NEIGHBORHOOD IMPACTS ON SUBURBAN HOUSING VALUES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
The Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Hadi Sucahyono, M.S.

The Ohio State University
2006

Dissertation Committee:
Professor Jean-Michel Guldmann, Co-Adviser
Professor Burkhard von Rabenau, Co-Adviser
Professor Philip A. Viton

Approved by

Co-Adviser
Graduate Program in
City and Regional Planning
ABSTRACT

Six categories of variables are used to explain the variations of single-family suburban housing values: housing characteristics and conditions, amenities, accessibilities, socio-economic characteristics, existing land use, and land-use regulations (zoning and comprehensive planning). Three circular buffers surrounding each single-family housing unit are utilized as neighborhood units: adjacent, medium, and large neighborhoods, with radii of 100 meters, 400 meters, and 1 mile, respectively. Socio-economic characteristics, existing land use, and zoning characteristics are converted into neighborhood units by using allocation factors based on the intersection of Census blocks and parcels data using Geographical Information Systems (GIS). Delaware County, Ohio is chosen as the case study, because it has been the fastest growing county in Central Ohio and the tenth fastest growing county in the U.S. A total number of 3144 single-family housing units are selected. The effects of neighborhood characteristics on housing values are analyzed in the individual neighborhood models and the multi-neighborhood model. The Box-Cox transformation with two parameters is applied: one for the dependent variable ($\lambda$), and one for all independent variables ($\mu$). The maximum log likelihood of the integrated model is achieved with the value of $\lambda = 0.30$ and the value of $\mu = 0.02$. The mean elasticity is used to compare the effects of each variable on housing values. We find that floor area, fire places, basement, and garage capacity have positive
effects and increase housing values. The average price of housing in the neighborhood, the share of middle-income people, the share of people who drive cars to work, the share of people who work outside the county, and the share of people who work in non-agriculture activities are significant, which characterizes typical suburbanization. The share of existing residential land use and the share of land zoned for low- and medium-density residences increase housing values, due to the certainty of compatible land development. The share of land zoned for agriculture and open space decreases housing values in the suburbs, due to its restriction on development. Homeowners are willing to pay significantly more to live in an area which is comprehensively planned.
Dedicated to My Family,
Particularly to My Mother
Who passed away on May 15, 2006
ACKNOWLEDGMENT

I would like to express my deepest gratitude to my Advisers, Professor Jean-Michel Guldmann and Professor Burkhardt von Rabenau. Without them, I would never have completed my Ph.D. course of study. My appreciation also goes to Professor Philip A. Viton, a member of my Committee, and to Professor Kenneth Pearlman, who served on my Candidacy examination. I also thank to Ellen Wallace, who has helped me navigate the administration during my five years at The Ohio State University.

Of course, million thanks go to all my family members, for their unlimited support for me. Without them, I would have felt lonely in my long journey towards the Ph.D.

My thanks also go to Mr. Djoko Kirmanto, Minister of Public Works, Republic of Indonesia, who has always given me challenging tasks and has supported my pursuing a Ph.D. degree.
VITA

EDUCATION

2001 – 2006  Ph. D. Student in City and Regional Planning.
              The Ohio State University, Knowlton School of Architecture, Graduate
              Program in City and Regional Planning, Columbus, Ohio.

              Georgetown University, Graduate Public Policy Program, Washington, D.C.

1980 – 1985  Bachelor of Science in City and Regional Planning.
              Bandung Institute of Technology, Department of City and Regional
              Planning, Bandung, Indonesia.

EMPLOYMENT

2000 - 2001  Head of the Planning and Programming Section.
              Directorat General of Housing Development, Ministry of Public Works,
              Jakarta, Indonesia.

              West Kalimantan Province (Borneo), Ministry of Public Works, Indonesia.

              Ministry of Public Works, Jakarta, Indonesia.

1997 – 1998  Head of Multilateral Section.
              Directorat General of Human Settlements, Ministry of Public Works,
              Jakarta, Indonesia.

INTERNSHIPS

1991    New Town Development Project, Amsterdam, the Netherlands.
1990    Housing Development Project, Montgomery Inc., San Francisco, California.

SELECTED ARTICLES


FIELDS OF STUDY

Major Field:   City and Regional Planning

Minor Fields:  Planning for Housing
               Urban Economics
               Neighborhood Development
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Vita</td>
<td>vi</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xi</td>
</tr>
</tbody>
</table>

Chapters:

1. **Introduction** ........................................................................... 1

2. **Literature Review** ................................................................... 4

   2.1 Suburbanization and Sprawl .................................................. 4
   2.2 Hedonic Price Approach ....................................................... 10
   2.3 Application of Geographic Information Systems (GIS) to Housing Studies ................................................................. 16
   2.4 Zoning and Comprehensive Planning ........................................ 19
   2.5 Summary .................................................................................. 22

3. **Modeling Methodology** ............................................................. 23

   3.1 Theoretical Background ....................................................... 23
   3.2 The Neighborhood Model ....................................................... 25

4. **Data Sources and Data Processing** .......................................... 30

   4.1 Geographical Area .................................................................. 30
   4.2 Regional Overview of Delaware County .................................... 32
   4.3 Local Overview of Delaware County ........................................ 34
4.4 Data Sources ................................................................. 35
   4.4.1 Parcels Data ......................................................... 38
   4.4.2 Census 2000 STF1 Data File ................................. 40
   4.4.3 Census 2000 STF3 Data File ................................. 40
   4.4.4 Census Transportation Planning Package (CTPP) 2000 .... 40
   4.4.5 Zoning and Comprehensive Planning Data ................. 41
   4.4.6 GIS Data of Infrastructures and Amenities ............... 41
4.5 Data Processing .......................................................... 43
   4.5.1 Housing Characteristics and Conditions .................. 60
   4.5.2 Amenity and Accessibility ..................................... 61
   4.5.3 Socio-Economic Characteristics of Neighborhoods .......... 64
      4.5.3.1 Overview .................................................. 64
      4.5.3.2 Creation and Overlaying of Buffers .................. 65
      4.5.3.3 Allocation Factors ...................................... 65
      4.5.3.4 Conversion of Census Blocks and Block Groups Data into Buffer Data ........................................... 70
   4.5.4 Zoning and Comprehensive Planning ......................... 70

5. Model Specification ....................................................... 72
   5.1 The Empirical Neighborhood Model ............................ 72
      5.1.1 Housing Characteristics and Condition Variables ....... 73
      5.1.2 Amenity Variables .......................................... 81
      5.1.3 Accessibility Variables .................................. 83
      5.1.4 Socio-economic Variables .................................. 84
      5.1.5 Land-use Variables ......................................... 87
      5.1.6 Zoning and Comprehensive Planning Variables .............. 88
   5.2 Functional Form .................................................... 89

6. Empirical Results ......................................................... 92
   6.1 Individual Neighborhood Model .................................. 92
   6.2 Multi-Neighborhood Model ....................................... 107
   6.3 Elasticity of Variables ........................................... 115

7. Conclusions ............................................................... 119

Bibliography ................................................................. 124

Appendices:
A. Parcels Data ............................................................. 135
B. Master Data ............................................................. 138
C. SAS Programs ........................................................... 141
D. Zoning of Townships/Municipalities ............................... 186


LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Population and Income in Delaware County</td>
</tr>
<tr>
<td>5.1</td>
<td>Variables in the Hedonic Price Model</td>
</tr>
<tr>
<td>5.2</td>
<td>Descriptive Statistics of All Variables</td>
</tr>
<tr>
<td>6.1</td>
<td>Empirical Results of Individual Neighborhoods</td>
</tr>
<tr>
<td>6.2</td>
<td>Multi-Neighborhood Model</td>
</tr>
<tr>
<td>6.3</td>
<td>Multi-Neighborhood Model with Box-Cox Transformations</td>
</tr>
<tr>
<td>6.4</td>
<td>Elasticity of Variables</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Concept of Neighborhood Model</td>
<td>29</td>
</tr>
<tr>
<td>4.1</td>
<td>Delaware County and the State of Ohio</td>
<td>31</td>
</tr>
<tr>
<td>4.2</td>
<td>Annual Population Growth Rates in Central Ohio</td>
<td>33</td>
</tr>
<tr>
<td>4.3</td>
<td>Delaware County and Its Townships/Municipalities</td>
<td>36</td>
</tr>
<tr>
<td>4.4</td>
<td>Distribution of Parcels in Delaware County</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>Census Blocks and Block Groups in Delaware County</td>
<td>42</td>
</tr>
<tr>
<td>4.6</td>
<td>Distribution of Cemeteries in Delaware County</td>
<td>44</td>
</tr>
<tr>
<td>4.7</td>
<td>Electric Transmission Lines in Delaware County</td>
<td>45</td>
</tr>
<tr>
<td>4.8</td>
<td>Gas Lines in Delaware County</td>
<td>46</td>
</tr>
<tr>
<td>4.9</td>
<td>Hospitals in Delaware County</td>
<td>47</td>
</tr>
<tr>
<td>4.10</td>
<td>Mobile Home Parks in Delaware County</td>
<td>48</td>
</tr>
<tr>
<td>4.11</td>
<td>Regional Shopping Malls in Delaware County</td>
<td>49</td>
</tr>
<tr>
<td>4.12</td>
<td>Railroads in Delaware County</td>
<td>50</td>
</tr>
<tr>
<td>4.13</td>
<td>Rivers in Delaware County</td>
<td>51</td>
</tr>
<tr>
<td>4.14</td>
<td>Schools in Delaware County</td>
<td>52</td>
</tr>
<tr>
<td>4.15</td>
<td>Town halls in Delaware County</td>
<td>53</td>
</tr>
<tr>
<td>4.16</td>
<td>Fire Stations in Delaware County</td>
<td>54</td>
</tr>
<tr>
<td>4.17</td>
<td>Supermarkets in Delaware County</td>
<td>55</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.18</td>
<td>Industrial Sites in Delaware County</td>
<td>56</td>
</tr>
<tr>
<td>4.19</td>
<td>Banks in Delaware County</td>
<td>57</td>
</tr>
<tr>
<td>4.20</td>
<td>Golf Courses in Delaware County</td>
<td>58</td>
</tr>
<tr>
<td>4.21</td>
<td>Post Offices in Delaware County</td>
<td>59</td>
</tr>
<tr>
<td>4.22</td>
<td>Single-Family Housing Units Sold in 2000</td>
<td>62</td>
</tr>
<tr>
<td>4.23</td>
<td>Three Types of Buffers/Neighborhoods</td>
<td>66</td>
</tr>
<tr>
<td>4.24</td>
<td>Parcel’s Centroid and Buffers</td>
<td>67</td>
</tr>
<tr>
<td>4.25</td>
<td>Intersection of Residential Parcels, Blocks and Buffers</td>
<td>69</td>
</tr>
<tr>
<td>D.1</td>
<td>Zoning of Ashley Municipality</td>
<td>187</td>
</tr>
<tr>
<td>D.2</td>
<td>Zoning of Berkshire Township</td>
<td>188</td>
</tr>
<tr>
<td>D.3</td>
<td>Zoning of Berlin Township</td>
<td>189</td>
</tr>
<tr>
<td>D.4</td>
<td>Zoning of Brown Township</td>
<td>190</td>
</tr>
<tr>
<td>D.5</td>
<td>Zoning of the City of Delaware</td>
<td>191</td>
</tr>
<tr>
<td>D.6</td>
<td>Zoning of Concord Township</td>
<td>192</td>
</tr>
<tr>
<td>D.7</td>
<td>Zoning of Delaware Township</td>
<td>193</td>
</tr>
<tr>
<td>D.8</td>
<td>Zoning of Galena Municipality</td>
<td>194</td>
</tr>
<tr>
<td>D.9</td>
<td>Zoning of Genoa Township</td>
<td>195</td>
</tr>
<tr>
<td>D.10</td>
<td>Zoning of Harlem Township</td>
<td>196</td>
</tr>
<tr>
<td>D.11</td>
<td>Zoning of Kingston Township</td>
<td>197</td>
</tr>
<tr>
<td>D.12</td>
<td>Zoning of Liberty Township</td>
<td>198</td>
</tr>
<tr>
<td>D.13</td>
<td>Zoning of Marlboro Township</td>
<td>199</td>
</tr>
<tr>
<td>D.14</td>
<td>Zoning of Orange Township</td>
<td>200</td>
</tr>
<tr>
<td>D.15</td>
<td>Zoning of Oxford Township</td>
<td>201</td>
</tr>
<tr>
<td>D.16</td>
<td>Zoning of Porter Township</td>
<td>202</td>
</tr>
<tr>
<td>D.17</td>
<td>Zoning of Powell Municipality</td>
<td>203</td>
</tr>
<tr>
<td>D.18</td>
<td>Zoning of Radnor Township</td>
<td>204</td>
</tr>
<tr>
<td>D.19</td>
<td>Zoning of Scioto Township</td>
<td>205</td>
</tr>
<tr>
<td>D.20</td>
<td>Zoning of Shawnee Hills Municipality</td>
<td>206</td>
</tr>
<tr>
<td>D.21</td>
<td>Zoning of Sunbury Municipality</td>
<td>207</td>
</tr>
<tr>
<td>D.22</td>
<td>Zoning of Thompson Township</td>
<td>208</td>
</tr>
<tr>
<td>D.23</td>
<td>Zoning of Trenton Township</td>
<td>209</td>
</tr>
<tr>
<td>D.24</td>
<td>Zoning of Troy Township</td>
<td>210</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Most single-family housing studies have, in the past, attempted to identify the effects of housing physical characteristics on housing values. Efforts to expand the scope of these studies to include neighborhood characteristics have been limited, due to the unavailability of suitable data at the neighborhood level, and also due to limited tools of analysis. Most of these studies were not conducted within the broader scope of suburban development and the role of land-use spatial policies, such as zoning and comprehensive planning. This research extends earlier studies on single-family housing by considering variables related to neighborhood characteristics, including amenities, accessibilities to urban activities, socio-economic characteristics of the surrounding population, existing land uses, and comprehensive planning and zoning policies.

Using the hedonic price method, this research develops neighborhood variables based on three circular buffers that surround individual single-family housing units. These three neighborhoods are classified as adjacent, medium, and large neighborhoods, with radii of 100 meters, 400 meters, and 1 mile, respectively. Socio-economic characteristics, existing land uses, and land-use regulations (zoning) are measured based on these neighborhoods. This is in contrast to past hedonic price studies by Geoghegan et al. (1997), Irwin et al. (2002), Dubin and Sung (1990), Li and Brown (1980), Dinan and
Miranowski (1989), and Cassel and Mendehlson (1985), who used jurisdictional boundaries, ZIP codes, Census tracks and block groups to capture neighborhood characteristics. In order to compute neighborhood variables, allocation factors are developed using Geographical Information Systems (GIS), based on the number of bedrooms in the intersection areas of buffers, Census blocks and residential areas. This process combines micro-level parcels data and Census 2000 data.

This research aims at explaining the impacts of neighborhood characteristics on single-family housing values, with a focus on suburban housing development, and at providing inputs for land-use policy development, as implemented through zoning and comprehensive planning. These insights should be useful for controlling suburban sprawl, particularly in fast growing counties surrounding metropolitan areas. Furthermore, this research should provide inputs for formulating policies on the distribution of public amenities and infrastructures in the suburbs.

A functional form based on the Box-Cox transformation is used and estimated by maximizing the log-likelihood function. The transformation is based on two parameters: one for the dependent variable and one for all the independent variables.

The housing market in Delaware County, Ohio, is chosen as the case study, because this county has been the fastest growing county in Central Ohio. It is listed among the top ten growing counties in the United States. The annual population growth rate in this county between 1990 to 2000 was 5.09 %, about five times the average growth rate in other Central Ohio counties, and about four times the national rate.

This research intends to answer the following questions:
- What are the major physical characteristics of single-family housing in the suburbs that affect housing values?
- Does closeness to amenities, infrastructures, and public utilities have positive effects on housing values?
- What suburban amenities affect housing values?
- Does accessibility to urban activities significant affect housing values?
- What are the significant socio-economic factors that increase single-family housing values in the suburbs?
- Which existing land uses contribute to housing values in the suburbs?
- What is the impact of zoning and comprehensive planning on housing values?

This dissertation is organized into seven chapters. After the introduction in Chapter 1, Chapter 2 reviews the literature. Chapter 3 presents the modeling methodology and the neighborhood concept. Chapter 4 describes the data sources and data processing. Chapter 5 develops the specification of the empirical neighborhood model. Chapter 6 presents the empirical results, and Chapter 7 concludes and discusses the implications of the results for public policy and future research.
CHAPTER 2

LITERATURE REVIEW

The literature review focuses on four general streams of research: (1) suburbanization/sprawl, (2) hedonic price modeling, (3) geographic information system/GIS, and (4) zoning and planning. The literature on suburbanization/sprawl covers both theory and empirical evidence, primarily in the U.S. and at the local level. The second stream deals with the theory and application of hedonic price modeling for housing, which is viewed as a bundle of physical and neighborhood characteristics. The stream on Geographical Information Systems (GIS) explores the benefits of applying GIS to housing and neighborhood research. The fourth stream presents the theory and empirical evidence on the effects of zoning and comprehensive planning practices.

