021606	2.57	0.010	.001325	.0097951
areu1_dis	12.75976	11.55897	1.10	0.270	-9.897105	35.41663
areu1_dissq	-.0218542	.0051941	-4.21	0.000	-.0320352	-.0116733
areu1_az	17.7359	5.03373	3.52	0.000	7.869232	27.60256
average	-3.052841	3.569389	-0.86	0.392	-10.04924	3.943556
avesq	-.0000402	.0017446	-0.02	0.982	-.0034597	.0033794
trafficnoi	-13620.35	734.2196	-18.55	0.000	-15059.5	-12181.2
visibility	-4722.531	1434.203	-3.29	0.001	-7533.728	-1911.334
seg_tnodes	-2679.165	1790.422	-1.50	0.135	-6188.591	830.2601
segtnodXrnet	-4.407904	1.593028	-2.77	0.006	-7.530415	-1.285393

```
  seg_unodea |   68.73778   9.199438     7.47   0.000     50.70587     86.7697
segunodaXr~t |  -.0130215   .0086522    -1.50   0.132    -.0299808    .0039378
     density |  -3916.571   644.2729    -6.08   0.000    -5179.417   -2653.725
   densitysq |   54.11164   9.049936     5.98   0.000     36.37277    71.85052
densityXar~1 |   1.093231   .2488908     4.39   0.000     .6053773    1.581084
   nonresmix |   1350.775   329.1975     4.10   0.000     705.5111    1996.038
 nonresmixsq |  -8.762306   6.980861    -1.26   0.209    -22.44556    4.920952
nonresXareu1 |  -.5930198   .2026094    -2.93   0.003    -.9901567   -.1958829
       _cons |   -4928935     878495    -5.61   0.000     -6650882    -3206988
------------------------------------------------------------------------------

.
. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2, 16197) =    7.09
            Prob > F =    0.0008


. test areu1_dis areu1_dissq

 ( 1)  areu1_dis = 0
 ( 2)  areu1_dissq = 0

       F(  2, 16197) =   15.79
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2, 16197) =    4.14
            Prob > F =    0.0160

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2, 16197) =   18.60
            Prob > F =    0.0000

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2, 16197) =   18.86
            Prob > F =    0.0000
```

Appendix C: West greater than 1400 feet

Regression with robust standard errors

Number of obs	= 2818
F(64, 2753)	= 99.03
Prob > F	= 0.0000
R-squared	= 0.7733
Root MSE	= 72264

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-13849.98	3790.718	-3.65	0.000	-21282.92	-6417.043
trendsq	3015.858	535.247	5.63	0.000	1966.332	4065.384
trendcu	-79.1952	22.18063	-3.57	0.000	-122.6876	-35.70285
sqfttotliv	28.64973	11.04136	2.59	0.010	6.999536	50.29993
sqftsq	.0048499	.0030226	1.60	0.109	-.0010769	.0107768
bedrooms	-749.2847	6125.739	-0.12	0.903	-12760.79	11262.22
sqftxbedroom	-.0255834	3.182908	-0.01	0.994	-6.266712	6.215545
bathrooms	13213.99	3291.499	4.01	0.000	6759.933	19668.05
age	-2329.278	363.2982	-6.41	0.000	-3041.642	-1616.913
agesq	19.09309	3.010341	6.34	0.000	13.19034	24.99585
condition	12174.81	2624.976	4.64	0.000	7027.688	17321.93
sqftlot	4.153433	1.210487	3.43	0.001	1.779879	6.526987
sqftlotsq	.0001074	.0000356	3.01	0.003	.0000375	.0001772
noviewd	-57757.46	5196.907	-11.11	0.000	-67947.69	-47567.22
wfntlocati	56643.83	17109.16	3.31	0.001	23095.76	90191.91
itbs_read	1834.35	432.8661	4.24	0.000	985.5747	2683.125
dis_bofa	-136.8999	44.14193	-3.10	0.002	-223.4545	-50.34524
dis_bofasq	.0008174	.0005136	1.59	0.112	-.0001897	.0018245
az_bofa	6029.08	7559.531	0.80	0.425	-8793.844	20852
dis_mic	-141.2994	70.60106	-2.00	0.045	-279.7358	-2.86301
dis_micsq	.0016781	.0008203	2.05	0.041	.0000697	.0032866
az_mic	64061.19	34448.19	1.86	0.063	-3485.711	131608.1
dis_xway	-17.63692	4.156873	-4.24	0.000	-25.78783	-9.486019
dis_xwaysq	.0003395	.0003404	1.00	0.319	-.000328	.001007
az_xway	-61.49662	36.05875	-1.71	0.088	-132.2016	9.208312
apt1_dis	33.04124	9.736231	3.39	0.001	13.95018	52.13229
apt1_dissq	-.0038983	.0037004	-1.05	0.292	-.0111542	.0033576
apt1_az	-45.90471	16.42351	-2.80	0.005	-78.10835	-13.70107
cult1_dis	6.459544	10.61491	0.61	0.543	-14.35444	27.27353
cult1_dissq	-.010612	.0038548	-2.75	0.006	-.0181707	-.0030534
cult1_az	-31.90876	17.80153	-1.79	0.073	-66.81445	2.996935
govt1_dis	-12.08537	11.07841	-1.09	0.275	-33.80821	9.637473
govt1_dissq	.0013292	.0016363	0.81	0.417	-.0018793	.0045377
govt1_az	59.99689	35.69702	1.68	0.093	-9.99875	129.9925
hotel1_dis	-9.146663	6.76581	-1.35	0.177	-22.41324	4.119913
hotel1_dissq	.0007861	.0005731	1.37	0.170	-.0003376	.0019098
hotle1_az	3.710947	28.61227	0.13	0.897	-52.39274	59.81464
off1_dis	35.14015	11.01916	3.19	0.001	13.53351	56.7468
off1_dissq	-.0108098	.0032795	-3.30	0.001	-.0172404	-.0043793
off1_az	28.30173	22.38388	1.26	0.206	-15.58916	72.19262
hops1_dis	-9.904204	7.25374	-1.37	0.172	-24.12753	4.319119
hops1_dissq	.0009668	.0006914	1.40	0.162	-.0003889	.0023226
hosp1_az	37.52984	26.55635	1.41	0.158	-14.54254	89.60222
ind1_dis	3.154301	17.05516	0.18	0.853	-30.2879	36.5965
ind1_dissq	-.0025839	.0079055	-0.33	0.744	-.0180851	.0129174
ind1_az	4.441816	21.21474	0.21	0.834	-37.15659	46.04022
sch1_dis	-5.73106	12.46567	-0.46	0.646	-30.17408	18.71196
sch1_dissq	.0056802	.0034932	1.63	0.104	-.0011693	.0125298
sch1_az	40.62457	23.70516	1.71	0.087	-5.857136	87.10627
r_net1	-66.29133	16.31726	-4.06	0.000	-98.28664	-34.29603
r_net1sq	.0082938	.0018288	4.54	0.000	.0047079	.0118796
average	-4.196944	8.684553	-0.48	0.629	-21.22584	12.83195
avesq	.0019515	.0033598	0.58	0.561	-.0046365	.0085394
trafficnoi	-15576.13	3071.17	-5.07	0.000	-21598.16	-9554.104
visibility	13946.75	14163.11	0.98	0.325	-13824.65	41718.14
seg_tnodes	-10479.46	11801.2	-0.89	0.375	-33619.56	12660.65
segtnodXrnet	6.430195	4.204929	1.53	0.126	-1.81494	14.67533
seg_unodea	-900.0013	220.975	-4.07	0.000	-1333.295	-466.7077
segunodaXr~t	.5464004	.1217201	4.49	0.000	.3077285	.7850724
density	4747.846	4839.233	0.98	0.327	-4741.049	14236.74

```
   densitysq |   -111.008   86.80361    -1.28   0.201    -281.2148    59.19874
densityXar~1 |   .3465579   .4052266     0.86   0.393    -.4480209    1.141137
   nonresmix |   2061.642   1313.708     1.57   0.117    -514.3111    4637.596
 nonresmixsq |  -63.91096    32.6941    -1.95   0.051    -128.0184    .1964851
       _cons |   -1786499    3469377    -0.51   0.607     -8589344     5016347
------------------------------------------------------------------------------

. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  2753) =   10.36
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  2753) =    0.19
            Prob > F =    0.8239

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2,  2753) =    1.74
            Prob > F =    0.1758

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2,  2753) =    2.51
            Prob > F =    0.0815
```