2.1. Suburbanization and Sprawl

Sprawl is a common phenomenon in the U.S. However, most discussions on sprawl begin with the admission that sprawl has not been well-defined (Frank, 2000). To answer the question “What is sprawl?”, different definitions have been used. For example, Ewing (1997) defines sprawl as low density, strip development, scattered development, and leapfrog development. Meanwhile, Gordon and Richardson (1997) define sprawl as low density, dispersed, decentralized, polycentric, and suburban. Fulton
et al. (2001) use the terms “sprawling” and “densifying” to contrast two types of development in metropolitan areas. They argue that sprawling happens when land is being consumed at a faster rate than population growth, while densifying implies that population is growing more rapidly than land consumed for urbanization. Batty, Xie, and Sun (1999) view sprawl as a growth involving two parameters: availability of land and aging process. They argue that new development (suburbanization) dominates the first stage of development, provided by land availability in fringe areas. In the second stage, new developments decrease, due to limited undeveloped land, and are followed by increasing redevelopment in central cities. Longley, Batty, and Chin (1999) explicitly define suburbanization as sprawl, and link it to urban transition or urban growth pattern, where the first phase of growth taken place in the core of the city and the second phase is suburbanization. They also note that there is little agreement on the characteristics, causes, and impacts of sprawl. The concept of sprawl from Ewing is different from that of Gordon and Richardson, since Ewing argues that sprawl is a matter of degree of development. Ewing also argues that it is the impacts of development that render development patterns undesirable, not the patterns themselves. He proposes two sprawl indicators: poor accessibility and lack of functional open space. Poor accessibility happens because residences are far from out-of-home activities, there is poor residential accessibility, and out-of-home activities are far from one another. Lack of functional open space happens because abundant open spaces are generally all in private hands and unavailable for public use. The term “functional” applied to open space means that the space performs some useful public role in containing development. Ewing and Gordon and Richardson also have different views on the causes of sprawl. Gordon and
Richardson argue that market failures encourage sprawl, including subsidies to automobile (encouraging long-distance driving) and local land-use regulations that discourage higher densities and mixed uses. Like Gordon and Richardson, Giulano (1995) argues that the price of private vehicles in the United States is very low and private-vehicle users do not pay directly for the pollution they generate, or the congestion they impose on other travelers. Giulano notes that gasoline taxes, vehicle-registration fees, drivers-license fees, and automobile taxes in the U.S. are lower than anywhere in the developed world, which encourages commuting and sprawl. In contrast, Ewing argues that land market imperfections induce sprawl. The rate of land appreciation is uncertain, causing land speculation and sprawl. Despite having different views on the causes of sprawl, Gordon and Richardson and Ewing have the same view that suburban development is caused by consumer preferences for living in low-density areas and having new single-family housing. They also agree that suburban economic development does not take evenly across the outlying area, but rather is centered in key or magnet areas, where economies of agglomeration arise and growth is rapid.

Two theories have been proposed to explain suburbanization in metropolitan areas (Mieszkowski and Mills, 1993): “natural evolution theory” and “flight from blight”. The first theory explains suburbanization as a result of growing population and household size, and of the rising income of people who used to live in central city areas and now seek new places in the suburbs. This shift has been encouraged by the massive development of freeways and the relatively cheap cost of commuting from suburbs to inner cities. The growth of automobile ownership has also influenced suburbanization. In addition, land in the suburbs or fringe areas is relatively cheaper than in the centers of
cities, which makes affluent people buy more land and build larger homes in the suburbs. Suburbanization is the natural result of built-up central cities where growth can no longer take place, and of rising incomes that allow people to buy bigger homes, larger lots, and the necessary cars. Using the case of Portland, Oregon, Davis et al. (1994) argue that people choose suburban residences on the basis of two preferences: the best or the most affordable house, and the looks/design of the neighborhood.

The second theory explains the growth of the suburbs as a result of the “flight” of people from inner cities, so as to stay away from crimes, racial tension, tax problems, low quality of public services, and low quality of environmental conditions. Suburbanization leads affluent central cities residents to migrate to the suburbs. Choosing suburban areas as a new place for living is linked to the desire of owning a single-family home, the need for an adequate environment for raising a family, the strong desire to have more privacy, and the appeal of rural ambiance (Longley, Batty, and Chin, 1999). The reasons for “White flight” can be grouped into three categories: (1) negative externalities associated to physical conditions in central cities; (2) negative externalities associated to socio-economic conditions or the aversion of living close to poor people and crime; and (3) government attempts at income redistribution at the local level, which people can avoid by moving.

Suburbanization/sprawl is sometimes thought as undesirable because of its negative consequences. Carruthers (2002) and Razin (1998) maintain that suburbanization/sprawl is undesirable because of its economic, social, and environmental disadvantages, such as excessive commuting and infrastructure costs, promotion of socio-economic segregation, and high consumption of natural open space. The costs of
infrastructure and commuting increase because of scattered and discontinuous residential locations. Segregation exists when affluent people migrate to reside in the suburbs and prefer to live in neighborhoods with people with similar income, race, education, and ethnicity. Natural open spaces are in jeopardy when suburbanization is uncontrolled. Burchell and Listokin (1995) conducted a study to compare the investment costs of infrastructure under sprawl and under managed growth, and found that sprawl cost is greater than the cost of managed growth/compact development. For example, the investment cost for roads in sprawl developments is 24% higher; the difference in investment costs of water and sewer infrastructure between these two types of developments is 7.6%. Sprawl or suburbanization also leads to the loss of resource lands. On average, Burchell and Listokin find sprawl development to cost about 10% more than compact development. Based on a study of the Los Angeles area, Ewing (1997) found that the loss of environmentally sensitive land is five times greater under sprawl, and the loss of prime farmland 66% greater. Sprawl has micro and macro effects. The micro effect comes in the form of green areas near residential areas (leap-frogging). This effect may have positive aspects, such as fresh and clean air, and negative aspects, such as bad odors and chemicals. The macro effects are related to the impact on the housing price gradient for the whole metropolitan region, since sprawl increases the amount of land the region uses, which increases the average commuting distance and housing prices throughout the metropolitan region.

There are several instruments used to control or reduce sprawl/suburbanization. Brueckner (2000) discusses development taxes, urban growth boundaries, and zoning/planning. A development tax can be charged on each acre of land converted from
agricultural to urban use. The magnitude of the tax is set equal to the value of the undeveloped land benefits that are lost when land is converted. By raising the cost of development, the tax retards the development process and slows the rate of urban expansion. The problems with this tax are (1) guessing the magnitude/value of undeveloped land, and (2) the risk of needlessly restricting development and shrinking housing supply. The urban growth boundary (UGB) is to limit development in designated areas on the urban fringe, and involves drawing a polygon around a city and prohibiting development beyond it. The problem with this instrument is similar to the problem of the development tax: guessing the exact extent of excessive expansion. Because UGB limits land available for development, this instrument may have severe effects on land/housing prices and unwarranted densities within the boundary line. Because UGB has the potential for excessively restricting city size, this policy should be used with great care. Razin (1998) points out that attempts to control urban sprawl are frequently based on land-use planning and zoning, which provide for a more aesthetic outcome than UGB. Razin also argues that the more feasible and effective way to confront urban sprawl is to modify the interests of the agents involved in land-use development, especially local governments. Planned growth and coordinated growth are also keys for confronting sprawl/suburbanization. Byun and Esparza (2005) note the effect of the fragmentation of local growth control on suburbanization and sprawl. They argue that localities’ enforcement of restrictive zoning and growth control may induce spillovers. The price effect of growth control, such as higher housing cost, may lead to an unfavorable housing market situation that forces prospective and existing residents to seek housing substitutes in neighboring localities. In addition, homebuilders have to move their residential
projects to other non-growth-control localities or less-control localities, because of profit consideration. These spillover effects may be considered as one process propelling suburbanization or sprawl.

Although there is no agreed-upon definition on sprawl, some studies have used the terms of sprawl and suburbanization interchangeably (Byun and Esparza, 2005). Unfortunately, sprawl and suburbanization have been mostly discussed and debated qualitatively, and have rarely been linked to quantifiable impacts on housing values, in particular in fringe areas which attempt to control sprawl through zoning.

2.2 **Hedonic Price Approach**

The hedonic price approach is based on consumer theory, which postulates that every good provides a bundle of characteristics or attributes (Garrod and Willis, 1999). The most common application of the hedonic price approach is related to the willingness to pay for housing. A house provides shelter, location, and access to different quantities and qualities of public services and environmental goods, such as open space. The method captures the individual’s willingness to pay to consume a particular good as a function of the goods’ characteristics. The price of a house which a potential buyer is willing to pay depends on a wide range of attributes, such as structural characteristics, public sector characteristics and local amenities. Data requirements in the hedonic price approach include structural data and local data. Structural data include details on the property being sold/structural information, and local data includes details on neighborhood and amenity. The general form of the price function is:

\[ P = f(S, N, A) \]
where:

\( P = \) Price of the house;
\( S = \) Structural characteristics of the house;
\( N = \) Neighborhood characteristics;
\( A = \) Amenities characteristics.

The functional forms commonly used in hedonic price models are the linear, semi-log, and log-linear forms. For example, Nelson, Genereux, and Genereux (1997) study the impacts of externalities on housing prices, and utilize the following housing price function:

\[
V_i(t) = a_0(t) + \sum_{j=1}^{n} b_j(t) X_{ij}(t) + b_k L_{Fi} + e_i,
\]

where \( V_i(t) = \) value of house \( i \) at time \( t \); \( X_{ij}(t) = \) contribution of individual attribute \( j \) to house \( i \) value at time \( t \); \( L_{Fi} = \) distance of the house to externalities; \( e_i = \) error term. The first derivative measures the influence of the negative externalities on the house value:

\[
\frac{d V_i(t)}{d L_{Fi}} > 0
\]

In the housing market, households seek to avoid nuisances. They will pay more for sites that are further from externalities or not affected by nuisances, and they will pay less for those closer to externalities or affected by nuisances.

Dunford, Marti, and Mittelhammer (1985) study housing in the urban fringe market and use a semi-log form:

\[
\ln P = \beta_0 + \sum_{j=1}^{m} \beta_j X_j + \epsilon
\]

where:

\( P = \) parcel’s price
\( X = \) explanatory variables, such as location characteristics.
With this functional form, a \( k \) unit increase in an explanatory variable \( X_j \) leads to an increase in the parcel price of \( \{(e^{β_j} - 1) 100 \} \) percent.

Garrod and Willis (1999) argue that economic theory imposes no restrictions on the form of the hedonic price function, and that there is no single “best” hedonic price function for all purposes. The best form is based on the statistical significance of the estimated coefficients and the goodness of fit of the model. The main advantage of the linear, semi-log and log-linear functional forms is the transparency of the relationships between the marginal attribute prices and the parameters of the function. The linear function implies constant marginal attribute prices, while the semi-log function implies that marginal attribute prices are proportional to the price of housing.

Another approach to the functional form in hedonic price models is the Box-Cox transformation. This model assumes that some transformation of the dependent variable is normally distributed and linearly related to some set of transformations of the independent variables. This model is applicable to continuous dependent and independent variables, but not binary variables. Dinan and Miranowski (1989) describe a Box-Cox model as follows:

\[
P^{(θ)} = α_o + \sum α_i h_i^{(λ_i)} + u
\]

where:

\[
P^{(θ)} = (P^θ − 1) / θ \quad θ \neq 0
\]

\[
P^{(θ)} = \ln P \quad θ \to 0
\]

\[
h_i^{(λ_i)} = (h_i^{λ_i} − 1) / λ_i \quad λ_i \neq 0
\]

\[
h_i^{(λ_i)} = \ln h_i \quad λ_i \to 0
\]
P = Price of the house;
hi = Housing characteristics i;
If $\theta = 1$ and $\lambda = 1$, then the model is reduced to a linear form;
If $\theta \to 0$ and $\lambda \to 0$, then the model is reduced to a log-linear form;
If $\theta \to 0$ and $\lambda = 1$, then the model is reduced to a semi-log form.

Cassel and Mendehlson (1985) note several drawbacks in using the Box-Cox model, due to the large number of coefficients to be estimated, which may reduce the accuracy of any single coefficient, and to the nonlinear transformation, which results in complex and cumbersome estimates of slopes and elasticities. They also note that this functional form is not suited to any data set containing negative numbers.

Linneman (1980) suggests that one potential problem in hedonic price models is that of coefficient bias because of omitted variables. Another potential problem is the multicollinearity among closely linked variables. Multicollinearity may be investigated by looking at the sensitivity of the coefficients to the omission of other variables. Maddala (1977), Kmenta (1971), and Kennedy (1985) argue that OLS coefficient estimates are statistically unbiased, even when multicollinearity is strong, but the estimates may be imprecise and unstable. If multicollinearity affects some unimportant explanatory variables, they suggest excluding the variables from function. To deal with this potential problem, they suggest computing coefficient correlations matrices for all explanatory variables prior to regressions. The degree of inter-correlation among variables will help to determine which variables to drop. Gatzlaff and Ling (1994) mention the problem of data availability, as structural and locational characteristics are difficult and expensive to obtain. In addition, omitted characteristics that affect house prices are potential problems in estimation. They also notice that two major structures for
hedonic modeling are the “explicit time variable” model and the “strictly” cross-sectional model, and the major functional forms are linear, semi-log, exponential, and double-log. Because theory offers no guidance for selecting the preferred model, Gatzlaff and Ling suggest that models be evaluated on the basis of “best fit”, or $R^2$.

Dubin and Sung (1990) argue that neighborhood quality is an important element of the housing bundle, and conduct a hedonic study in Baltimore, using neighborhood characteristics grouped into three categories: socio-economic status, municipal services, and racial composition. They use a semi-log form, which is deemed more appropriate for owner-occupied housing. Besides housing structural characteristics, they focus on neighborhood socio-economic characteristics and race composition. The model used is follows:

$$Y = X\beta + Z_o\gamma_0 + Z_1\gamma_1 + \epsilon$$

where:

- $Y$ = sales price of houses (log)
- $X$, $Z_o$, and $Z_1$ represent vectors of explanatory variables of structural characteristics, socio-economic characteristics, and public services. They find that structural characteristics, such as the number of rooms, size of garage, availability of fireplace, and number of bathrooms, are significant and positively contribute to housing prices. In addition, socio-economic characteristics such as income level and the share of White population also positively contribute to housing prices.

Green (1999) studies the housing market in suburban Wisconsin, using the semi-log form. He derives the functional form of the housing market as the reduced form of demand and supply:
Demand for housing:

\[ Q_d = \alpha_d + \gamma_{1d} P_d + \beta_{1d} \text{Population} + \beta_{2d} \text{Income} + \beta_{3d} \text{Age} + \beta_{4d} \text{Household Type} + \beta_{5d} \text{Race} + \varepsilon_1; \]

Supply of housing:

\[ Q_s = \alpha_s + \gamma_{1s} P_s + \beta_{1s} \text{Wages} + \beta_{2s} \text{Materials} + \beta_{3s} \text{Land} + \varepsilon_2; \]

Reduced form:

\[ P = \pi_1 + \pi_2 \text{Population} + \pi_3 \text{Income} + \pi_4 \text{Age} + \pi_5 \text{Household Type} + \pi_6 \text{Race} + \pi_7 \text{Wages} + \pi_8 \text{Materials} + \pi_9 \text{Land} + \mu. \]

The dependent variable used in this study is the natural logarithm of housing price. Education, age, and marital status are among the significant factors contributing to housing prices in suburban Wisconsin.

The most recent study utilizing the hedonic price approach with linear and semi-log forms are Lynch and Rasmussen (2004) and Mathur, Waddel, and Blanco (2003). Lynch and Rasmussen study the impact of proximity to local public services on housing prices. They regress the log of housing price on structural characteristics, local public services characteristics, and neighborhood characteristics. They create the following neighborhood variables: population share of age 55 and over, percentage of Black people, percentage of homeowners, median household income, and percentage of white-collar people. The results show that housing prices are bid up when the house is surrounded by households with a relatively high socio-economic status.

Hedonic studies for housing have, in the past, commonly focused on the structural characteristics of housing, including: floor area, number of bathrooms, number of bedrooms, house age, garage capacity, number of fireplaces, availability of basement,
number of rooms, and number of stories. The neighborhood characteristics were mostly in the forms of socio-economic characteristics, generally at the level of Census tracks [Li and Brown (1980), McDonald (1980), Dinan and Miranowski (1989), Cobb (1984)], block groups [Dubin and Sung (1990), Kestens et al. (2004), Geoghegan et al. (1997)], transportation analysis zone [McMillen and Smith (2003)], and ZIP codes [Goodman et al. (2003)]. Besides using data taken from the Census, these studies have also gathered data from real estate agents, city’s Assessment Office, and parcels data. The geographic study area is generally a single city area [Nelson et al. (1997), Kestens et al. (2004), Singell and Lillydahl (1990), Thorsnes (2000), McMillen and McDonald (1991), Dubin and Sung (1990), Dinan and Miranowski (1989), Cobb (1984), and Jud (1980)], a county [Dunford et al. (1985), Green (1999), Geoghegan et al. (1997)], or involve a national coverage [Linneman (1980), McDonald and McMillen (2000)].

One drawback of past hedonic studies is the limited use of neighborhood characteristics associated with suburbanization and growth control. In addition, the characteristics of amenities have generally been captured by using dummy variables [Follain and Malpezzi (1981), Bell and Bockstael (1997), Wallace (1988), Linneman (1980), Tse (2002)].

2.3 Application of Geographic Information Systems (GIS) to Housing Studies

The hedonic price model in housing research is commonplace; however, it is rarely combined with spatial data gathered with Geographical Information Systems (GIS). Geoghegan, Wainger and Bockstael (1997) argue that GIS used in housing studies provide an obvious and appealing means of organizing and storing information that is
spatially based, such as land/housing records, natural resource features, and public infrastructure location. Another advantage of GIS is to make distances more accurately computed. In housing research, what surrounds a property has a major influence on housing value. The analysis of the pattern of land use and amenities surrounding the property is important, and can be captured with GIS applications. Geoghegan, Wainger and Bockstael (1997) conduct an analysis of a region within a 30-mile radius from Washington D.C., using GIS tools. One of their findings is that the marginal value of open space changes under different landscape patterns. One drawback of this study is the limited analysis of only open space land, without considering the complete land-use composition around a housing unit. Also, the characteristics of the neighborhood around an individual housing unit are derived from Census block group data.

Open space provided by agricultural land, forests, and water may be expected to contribute positively to property values. However, much of agriculture in the community is shade-grown plants, which may not yield the positive amenities provided by open fields and pastures. Paterson and Boyle (2002) examine the importance of visibility of fields and pastures in residential selling price. Using GIS, they calculate visibility measures in a 360-degree circle around each property, including the percentage of visible area overall within a one-kilometer radius and the percentage of each visible land use. They find that visibility of agriculture and open space areas is an important determinant of residential prices.

Irwin and Bell (2002) analyze land-use change in the rural-urban fringe, using GIS applications. They create donut buffers for analyzing conversion of farmland and public open space in Maryland. The rural-urban fringe is defined as the exurban area that
extends beyond the urban and suburban area. Their study uses a binary variable of land conversion as the dependent variable, and introduces the minimum-lot variable as a proxy for growth management. Distance to cities is used as a proxy for accessibility, and is measured over the road network. Land uses within the buffers are grouped into two categories, open space and high density urban development, which includes all land for commercial, industrial and multi-family residential. Like the study by Geoghegan et al. (1997), the drawback of this study is the limited analysis of only open space land, without considering the whole composition of existing land uses and zoning.

Din, Hoesli and Bender (2001) argue that GIS have made possible the development of databases that can be used to better measure environmental characteristics. Their environmental parameters refer to the quality of the neighborhood and the quality of the location within a neighborhood.

Another benefit of applying GIS in spatial economics and real estate research is demonstrated by Clapp (1997). He argues that GIS is a powerful tool for supporting research because of its capability of storing and manipulating large data sets on spatial relationships. GIS can compute a matrix of distances among points in a database by measuring straight-line distance among features.

The common drawback of past studies with GIS application is the measurement of distances by using road networks without weighting it by types of roads. Because roads range from highways to local roads, ignoring this variability may lead to inaccurate measurements. In addition, past housing studies with GIS application have rarely considered the whole composition of land uses and land zoning in the neighborhood.
2.4 Zoning and Land-Use Planning

Growth control or growth management techniques include housing permit, population growth cap, urban growth boundaries, urban service limit lines, and residential zoning (Byun et al., 2005). Among the main reasons for growth control is rural land preservation and population growth containment. McDonald and McMillen (2004) mention as a motive for development control/zoning the promotion of general welfare by separating land uses so as to mitigate negative external effects. Zoning corrects the market allocation of land that is inefficient because of non-internalized externalities. With zoning, a community achieves three important values:

- Social use value: a good quality of life, such as supportive neighborhoods.
- Market value: allocation of parcels of land to the “highest and best use”.
- Ecological value: a preservation of nature.

One major advantage of zoning is the protection of home values, especially in the suburbs, from the threat of externalities or incompatible land uses settling in the neighborhood. Green (1999) mentions that zoning serves principally to protect property owners from the negative externalities of new developments, so that residential properties are sited next to other residential properties, and commercial enterprises are sited near other commercial enterprises. When zoning is adopted, it is no longer uncertain whether the nearby track of undeveloped land will be developed with an incompatible land use (Fischel, 2004). Tse (2001) notes that zoning has a dual character: (i) assigning exclusive property rights, and (ii) as a planning instrument that attenuates private property rights over the most valuable uses of land. Crone (1983) relates zoning to remedies for
achieving a Pareto-optimal allocation of resources, particularly when externalities are present in the economy, by defining property rights more precisely.

There are three effects from zoning: (i) increase/decrease in housing prices, (ii) residential segregation and (iii) spillover development. Zoning/land use regulations may increase housing prices because of an increasing housing cost following an increase of housing services and the quantity of services. Increasing housing prices may also result from an increase in the demand for housing or a decrease in the supply of housing. If zoning results in the neighborhood is becoming less congested, less polluted, and with less externalities, then demand for housing increases. If demand for housing in the suburbs increases, then there is a positive net social benefit, even if low income people are excluded from the jurisdiction. Spillover development may happen because restriction of potential development in particular zoning districts, that makes developers and prospective homeowners build and buy housing in adjacent neighborhood areas with fewer restrictions. It may also happen because of the high price of housing in the restricted zoning districts, which pushes people to buy housing in other areas with cheaper prices.

Zoning intervention is designed to reduce market failures. Uncertainty of land uses and lack information in the market are among the sources of market failures. Zoning provides this information and it reduces development cost. Knowing the zoning of a particular area, with the number of housing units permitted, will lead to correct planning for infrastructure systems and land development. Since higher density is associated with increased pollution and congestion, zoning can restrict the number of housing units to an optimal level, making the neighborhood more desirable and raising housing prices.
Zoning/local land-use regulations may contribute to income and racial segregation, particularly when zoning strongly increases housing prices, and decreases housing affordability for low-income populations. Ihlanfeldt (2004) argues that this happens in local land-use regulation, which tends to exclude lower-income households from suburban communities, particularly through adoption of minimum-lot sizes for single-family housing. Large lots inflate the cost of owner-occupied housing, making it difficult for low and moderate income households to buy houses in these communities. The desire to exclude low-income households in suburban neighborhoods may be due to fear of crimes and the poor maintenance of low-income houses, which may decrease neighborhood values.

Jud (1980) uses dummy variables as proxies for residential zoning, and finds that residential zoning classification exerts a positive and very significant effect on residential property values. He argues that certainty of residential land use provides owners with the security that other nearby properties will continue to be used solely for residential use. Meanwhile, Quigley and Rosenthal (2005) report that zoning impacts on prices have been mixed, sometimes positive, sometimes negative, and sometimes completely insignificant. They imply that, in some cases, local land regulation/zoning is symbolic, ineffectual, or only weakly enforced.

McMillen and McDonald (1991) note the endogenous character of zoning, whereby zoning is a function of lagged land use and relative land values. They use a logit model to analyze determinants of zoning by using estimated land values in each tract and land-use category. One drawback of this study is the use of land values estimated by local governments, instead of the actual market values.
To capture the role of zoning, past studies have generally used the minimum-lot size as a proxy for zoning [(Wheaton (1993), Green (1999), Irwin et al. (2002), and Rolleston (1987)) and binary variables as proxies for residential and non-residential zoning [Wallace (1988), Feitelson (1993)], which does not capture the whole zoning pattern in a neighborhood.

### 2.5 Summary

In view of the limited neighborhood characteristics associated with sprawl in past studies, it is important to use more comprehensive variables to capture neighborhood effects, sprawl, density, planning, and zoning impacts on housing prices. The physical characteristics of housing, which were commonly used in past hedonic studies, are still important and must be included in any research, but neighborhood characteristics should be addressed to include more detailed characteristics of amenities, accessibilities, socio-economic condition, existing land uses, and zoning.
CHAPTER 3

MODELING METHODOLOGY

The purpose of this research is to empirically analyze neighborhood impacts on housing values. A theoretical background is first formulated by referring to the hedonic price model developed by Rosen (1974) and Brown and Rosen (1982). The theoretical background provides guidance to the selection of the variables in the model. In addition to housing characteristics, commonly used in the past, the model includes neighborhood characteristics that have never been used. To specify these neighborhood characteristics, three types of neighborhoods are considered in the model. A generalized functional form is then identified to explain neighborhood impacts on housing values.

3.1 Theoretical Background

Rosen (1974) and Brown and Rosen (1982) argue that goods are valued for their utility-bearing attributes or characteristics. Each good is described by \( n \) objectively-measured characteristics, represented by a vector \( z = (z_1, z_2, \ldots, z_n) \), with \( z_i \) measuring characteristics \( i \) of the good. Each good offers buyers distinct packages of characteristics. The markets of goods implicitly reveal a price function \( p(z) = p(z_1, z_2, \ldots, z_n) \), relating price and good characteristics.
Product differentiation implies that a wide variety of alternative packages are available in the market. The demand for products or goods with characteristics $z$ is denoted by $Q^d(z)$, while $Q^s(z)$ is defined as the market supply with those attributes. In market equilibrium, $Q^d(z)$ equals $Q^s(z)$, with $p(z)$ as the equilibrium implicit price with characteristics $z$. Hedonic prices are defined as the implicit prices of attributes and the specific amounts associated with them. Econometrically, implicit prices are estimated by regression analysis (product price regressed on characteristics).

Garrod (1999) argues that the most common application of the hedonic price model is in relation to the public willingness to pay for housing. Each property is assumed to constitute a distinct combination of attributes, which determine the price or buyers’ willingness to pay. The price of a housing unit is dependent upon the availability and level of a wide range of attributes, such as structural characteristics, local socio-economic characteristics, and local amenities. The hedonic price function is defined as a reduced form equation accounting for both demand and supply factors.

Torrens and Alberti (2000) argue for the importance of accessibility in the analysis of suburban development. They note the two important ways through which accessibility impacts development: residential accessibility and destination accessibility. Residential accessibility is related to the distances faced in reaching working, recreational, and shopping places. Destination accessibility is related to the fact that destination opportunities themselves are spatially separated from each other. Three methods by which accessibility may be quantified have been derived from the field of transportation, economics, and regional science: cumulative opportunities measures, utility-based measures, and gravity-based measures. The first method (cumulative
opportunities measures) counts the number of opportunities that can be visited within a given travel time. Such measure provides an indication of the volume of potential destinations or activities available, rather than the distance to these opportunities. The second method (utility-based) measures the accessibility derived from spatial choice and decision theory. It determines the utility of adopting one decision from a set of available choices, and it weights up the utility value of trip choices available within a given distance from a location. The utility-based measures are commonly calculated with logit models. The third method (gravity-based), related to the Newtonian physics concept, measures spatial interactions and predicts the size and direction of spatial flows. The gravity-based accessibility is based on the idea that the interaction \( C_{ij} \) is proportional to the capacity of activities at the origin \( i \) to generate trips \( W_i \) and at the destination \( j \) to attract trips \( W_j \), and inversely related to the distance between those activities \( D_{ij} \) with some weighting mechanism \( (\alpha) \) used to discourage trips over long distances. This gravity-based accessibility can be written as follows:

\[
C_{ij} = \frac{W_i W_j}{(D_{ij})^\alpha}
\]

The accessibility of activity \( i \) to all activities \( j \) is then:

\[
A_i = \sum_j \frac{W_j}{(D_{ij})^\alpha}
\]

3.2 The Neighborhood Model

The value of single-family housing units is affected primarily by housing characteristics and conditions, represented by vector \( h = (h_1...h_j...h_n) \), where \( h_j \) is the characteristics and condition \( j \) of the house. Another major factor contributing to housing
value is the bundle of amenities, represented by vector \( a = (a_1 ... a_k ... a_K) \), where \( a_k \) represents amenity \( k \). Closeness or distance to amenities affects housing values through the benefits and negative externalities generated by the amenities. Furthermore, accessibility to urban activities also impacts housing values. These are represented by the vector \( c = (c_1 ... c_m ... c_M) \), where \( c_m \) is the accessibility to activity \( m \).

The amenity concept used in this research is different from the accessibility concept. In the case of accessibility, choices are considered and the complete opportunity set is weighted by distance. The accessibility index, measured with a gravity-based formula, is used to analyze the impact of the complete set of opportunities on housing values (e.g. employment and shopping). In the case of amenities, the full set of options is considered to have little value, and the focus is set on the nearest opportunity, reflecting the importance (positive or negative) of proximity to such amenities as schools, hospitals and railroads. Unlike past hedonic models which measured distances to amenities by using the road network, this research measures straight-line distances between housing parcel centroids and amenities, a better measure of proximity.

Neighborhood socio-economic characteristics, existing land uses, and government land-use regulations (comprehensive planning and zoning) are also strong determinants of housing values. Most studies in the past have generally used dummy variables as proxies for zoning availability. This research captures the broader role of zoning by using the detailed composition of land zoned for major land uses (residential, commercial, industrial, and agricultural). It is also important to analyze zoning in tandem with another local land-use regulation (comprehensive planning), which also attempts to control sprawl and suburbanization. Spatial analyses of zoning and comprehensive planning in
individual townships and municipalities will provide detailed estimation of the effects of fragmented local land-use regulation on housing values.

To capture neighborhood characteristics, let $S_{pi}$ be the socio-economic characteristic $p$ for house $i$, $L_{qi}$ the land-use characteristic $q$ for house $i$, and $Z_{ri}$ the zoning characteristic $r$ for house $i$. These variables affect housing values locally, within a certain distance from the housing unit. Therefore, these factors need to be captured in a particular neighborhood unit. In order to explain the impacts of neighborhood characteristics on housing values, the use of neighborhood units as circular buffers surrounding the house is more appropriate than other conventional neighborhood units, such as ZIP codes, Census tracts, block groups and blocks.

Applications of GIS in past hedonic studies have used data from the Census, particularly on a track and block group basis, without any spatial conversion. Using GIS, this research will create buffers as neighborhood units, and convert data from block and block groups into these buffers. The advantage of this conversion is the use of similar spatial units, as compared to the irregular shapes of blocks or block groups. Using different circular neighborhoods allows for differentiating the impacts of different neighborhoods. Three circular neighborhoods, with different radii and surrounding each housing unit, are used to analyze the impacts of neighborhood characteristics on housing values: adjacent neighborhood with a radius of 100 meters, medium neighborhood with a radius of 400 meters, and large neighborhood with a radius of 1 mile. These ranges represent (1) the normal distance for visibility in the neighborhood, (2) the normal walking distance within the neighborhood, and (3) the minimum distance for people to rely on cars. Let $n$ be the index of neighborhood.
The hedonic function is then formulated as follows:

\[ V_i = F (H_i, A_i, C_i, S_i, L_i, Z_i) \] (3)

where:

- \( V_i \) = Price of housing unit \( i \),
- \( H_i = (H_{ij}) \) = vector of housing attributes \( j \) for housing unit \( i \),
- \( A_i = (A_{ik}) \) = vector of amenity attributes \( k \) for housing unit \( i \),
- \( C_i = (C_{im}) \) = vector of accessibility attributes \( m \) for housing unit \( i \),
- \( S_i = (S_{pin}) \) = vector of socio-economic characteristics \( p \) in buffer \( n \) surrounding housing unit \( i \),
- \( L_i = (L_{qin}) \) = vector of land-use characteristics \( q \) in buffer \( n \) surrounding housing unit \( i \),
- \( Z_i = (Z_{rin}) \) = vector of zoning characteristics \( r \) in buffer \( n \) surrounding housing unit \( i \).

Figure 3.1 presents a diagram of the neighborhood model.
Figure 3.1 The Neighborhood Model
CHAPTER 4
DATA SOURCES AND PROCESSING

This chapter describes the geographical context of the research, data sources, and data processing. The reasons for choosing the geographical area are first presented, followed by regional and local overviews of this area. Data sources are next described. Finally, the procedures and tools used in converting the raw data into the variables used in this research are discussed.

4.1 Geographical Area

Delaware County, Ohio, is selected as the geographical area for this research. It is located just north of Franklin County and the City of Columbus, and is one of the seven counties that make up the Columbus Metropolitan Area (Franklin, Delaware, Licking, Fairfield, Madison, Union, and Pickaway). Figure 4.1 presents the location of Delaware County within the State of Ohio.

There are three reasons for choosing Delaware County. First, it has been the fastest growing county in Ohio and the tenth fastest growing county in the United States. Second, this county has many rapid-growing townships and municipalities, which provide an ideal context to test the effects of neighborhood characteristics and local land-use regulations (zoning and comprehensive planning) on housing values in the suburbs.
Third, Delaware County has been a leading county in the U.S. for developing a comprehensive land information system (DALIS/Delaware Appraisal Land Information System), which provides a wealth of data.

The major advantage of using this county as the geographical area for this research resides in the detailed characteristics of housing and neighborhoods. Micro-level data, such as parcels and master field data, provides very detailed characteristics for individual houses, parcels, land uses, and amenities, as compared to macro-level data (e.g. regional or national).

4.2 Regional Overview of Delaware County

Delaware has been the fastest growing county in Ohio over the 1990-2000 period with an annual population growth rate of 5.1 %, and has been the tenth fastest growing county in the United States. During the same period, the annual population growth rates of the other Central Ohio counties of Franklin, Fairfield, Licking, Union, Pickaway, and Madison were only 1.07 %, 1.72 %, 1.27 %, 2.50 %, 0.89 %, and 0.89 %, respectively. From 1990 to 2000, the overall annual population growth in Central Ohio, the State of Ohio, and the United States were 1.39 %, 0.46 %, and 1.24 %, respectively. From year 2001 to year 2002, Delaware County growth rate was still the highest in Central Ohio. Figure 4.2 presents a graphical comparison of these growth rates in Central Ohio.

The high population growth rates in Delaware County have affected the conversion of land from agricultural usage to non-agricultural usage, as demonstrated by taxable land values. During the period 1994-2001, the taxable value of agricultural land
Figure 4.2  Annual Population Growth Rates in Central Ohio
decreased by 2.9%. In contrast, the taxable value of residential land increased by 120%, and the taxable value of commercial land increased by 141%.