Appendix C: East less than 1400 feet

Regression with robust standard errors

Number of obs = 1482
F(67, 1413) = .
Prob > F = .
R-squared = 0.7169
Root MSE = 1.0e+05

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-29523.34	8188.667	-3.61	0.000	-45586.59	-13460.09
trendsq	4697.712	1123.907	4.18	0.000	2493.007	6902.418
trendcu	-113.1766	45.58032	-2.48	0.013	-202.589	-23.76423
sqfttotliv	34.83926	54.61272	0.64	0.524	-72.29146	141.97
sqftsq	.0188049	.0158303	1.19	0.235	-.0122485	.0498583
bedrooms	20316.21	17062.77	1.19	0.234	-13154.87	53787.29
sqftxbedroom	-13.81431	8.741395	-1.58	0.114	-30.96182	3.333198
bathrooms	2005.586	7546.468	0.27	0.790	-12797.9	16809.07
age	-2801.636	623.2479	-4.50	0.000	-4024.227	-1579.045
agesq	25.6264	5.405342	4.74	0.000	15.02304	36.22976
condition	800.1304	4204.131	0.19	0.849	-7446.879	9047.14
sqftlot	1.376908	1.860257	0.74	0.459	-2.272255	5.026072
sqftlotsq	.00002	.000041	0.49	0.626	-.0000604	.0001003
noviewd	-21458.07	10757.2	-1.99	0.046	-42559.87	-356.2633
wfntlocati	89539.61	16795.39	5.33	0.000	56593.03	122486.2
itbs_read	2899.992	794.2088	3.65	0.000	1342.037	4457.947
dis_bofa	-719.5777	139.1082	-5.17	0.000	-992.4585	-446.6969
dis_bofasq	.0076677	.0015516	4.94	0.000	.004624	.0107113
az_bofa	-164648.7	49572.09	-3.32	0.001	-261891.5	-67405.93
dis_mic	-329.6648	81.95146	-4.02	0.000	-490.4244	-168.9052
dis_micsq	.0031546	.0011821	2.67	0.008	.0008357	.0054734
az_mic	-14525.51	4245.261	-3.42	0.001	-22853.2	-6197.816
dis_xway	-27.07269	15.52416	-1.74	0.081	-57.52556	3.380189
dis_xwaysq	-.0017968	.0017539	-1.02	0.306	-.0052372	.0016437
az_xway	44.91891	181.4596	0.25	0.805	-311.0402	400.878
apt1_dis	-14.89538	26.59468	-0.56	0.576	-67.06469	37.27392
apt1_dissq	.0074089	.0139933	0.53	0.597	-.0200409	.0348587
apt1_az	25.93445	32.4519	0.80	0.424	-37.72463	89.59353
cult1_dis	25.03016	20.51839	1.22	0.223	-15.21963	65.27994
cult1_dissq	-.0025514	.0091665	-0.28	0.781	-.0205328	.0154299
cult1_az	-22.43819	38.94612	-0.58	0.565	-98.83663	53.96025
govt1_dis	26.56915	11.49657	2.31	0.021	4.016965	49.12133
govt1_dissq	-.0055814	.0015482	-3.61	0.000	-.0086184	-.0025443
govt1_az	-37.40828	39.75319	-0.94	0.347	-115.3899	40.57334
hotel1_dis	16.509	20.10577	0.82	0.412	-22.93137	55.94938
hotel1_dissq	-.0002884	.0012023	-0.24	0.810	-.0026469	.0020701
hotle1_az	53.67059	39.52713	1.36	0.175	-23.86757	131.2087
off1_dis	13.44511	19.57205	0.69	0.492	-24.94829	51.83851
off1_dissq	-.0020446	.0046906	-0.44	0.663	-.0112458	.0071567
off1_az	7.196954	36.09647	0.20	0.842	-63.61147	78.00538
hops1_dis	-41.11878	24.03069	-1.71	0.087	-88.25844	6.020879
hops1_dissq	.000965	.0012526	0.77	0.441	-.0014922	.0034221
hosp1_az	1.924766	38.24228	0.05	0.960	-73.09299	76.94252
ind1_dis	93.29854	35.71518	2.61	0.009	23.23806	163.359
ind1_dissq	-.0469424	.0237394	-1.98	0.048	-.0935106	-.0003741
ind1_az	-9.369901	30.61619	-0.31	0.760	-69.42798	50.68817
sch1_dis	-49.94893	25.56986	-1.95	0.051	-100.1079	.2100407
sch1_dissq	.0146429	.0067852	2.16	0.031	.0013328	.0279531
sch1_az	-20.68234	55.88492	-0.37	0.711	-130.3087	88.94399
r_net1	-107.893	117.3798	-0.92	0.358	-338.1505	122.3645
r_net1sq	.000551	.0025973	0.21	0.832	-.004544	.005646
areu1_dis	-32.37104	55.91264	-0.58	0.563	-142.0517	77.30966
areu1_dissq	-.0055531	.023072	-0.24	0.810	-.0508121	.0397059
areu1_az	13.11576	31.03802	0.42	0.673	-47.76978	74.00131
average	19.98465	25.28756	0.79	0.429	-29.62054	69.58984
avesq	-.0006588	.0053388	-0.12	0.902	-.0111317	.0098141
trafficnoi	-7115.018	5050.626	-1.41	0.159	-17022.55	2792.513
visibility	11011.16	11261.12	0.98	0.328	-11079.15	33101.47
seg_tnodes	-1099730	573117.5	-1.92	0.055	-2223982	24522.94
segtnodXrnet	99.31432	123.105	0.81	0.420	-142.1739	340.8025

```
   seg_unodea |   19010.84      21610     0.88   0.379    -23380.29    61401.96
segunodaXr~t |   2.782276   6.204009     0.45   0.654    -9.387782    14.95233
     density |   -6312.56   17961.19    -0.35   0.725    -41546.02     28920.9
   densitysq |  -427.3428   491.4589    -0.87   0.385     -1391.41    536.7249
densityXar~1 |   5.909322   4.184656     1.41   0.158    -2.299485    14.11813
   nonresmix |  -22969.87   9295.775    -2.47   0.014    -41204.87   -4734.865
 nonresmixsq |   834.5336   250.7406     3.33   0.001     342.6698    1326.397
nonresXareu1 |  -6.917123    3.12475    -2.21   0.027    -13.04677    -.787476
       _cons |   6.41e+07   1.58e+07     4.05   0.000     3.30e+07    9.51e+07
------------------------------------------------------------------------------

. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  1413) =    0.52
            Prob > F =    0.5976


. test areu1_dis areu1_dissq

 ( 1)  areu1_dis = 0
 ( 2)  areu1_dissq = 0

       F(  2,  1413) =    0.28
            Prob > F =    0.7581

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  1413) =    1.10
            Prob > F =    0.3333

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2,  1413) =    4.64
            Prob > F =    0.0098

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2,  1413) =    7.04
            Prob > F =    0.0009
```