The increasing number of net in-migrants, building permits, and commuters further demonstrate the rapid development of Delaware County. From 1990 to 2000, net migration to Delaware County increased by 35,118 people. In contrast, net migration to Franklin County decreased by 15,469 people. The total number of building permits in Delaware County increased by 60% between 1995 and 2001. Interestingly, most of this development took place in the townships directly adjacent to the City of Columbus and Franklin County, especially Liberty, Orange, and Genoa Townships. The total number of commuters from Delaware County to Franklin County has also increased substantially from 1990 to 2000, by 15,254 people.

These indicators show that the process of development and suburbanization has been rapid in Delaware County. The rapid population growth affects demand for housing, infrastructures, and public services. It also affects the transition from agriculture to urban land use, increasing the demand for land zoned for non-agriculture usage. This trend also means that a large share of buildings sales represents first-time sales by developers, with much new housing located next to natural areas.

4.3 Local Overview of Delaware County

The total 2000 population of Delaware County was 109,989, with most of it living in the southern part of the county, particularly Orange, Genoa, Liberty, Concord, and Harlem townships. In 2000, the total population of these five townships made up 37.08%
of the total population of the county. From 1990 to 2000, the population growth in these five townships was relatively higher than in other county townships: Orange, Genoa, and Liberty townships grew by 228.95 %, 178.63 %, and 142.27 %, respectively. The population growth in the northern part of the county was lower. For example, Marlboro Township grew by 6.57 %. Figure 4.3 presents Delaware County and its townships/municipalities.

The 2000 median household income in the townships in the southern part of the county was also higher when compared to other county township, with $115,904, $94,167 and $89,787 for Powell, Genoa, and Liberty, respectively. The median income in the northern part of the county is lower. For example, Marlboro and Ashley median incomes were $29,514 and $39,239, respectively. Table 4.1 presents 1990 and 2000 population and income data change for each township/municipality in the county.

4.4 Data Sources

Data were gathered from different sources. The basic data capturing housing characteristics, housing conditions and housing amenities were drawn from the Auditor’s Office database, including parcels in GIS shape files format. Socio-economic and demographic data were drawn from Census 2000 STF1, STF3, and CTPP (Census Transportation Planning Package) files produced by Bureau of the Census. Zoning and comprehensive planning data were obtained from the Delaware County Regional Planning Commission.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Berkshire</td>
<td>1,713</td>
<td>1,946</td>
<td>13.60</td>
<td>70,663</td>
</tr>
<tr>
<td>Berlin</td>
<td>1,978</td>
<td>3,313</td>
<td>67.49</td>
<td>69,028</td>
</tr>
<tr>
<td>Brown</td>
<td>1,164</td>
<td>1,290</td>
<td>10.82</td>
<td>63,456</td>
</tr>
<tr>
<td>Concord</td>
<td>3,363</td>
<td>4,088</td>
<td>21.56</td>
<td>79,169</td>
</tr>
<tr>
<td>Delaware</td>
<td>1,607</td>
<td>1,559</td>
<td>-2.99</td>
<td>60,372</td>
</tr>
<tr>
<td>Genoa</td>
<td>4,053</td>
<td>11,293</td>
<td>178.63</td>
<td>94,167</td>
</tr>
<tr>
<td>Harlem</td>
<td>3,391</td>
<td>3,762</td>
<td>10.94</td>
<td>55,080</td>
</tr>
<tr>
<td>Kingston</td>
<td>1,136</td>
<td>1,063</td>
<td>41.11</td>
<td>68,750</td>
</tr>
<tr>
<td>Liberty</td>
<td>3,790</td>
<td>9,182</td>
<td>142.27</td>
<td>89,787</td>
</tr>
<tr>
<td>Marlboro</td>
<td>213</td>
<td>227</td>
<td>6.57</td>
<td>29,514</td>
</tr>
<tr>
<td>Orange</td>
<td>3,789</td>
<td>12,464</td>
<td>228.95</td>
<td>74,612</td>
</tr>
<tr>
<td>Oxford</td>
<td>901</td>
<td>854</td>
<td>-5.22</td>
<td>47,100</td>
</tr>
<tr>
<td>Porter</td>
<td>1,345</td>
<td>1,696</td>
<td>26.10</td>
<td>70,949</td>
</tr>
<tr>
<td>Radnor</td>
<td>1,156</td>
<td>1,335</td>
<td>15.48</td>
<td>55,089</td>
</tr>
<tr>
<td>Scioto</td>
<td>1,698</td>
<td>2,122</td>
<td>24.97</td>
<td>54,706</td>
</tr>
<tr>
<td>Thomson</td>
<td>582</td>
<td>558</td>
<td>-4.12</td>
<td>57,639</td>
</tr>
<tr>
<td>Trenton</td>
<td>1,906</td>
<td>2,137</td>
<td>12.12</td>
<td>62,500</td>
</tr>
<tr>
<td>Troy</td>
<td>1,652</td>
<td>2,021</td>
<td>22.34</td>
<td>51,951</td>
</tr>
<tr>
<td>City of Delaware</td>
<td>20,030</td>
<td>25,243</td>
<td>26.03</td>
<td>46,030</td>
</tr>
<tr>
<td>Galena</td>
<td>361</td>
<td>305</td>
<td>-15.51</td>
<td>46,250</td>
</tr>
<tr>
<td>Sunbury</td>
<td>2,046</td>
<td>2,630</td>
<td>28.54</td>
<td>46,477</td>
</tr>
<tr>
<td>Shawnee Hills</td>
<td>423</td>
<td>419</td>
<td>-0.95</td>
<td>52,222</td>
</tr>
<tr>
<td>Powell</td>
<td>2,154</td>
<td>6,247</td>
<td>190.02</td>
<td>115,904</td>
</tr>
<tr>
<td>Ashley</td>
<td>1,059</td>
<td>1,216</td>
<td>14.83</td>
<td>39,239</td>
</tr>
</tbody>
</table>

Source: Census 2000

Table 4.1 Population and Income in Delaware County
4.4.1 Parcels Data

Parcels data from the Auditor’s Office, Delaware County, describe the land-use, housing characteristics, and ownership status of all 64,283 parcels in the county in the year 2000. The parcels data are represented by polygon shape files associated to tables in dbf format, appropriate for operations with ArcGIS software.

The parcels are classified by land use; for example, agriculture land uses are classified as Class 101 and single-family houses as Class 510 to Class 515. The major land-use categories are open space, agricultural, residential, commercial and industrial. Detailed information includes parcel shape, ownership, and transaction record. All transaction/sales are specified by Sales Date and Sales Amount. The data also include the township/municipality location of each parcel.

The strength of parcels data resides in the detailed housing characteristics, such as floor area, total numbers of rooms, bedrooms, and bathrooms, number of stories and garage capacity. This data also include information on features such as fireplace and basement. The distribution of parcels in Delaware County is illustrated in Figure 4.4.

In addition to parcels data, the master field data file from the Auditor’s Office also provides detailed information about land use and building use. The strength of the master field data resides in detailed informations on building use, based on the buildings field survey. Each individual building is referenced by its geographic coordinates, and the master field file also compiles the photos of the buildings. In this research, both the parcels and master field data files are used to locate single-family houses and amenities. Detailed descriptions of the parcels data and the master field coverage data are presented in Appendices A and B.
Figure 4.4 Parcels in Delaware County
4.4.2 Census 2000 STF1 Data File

The Census 2000 STF1 file is one of the main sources of socio-economic data for this research. It includes information on housing, demography, and other socio-economic characteristics of neighborhoods, available at the Census block level (the smallest Census geographical unit), which is sufficiently detailed geographically to allow for data conversion from blocks to circular buffers. The format of these data file is comma delimited, which is convenient for processing using the SAS statistical software. There were 2004 Census blocks in Delaware County in 2000.

4.4.3 Census 2000 STF3 Data File

In contrast to the block basis of STF1, the smallest geography in the Census 2000 STF3 file is the Census block groups (BG), which is made of a set of blocks. However, while only available at the block group level, this file includes much more detailed and varied socio-economic and demographic data than STF1. STF3 data are used to complement STF1 data when data are not available at the block level. The data format is the same as for STF1, and the data are also processed with SAS.

4.4.4 Census Transportation Planning Package (CTPP) 2000

The CTPP 2000 database is also an important data source for this research. In particular, it provides information about employment by place of residence and by place of work. CTPP Part 2 file, which provides workers/employment data by place of work in each block group, is to be used. STF1 and STF3 do not provide this information. The
employment data pertain to all block groups in the Columbus metropolitan area (Franklin, Delaware, Licking, Fairfield, Madison, Union, and Pickaway Counties).

4.4.5 Zoning and Comprehensive Plan Data

Zoning and comprehensive plan data were gathered from the Delaware County Regional Planning Commissions. These data are available in shape files and dbf tables. Zoning data presents the allocation of future land uses in each jurisdiction/township and municipality, including agricultural, residential, commercial, industrial, and open space land. All twenty-four townships and municipalities in Delaware County have formulated a zoning plan, but only ten of them have formulated a comprehensive plan.

Another set of data obtained from the Delaware County Regional Planning Commissions includes shape files for the boundaries of census tracts, census block groups, and census blocks in the county. Delaware County has 24 census tracts, 71 block groups, and 2004 blocks. The boundary data files will be useful in overlay processes. Figure 4.5 shows Census blocks and block groups in Delaware County.

4.4.6 GIS Data on Infrastructures and Amenities

The DALIS (Delaware Appraisal Land Information System) project from the Auditor’s Office provides data about amenities and infrastructures, that are represented by polygon shape files and associated tables in dbf format. These data include: cemetery, electric lines, gas lines, hospital, mobile homes, regional shopping malls, railroads, rivers, schools, town halls, fire stations, supermarkets, industrial sites, banks, golf courses, and
Figure 4.5  Census Blocks and Block Groups

Legend
- Block Group
- Block

0  2.5  5  10 Miles
post offices within the county. Figures 4.6 to 4.21 present these infrastructures and amenities.

4.5 Data Processing

This section presents the process of converting the raw data from different sources into six categories of variables: housing characteristics and conditions, amenities, accessibilities, socio-economic characteristics, land uses, and zoning. The difference between amenities and accessibilities variables has been explained in Chapter 3.

The processing tasks principles are as follows:

1. Housing characteristics and conditions variables are drawn from parcels data using SAS. The parcel identification numbers and the land-use class of a parcel are the basis for extracting these variables from the parcels file.

2. Amenity and accessibility variables are extracted from the master field data file, which provides location and size data. To measure distances to amenities, all amenity parcels and all housing parcels are converted into a centroid shape file. ArcGIS tools are used to compute distances between these centroids. For amenity variables, distances are computed as shortest distances. In the case of accessibility to employment centers, distances are computed from the housing unit to the centroid of a Census block group, since employment data is only available at the block group level.

3. Census block socio-economic characteristics are converted into circular buffers by using allocation factors. Similar processing applies to Census block group data.
Figure 4.7 Electric Transmission Lines in Delaware County
Figure 4.10 Mobile Home Parks in Delaware County
Figure 4.14  Schools in Delaware County
Figure 4.15  Town Halls in Delaware County
Figure 4.16  Fire Stations in Delaware County
Figure 4.17  Supermarkets in Delaware County
Figure 4.19  Banks in Delaware County
Figure 4.20 Golf Courses in Delaware County
4. Buffers are created using ArcGIS tools, with radii of 100 m, 400 m, and 1 mile, surrounding each housing unit. The radius of each circular buffer is measured from the centroid of the housing parcel.

5. Overlaying medium and large buffers with land-use and zoning shape files must be done in several steps, by month of sales, due to the large size of these files. Similarly, the process of overlaying zoning shape files must be done gradually.

One problem in data processing was the lack of information for areas outside Delaware County. This may impact amenity variables, because housing units close to Delaware County border do not have records on the nearest amenities located in Franklin County. However, this is not much a problem for amenity variables, since amenities in Delaware County are distributed relatively close to the selected housing units. Data show that the amenities included in this research have an average distances to the selected units of less than 2 miles. It is assumed that these amenities are appropriate substitutes for amenities in Franklin County.

The unavailable data for areas across the border of Franklin County is not much of a problem for socio-economic characteristics, land-use, and zoning variables, since public services (e.g. tax districts and zoning) for the selected housing units are under Delaware County jurisdiction.

4.5.1 Housing Characteristics and Conditions

Housing characteristics and conditions data are mainly extracted from the parcels file, including housing sale price (SALE), number of stories (STOR), number of rooms
(ROOM), number of bathrooms (BATH), total floor area of the house (FLOR), age of the house (AGES), number of years since remodeling of the house (MODE), total lot size (LOTS), availability of basement (BASE), number of fireplaces (FIRE), and garage capacity (GARA). The BASE variable is defined as equal to one if a basement is available in the house, and to zero otherwise. All data extracted from the parcels file are identified by the parcel number.

In 2000, there were 64,283 parcels, and 3144 of them had single-family housing that was sold between January and December 2000. These parcels are classified with codes 510-515. All these 3144 single-family housing units have been selected for the estimation of the hedonic price function. Figure 4.22 presents the distribution of the single-family housing units sold in 2000.

4.5.2 Amenity and Accessibility Variables

Amenity and accessibility variables have been computed based on the parcels data, master field data, and amenities data files from the Auditor’s Office. Amenities in the master field file are represented by building use in each parcel. The computation of the shortest linear distance of any individual housing unit to amenities was implemented with the Geoprocessing Wizards tools of ArcGIS. This distance is measured from the centroid of a parcel to the centroid of an amenity. The unit measurement is miles.

The shortest distance is used to measure the impacts of individual amenities, which are assumed to directly contribute to housing values due to the function of amenities, no matter the size of the amenities. For example, adjacency to railroad, to school, and to mobile-homes is assumed to affect single-family housing values.
Figure 4.22  Single-Family Housing Units Sold in 2000
The following amenities have had their distances to housing units computed: Cemetery (DICE), Electric Lines (DIEL), Gas Lines (DIGA), Hospital (DIHO), Mobile Homes (DIMH), Regional Shopping Mall Polaris (DIPO), Railroad (DIRA), River (DIRI), Schools (DISC), Town Hall (DITO), and Downtown Columbus (DIDW).

Places of works and employment centers are important factors in the decision of homeowners to reside in a particular place. Some homeowners may put a premium in residing close to their place of work, but others may not care. In order to account for employment centers in a regional context, an employment accessibility index has been computed, which measures the relative accessibility of a house to employment centers. Data from the CTPP are used to compute this index, based on a gravity formula applied to all employment centers in the Columbus Metropolitan Area, with:

\[ W_h = \sum_s \frac{E_s}{(D_{hs})^\alpha}; \]

where:

- \( W_h \) = Accessibility index for house \( h \) to all census block groups \( s \),
- \( E_s \) = Total employment in census block group \( s \),
- \( D_{hs} \) = Distance from house \( h \) to census block group \( s \),
- \( \alpha \) = distance exponent.

The employment \( (E_s) \) represents the aggregate employment in each block group \( s \) within Franklin, Delaware, Union, Licking, Fairfield, Madison, and Pickaway counties. The distance \( (D_{hs}) \) measures the straight-line distance between the centroid of housing parcel \( h \) and the centroid of block groups \( s \). The distance exponent \( \alpha \) may take any of the four values: 0.5, 1.0, 1.5, and 2.0.
Accessibility indices to Supermarkets (ASSU), Fire Stations (ASFI), Banks (ASBA), Golf Courses (ASGO), Post Offices (ASPO), and Industrial sites (ASIN), have been computed as follows:

\[ A^I_h = \sum_{j \in J} \frac{M_j}{(D_{hj})^\alpha} ; \]

where:

- \( A^I_h \) = Accessibility index for house \( h \) to amenity type \( J \),
- \( M_j \) = Scale of amenity \( j \),
- \( D_{hj} \) = Distance from house \( h \) to amenity \( j \),
- \( \alpha \) = distance exponential,
- \( j \) = Specific amenity of type \( J \).

The scale of amenity \( j \) (\( M_j \)) is measured as the area of the parcel with the amenity. Distances are computed as straight-line distances between parcel centroids. The distance exponent \( \alpha \) may take the values: 0.5, 1.0, 1.5, and 2.0.

### 4.5.3 Socioeconomic Characteristics of Neighborhood Buffers

#### 4.5.3.1 Overview

Socio-economic variables are used to characterize the neighborhood surrounding any single-family housing unit. These neighborhood factors are assumed to impact housing values. Three different kinds of neighborhood are considered: adjacent/immediate neighborhood, medium neighborhood, and large neighborhood. These three neighborhoods are represented by three circular buffers surrounding each housing unit, with radii of 100 meters, 400 meters, and 1 mile.

The use of these buffers differentiates this research from research using administrative boundaries/jurisdictions to define neighborhoods. A major advantage of
using such buffers is to define comparable neighborhood environments. Another advantage is to allow for scale diversity in neighborhood impacts on housing values.

4.5.3.2 Creation and Overlaying of Buffers

All buffers have been created with the ArcGIS 9 software, including ArcMap, ArcCatalog and Geoprocessing Wizards. In total, there are 3144 buffers of each type, as illustrated in Figure 4.23. Circular buffers created around the parcel’s centroid are depicted in Figure 4.24.

To compute the socio-economic characteristics of buffers, their coverages (in shape files format) were overlaid over other coverages by using ArcGIS. The intersected areas become parts of the buffer. For example, intersecting a 100 meter buffer with the parcels coverage results in capturing all parcels within the buffer.

The information file of the intersection coverage is processed with SAS, leading to summary characteristics for the whole buffer. Summarizing the intersected areas with regard to land-use class results in the proportions of land uses within each buffer: Agricultural (PAGR), Residential (PRES), Commercial (PCOM), Industrial (PIND), and Open Space (POSP). Summarizing all intersected areas with regard to the market values of housing units results in an Average Price of housing units in each buffer (AVPR).

4.5.3.3 Allocation Factors

The conversion of Census 2000 data in census blocks (STF1) and census block groups (STF3) format to neighborhood buffers requires the use of allocation factors.
Figure 4.23: Three Types of Buffers/Neighborhoods
Figure 4.24  Parcel's Centroid and Buffers
Two types of allocation factors can be computed: one based on area proportion and one based on number of bedrooms. The factors based on number of bedrooms is retained in conjunction with Census data, as it is more accurate for converting blocks or block groups data into buffers.

The coverages of residential single-family and multi-family housing units are intersected with Census block coverage (STF1) using ArcGIS. The resulting coverage is then intersected with the buffer coverage. These two intersection steps result in basic polygons that capture all the residential, census block, and buffer data. These data are used to compute the allocation factors. A similar process is used for Census block group data (STF3). The result of intersecting residential land use, Census blocks, and buffers is presented in Figure 4.25.

The allocation factors are computed as follows in the case of Census blocks:

A. **Allocation Factors based on Area**:

\[ F_{ij} = \frac{M_{ij}}{N_j} \quad \Sigma_i F_{ij} = 1 \]  

where:

- \( F_{ij} \) = Allocation factor based on area for buffer \( i \) and census block \( j \),
- \( M_{ij} \) = Intersection area of residential land use, census block \( j \), and buffer \( i \),
- \( N_j \) = Total area of census block \( j \).

B. **Allocation Factors based on the Number of Bedrooms**:

\[ F_{ks} = \frac{P_{ks}}{R_s} \quad \Sigma_s F_{ks} = 1 \]  

where:

- \( F_{ks} \) = Allocation factor based on the number of bedrooms for buffer \( k \) and census block \( s \),
- \( P_{ks} \) = Total number of bedrooms in the intersection area of residential land use, census block \( s \), and buffer \( k \),
- \( R_s \) = Total number of bedrooms in census block \( s \).
The computation of allocation factors for STF3 data is done using similar formulas.

4.5.3.4 Conversion of Census Block and Block Group Data to Buffer Data

Using the allocation factors $F_{ij}$ based on the number of bedrooms, the conversion of data $X_j$ for block $j$ to data $Y_i$ for buffer $i$ is done as follows:

$$Y_i = \sum_j F_{ij} X_j$$

As an example, assume that $X_j$ represents the Black population of Census block $j$. The product $F_{ij}X_j$ represents the share of this population located within buffer $i$. This product is then summed up over all Census blocks that intersect with buffer $i$ to obtain the Black population in buffer $i$, $Y_i$. A similar process is implemented to allocate block group data to buffers.

4.5.4 Zoning and Comprehensive Planning

Zoning and Comprehensive Planning are used as proxies for local land-use regulations. Although all townships and municipalities have developed zoning documents, only ten (Berkshire, Berlin, Brown, Concord, Genoa, Liberty, Porter, Shawnee Hills, Trenton, and Troy) have formulated comprehensive plans for their areas.

To convert townships/municipalities zoning data into buffers, the zoning maps, in shape file format, were intersected with the 100-meter, 400-meter, and 1-mile buffers around each of the 3144 properties, using ArcGIS. Then, buffer summarizing was done using SAS. This computation process resulted in the proportion of land zoned for Agricultural (PZAG), Residential with Low Density (PZLO), Residential with Medium
Density (PZME), Residential with High Density (PZHI), Commercial (PZCO), Industrial (PZIN), Open Space (PZOP), and Others (PZOT) for all three buffers separately.

Since comprehensive plans are available only in ten townships/municipalities, a binary variable is used as a proxy, with a value of 1 for those townships/municipalities that have formulated a comprehensive plan, and 0 for those that have not.
CHAPTER 5

MODEL SPECIFICATION

5.1 The Empirical Neighborhood Model

A major goal of this research is to analyze neighborhood impacts on housing values in suburban developments, using the hedonic price method. The concept of neighborhoods of different sizes (circular buffers with 100 m, 400 m, and 1 mile radii) is used, and it is assumed that housing values are a function of housing characteristics and conditions, amenities, accessibilities, socio-economic characteristics, land use, and regulation factors. The concept of variable neighborhood captures adjacent, medium-range, and longer-range effects, particularly as applied to socio-economic characteristics, existing land use, and zoning.

The estimated hedonic model is as follows:

\[
\ln V_i = \beta_0 + \beta_j H_{ij} + \beta_k A_{ik} + \beta_m C_{im} + \beta_p S_{pin} + \beta_q L_{qin} + \beta_r Z_{rin} + \varepsilon_i.
\]

where:

- \( \ln V_i \) = logarithm of sales price/transaction price of house \( i \),
- \( H_{ij} \) = characteristics and condition of attribute \( j \) for house \( i \),
- \( A_{ik} \) = amenity attribute \( k \) of house \( i \),
- \( C_{im} \) = accessibility attribute \( m \) for house \( i \),
- \( S_{pin} \) = socio-economic characteristics \( p \) of buffer \( n \) surrounding house \( i \),
- \( L_{qin} \) = land-use characteristics \( q \) of buffer \( n \) surrounding house \( i \),
- \( Z_{rin} \) = zoning characteristics \( r \) of buffer \( n \) surrounding house \( i \),

72
\( \beta_0, \beta_1, \beta_k, \beta_m, \beta_p, \beta_q, \beta_r \) = parameters to be estimated, 
\( \varepsilon_i \) = random error.

The following sections describe the specific variables, the rationale for selecting those variables, and their expected impacts on housing values. Table 5.1 presents the variables used in this research, combining common variables used in past studies with new variables characterizing the neighborhoods. Descriptive statistics for all the variables used in this research are presented in Table 5.2.

5.1.1 Housing Characteristics and Conditions Variables

Housing characteristics and condition variables include: availability of basements (BASE), age of housing unit (AGE), number of years since remodeling the house (REM), floor area (FLOR), number of full bathrooms (FBAT), number of fireplaces (FIRE), number of half bathrooms (HBAT), number of rooms (ROOM), garage capacity (GARA), and number of stories (STOR).

The availability of basement is a binary variable, which has a value of 1 if a basement is available, and 0 if not. The expected sign is positive: the availability of basement increases housing values.

The age of a housing unit and the number of years since remodeling the house are two proxies for housing condition. It is expected that newly remodeled houses carry a premium. There is no presumption about the sign of the age variable. A negative sign would indicate that newer housing units have a premium, and a positive sign that older houses have a premium, presumably because of their historical values.

The floor area, number of full bathrooms, number of fireplaces, number of half bathrooms, number of rooms, garage capacity and number of stories, are other proxies for
<table>
<thead>
<tr>
<th>Number</th>
<th>Variables</th>
<th>Definition</th>
<th>Author</th>
<th>GIS</th>
<th>Buffer Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PZAG</td>
<td>Land-Use Regulation</td>
<td>Percentage of Land Zoned for Agricultural and Open Space (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>PZRE</td>
<td>Land-Use Regulation</td>
<td>Percentage of Land Zoned for Low and Medium Density Residential (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>PZNR</td>
<td>Land-Use Regulation</td>
<td>Percentage of Land Zoned for Commercial, Industrial and Other Use (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>PLAN</td>
<td>Land-Use Regulation</td>
<td>Availability of Comprehensive Plan (0 or 1)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>PAGR</td>
<td>Existing Land Uses</td>
<td>Percentage of Agric. and Open Space Area (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>PRES</td>
<td>Existing Land Uses</td>
<td>Percentage of Residential Area (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>PCOM</td>
<td>Existing Land Uses</td>
<td>Percentage of Commercial Area (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
</tr>
<tr>
<td>10</td>
<td>REM</td>
<td>Housing Characteristics and Condition</td>
<td>Age of Remodeling the House (years)</td>
<td>Sucahyono (2006)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)

Table 5.1 Variables in the Hedonic Price Model
<table>
<thead>
<tr>
<th>Number</th>
<th>Variables</th>
<th>Definition</th>
<th>Author</th>
<th>GIS</th>
<th>Buffer Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cobb (1984)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kim et al. (2003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Crone (1983)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Singell &amp; Lillydahl (1990)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smith &amp; Kim (2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Nelson et al. (1997)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mathur et al. (2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>FBAT</td>
<td>Full Bathrooms (unit)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paterson &amp; Boyle (2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>HBAT</td>
<td>Half Bathrooms (unit)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paterson &amp; Boyle (2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>FIRE</td>
<td>Number of Fireplaces (unit)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paterson &amp; Boyle (2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kestens et al. (2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mathur et al. (2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bowen et al. (2001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Michaels &amp; Smith (1990)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ROOM</td>
<td>Number of Rooms (unit)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Linneman (1980)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dubin &amp; Sung (1990)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cassel &amp; Meddelsohn (1985)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kim et al. (2003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Linneman (1981)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Follain &amp; Malpezzi (1981)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Jud (1980)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>GARA</td>
<td>Garage capacity (car)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paterson &amp; Boyle (2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Brasington &amp; Hite (2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Jud (1980)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Li &amp; Brown (1980)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>STOR</td>
<td>Number of Stories (floor)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Kestens et al. (2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dubin &amp; Sung (1990)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 5.1: Continued

<table>
<thead>
<tr>
<th>Number</th>
<th>Variables</th>
<th>Definition</th>
<th>Author</th>
<th>GIS</th>
<th>Buffer Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>DICE</td>
<td>Closest Distance to Cemetery (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>19</td>
<td>DIEL</td>
<td>Closest Distance to Electric Lines (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>20</td>
<td>DIGA</td>
<td>Closest Distance to Gas Lines (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>21</td>
<td>DIHO</td>
<td>Closest Distance to Hospital (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>22</td>
<td>DIMH</td>
<td>Closest Distance to Mobile Home Parks (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>23</td>
<td>DIPO</td>
<td>Closest Distance to Regional Shopping Mall (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>24</td>
<td>DIRA</td>
<td>Closest Distance to Railroads (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>25</td>
<td>DIRI</td>
<td>Closest Distance to River (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>26</td>
<td>DISC</td>
<td>Closest Distance to School (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>27</td>
<td>DITO</td>
<td>Closest Distance to Townhall (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>28</td>
<td>DIDW</td>
<td>Closest Distance to Downtown Columbus (mile)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>29</td>
<td>ASEM</td>
<td>Accessibilities to Regional Employment Centers</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>30</td>
<td>ASFI</td>
<td>Accessibilities to Fire Stations</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>31</td>
<td>ASSU</td>
<td>Accessibilities to Supermarkets</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>32</td>
<td>ASIN</td>
<td>Accessibilities to Industrial Sites</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>33</td>
<td>ASBA</td>
<td>Accessibilities to Banks</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>34</td>
<td>ASGO</td>
<td>Accessibilities to Golf Courses</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>35</td>
<td>ASPO</td>
<td>Accessibilities to Post Offices</td>
<td>Sucahyono (2006)</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>36</td>
<td>TAXS</td>
<td>Property tax rate (%)</td>
<td>Sucahyono (2006)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
### Table 5.1: Continued

<table>
<thead>
<tr>
<th>Number</th>
<th>Variables</th>
<th>Definition</th>
<th>Author</th>
<th>GIS</th>
<th>Buffer Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>DENS</td>
<td>Net population density (person/acre)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>38</td>
<td>AVPR</td>
<td>Average Price of Housing in Neighborhood ($)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>39</td>
<td>PWOW</td>
<td>Percentage of White Population in Owner Occupied Housing (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>40</td>
<td>PBOW</td>
<td>Percentage of Black Population in Owner Occupied Housing (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>41</td>
<td>PAOW</td>
<td>Percentage of Asian Population in Owner Occupied Housing (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>42</td>
<td>WONG</td>
<td>Percentage of Population 16 Years and Over Who Works in Non Agricultural Activities (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>43</td>
<td>WOOU</td>
<td>Percentage of Population Who Works Outside the County (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>44</td>
<td>DCAR</td>
<td>Percentage of Population Who Drives a Car to Work (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>45</td>
<td>EDHS</td>
<td>Percentage of Population with High School Educational Attainment (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>46</td>
<td>EDBA</td>
<td>Percentage of Population with Bachelor's Degree (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>47</td>
<td>CIUS</td>
<td>Percentage of Homeowners who are U.S Citizen and born in the U.S (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>48</td>
<td>INCB</td>
<td>Percentage Households with Medium Income (%) (US $ 50,000 - US $ 99,000)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>49</td>
<td>INCC</td>
<td>Percentage Households with High Income (%) (US $ 100,000 and over)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>50</td>
<td>SIZB</td>
<td>Percentage of Households with more than 2 persons (%)</td>
<td>Sucahyono (2006)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
## Table 5.2 Descriptive Statistics of All Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Housing Characteristics and Conditions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SALE</td>
<td>3144</td>
<td>226935</td>
<td>172961</td>
<td>1156</td>
<td>5900000</td>
</tr>
<tr>
<td>BASE</td>
<td>3144</td>
<td>0.94466</td>
<td>0.22869</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>AGE</td>
<td>3144</td>
<td>11.81075</td>
<td>23.76300</td>
<td>1.0000</td>
<td>165.0000</td>
</tr>
<tr>
<td>REM</td>
<td>3144</td>
<td>9.86641</td>
<td>20.19469</td>
<td>1.0000</td>
<td>146.0000</td>
</tr>
<tr>
<td>FLOR</td>
<td>3144</td>
<td>2345</td>
<td>859.27370</td>
<td>748.0000</td>
<td>10106</td>
</tr>
<tr>
<td>HBAT</td>
<td>3144</td>
<td>0.79835</td>
<td>0.45969</td>
<td>0.0000</td>
<td>3.0000</td>
</tr>
<tr>
<td>FIRE</td>
<td>3144</td>
<td>0.83206</td>
<td>0.47573</td>
<td>0.0000</td>
<td>5.0000</td>
</tr>
<tr>
<td>ROOM</td>
<td>3144</td>
<td>7.35623</td>
<td>1.47122</td>
<td>0.0000</td>
<td>16.0000</td>
</tr>
<tr>
<td>FBAT</td>
<td>3144</td>
<td>2.07156</td>
<td>0.64335</td>
<td>0.0000</td>
<td>6.0000</td>
</tr>
<tr>
<td>GARA</td>
<td>3144</td>
<td>2.05534</td>
<td>0.79478</td>
<td>0.0000</td>
<td>6.0000</td>
</tr>
<tr>
<td>STOR</td>
<td>3144</td>
<td>1.79246</td>
<td>0.39691</td>
<td>0.0000</td>
<td>2.8000</td>
</tr>
<tr>
<td><strong>2. Amenities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DICE</td>
<td>3144</td>
<td>1.02889</td>
<td>0.53205</td>
<td>0.02700</td>
<td>2.99300</td>
</tr>
<tr>
<td>DIEI</td>
<td>3144</td>
<td>1.18114</td>
<td>1.18360</td>
<td>0.00400</td>
<td>9.80400</td>
</tr>
<tr>
<td>DIGA</td>
<td>3144</td>
<td>1.97515</td>
<td>1.98999</td>
<td>0.00300</td>
<td>8.83800</td>
</tr>
<tr>
<td>DIHO</td>
<td>3144</td>
<td>9.55320</td>
<td>4.57717</td>
<td>0.13300</td>
<td>20.38200</td>
</tr>
<tr>
<td>DIMH</td>
<td>3144</td>
<td>1.93988</td>
<td>1.18366</td>
<td>0.00700</td>
<td>5.23900</td>
</tr>
<tr>
<td>DIPO</td>
<td>3144</td>
<td>6.64365</td>
<td>4.33725</td>
<td>0.41900</td>
<td>23.67860</td>
</tr>
<tr>
<td>DIRA</td>
<td>3144</td>
<td>2.59187</td>
<td>2.50771</td>
<td>0.17000</td>
<td>13.64500</td>
</tr>
<tr>
<td>DIRI</td>
<td>3144</td>
<td>1.12466</td>
<td>0.78929</td>
<td>0.00600</td>
<td>5.48100</td>
</tr>
<tr>
<td>DISC</td>
<td>3144</td>
<td>1.01412</td>
<td>0.98537</td>
<td>0.05300</td>
<td>8.11100</td>
</tr>
<tr>
<td>DITO</td>
<td>3144</td>
<td>2.12510</td>
<td>0.97077</td>
<td>0.08600</td>
<td>4.61300</td>
</tr>
<tr>
<td>DIDW</td>
<td>3144</td>
<td>20.49487</td>
<td>4.81501</td>
<td>14.94400</td>
<td>38.20800</td>
</tr>
<tr>
<td><strong>3. Accessibilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASEM3</td>
<td>3144</td>
<td>0.29728</td>
<td>0.11297</td>
<td>0.08270</td>
<td>0.50496</td>
</tr>
<tr>
<td>ASFI1</td>
<td>3144</td>
<td>47178</td>
<td>7248</td>
<td>31253</td>
<td>87918</td>
</tr>
<tr>
<td>ASSU1</td>
<td>3144</td>
<td>78591</td>
<td>18627</td>
<td>37758</td>
<td>149843</td>
</tr>
<tr>
<td>ASIN1</td>
<td>3144</td>
<td>437621</td>
<td>74689</td>
<td>285428</td>
<td>782517</td>
</tr>
<tr>
<td>ASBA1</td>
<td>3144</td>
<td>15153</td>
<td>3397</td>
<td>7536</td>
<td>35337</td>
</tr>
<tr>
<td>ASGO3</td>
<td>3144</td>
<td>31.90961</td>
<td>37.35144</td>
<td>3.79500</td>
<td>911.83100</td>
</tr>
<tr>
<td>ASP03</td>
<td>3144</td>
<td>0.30672</td>
<td>0.52337</td>
<td>0.02000</td>
<td>12.77300</td>
</tr>
<tr>
<td><strong>4. Socio-economic Characteristics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Adjacent Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PWOW1</td>
<td>3144</td>
<td>91.79012</td>
<td>14.14984</td>
<td>0.0000</td>
<td>100.0000</td>
</tr>
<tr>
<td>PBOW1</td>
<td>3144</td>
<td>62.91133</td>
<td>43.89549</td>
<td>0.0000</td>
<td>100.0000</td>
</tr>
<tr>
<td>PAOW1</td>
<td>3144</td>
<td>64.93357</td>
<td>45.39015</td>
<td>0.0000</td>
<td>100.0000</td>
</tr>
</tbody>
</table>