Appendix C: East greater than 1400 feet

Regression with robust standard errors

Number of obs	= 5258
F(64, 5193)	= 201.75
Prob > F	= 0.0000
R-squared	= 0.7411
Root MSE	= 68072

saleprice	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
trend	-3933.232	2900.383	-1.36	0.175	-9619.204	1752.739
trendsq	1458.08	448.7143	3.25	0.001	578.4114	2337.749
trendcu	-19.16788	20.06294	-0.96	0.339	-58.49968	20.16392
sqfttotliv	55.78006	21.35071	2.61	0.009	13.92368	97.63643
sqftsq	.008203	.0078616	1.04	0.297	-.007209	.0236149
bedrooms	14421.81	10493.02	1.37	0.169	-6148.926	34992.55
sqftxbedroom	-8.021068	5.1432	-1.56	0.119	-18.1039	2.061768
bathrooms	488.9991	2725.341	0.18	0.858	-4853.816	5831.814
age	-3695.265	320.2423	-11.54	0.000	-4323.075	-3067.455
agesq	31.93297	3.736269	8.55	0.000	24.60831	39.25763
condition	4135.966	2956.077	1.40	0.162	-1659.189	9931.121
sqftlot	1.370801	.2723793	5.03	0.000	.8368233	1.90478
sqftlotsq	-4.63e-06	1.55e-06	-2.99	0.003	-7.66e-06	-1.60e-06
noviewd	-48138.47	8228.274	-5.85	0.000	-64269.35	-32007.59
wfntlocati	58004.51	13986.38	4.15	0.000	30585.32	85423.71
itbs_read	1183.032	327.6953	3.61	0.000	540.6113	1825.453
dis_bofa	-21.52058	76.07903	-0.28	0.777	-170.6675	127.6263
dis_bofasq	-.0001184	.0008038	-0.15	0.883	-.0016943	.0014575
az_bofa	15668.64	26648.17	0.59	0.557	-36572.99	67910.27
dis_mic	-17.1424	31.46696	-0.54	0.586	-78.83088	44.54608
dis_micsq	.0010874	.0002744	3.96	0.000	.0005495	.0016253
az_mic	6843.581	1515.273	4.52	0.000	3873.009	9814.153
dis_xway	3.582372	4.202284	0.85	0.394	-4.655873	11.82062
dis_xwaysq	.0006081	.0003831	1.59	0.113	-.000143	.0013592
az_xway	66.93523	33.54988	2.00	0.046	1.163347	132.7071
apt1_dis	-17.4824	5.639959	-3.10	0.002	-28.53909	-6.425703
apt1_dissq	.0019465	.0012943	1.50	0.133	-.0005909	.0044839
apt1_az	26.52465	13.9563	1.90	0.057	-.8355747	53.88488
cult1_dis	-1.213523	6.401372	-0.19	0.850	-13.76291	11.33586
cult1_dissq	.0017724	.002028	0.87	0.382	-.0022033	.0057481
cult1_az	13.43076	13.20603	1.02	0.309	-12.45863	39.32014
govt1_dis	8.794994	4.787775	1.84	0.066	-.59106	18.18105
govt1_dissq	-.0000608	.0004608	-0.13	0.895	-.0009641	.0008426
govt1_az	50.637	22.6867	2.23	0.026	6.161518	95.11248
hotel1_dis	-1.545785	8.551088	-0.18	0.857	-18.30952	15.21795
hotel1_dissq	-.0001102	.0005954	-0.19	0.853	-.0012775	.0010571
hotle1_az	14.59766	25.81963	0.57	0.572	-36.01968	65.21501
off1_dis	-2.855867	4.64831	-0.61	0.539	-11.96851	6.256777
off1_dissq	-.0004063	.0009363	-0.43	0.664	-.0022419	.0014292
off1_az	1.59079	17.5962	0.09	0.928	-32.90517	36.08675
hops1_dis	24.52594	7.046865	3.48	0.001	10.71112	38.34076
hops1_dissq	-.0011721	.0004132	-2.84	0.005	-.0019822	-.000362
hosp1_az	-38.25232	22.04896	-1.73	0.083	-81.47758	4.972926
ind1_dis	-6.696113	7.591901	-0.88	0.378	-21.57943	8.187208
ind1_dissq	-.0021341	.0029911	-0.71	0.476	-.0079978	.0037297
ind1_az	44.77515	10.29667	4.35	0.000	24.58934	64.96095
sch1_dis	-4.294592	7.018511	-0.61	0.541	-18.05383	9.464644
sch1_dissq	.0029288	.002128	1.38	0.169	-.001243	.0071006
sch1_az	-8.576794	12.7354	-0.67	0.501	-33.54354	16.38995
r_net1	97.80937	24.25367	4.03	0.000	50.26196	145.3568
r_net1sq	-.0024013	.000601	-4.00	0.000	-.0035795	-.0012231
average	-11.93764	4.230578	-2.82	0.005	-20.23135	-3.643924
avesq	.001239	.0007066	1.75	0.080	-.0001463	.0026242
trafficnoi	-7228.725	2360.536	-3.06	0.002	-11856.37	-2601.08
visibility	20604.52	24482.71	0.84	0.400	-27391.89	68600.93
seg_tnodes	392869.8	162815.2	2.41	0.016	73683.6	712056.1
segtnodXrnet	-74.09595	28.42849	-2.61	0.009	-129.8278	-18.36414
seg_unodea	-25758.51	7006.515	-3.68	0.000	-39494.23	-12022.79
segunodaXr~t	3.291367	1.59344	2.07	0.039	.1675546	6.415179
density	-6055.152	5877.317	-1.03	0.303	-17577.17	5466.863

```
   densitysq |   266.4179   187.1974     1.42   0.155    -100.5678    633.4036
densityXar~1 |  -.7893697   .1937313    -4.07   0.000    -1.169165   -.4095748
   nonresmix |  -3987.955   3162.645    -1.26   0.207    -10188.07     2212.16
 nonresmixsq |   116.8296   84.62263     1.38   0.167    -49.06633    282.7256
       _cons |   -3826954    8744531    -0.44   0.662    -2.10e+07    1.33e+07
------------------------------------------------------------------------------

.
. test r_net1 r_net1sq

 ( 1)  r_net1 = 0
 ( 2)  r_net1sq = 0

       F(  2,  5193) =   19.19
            Prob > F =    0.0000

. test average avesq

 ( 1)  average = 0
 ( 2)  avesq = 0

       F(  2,  5193) =   12.90
            Prob > F =    0.0000

. test density densitysq

 ( 1)  density = 0
 ( 2)  densitysq = 0

       F(  2,  5193) =    1.98
            Prob > F =    0.1380

. test nonresmix nonresmixsq

 ( 1)  nonresmix = 0
 ( 2)  nonresmixsq = 0

       F(  2,  5193) =    1.14
            Prob > F =    0.3214
```

APPENDIX D

MEASURES OF NEIGHBORHOOD LAYOUT

MEASURES OF NEIGHBORHOOD LAYOUT

Ongoing research is developing methods of measuring the functional aspects of street configuration to compare different types of neighborhood layouts (Frank and Engelke 2005; Song 2005; Victoria Transport Policy Institute 2005). In the field of urban design, space syntax, a method based on graph theory, has also been used to measure neighborhood layout and street connectivity (Bafna 2003; Hillier 1996; Penn 2003; Peponis et al. 1996).

Types of Measures

Space syntax is a new method of measuring design properties of street patterns. Space syntax uses graph theory to develop an index of the properties of neighborhood layouts. This method has been implemented in architectural design studies of floor plans and pedestrian flows and has also been used as a significant independent variable explaining the extent to which neighborhood design affects neighborhood crime rates (Hillier 1999). The following simple examples to illustrate how show different street layouts are reflected in different space syntax index values.

The gridiron pattern in Figure 6 has a higher index for both Mean Connectivity and Mean Integration, two space syntax measures that describe layout for the neighborhood as a whole. Space syntax also measures the relative integration for each street. Integration is the same for each street in the pure grid pattern

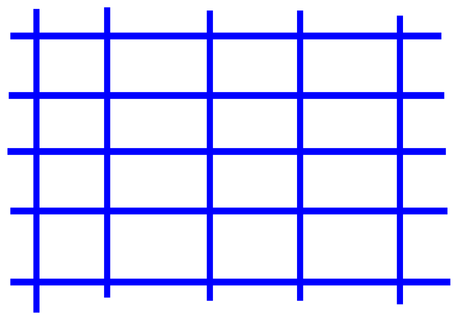

Figure 6: **Space Syntax**
Pure Grid Layout
Ten Streets
Mean Connectivity = 5
Mean Integration =2.75

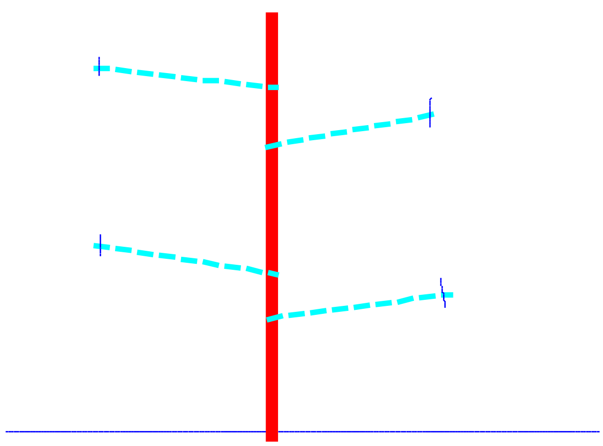

Figure 7: Space Syntax
Cul-de-sac Layout
Ten Streets
Mean Connectivity = 1.8
Mean Integration =1.05

in Figure 6, but this is not true for deviations from the gridiron pattern. Look at the cul-de-sac street pattern in Figure 7, as an example.

The street indicated by the heavy vertical solid line in the figure is more easily accessed from all other places and as such, it has greater relative integration than the other streets, depicted with thinner lines in the figure. In contrast, the space syntax algorithm indicates that all places are easily accessed in the pure grid system. The grid layout enhances convenience and enhances the portions of the basket of residential values arising from convenient access to retail and neighborhood services, but it reduces residential site privacy and may expose houses to more negative spillovers from retail sites. A cul-de-sac layout like Figure 7, on the other hand, naturally concentrates retail activities along the most accessible streets, increases residential site privacy, and increases protection from negative externalities from retail and other nonresidential activities. The cul-de-sac pattern, however, also reduces convenient access to retail sites in or near the neighborhood.