(continued)
Table 5.2: Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>WONG_1</td>
<td>3144</td>
<td>99.47327</td>
<td>1.03109</td>
<td>93.55000</td>
<td>100.00000</td>
</tr>
<tr>
<td>WOOU1</td>
<td>3144</td>
<td>67.53578</td>
<td>13.69137</td>
<td>22.04000</td>
<td>100.00000</td>
</tr>
<tr>
<td>DCAR1</td>
<td>3144</td>
<td>93.80625</td>
<td>3.78882</td>
<td>58.82000</td>
<td>100.00000</td>
</tr>
<tr>
<td>CIUS1</td>
<td>3144</td>
<td>96.46628</td>
<td>2.75584</td>
<td>88.35000</td>
<td>100.00000</td>
</tr>
<tr>
<td>EDHS_1</td>
<td>3144</td>
<td>20.25569</td>
<td>10.96784</td>
<td>0</td>
<td>57.75000</td>
</tr>
<tr>
<td>EDBA_1</td>
<td>3144</td>
<td>32.62415</td>
<td>11.34418</td>
<td>3.91000</td>
<td>51.33000</td>
</tr>
<tr>
<td>INCB1</td>
<td>3144</td>
<td>42.19615</td>
<td>11.84478</td>
<td>17.04000</td>
<td>65.17000</td>
</tr>
<tr>
<td>INCC1</td>
<td>3144</td>
<td>33.58171</td>
<td>20.47454</td>
<td>0</td>
<td>73.78000</td>
</tr>
<tr>
<td>DENS1</td>
<td>3144</td>
<td>6.17660</td>
<td>2.05864</td>
<td>0.10000</td>
<td>9.90000</td>
</tr>
<tr>
<td>AVPR1</td>
<td>3144</td>
<td>241804</td>
<td>126309</td>
<td>30914</td>
<td>1389150</td>
</tr>
<tr>
<td>TAXS</td>
<td>3144</td>
<td>47.48884</td>
<td>5.41017</td>
<td>35.16160</td>
<td>59.91680</td>
</tr>
<tr>
<td>SIZB_1</td>
<td>3144</td>
<td>50.91823</td>
<td>15.78062</td>
<td>0</td>
<td>100.00000</td>
</tr>
</tbody>
</table>

B. Medium Neighborhood

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWOW2</td>
<td>3144</td>
<td>91.22999</td>
<td>12.57575</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>PBOW2</td>
<td>3144</td>
<td>72.81194</td>
<td>37.26868</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>PAOW2</td>
<td>3144</td>
<td>75.25633</td>
<td>39.41682</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>INCB2</td>
<td>3144</td>
<td>42.08488</td>
<td>11.23423</td>
<td>17.04000</td>
<td>65.17000</td>
</tr>
<tr>
<td>INCC2</td>
<td>3144</td>
<td>33.54641</td>
<td>20.13290</td>
<td>0</td>
<td>73.78000</td>
</tr>
<tr>
<td>CIUS2</td>
<td>3144</td>
<td>96.48868</td>
<td>2.62739</td>
<td>88.35000</td>
<td>100.00000</td>
</tr>
<tr>
<td>EDHS_2</td>
<td>3144</td>
<td>20.32580</td>
<td>10.87125</td>
<td>0</td>
<td>48.25000</td>
</tr>
<tr>
<td>EDBA_2</td>
<td>3144</td>
<td>32.65574</td>
<td>11.23423</td>
<td>0</td>
<td>51.33000</td>
</tr>
<tr>
<td>WONG_2</td>
<td>3144</td>
<td>99.47629</td>
<td>0.99295</td>
<td>93.55000</td>
<td>100.00000</td>
</tr>
<tr>
<td>WOOU2</td>
<td>3144</td>
<td>67.42498</td>
<td>13.61521</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>DCAR2</td>
<td>3144</td>
<td>93.73018</td>
<td>3.55301</td>
<td>61.50000</td>
<td>100.00000</td>
</tr>
<tr>
<td>DENS2</td>
<td>3144</td>
<td>1.98668</td>
<td>1.65628</td>
<td>0.02000</td>
<td>15.56000</td>
</tr>
<tr>
<td>AVPR2</td>
<td>3144</td>
<td>234526</td>
<td>103642</td>
<td>75136</td>
<td>513682</td>
</tr>
<tr>
<td>TAXS</td>
<td>3144</td>
<td>47.48884</td>
<td>5.41017</td>
<td>35.16160</td>
<td>59.91680</td>
</tr>
<tr>
<td>SIZB_2</td>
<td>3144</td>
<td>50.91823</td>
<td>15.78062</td>
<td>0</td>
<td>100.00000</td>
</tr>
</tbody>
</table>

C. Large Neighborhood

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>PWOW3</td>
<td>3144</td>
<td>87.25801</td>
<td>11.82075</td>
<td>44.26000</td>
<td>99.47000</td>
</tr>
<tr>
<td>PBOW3</td>
<td>3144</td>
<td>74.32645</td>
<td>28.60357</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>PAOW3</td>
<td>3144</td>
<td>78.65064</td>
<td>29.97211</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>WONG_3</td>
<td>3144</td>
<td>99.48128</td>
<td>0.82423</td>
<td>93.55000</td>
<td>100.00000</td>
</tr>
<tr>
<td>WOOU3</td>
<td>3144</td>
<td>67.12147</td>
<td>13.60840</td>
<td>33.05000</td>
<td>82.40000</td>
</tr>
<tr>
<td>DCAR3</td>
<td>3144</td>
<td>93.68166</td>
<td>2.20156</td>
<td>82.24000</td>
<td>100.00000</td>
</tr>
<tr>
<td>CIUS3</td>
<td>3144</td>
<td>96.71496</td>
<td>2.03374</td>
<td>91.86000</td>
<td>100.00000</td>
</tr>
<tr>
<td>EDHS_3</td>
<td>3144</td>
<td>20.90549</td>
<td>10.33424</td>
<td>7.94000</td>
<td>47.32000</td>
</tr>
<tr>
<td>EDBA_3</td>
<td>3144</td>
<td>32.34484</td>
<td>10.56047</td>
<td>6.35000</td>
<td>50.06000</td>
</tr>
<tr>
<td>INCB3</td>
<td>3144</td>
<td>41.32228</td>
<td>8.50524</td>
<td>21.40000</td>
<td>65.17000</td>
</tr>
<tr>
<td>INCC3</td>
<td>3144</td>
<td>33.10440</td>
<td>18.32446</td>
<td>3.04000</td>
<td>67.14000</td>
</tr>
<tr>
<td>DENS3</td>
<td>3144</td>
<td>1.89458</td>
<td>1.34466</td>
<td>0.08000</td>
<td>7.48000</td>
</tr>
<tr>
<td>AVPR3</td>
<td>3144</td>
<td>229536</td>
<td>78976</td>
<td>75136</td>
<td>513682</td>
</tr>
<tr>
<td>TAXS</td>
<td>3144</td>
<td>47.48884</td>
<td>5.41017</td>
<td>35.16160</td>
<td>59.91680</td>
</tr>
<tr>
<td>SIZB_3</td>
<td>3144</td>
<td>49.01626</td>
<td>10.66409</td>
<td>25.65000</td>
<td>72.24000</td>
</tr>
</tbody>
</table>

(continued)
Table 5.2: Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Land Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Adjacent Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAGR_1</td>
<td>3144</td>
<td>29.08243</td>
<td>14.86990</td>
<td>0</td>
<td>92.62000</td>
</tr>
<tr>
<td>PRES1</td>
<td>3144</td>
<td>69.33669</td>
<td>15.72367</td>
<td>7.38000</td>
<td>100.00000</td>
</tr>
<tr>
<td>PCOM1</td>
<td>3144</td>
<td>1.47171</td>
<td>5.98569</td>
<td>0</td>
<td>55.65000</td>
</tr>
<tr>
<td>B. Medium Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAGR_2</td>
<td>3144</td>
<td>42.08633</td>
<td>16.15564</td>
<td>3.55000</td>
<td>97.14000</td>
</tr>
<tr>
<td>PRES2</td>
<td>3144</td>
<td>52.81183</td>
<td>16.16143</td>
<td>2.85000</td>
<td>96.45000</td>
</tr>
<tr>
<td>PCOM2</td>
<td>3144</td>
<td>4.43310</td>
<td>8.56318</td>
<td>0</td>
<td>66.52000</td>
</tr>
<tr>
<td>C. Large Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAGR_3</td>
<td>3144</td>
<td>53.53152</td>
<td>15.49169</td>
<td>22.36000</td>
<td>96.69000</td>
</tr>
<tr>
<td>PRES3</td>
<td>3144</td>
<td>35.25017</td>
<td>12.60612</td>
<td>3.31000</td>
<td>72.05000</td>
</tr>
<tr>
<td>PCOM3</td>
<td>3144</td>
<td>9.08518</td>
<td>10.84518</td>
<td>0</td>
<td>58.99000</td>
</tr>
<tr>
<td>6. Zoning and Comprehensive Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Adjacent Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN</td>
<td>3144</td>
<td>0.47901</td>
<td>0.49964</td>
<td>0</td>
<td>1.00000</td>
</tr>
<tr>
<td>PZAG_1</td>
<td>3144</td>
<td>24.24422</td>
<td>27.20878</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>PZRE_1</td>
<td>3144</td>
<td>57.20017</td>
<td>35.08524</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>PZNR_1</td>
<td>3144</td>
<td>15.86524</td>
<td>27.25897</td>
<td>0</td>
<td>96.38000</td>
</tr>
<tr>
<td>B. Medium Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN</td>
<td>3144</td>
<td>0.47901</td>
<td>0.49964</td>
<td>0</td>
<td>1.00000</td>
</tr>
<tr>
<td>PZAG_2</td>
<td>3144</td>
<td>25.51613</td>
<td>28.04918</td>
<td>0</td>
<td>100.00000</td>
</tr>
<tr>
<td>PZRE_2</td>
<td>3144</td>
<td>53.04901</td>
<td>32.14157</td>
<td>0</td>
<td>99.25000</td>
</tr>
<tr>
<td>PZNR_2</td>
<td>3144</td>
<td>17.72403</td>
<td>25.79662</td>
<td>0</td>
<td>93.96000</td>
</tr>
<tr>
<td>C. Large Neighborhood</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN</td>
<td>3144</td>
<td>0.47901</td>
<td>0.49964</td>
<td>0</td>
<td>1.00000</td>
</tr>
<tr>
<td>PZAG_3</td>
<td>3144</td>
<td>31.45424</td>
<td>27.74809</td>
<td>1.97000</td>
<td>100.00000</td>
</tr>
<tr>
<td>PZRE_3</td>
<td>3144</td>
<td>44.44808</td>
<td>27.08697</td>
<td>0</td>
<td>97.76000</td>
</tr>
<tr>
<td>PZNR_3</td>
<td>3144</td>
<td>21.16113</td>
<td>20.93897</td>
<td>0</td>
<td>86.37000</td>
</tr>
</tbody>
</table>
the structural characteristics of the house. The expected signs for floor area, number of full bathrooms, number of fireplaces, number of half bathrooms, and garage capacity are positive, implying that the bigger or the more luxurious the house, the higher the price. The expected signs of the number of rooms and the number of stories are negative. All else being equal, a housing unit with a larger number of rooms has necessarily smaller rooms, which may make the house uncomfortable. The floor area and garage capacity are important variables for housing units in the suburbs, as they characterize the specific demand for single-family units. It is assumed that people move to the suburbs to have larger houses.

There are some housing variables that have been commonly used in past studies, but are not included in this research due to irrelevancy of the variables and unavailability of data, such as availability of dishwashers [Kestens et al. (2004), Cobb (1984)], dining rooms [Dinan and Miranowski (1989)], and swimming pools [Li and Brown (1980), Dubin and Sung (1990)].

5.1.2 Amenity Variables

Amenity variables include public services, utilities and infrastructures. Eleven variables are developed to capture the positive and negative effects of distance to the following amenities: cemetery (DICE), electric lines (DIEL), gas lines (DIGA), hospitals (DIHO), mobile homes parks (DIMH), the regional shopping mall Polaris (DIPO), railroads (DIRA), rivers (DIRI), schools (DISC), town halls (DITO), and downtown Columbus (DIDW).
The distances to amenities from an individual single-family housing unit are used as proxies for adjacencies to these amenities. The distance is measured as the straight line distance from an individual housing unit to an amenity. In contrast to most past hedonic studies, this research explores the use of GIS to analyze the characteristics of amenities. Paterson and Boyle (2002) argue that GIS, in recent years, have facilitated the measurements of distance and size and shape of amenities. In comparison, past hedonic studies [Linneman (1980), Nelson (1978), and Cobb (1984)] generally measured distances to amenities by creating grid cells and using dummy variables for vicinity to amenities.

The expectation for the signs of electric lines and gas lines in the model are positive, indicating that greater distances to these amenities increase housing values, since adjacency to these amenities produces negative externalities. The expected signs of railroads and hospitals are also positive, implying that being adjacent to them reduces housing value, as railroads and hospitals generate contamination risk and traffic noise. The expected sign of the mobile-home variable is positive: greater distances to mobile home parks increase single-family housing values. Adjacency to cemetery, river, and school may have either a positive or a negative sign. A positive sign for the cemetery variable indicates that adjacency may create negative externalities (bad image), while a negative sign implies that this amenity provides benefits (quietness). A positive sign for the river variable implies that adjacency to rivers decreases housing value, due to flood risks, while a negative sign indicates that being adjacent to a river increases housing value due to such benefits as a pleasant scenery. A positive sign for the school variable indicates that distance to schools increases housing value, because schools create
negative externalities, such as noise and traffic congestion. A negative sign implies that being adjacent to a school carries a premium due to the proximity for families with children.

The distances to the regional shopping mall Polaris, to town halls, and to downtown Columbus are used as proxies to capture land rent, a major contributor to housing value. The expected signs of these three variables are negative: greater distances decrease single-family housing values as a result of a decreasing land rent. Typically, land rents in central business districts (CBDs) are steep, and become flatter with greater distance from CBDs. In suburban areas, the land rent around township centers, town halls, and regional shopping mall may perform a similar role to the land rent in central city CBDs.

5.1.3 Accessibility Variables

Seven variables characterizing geographical accessibilities are considered: employment centers (ASEM), banks (ASBA), fire stations (FIRE), supermarkets (ASSU), industrial sites (ASIN), golf courses (ASGO), and post offices (ASGO). Gravity indices are used to measure weighted accessibilities to these activities. These indices are directly related to the sizes of the activities and inversely related to the distances to them. The detailed formulas for these indices have been presented in Chapter 4.

The accessibility to employment centers may have either a positive or a negative effect. A positive effect implies that closeness to employment centers does matter to suburban homeowners and increases housing values, while a negative effect would indicate that closeness decreases housing values due to such externalities as traffic.
Accessibilities to fire stations and supermarkets are expected to have positive effects. A good access to fire stations provides a safe environment, and a good access to supermarkets provides convenience for homeowners in the suburbs for buying groceries and other daily commodities. Supermarkets may provide other benefits, such as job opportunities, increased tax revenues, and local economic growth.

The accessibility to industrial sites may have either a positive or a negative effect. A negative effect indicates that such sites generate negative externalities, such as solid waste and air pollution, while a positive effect implies that they provide benefits, such as job opportunities and public facilities. Accessibilities to banks and post offices are predicted to have negative effects, due to traffic and noise. It is assumed that people dependency on post offices has decreased as a result of technology innovations that makes people prefer to do business and sales transactions on line through the Internet.

5.1.4 Socio-economic Variables

Socio-economic characteristics are captured on a neighborhood basis, utilizing three different circular buffer sizes. This makes this research different from past hedonic studies. Geoghegan et al (1997), Irwin et al (2002), Dubin and Sung (1990), Li and Brown (1980), Dinan and Miranowski (1989), Cassel and Mendehlson (1985) used jurisdictional boundaries, ZIP codes, census tracks and census block groups to capture socio-economic characteristics in their models.

Conversion of socio-economic characteristics into circular buffers is intended to accurately capture neighborhood characteristics. Fourteen socio-economic variables are considered: proportion of White people (PWow), proportion of Black people (PBow),
The expected effect of the proportion of White people is positive, implying that a higher proportion of White in the neighborhood increases housing values. The proportions of Blacks and Asians may have either a positive or a negative effect. A positive effect implies that homeowners who are minorities increase housing value, while a negative effect implies the opposite. Variables associated with race composition are intended to test the effect of segregation of homeowners in the suburbs, which commonly happens in the U.S. The proportions of Blacks and Asians are used because of their substantial shares in Ohio.