One problem with space syntax is that the integration and connectivity measures are complicated, are not intuitive, and somewhat controversial (Bafna 2003; Rotti 2004). An alternative and more intuitively appealing measure of street connectivity is the ratio of total street intersections to total street segments. Figure 8 provides an example of this method of measuring street connectivity. Generally, a ratio of 1.4 or greater indicates a well connected community (Victoria Transport Policy Institute 2005). Applying this method to the street patterns in Figures 6 and 7

yields the same qualitative conclusions as space syntax: The gridiron street pattern exhibits greater connectivity than the cul-de-sac pattern.

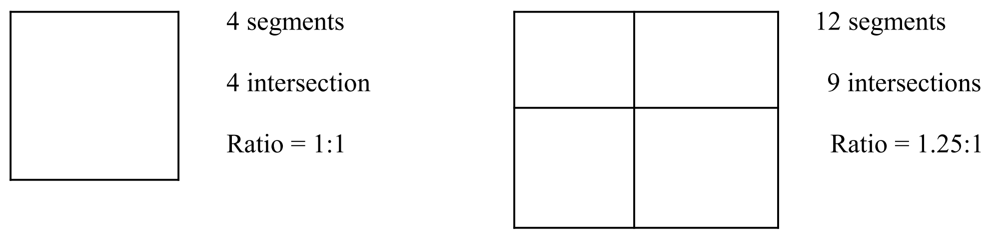

Figure 8: Street segment/intersection ratios

While the ratios of segments to intersections is straight-forward, space syntax is more complicated. But, space syntax does account for the fact that some individual can be curves, whereas the street segment/intersection ratio does not. Space syntax analysis of street (and other travelway) layout is based on the notion that street layouts are lines that compose a graph with inherent patterns of connectivity. The patterns of connectivity may be simple and direct or complex and indirect, or anywhere in between. Generally, street intersections are analogous to nodes in graph and streets themselves are analogous to the lines or "edges" of a graph. This analogy leads to use of some elements of graph theory to analyze layouts, with analysis leading to measures – or indexes – of ease of movement and access in the layout.

Method of Space Syntax

Before beginning a discussion of the method of space syntax, it must be mentioned that the method has been automated with the AXWOMAN extension to the ARCVIEW GIS software (Jiang, 1999).

The method of space syntax begins with representing a layout as an adjacency matrix. A simple grid layout and its accompanying adjacency matrix are presented in Figure 9 and Table 15.

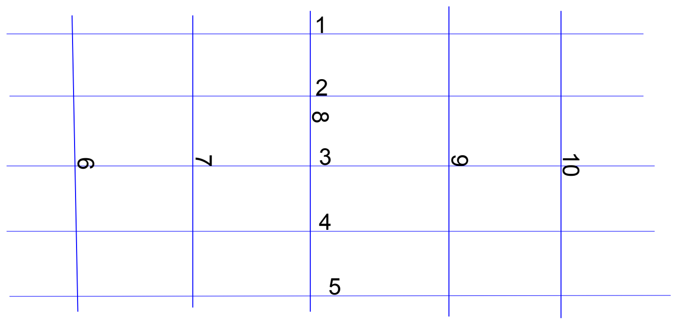

Figure 9: Simple Grid Layout

Table 15: Adjacency Matrix for Figure 9

	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	1	1	1	1	1
2	0	0	0	0	0	1	1	1	1	1
3	0	0	0	0	0	1	1	1	1	1
4	0	0	0	0	0	1	1	1	1	1
5	0	0	0	0	0	1	1	1	1	1
6	1	1	1	1	1	0	0	0	0	0
7	1	1	1	1	1	0	0	0	0	0
8	1	1	1	1	1	0	0	0	0	0
9	1	1	1	1	1	0	0	0	0	0
10	1	1	1	1	1	0	0	0	0	0

The adjacency matrix is used to produce several measures descriptive of the street (graph) layout. Connectivity, control, depth, and integration are the most frequently used measures in space syntax analysis.

<u>Connectivity</u>. The connectivity of a line (street) is the number of intersections – connections - with that street. Summing the rows of the adjacency matrix yields connectivity values for all streets:

Connectivity: $c_i = \sum a_{ij}$

Control. Control is a measure of the extent to which a given node (intersection) controls access to adjacent intersections. If, for example, the only way to get to Node B is to travel through Node A, then Node A controls Node B entirely. On the other hand, if there are other routes to Node B, then A's control is less. Control is inversely proportional to connectivity.

Control: $ctrl = \sum a_{ij}\ 1/c_{ij}$

Control for the i^{th} intersection is computed by multiplying its adjacency vector by the reciprocal of the connectivity values and summing the products. The products for directly connected nodes will equal the reciprocals of connectivity while they will be "0" for nodes that are not connected.

Note: Mean connectivity always equals 1. This property provides a useful check on overall matrix computations.

Depth. In graph theory "distance" is the same as the "depth" function in space syntax. Depth, or distance, between two nodes (intersections) is computed by counting the number of "edges" (street segments) along the shortest path between one node and another. Total depth for a given node is the sum of the depths to all other nodes in the graph. As depth increases, ease of access decreases. Note that neither graph theory nor space syntax use spatial distance measures, e.g. 10' 6", but use only the number of changes in direction as turns define line segments

Integration. The most widely used measure in space syntax id integration. Integration is based on the total depth from each node in a graph to all other nodes: the "mean depth":

Mean Depth: $MD = \sum d_{ij}/(n-1)$; using "n-1" to account for the "origin node."

Nodes with lower MD (mean depth) are more integrated or more generally accessible.

In practice, space syntax frequently converts mean depth to "Relative Asymmetry":

Relative Asymmetry: $RA = 2(MD - 1)/(n-2)$

Relative Asymmetry is said to normalize the mean depth relative to minimum and maximum mean depth in a graph.

Numerator = the difference between the observed mean depth and the minimum possible mean depth.

Denominator = the difference between the minimum and maximum.

Some practitioners claim that relative asymmetry is influences by the total number of nodes in a graph. "Real Relative Asymmetry" (RRA) is designed to accout for this type bias:

$RRA = RA/D$, where D is:

$D = 2(n(\log_2((n+2)/3)-1)+1/(n-1)(n-2)$

"The initial justification for the additional manipulation was opaque. Nor is it clear the RRA allows a more meaningful comparison … than RZ, or even MD." (Neiman, p.6 citing O'Brien)

None-the-less, RRA is the measure of integration used in available software, Axwoman and Axman. Actually, the integration measure used in both these packages is the reciprocal of RRA. The non-transformed RRA varies inversely with integration; as RRA increases. The reciprocal provides a more intuitive direct relationship.