The average price of single-family housing units in the neighborhood is expected to have a positive effect on individual housing values; The higher the average price of housing in the neighborhood, the higher the price of the individual house.

The proportions of middle-income and high-income people are intended to capture the effects of wealth on housing values. Middle-income and high-income people
are assumed to make up the major share in the suburbs, and the share of low-income population is not included in the model. The existence of low-income people has been represented by such variables as adjacency to mobile homes and high-school educational attainment. Middle-income people in the model are defined as those with income from $50,000 to $100,000. High-income people are those with income of more than $100,000. It is expected that the proportions of middle-income and high-income people have positive effects on housing values.

Besides income level, the educational level is also an important variable, a proxy for the level of homeowners’ wealth. Two variables are used to measure the educational level: the proportion of the population with only high-school educational attainment, and the proportion of the population with bachelor degrees. The expected effect of the proportion of population with only a high-school educational attainment is negative, while the expected effect of a bachelor degree is positive. The proportions of population with master’s degree and with an educational level lower than high-school are not included, as they are very small.

The property tax rate is used as a proxy for governmental socio-economic policy in the area. Property taxes are used by governments to internalize externalities and to collect revenues for public services. It is assumed that this variable has a negative sign in the model, because an increase in the property tax rate reduces housing values. Locations with low property tax rates attract demand for housing, and therefore increase housing values.

The share of households with more than two persons is intended to capture homeowners with families, who are typically residing in the suburbs. This variable is
expected to have a positive sign, because of the preferences of homeowners with families to live in communities with the same characteristics.

Three specific socio-economic variables are employed to capture the effect of suburbanization on housing values: the proportion of people who work in non-agricultural activities, the proportion of people who drive cars to work, and the proportion of people who work outside the county. The expected signs of these three variables are positive.

5.1.5 Land-Use Variables

In order to capture the effect of existing land uses on housing values, three variables are used: the share of agricultural land use and open space, the share of residential land use, and the share of commercial land use. Agriculture and open space are classified into one group, as undeveloped area. Residential land use captures all residential uses, including multi-family housing. Industrial land use is not included, since it has been represented by the accessibility to industrial sites. Commercial land use is included to capture the impacts of existing developments on housing values in the suburbs.

Agriculture and open space may have either a negative or a positive impact on housing values. A positive effect points to such benefits as a good view in the neighborhood. A negative effect points to externalities, such as bad odor and chemical pollution. The negative impact is assumed to particularly take place in adjacent neighborhoods.
Residential land use is expected to have a positive sign, implying benefits to the neighborhood, such as agglomeration effects and provision of public services. Commercial land use is employed to capture the effect of business areas in the neighborhood. It is expected that an increase of commercial land use leads to increasing housing values, due to increasing land rents in business districts. Although commercial land uses may generate externalities, such as traffic, the magnitude of commercial land use in the suburbs, which is not associated with high rise buildings, provides an aggregate positive effect to the neighborhood.

Other land-use variables have been used in past studies, such as diversity and fragmentation indices [(Geoghegan et al. (1997)] and tree density in residential areas [Kestens et al. (2004)]. These ecological variables were not considered in this research.

5.1.6 Zoning and Comprehensive Planning Variables

Zoning and comprehensive planning variables are important factors capturing future spatial development in the neighborhood. Four variables are considered: availability of comprehensive planning (PLAN), proportion of land zoned for agricultural and open space (PZAG), proportion of land zoned for low and medium residential land uses (PZRE), proportion of land zoned for commercial, industrial and other uses (PZNR).

The comprehensive planning variable used in the model is a binary variable. It has a value of 1 if comprehensive planning has been enacted in the neighborhood, and 0 if not. Comprehensive planning guides multi-sector development in the neighborhood, including the maximum density of residential activities in the future. This characteristic makes comprehensive planning differ from zoning, which mainly allocates future land
uses. It is expected that comprehensive planning has a positive effect on housing value, implying that availability of such planning increases housing values as a result of the certainty surrounding future developments in the neighborhood.

Land zoned for agriculture and open space in the neighborhood may have either a positive or a negative effect. A positive effect indicates benefits, such as better environmental conditions. A negative effect implies disadvantages in restricting development in the future.

Land zoned for low and medium residential use is expected to have a positive effect on housing value, because of the certainty of similar development in the future. This certainty is a major benefit of zoning for homeowners and developers. Land zoned for low and medium residential use provides certainty on the supply side of housing markets. Since the model is focused on single-family housing, the proportion of land zoned for multi-family housing units is not included.

The proportion of land zoned for non-residential development (commercial, industrial, and other uses) is expected to have a positive impact on single-family housing value because of increasing land rents in the neighborhood, and also because of the certainty of future development with compatible land uses.

5.2 Functional Form

Several functional forms are considered: linear, semi-log, double-log, and Box-Cox transformation. The main advantage of the linear, semi-log and log-linear functional forms is the transparency of the relationships between the marginal attributes prices and the parameters of the function. The linear function implies constant marginal attributes.
In the semi-log model, the dependent variable is transformed into the natural logarithm form. The semi-log function implies that the marginal attributes prices are proportional to the price of housing. In the double-log form, both dependent and independent variables are transformed into their natural logarithm forms. However, the independent variables that contain 0 values are not transformed into natural logarithm forms. A cross-sectional model is used, with single-family housing sales data for the year 2000, which has been selected to match the availability of Census 2000 data to be converted into neighborhoods characteristics.

Guldmann (1985) has used the Box-Cox transformation to derive the best model explaining the data, by maximizing the log-likelihood function. There are two parameters used in the transformation model: one for the dependent variables and one for all the independent variables.

With the Box-Cox estimation procedure, the empirical model can be rewritten as follows:

$$ V (\theta) = \beta_0 + \beta_j H_{ij} (\lambda) + \beta_k A_{ik} (\lambda) + \beta_m C_{im} (\lambda) + \beta_p S_{pin} (\lambda) + \beta_q L_{qin} (\lambda) + $$

$$ \beta_r Z_{rin} (\lambda) + \varepsilon $$

where $\varepsilon$ is assumed to be a normally distributed error term, with $E(\varepsilon) = 0$ and $E(\varepsilon \varepsilon') = \sigma^2 I$, and the dependent and independent variables are related to the corresponding original variables by the Box-Cox transformation defined by:

$$ X(\lambda) = (X^\lambda - 1) / \lambda, \quad \lambda \neq 0, $$

$$ X(\lambda) = \ln X, \quad \lambda \rightarrow 0. $$

90
The optimal values of the parameters $\theta$ and $\lambda$ maximize the log-likelihood function:

$$L_{\text{max}}(\theta, \lambda) = -\frac{N}{2} \left[ \ln(2\pi) + 1 \right] -\frac{N}{2} \ln[\sigma^2(\theta, \lambda)] + (\theta-1) \sum_{i=1}^{N} \ln V_i,$$

where $N$ is the sample size, $\sigma^2(\theta, \lambda)$ is the error variance estimated from the regression of $V(\theta)$ on the transformed independent variables, and $V_i$ is the $i$th observation of the dependent variable.
CHAPTER 6

EMPIRICAL RESULTS

Two models are estimated to analyze the impacts of housing and neighborhood characteristics on housing values in Delaware County. The first, termed “the individual neighborhood model”, focuses on each of the three neighborhoods separately. The resulting regression models point to the significant variables in each neighborhood case. The second, termed “the multi-neighborhood model”, combines all the significant variables uncovered in the individual models. The Box-Cox transformation is applied to the variables of the multi-neighborhood model, and the maximum log-likelihood form is selected. An elasticity analysis is conducted, and the policy implications of the results are discussed.

6.1 Individual Neighborhood Model

This model includes ten variables representing housing characteristics and conditions, eleven variables representing amenities, seven accessibility variables, fifteen socio-economic variables, three land-use variables, and four variables representing zoning and comprehensive planning, thus a total of fifty variables assumed to be relevant in explaining suburban housing values.
The ten housing variables include: availability of basement (BASE), number of full bathrooms (BXFBAT), number of fireplaces (BXFIRE), number of half bathrooms (BXHBAT), floor area (LFLOR), number of stories (LSTOR), age of the housing unit (LAGE), time elapsed since remodeling the house (LREM), number of rooms (LROOM), and garage capacity (BXGARA).

The eleven amenity variables include: closest distances to cemeteries (LDICE), electric lines (LDIEL), gas lines (LDIGA), hospitals (LDIHO), mobile home parks (LDIMH), regional shopping malls (LDIPO), railroads (LDIRA), rivers (LDIRI), schools (LDISC), town halls (LDITO), and downtown Columbus (LDIDW).

The seven accessibility variables include: accessibility to employment centers (LASEM), fire stations (LASFI), supermarkets (LASSU), industrial sites (LASIN), golf courses (LASGO), and post offices (LASPO).

To capture zoning characteristics, three variables are included in this model: proportion of land zoned for agricultural and open space (LPZAG or BXPZAG), proportion of land zoned for low and medium density residential use (BXPZRE), and proportion of land zoned for commercial, industrial and other uses (BXPZNR).

Existing land-use variables include: proportions of land used for agricultural activities (LPAGR or BXPAGR), of land used for residential activities (LPRES), and of land used for commercial activities (BXPCOM).

The socio-economic variables are related to racial composition, educational attainment, income level, and household size. To capture suburbanization factors, the following variables are also included: proportions of people working in non-agriculture
sectors (LWONG), of people who drive cars to work (LDCAR), and of people who work outside the county (LWOOU).

The model uses ordinary least squares estimation, with housing sales price in natural logarithmic form as the dependent variable. This transformation has been widely used in past hedonic studies of various housing markets in different geographical locations, and it generally provides more robust results (Mathur et al., 2003). The same transformation is applied to the independent variables with non-zero continuous values. A Box-Cox transformation with the value of 0.1 is applied to those independent variables which have zero values. To distinguish, a variable name preceded by “L” represents a logarithmic transformation, and a variable name preceded by “BX” represents a Box-Cox transformation. The same variable may be subject to both transformations, depending upon the neighborhood (e.g. LPZAG and BXPZAG). The model is estimated with 3144 observations.

Table 6.1 reports the estimation results. As predicted, almost all the housing variables are significant and have the expected signs (effects). The coefficients of floor area (LFLOR), number of fireplaces (BXFIRE), garage capacity (BXGARA), age of house (LAGE), and availability of basement (BASE), are positive and significant at the 1% level in all neighborhood cases. Thus, the price of a house increases with its living area, number of fireplaces, age, garage capacity, and basement availability. These results are consistent with the results of past hedonic studies (Kestens et al., 2004; Mathur et al., 2003; Kim et al., 2003; Paterson and Boyle, 2002; Din et al., 2001; Hughes and Turnbull, 1996; Singell and Lillydahl, 1990; Crone, 1983; Jud, 1980). The significant positive effect of the age of the house suggests that older houses carry a premium, possibly
<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Immediate Neighborhood</th>
<th>Medium Neighborhood</th>
<th>Large Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>-18.8544***</td>
<td>-26.7042***</td>
<td>-66.8205***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.99)</td>
<td>(-3.67)</td>
<td>(-6.11)</td>
</tr>
</tbody>
</table>

### Housing Characteristics & Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Immediate Neighborhood</th>
<th>Medium Neighborhood</th>
<th>Large Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE</td>
<td>Basement</td>
<td>0.1331***</td>
<td>0.1562***</td>
<td>0.1599***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.40)</td>
<td>(3.90)</td>
<td>(3.91)</td>
</tr>
<tr>
<td>LAGE</td>
<td>Age of House</td>
<td>0.1416***</td>
<td>0.1231***</td>
<td>0.0836***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(7.64)</td>
<td>(6.57)</td>
<td>(4.36)</td>
</tr>
<tr>
<td>LREM</td>
<td>Age of Remod. House</td>
<td>0.0040</td>
<td>-0.0029</td>
<td>0.0190</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.22)</td>
<td>(-0.16)</td>
<td>(0.99)</td>
</tr>
<tr>
<td>LFLOR</td>
<td>Floor Area</td>
<td>0.2949***</td>
<td>0.3476***</td>
<td>0.5155***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.49)</td>
<td>(6.77)</td>
<td>(10.63)</td>
</tr>
<tr>
<td>BXFBAT</td>
<td>Full Bathrooms</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.11)</td>
<td>(0.14)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>BXFIRE</td>
<td>Fireplaces</td>
<td>0.0009***</td>
<td>0.0009***</td>
<td>0.0010***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.12)</td>
<td>(4.08)</td>
<td>(4.16)</td>
</tr>
<tr>
<td>BXHBAT</td>
<td>Half Bathrooms</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.54)</td>
<td>(1.07)</td>
<td>(1.30)</td>
</tr>
<tr>
<td>LROOM</td>
<td>Total Rooms</td>
<td>-0.0301</td>
<td>0.0085</td>
<td>0.0271</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.43)</td>
<td>(0.12)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>BXGARA</td>
<td>Garage Capacity</td>
<td>0.0011***</td>
<td>0.0016***</td>
<td>0.0019***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.81)</td>
<td>(4.02)</td>
<td>(4.80)</td>
</tr>
<tr>
<td>LSTOR</td>
<td>Number of Stories</td>
<td>0.0217</td>
<td>0.0318</td>
<td>0.0671</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.49)</td>
<td>(0.71)</td>
<td>(1.47)</td>
</tr>
</tbody>
</table>

(continued)
Table 6.1: Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Immediate Neighborhood</th>
<th>Medium Neighborhood</th>
<th>Large Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Amenities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDICE</td>
<td>Dist. Cemetery</td>
<td>-0.0240 (-1.51)</td>
<td>-0.0212 (-1.31)</td>
<td>-0.0076 (-0.43)</td>
</tr>
<tr>
<td>LDIEL</td>
<td>Dist. Electric Lines</td>
<td>0.0132 (1.25)</td>
<td>0.0201* (1.86)</td>
<td>0.0071 (0.59)</td>
</tr>
<tr>
<td>LDIGA</td>
<td>Dist. Gas Lines</td>
<td>0.0012 (0.11)</td>
<td>-0.0078 (-0.71)</td>
<td>-0.0141 (-1.12)</td>
</tr>
<tr>
<td>LDIHO</td>
<td>Dist. Hospital</td>
<td>0.2910*** (6.88)</td>
<td>0.3091*** (6.63)</td>
<td>0.2323*** (4.70)</td>
</tr>
<tr>
<td>LDIMH</td>
<td>Dist. Mobile Homes</td>
<td>0.0158 (1.01)</td>
<td>0.0046 (0.28)</td>
<td>0.0106 (0.59)</td>
</tr>
<tr>
<td>LDPOI</td>
<td>Dist. Reg.Shop.Mall</td>
<td>-0.1525*** (-4.22)</td>
<td>-0.0732* (-1.89)</td>
<td>-0.1491*** (-3.14)</td>
</tr>
<tr>
<td>LDIRA</td>
<td>Dist. Railroads</td>
<td>-0.0130 (-0.92)</td>
<td>0.0042 (0.28)</td>
<td>0.0652*** (3.88)</td>
</tr>
<tr>
<td>LDIRI</td>
<td>Dist. River</td>
<td>0.0238** (1.99)</td>
<td>0.0278** (2.21)</td>
<td>0.0318** (2.24)</td>
</tr>
<tr>
<td>LDISC</td>
<td>Dist. School</td>
<td>-0.0282* (-1.77)</td>
<td>-0.0552*** (-3.26)</td>
<td>0.0078 0.43</td>
</tr>
<tr>
<td>LDITO</td>
<td>Dist. Town Hall</td>
<td>-0.0228 (-1.32)</td>
<td>-0.0122 (-0.67)</td>
<td>-0.0728*** (-3.60)</td>
</tr>
<tr>
<td>LDIDW</td>
<td>Dist. Down. Col.</td>
<td>0.3038 (1.32)</td>
<td>0.2094 (0.80)</td>
<td>0.3023 (1.09)</td>
</tr>
</tbody>
</table>