Integration3. "Integration" analysis is applied across an entire map or layout. It is called a global measure. "Integration3" is often used as a local measure. Here,

depth does not account for the number of steps to all other nodes in the map, but only to those with three steps instead of all steps (local integration does not have to be “3”, that is the measure used in Axman and Axwoman). Pedestrian movement rates are more highly correlated to the local integration than global integration (Jiang 1999).

APPENDIX E

SENSITIVITY TO COLLINEARITY

SENSITIVITY TO COLLINEARITY

Collinearity is an unavoidable problem in hedonic analysis. The following tables show the effect of collinearity on the variables of interest: travel distance and straight line distance from residences to retail sites. The first table is a partial correlation matrix showing correlations between the two variables of interest and basic control variables. Note that most correlations are reasonably low. Of the 27 control variables in the table, only four exhibit correlations with the variables of interest greater than 0.4 or less than negative 0.4. Distance to apartments, offices, and industrial sites is correlated with distance to retail sites. Visibility of retail sites is, obviously, negatively correlated with distance. Five more show correlation at the 0.2 to 0.4 (or negative 0.2 to negative 0.4) level. Lot size tends to increase with distance from retail sites, average distance between retail sites tends to increase with distance to the nearest retail site, the ratio of cul-de-sacs is negatively correlated with distance to retail sites, and density decreases with distance to retail.

The second table shows the effect on the coefficients, robust standard errors, and t scores of the variables of interest as more independent variables are added to regressions. F tests of joint significance of the variables of interest and their squared terms are also included in the table. The first set of variables regressed are simply the structural variables, the general location variables (distance to the CBDs and expressways) and the variables of interest. The next three are the increasingly

complex sets reported in Appendices A, B, and C. The results are robust as additional variables are added. The robust standard errors on the variables of interest are stable throughout except for the travel distance and straight line distance variables (but not their squared terms) for the full set of variables added with Appendix C. here the robust standard errors increase slightly.

Table 16: Partial Correlation Matrix
Variables of Interest

	r_net1	areu1_dis
r_net1	1	
areu1_dis	0.8739	1
sqfttotliv	0.1279	0.0763
bedrooms	0.042	0.0199
bathrooms	0.0782	0.0382
age	-0.125	-0.0785
condition	-0.0379	-0.0246
sqftlot	0.2617	0.2032
noviewd	-0.1554	-0.099
wfntlocati	-0.0192	-0.0149
itbs_read	0.136	0.0981
dis_bofa	0.1476	0.098
dis_mic	-0.1237	-0.065
dis_xway	0.1445	0.1101
apt1_dis	0.4514	0.4563
cult1_dis	0.1965	0.1581
govt1_dis	0.2762	0.2103
hotel1_dis	0.195	0.1275
off1_dis	0.4748	0.4865
hops1_dis	0.1517	0.099
ind1_dis	0.4424	0.5022
sch1_dis	0.1308	0.1136
average	0.2767	0.217
trafficnoi	-0.123	-0.0672
visibility	-0.3616	-0.4109
seg_tnodes	-0.0043	-0.0289
seg_unodea	-0.2086	-0.1721
density	-0.2045	-0.1871
nonresmix	-0.0244	-0.0113

Table 17: Comparison of Variables of Interest Among Models

		r_net1	r_net1sq	Joint F Test	areu1_dis	areu1_dissq	Joint F Test
Structure + CDB, x-way							
	Coefficient	-3.609	0.006		58.175	-0.040	
	Robust Std Err	6.4711	0.0021		9.2815	0.0052	
	P>|t|	0.577	0.006		0.000	0.000	
	F			13.01			29.730
	Prob>F			0.0000			0.0000
Appendix A							
	Coefficient	-2.471	0.004		40.635	-0.035	
	Robust Std Err	6.4224	0.0021		9.2644	0.0051	
	P>|t|	0.700	0.046		0.000	0.000	
				8.400			31.680
				0.0002			0.0000
Appendix B							
	Coefficient	-14.190	0.006		37.753	-0.026	
	Robust Std Err	6.4245	0.0021		9.3339	0.0052	
	P>|t|	0.027	0.007		0.000	0.000	
				3.920			13.660
				0.0199			0.0000
Appendix C							
	Coefficient	-4.657	0.006		10.304	-0.022	
	Robust Std Err	7.2527	0.0022		11.6238	0.0052	
	P>|t|	0.521	0.008		0.375	0.000	
				7.660			17.280
				0.0005			0.0000

References

Adair, A., McGreal, S., Smyth, A., Cooper, J., & Ryley, T. (2000). House Prices and Accessibility: The Testing of Relationships Within the Belfast Urban Area. *Housing Studies, 15*(5), 699-716.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht, Netherlands: Kluwer.

Bafna, S. (2003). Space Syntax: A Brief Introduction to Its Logic and Analytical Techniques. *Environment and Behavior, 35*(1), 17-29.

Bartik, T. J. (1987). The Estimation of Demand Parameters in Hedonic Price Models. *Journal of Political Economy, 95*(11), 81-88.

Batalhone, S., Nogueira, J., & Mueller, B. (2002). *Economics of Air Pollution: Hedonic Model and Smell Consequences of Sewage Treatment Plants in Urban Areas*. Brasilia: University of Brasilia.

Batemen, I., Day, B., Lake, I., & Lovett, A. (2001). *The Effect of Road Traffic on Residential Property Values: A Literature Review and Hedonic Pricing Study*. Norwich: University of East Anglia; Economic and Social Research Council; University College London.

Bowen, W. M., Mikelbank, B. A., & Prestegaard, D. M. (2001). Theoretical and Empirical Considerations Regarding Space in Hedonic Housing Price Model Applications. *Growth and Change, 32*(4), 466-490.

Bowes, D. R., & Ihlanfeldt, K. R. (2001). Identifying the Impacts of Rail Transit Stations on Residential Property Values. *Journal of Urban Economics, 50*(1), 1-25.

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society, Series B (Methodological), 26*(3), 211-252.

Boykin, J. H., & Ring, A. A. (1993). *The Valuation of Real Estate (4th ed)*. Englewood Cliffs, N.J.: Prentice-Hall.

Butler, R. V. (1982). The Specification of Hedonic Indexes for Urban Housing. *Land Economics, 58*(1), 96-108.

Can, A. (1990). The Measurement of Neighborhood Dynamics in Urban House Prices. *Economic Geography, 66*(3), 254-272.

Cao, T. V., & Cory, D. C. (1981). Mixed Land Uses, Land Use Externalities, and Residential Property Values: A Reevaluation. *Annuals of Regional Science, 16*, 1-24.

Carter, C. G., & Haloupek, W. J. (2000). Spatial Autocorrelation in a Retail Context. *International Real Estate Review, 3*(1), 34-48.

Clapp, J. M., Rodriguez, M., & Thrall, G. (1997). How GIS Can Put Urban Economics on the Map. *Journal of Housing Economics, 6*(4), 369-386.

Clark, D. E., & Herrin, W. E. (2000). The Impact of Public School Attributes on Home Sale Prices in California. *Growth and Change, 31*(3), 385-407.

Crafts, J. M., MAI, SRA. (1998). Impact of Commercial Development on Adjacent Residential Properties. *Appraisal Journal, 65*(1), 6-11.

Craig, S. G., Kohlhase, J. E., & Papell, D. H. (1991). Chaos Theory and Microeconomics: An Application to Model Specification and Hedonic Estimation. *Review of Economics and Statistics, 73*(2), 208-215.

Crane, R. (2000). The Impacts of Urban Form on Travel: A Critical Review. *Journal of Planning Literature, 15*(1), 3-23.

Crecine, J. P., Davis, O. A., & John E. Jackson. (1967). Urban Property Markets: Some Empirical Results and Their Implications for Municipal Zoning. *Journal of Law and Economics, 10*, 79-99.

Cropper, M. L., Deck, L. B., & McConnell, K. E. (1988). On the Choice of Functional Form for Hedonic Price Functions. *The Review of Economics and Statistics, 70*(4), 668-675.

Cummings, P., & Landis, J. (1993). *Relationships Between Affordable Housing Developments and Neighboring Property Values*. Berkeley: University of California at Berkeley.

Davis, D. E. (2003). *GIS for Everyone, 3rd edition*: ESRI Press;.

Des Rosiers, F., Thériault, M., & Kestens, Y. (2003). *Measuring Accessibilty to Urban Services.* Paper presented at the 10th European Real Estate Society Conference, Helsinki, Finland.

Duany, A., Plater-Zybek, E., & Speck, J. (2000). *Surban Nation: the Rise of Sprawl and the Decline of the American Dream*. New York: West Point Press.

Dubin, R. A. (1988). Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms. *Review of Economics and Statistics, 70*(3), 466-474.

Espey, M., & Lopez, H. (2000). The Impact of Airport Noise and Proximity on Residential Property Values. *Growth and Change, 31*(Summer 2000), 408-419.

Follain, J. R., & Jimenez, E. (1985). Estimating the Demand for Housing Characteristics: A Survey and Critique. *Regional Science and Urban Economics, 15*(1), 77-107.

Franklin, J. P., & Waddell, P. (2003, July 31, 2002). *A Hedonic Regression of Home Prices in King County, Washington, using Activity-Specific Accessibility Measures.* Paper presented at the TRB 2003 Annual Meeting, Washington, D.C.

Freeman, A. M. (1971). Air Pollution and Property Values. *The Review of Economics and Statistics, 53*(4), 415-416.

Freeman, A. M. (1993). *The Measurement of Environmental and Resource Values : Theory and Methods*. Washington, D.C.: Resources for the Future.

Freeman, A. M. (1999). *The Measurement of Environmental and Resource Values: Theory and Methods*. Washington, D.C.: Resources for the Future.

Frew, F., & Judd, G. D. (2003). Estimating the Value of Apartment Buildings. *Journal of Real Estate Research, 23*(1), 77-86.

Garreau, J. (1991). *Edge Cities: Life on the New Frontier*: Doubleday.

Geohegan, J., Lisa A, W., & Bockstael, N. E. (1997). Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS. *Ecological Economics, 23*, 251-264.

Goodman, A. C., & Thibodeau, T. G. (1995). Age-Related Heteroskedasticity in Hedonic House Price Equations. *Journal of Housing Research, 6*(1), 25-42.

Grass, R. G. (1992). The Estimation of Residential Property Values Around Transit Station Sites in Washington, D.C. *Journal of Economics and Finance, 16*(2), 139-146.

Grether, D. M., & Mieszkowski, P. (1980). The Effects of Nonresidential Land Uses on the Prices of Adjacent Housing: Some Estimates of Proximity Effects. *Journal of Urban Economics, 8*(1), 1-15.

Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Changes. In *The Price Statistics of the Federal Government*. Princeton, New Jersey: National Bureau of Economic Research.

Gujarati, D. N. (1995). *Basic Econometrics* (3rd ed.). New York: McGraw-Hill, Inc.

Hillier, B. (1996). *Space is the Machine*. Cambridge: Cambridge University Press.

Hillier, B. (1999). The Hidden Geometry of Deformed Grids: or, Why Space Syntax Works, When It Looks as Though It Shouldn't. *Environment and Planning B: Planning and Design, 26*, 169-191.

Hillier, B., & Hanson, J. (1984). *The Social Logic of Space*. Cambridge: Cambridge University Press.

Hosch, S. T., & Koehlinger, D. V. (1997). How Residential Properties are Impacted by Adjacent Commercial Uses. *Valuation Viewpoint, 1*(4), www.shenehon.com/library/valuation_viewpoint/resident.html.

Hotaling, L., & Hooper, R. (2001, September 13, 2001). TWSA to Move Forward with Purchase of Morris Property. *Sylva Herald*.

Hughes, W. T., Jr., & Sirmans, C. F. (1992). Traffic Externalities and Single-Family House Prices. *Journal of Regional Science, 32*(4), 487-500.

Hulten, C. R. (2003). Price Hedonics: A Critical Review. *FRBNY Economic Policy Review, 9*(3), 5-15.

Jenness, J. (2004a). *Nearest Features, with Distances and Bearings (v. 3.8a)*. Retrieved August 31, 2004, from http://www.jennessent.com/arcview/nearest_features.htm

Jenness, J. (2004b). *Distance Matrix (dist_mat_jen.avx) extension for ArcView 3.x, v.1.4*. Retrieved September 11, 2004

Jo, S.-K. (1996). *A Balance Between Pedestrian and Vehicular Movement in Relation to Street Configuration.* Unpublished Doctoral Dissertation, Georgia Institute of Technology, Atlanta.

Jones, L. E. (1988). The Characteristics Model, Hedonic Prices, and the Clientele Effect. *Journal of Political Economy, 96*(3), 551-567.

Kahn, S., H., & Case, F. E. (1977). *Real Estate Appraisal and Investment (2nd ed.)*. New York: Ronald Press.

Kain, J. F., & Quigley, J. M. (1970). Measuring the Value of Housing Quality. *Journal of the American Statistical Association, 65*(320), 532-548.

Kennedy, P. (1998). *A Guide to Econometrics* (4th ed.). Cambridge, Massachusetts: The MIT Press.

Lake, I. R., Lovett, A. A., Bateman, I. J., & Langford, I. H. (1998). Modeling Environmental Influences on Property Prices in an Urban Environment. *Computers, Environment and Urban Systems, 22*(2), 121-136.

Lake Washington School District #414: Elementary School Boundary Map. Retrieved August 10, 2004, from http://www.lkwash.wednet.edu/lwsd/images/schools/BoundariesEL.gif

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

LeSage, J. P. (1998). *Spatial Econometrics*. University of Toledo, Department of Economics.

Li, M. M., & Brown, J. (1980). Micro-Neighborhood Externalities and Hedonic Housing Prices. *Land Economics, 56*(2), 125-141.

Mahan, B. L., Polasky, S., & Adams, R. M. (2000). Valuing Urban Wetlands: A Property Price Approach. *Land Economics, 76*(1), 100-113.

Malpezzi, S. (2002). *Hedonic Price Models: A Selective and Applied Review*.Unpublished manuscript, Madison, Wisconsin.

MaRous, M. S. (1996). Low-income Housing in Our Backyards: What Happens To Residential Property Values? Appraisal Journal , Jan96, Vol. 64 Issue 1, p. 27. *Appraisal Journal, 64*(1), 27-34.

McMillen, D. P., & McDonald, J. F. (1989). Selectivity Bias in Urban Land Value Functions. *Land Economics, 65*(4), 341-351.

McMillen, D. P., & McDonald, J. F. (1991). Urban Land Value Functions With Endogenous Zoning. *Journal of Urban Economics, 29*, 14-27.

Miller, R. L. (1999). *Economics Today:1999-2000 Edition*. Reading, Massachusetts: Addison-Wesley.

Mills, E. S. (1979). Economic Analysis of Urban Land-Use Controls. In P. Miezkowski & M. Straszheim (Eds.), *Current Issues in Urban Economics* (pp. 511-541). Baltimore: Johns Hopkins University Press.

Muth, R. F. (1969). *Cities and Housing*. Chicago: University of Chicago Press.

Neighborhood Attendance Reference Areas. Retrieved September 12, 2004, from http://www.seattleschools.org/area/facilities-plan/geo/refarea/ref_area.xml

Neiman, F. D. (2003). *A Very Brief Introduction to Space Syntax Analysis*: University of Virginia, Anthropology 589a.

Nelson, A. C. (1986). Using Land Markets to Evaluate Urban Containment Programs. *Journal of the American Planning Association52*(2), 156-171.

Nelson, A. C. (1993). Disamenity Influences of Edge Cities on Exurban Land Values: A Theory with Empirical Evidence and Policy Implications. *Urban Studies 30*(10), 1683-1690.

Nelson, A. C., & McClesky, S. J. (1990). Improving the Effects of Elevated Transit Stations on Neighborhoods. *Transportation Research Record 1266* 173-180.

Orford, S. (2002). Valuing Locational Externalities: A GIS and Multilevel Modeling Approach *Environment and Planning B, 29*, 105-127.

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

Patterson, R. W., & Boyle, K. J. (2002). Out of Sight, Out of Mind? Using GIS to Incorporate Visibility in Hedonic Property Value Models . *Land Economics, 78*(3), 417-425.

Penn, A. (2003). Space Syntax and Spatial Cognition or Why the Axial Line? *Environment and Behavior, 35*(1), 30-65.

Peponis, J., Ross, C., Rashid, M., & Kim, S. H. (1996). Regularity and Change in Urban Space: A Syntactic Analysis of Movement and Co-Presence in Atlanta. *Ekistics, 63*(376-378), 4-17.

Raun, A. (1997, August 19, 1997). Panel Rejects Plan for Lake Rezoning: Neighbors Oppose Commercial Plans. *Hastings Tribune*.

Ridker, R. G., & Henning, J. A. (1967). The Determinants of Residential Property Values with Special Reference to Air Pollution. *The Review of Economics and Statistics, 49*(2), 246-257.

Rodriguez, M., Sirmans, C. F., & Parks, A. P. (1995). Using Geographic Information Systems to Improve Real Estate Analysis. *The Journal of Real Estate Research, 10*(2), 163-173.

Rogers, W. (2000). Errors in Hedonic Modeling Regressions: Compound Indicator Variables and Omitted Variables. *Appraisal Journal, 68*(2).

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

School Guide. Retrieved September 12, 2004, from http://schoolguide.seattletimes.nwsource.com/

Sheppard, S. (1999). Chapter 41: Hedonic Analysis of Housing Markets. In P. Cheshire & E. S. Mills (Eds.), *Handbook of Regional and Urban Economics, Volume 3, Applied Urban Economics* (Vol. Volume 3, pp. 1595-1635). New York: Elsevier.

Song, Y., & Knaap, G.-J. (2003a). New Urbanism and Housing Values: A Dissaggregate Assessment. *Journal of Urban Economics, 54*, 218-238.

Song, Y., & Knapp, G.-J. (2003b). *The Effects of New Urbanism on Housing Values: A Quantitative Approach*. College Park, MD: National Center for Smart Growth Research and Education, University of Maryland.

Victoria Transport Policy Institute. (2003). Location Efficient Development and Mortgages. In *TDM Encyclopedia*.

Wallace, N. E. (1986). The Market Effects of Zoning Undeveloped Land: Does Zoning Follow the Market? *Journal of Urban Economics, 23*, 307-326.

Whitehead, J. C. (1999). http://core.ecu.edu/econ/whiteheadj/5000/ch10b/hpm.htm.

Wilhelmsson, M. (2000). The Impact of Traffic Noise on the Values of Single-family Houses. *Journal of Environmental Planning & Management, 43*(6), 799-815.

Wiltshaw, D. G. (1996). Spatial Autocorrelation and Valuation Accuracy: A Neglected Methodological Problem. *Journal of Property Research, 13*, 275-286.

Woolridge, J. M. (2000). *Introductory Econometrics*: South-Western College Publishing.