(continued)
Table 6.1: Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Immediate Neighborhood</th>
<th>Medium Neighborhood</th>
<th>Large Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accessibilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LASEM</td>
<td>Access. Employments</td>
<td>-0.0120 (-0.16)</td>
<td>0.1141 (0.29)</td>
<td>0.0079 (0.10)</td>
</tr>
<tr>
<td>LASFI</td>
<td>Access. Fire Stations</td>
<td>0.4247*** (3.38)</td>
<td>0.4447*** (3.66)</td>
<td>0.4061*** (2.93)</td>
</tr>
<tr>
<td>LASSU</td>
<td>Access. Supermarkets</td>
<td>0.5291*** (2.96)</td>
<td>0.5736*** (3.01)</td>
<td>0.5226** (2.55)</td>
</tr>
<tr>
<td>LASIN</td>
<td>Access. Industrial Sites</td>
<td>1.1364*** (5.81)</td>
<td>1.4931*** (7.18)</td>
<td>1.1942*** (5.51)</td>
</tr>
<tr>
<td>LASBA</td>
<td>Access. Banks</td>
<td>-1.1595*** (-6.14)</td>
<td>-1.2180*** (-6.38)</td>
<td>-1.2162*** (-5.90)</td>
</tr>
<tr>
<td>LASGO</td>
<td>Access. Golf Courses</td>
<td>0.0062 (0.29)</td>
<td>0.0197 (0.89)</td>
<td>0.0138 (0.58)</td>
</tr>
<tr>
<td>LASPO</td>
<td>Access. Post Offices</td>
<td>-0.0442*** (-2.61)</td>
<td>-0.0208 (-1.13)</td>
<td>-0.0582*** (-2.61)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Socio-Economic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTXAS</td>
<td>Property Tax</td>
<td>0.0347 (0.32)</td>
<td>-0.0863 (-0.74)</td>
<td>-0.4539*** (-3.37)</td>
</tr>
<tr>
<td>LDENS</td>
<td>Net. Pop. Density</td>
<td>0.1990*** (9.91)</td>
<td>0.0405** (2.30)</td>
<td>0.1777*** (5.32)</td>
</tr>
<tr>
<td>LAVPR</td>
<td>Ave. Price Neigh.Hou.</td>
<td>0.6850*** (17.61)</td>
<td>0.6032*** (13.33)</td>
<td>0.5919*** (7.30)</td>
</tr>
<tr>
<td>BXPWOW</td>
<td>Pop. White of Own.</td>
<td>0.0007 (0.23)</td>
<td></td>
<td>0.1432** (1.98)</td>
</tr>
<tr>
<td>LPWOW</td>
<td>Pop. White of Own.</td>
<td></td>
<td>0.1432** (1.98)</td>
<td>0.5143*** (2.96)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Immediate Neighborhood</th>
<th>Medium Neighborhood</th>
<th>Large Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>BXPBOW</td>
<td>Pop. Black of Own.</td>
<td>0.0005**</td>
<td>0.0014***</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.50)</td>
<td>(4.10)</td>
<td>(1.45)</td>
</tr>
<tr>
<td>BXPAOW</td>
<td>Pop. Asian of Own.</td>
<td>0.0003</td>
<td>0.0010***</td>
<td>0.0017***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.48)</td>
<td>(3.31)</td>
<td>(3.73)</td>
</tr>
<tr>
<td>LWONG</td>
<td>Work Non Agriculture</td>
<td>0.2332</td>
<td>0.8156</td>
<td>8.1424***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.20)</td>
<td>(0.63)</td>
<td>(3.92)</td>
</tr>
<tr>
<td>LWOOU</td>
<td>Work Outside County</td>
<td>0.4119***</td>
<td>0.4129***</td>
<td>-0.0166</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4.08)</td>
<td>(3.79)</td>
<td>(-0.10)</td>
</tr>
<tr>
<td>LDCAR</td>
<td>Driving Car to Work</td>
<td>-0.2588</td>
<td>0.1779</td>
<td>1.5192**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.07)</td>
<td>(0.64)</td>
<td>(2.53)</td>
</tr>
<tr>
<td>LCIUS</td>
<td>U.S. Citizen</td>
<td>0.3712</td>
<td>0.0301</td>
<td>1.6729</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.85)</td>
<td>(0.06)</td>
<td>(1.45)</td>
</tr>
<tr>
<td>BXEDHS</td>
<td>Edu. High School</td>
<td>0.0015</td>
<td>-0.0012</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.40)</td>
<td>(-1.03)</td>
<td></td>
</tr>
<tr>
<td>LEDHS</td>
<td>Edu. High School</td>
<td></td>
<td></td>
<td>-0.0769</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(-0.96)</td>
</tr>
<tr>
<td>LEDBA</td>
<td>Edu. Bachelor Degree</td>
<td>0.0496</td>
<td>0.0633</td>
<td>0.0534</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.06)</td>
<td>(1.20)</td>
<td>(0.57)</td>
</tr>
<tr>
<td>LINCB</td>
<td>Middle Income</td>
<td>0.0774*</td>
<td>0.0707</td>
<td>0.0170</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.80)</td>
<td>(1.42)</td>
<td>(0.17)</td>
</tr>
<tr>
<td>LINCC</td>
<td>High Income</td>
<td>-0.0339</td>
<td>-0.1083***</td>
<td>-0.1557**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.08)</td>
<td>(-3.04)</td>
<td>(-2.57)</td>
</tr>
<tr>
<td>BXSIZB</td>
<td>Household Size &gt;2</td>
<td>-0.0014</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSIZB</td>
<td>Household Size &gt;2</td>
<td>0.0455</td>
<td></td>
<td>-0.0333</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.29)</td>
<td></td>
<td>(-0.38)</td>
</tr>
</tbody>
</table>

(continued)
Table 6.1: Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Immediate Neighborhood</th>
<th>Medium Neighborhood</th>
<th>Large Neighborhood</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Land Use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BXPAGR</td>
<td>Exist.Agric.Land Use</td>
<td>-0.00168** (-2.40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPG AGR</td>
<td>Exist.Agric.Land Use</td>
<td></td>
<td>0.0265 (0.78)</td>
<td>0.0139 (0.22)</td>
</tr>
<tr>
<td>LPRES</td>
<td>Exist.Resid.Land Use</td>
<td>0.1507*** (4.62)</td>
<td>0.0694* (1.85)</td>
<td>0.0621 (1.44)</td>
</tr>
<tr>
<td>BXPCOM</td>
<td>Exist.Comm.Land Use</td>
<td>0.0001 (0.37)</td>
<td>0.00002 (0.09)</td>
<td>0.0025*** (4.90)</td>
</tr>
<tr>
<td><strong>Zoning &amp; Planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLAN</td>
<td>Comprehensive Plan</td>
<td>0.0490* (1.76)</td>
<td>-0.0157 (-0.55)</td>
<td>0.0480 (1.52)</td>
</tr>
<tr>
<td>BXPZAG</td>
<td>Zoned Agr.and Op.Sp.</td>
<td>-0.0010** (-2.43)</td>
<td>-0.0006 (-1.46)</td>
<td></td>
</tr>
<tr>
<td>LPZAG</td>
<td>Zoned Agr.and Op.Sp.</td>
<td></td>
<td></td>
<td>-0.0252 (-1.13)</td>
</tr>
<tr>
<td>BXPZRE</td>
<td>Zoned Res.Low and Med.</td>
<td>0.0010*** (4.12)</td>
<td>0.0014*** (4.17)</td>
<td>-0.0001 (-0.24)</td>
</tr>
<tr>
<td>BXPZNR</td>
<td>Zoned Com,Ind, &amp; Other</td>
<td>0.0007*** (3.06)</td>
<td>-0.0001 (-0.73)</td>
<td>0.0003 (0.76)</td>
</tr>
<tr>
<td>Number of observations</td>
<td></td>
<td>3144</td>
<td>3144</td>
<td>3144</td>
</tr>
<tr>
<td>R-square</td>
<td></td>
<td>0.5304</td>
<td>0.5113</td>
<td>0.4864</td>
</tr>
</tbody>
</table>

Note:
Figures in parentheses are \(t\)-values;
*** Significant at the 1% level;
** Significant at the 5% level;
* Significant at the 10% level.
because of historic ambience and identity. This is consistent with the results in Li and Brown (1980) and Cobb (1984). The strong significance of the floor area and garage capacity variables implies that larger suburban houses with multi-car garage capacity carry a premium, as is typical for suburban homeowners.

Electric transmission lines (LDIEL) and railroads (LDIRA) have negative impacts on single-family house prices. The closer the house is to these infrastructures the lower the price, probably because of such threats as high voltage power and noise. Hospital proximity is significant in all neighborhood cases; the farther the hospital, the higher the price, possibly because of such externalities as traffic, noise, and contamination/toxic waste.

Proximity to rivers has a significant and negative effect in all neighborhood types: the greater the distance, the higher the price. The negative impact may be due to the risk of flooding and the undesirable swampy sites associated with many streams and brooks. In addition, locating homes close to rivers may also be undesirable due to restrictions on development. This result differs from Paterson and Boyle (2002), who found that proximity to rivers is not significant in explaining housing values.

Proximity to town halls (LDITO) carries a premium, because a town hall is generally located at the center of a township, where the land rent is bid up. Town halls may be equivalent to central business districts in big cities. Closeness to regional shopping malls (LDIPO), located in the southern part of the county, carries a premium in all neighborhood types. This is not surprising: the greater the distance from CBDs, the lower the rent. (Mills, 1981; O’Sullivan, 2003). This result is also consistent with Davis et al. (1994) and Nelson and Sanchez (1999), who argue that the land bid rent declines
along the urban-suburban-exurban axis. Regional shopping malls provide convenience for shopping and other public services/entertainments, and single-family houses located closer to the regional shopping malls in the southern part of Delaware County are more expensive than those in the northern part of the county.

The accessibilities of single family houses to fire stations, industrial sites, and supermarkets produce significant positive impacts on house prices in all neighborhood types: the closer these services, the higher the prices. Accessibility to fire stations provides safety for neighborhoods, and accessibility to supermarkets provides convenience for one-stop shopping. The case of industrial sites is somewhat surprising. The positive impact may be related to job availability in light industries, which do not produce toxic wastes.

The accessibility to banks has a significant and negative impact, possibly because banks generate the risks of traffic and crimes. People’s preference to reside closer to banks may be low, because banking services are also available in shopping centers, where people can do one-stop shopping and business transactions at the same time. People may also not prefer to reside closer to banks because of the availability of banking services through the Internet.

The accessibility to post offices has also a negative impact on housing values, maybe because post offices generate traffic. People have also become less dependent on traditional mailing, since many business activities and transactions can be processed online with Internet access. Ewing (1997) argues that technology innovation, particularly in telecommunications, is among the reasons that make people living dispersed in the suburbs.
The socio-economic characteristics of neighborhoods do matter as determinants of housing values. The average price of housing in the neighborhood (LAVPR) has a positive and strongly significant (1% level) impact in all neighborhood types. This suggests that the price for a single-family house increases whenever the average price of houses in the neighborhood increases. Likewise, the price of a given house may decrease if the average price of houses in the neighborhood decreases. Therefore, high-income people tend to exclude low-income houses in their neighborhood in order to keep their high price levels. Fischel (2004) argues that maintaining exclusiveness of housing in the suburbs has become the main concern of homeowners, and directs them to exclude low-income housing by enacting zoning ordinances.

Ihlanfeldt (2004) argues that the proportion of high-income people increases house values, while the proportion of Blacks lower housing values in the neighborhood. Other studies, conducted by McDonald and McMillen (2004), Green (1999), and Thorsnes (2000), found that property values can be increased by excluding certain groups. They also found that a higher share of White households has positive impacts on housing values.

The empirical results show that the percentage of White homeowners (LPWOW) increases housing value, as shown in the results for the medium and large neighborhoods. The proportions of Blacks (BXPBOW) and Asians (BXPAOW) are also significant and have positive impacts on housing values. The willingness of homeowners to pay to reside in neighborhoods with minorities may be related to the higher income of these groups, as indicated by the significance of such variables as the share of the middle-income population (LINCB) and the average price of housing in the neighborhood (LAVPR).
These results imply that homeowners prefer to reside in neighborhoods with other wealthy people, regardless of their race. This result is consistent with Michaels and Smith (1990), who find a significant and positive impact of the Black population on housing values in Cambridge, Massachusetts.

Population density (LDENS) is significant and has a positive effect on housing values. In addition to being a proxy for availability of services, this variable also reflects development intensity. This result is consistent with Rolleston (1987), who argues that communities experiencing greater growth in fiscal capacity and those in paths of development generally accommodate those growth pressures by permitting denser residential development.

Other socio-economic characteristics of suburbs which are found significant in this study are the proportion of people who are driving cars to work (LDCAR), the proportion of people who work in non-agricultural activities (LWONG), and the proportion of people who work outside the county (LWOOU). These three variables are positive and significant. Suburban residents’ dependency on cars is generally high because of limited public transportation, and therefore cars are needed for commuting to jobs locations. Daniels (1999) describes suburbs as places where city and country life collides: typical suburban dwellers commute to work in offices and enjoy single-family homes with a feeling of space.

Assessing the impacts of existing land uses and zoning is of particular interest to this research. The empirical results suggest that existing land uses have an impact on single-family housing values. Existing agricultural land and open space (BXPAGR) have significant and negative impacts on housing values. This is surprising, as one generally
assumes that people tend to move out from central cities to suburbs in order to live in ‘greener’ environments. In addition, undeveloped ‘greenfield’ sites outside central cities are often cheaper to be developed as residential sites than ‘brownfield’ sites within cities that had previously been used for other purposes (Squires, 2002). One may also think that people may enjoy living in suburbs with a higher share of agricultural land and open space, because of the cleaner air environment. The negative impacts of agricultural land may be related to the negative impacts from farming, such as bad odors and chemical contamination of the land. Also, agricultural lands and open spaces in the suburbs are mostly in private hands, which makes it difficult to consider them as public benefits. The negative impacts of agriculture land and open space may also be related to restrictions on housing development. In any case, this result differs from Geoghegan et al. (1997) and Kestens (2004), who found that existing open space has both positive and negative impacts.

In contrast to agricultural land, existing residential (LPRES) and commercial (BXPCOM) lands have significant and positive impacts on housing values. Existing residential land use provides some certainty that the area will remain a residential development site. Also, developed areas carry smaller risks than undeveloped areas, and also have superior infrastructures. Commercial land provides convenience and amenities, which positively impacts the neighborhood. The Columbus Dispatch, on February 26, 2006, reported that housing development in southern Delaware County had grown rapidly because of its retail draw and solid housing stock, and that the Powell municipality, in southern Delaware County, is listed currently as the 18th best place to live in the U.S.
The availability of comprehensive planning (PLAN) is found to be significant and has a positive effect on housing prices, implying that homeowners are willing to pay more to live in an area which is comprehensively planned. A comprehensive plan has a positive impact on housing values because it provides certainty for future development in the neighborhood and also guides the development of the area in terms of infrastructures, which are important for developers and potential homeowners. Comprehensive planning corrects for market imperfections by providing information and regulations on land development. This finding differs from Quigley and Rosenthal (2005) who report that local land regulation is symbolic, ineffectual, or only weakly enforced.

Like comprehensive planning, zoning does matter. The proportion of land zoned for agriculture and open space (BXPZAG) negatively impacts housing values in the immediate neighborhood, possibly because of the implied restriction on land development, and because of such externalities as bad odor and contamination by chemicals used in fertilizers and pesticides. Fiscal impacts may also be a factor, because development will have to be financed by homeowners. Kestens (2004) notices that agriculture may have negative impacts because the lack of urbanization lowers the level of services in the neighborhood.

In contrast, the proportion of land zoned for low and medium residential density has a significant and positive impact on housing values in the immediate and medium neighborhoods. This may be linked to the certainty of residential land use in the future. Once zoning is adopted, it is certain that nearby lands will be compatible with current residential land use. Past studies, such as Fischel (2004) and Jud (1980), found that if a particular piece of land is zoned residential, it is likely that the surrounding properties are
also zoned for residential use, providing the certainty to developers and homeowners that this land will be developed in the same residential use in the future. This certainty increases demand for low and medium density housing and also increases housing supply from developers, leading to higher housing values. More land zoned for residential use also increases housing prices by producing amenity effects that heighten the demand for these types of housing. It is also possible that such zoning enhances property values because it restricts multi-family housing development, such as apartments and mobile homes.

Like zoning for residential use, the proportion of land zoned for commercial and industrial uses also generates positive impacts on housing values, possibly because business is welcome in the suburbs if its effects are not too noxious, and because such activities can pay local taxes in excess of the additional costs of required local services. Commercial activities increase opportunities for shopping and employment, which increases local convenience and economic growth. Fischel (2004) notes that industrial zoning is welcome in neighborhoods because it leads to payment of property taxes and because it mitigates spillover effects that might reduce home values. Technological innovations may also help reduce the externalities of commercial and industrial activities, thus allowing a focus on the more positive impacts.

Table 6.1 also shows that socio-economic, land-use and zoning variables have different ranges of spatial spillover effects. The average price of housing in a neighborhood (LAVPR) is significant in all three neighborhoods with, however, coefficients that decrease as the neighborhood becomes larger. This implies that homeowners believe the immediate neighborhood has the strongest effect in increasing
housing values, and are very much concerned by housing prices right next to their houses. The middle-income variable (LINC) is significant in the immediate neighborhood, but not so in the larger neighborhoods, suggesting that homeowners are concerned by the wealth of their closest neighbors. This is a reasonable concern, since a homeowner’s wealth is associated with his/her capability for appropriate housing maintenance, which impacts housing prices. The coefficient of the existing residential land-use share (LPRES) decreases as the neighborhood becomes larger. This suggests that homeowners believe that close proximity of residential land use is the more important factor. The existing agricultural land-use share (BXPAGR), the share of land zoned for agricultural (BXPZAG), and the share of land zoned for commercial and industrial activities (BXPZNR) are all significant in the close neighborhood only, confirming the importance for homeowners of the land-use structure in their immediate surroundings.

### 6.2 Multi-Neighborhood Model

Unlike the previous model, which focuses on individual neighborhoods separately, this model combines the characteristics of all neighborhoods/buffers into one regression model. The variables used in this multi-neighborhood model are selected from the set of significant variables in each individual neighborhood by considering the t-values of the independent variables. For example, the variable LPWOW_3 is selected because it has a greater t-value than LPWOW_2, while BXPWOW (for buffer 1) is not selected because of its insignificant coefficients. While the total number of variables included in the multi-neighborhood model is smaller, the selected variables still represent the six major categories of factors: housing characteristics and conditions, amenities,
accessibilities, socio-economic characteristics, existing land use, and comprehensive planning and zoning.

Table 6.2 reports the results of the estimation of the multi-neighborhood model, which consists of thirty variables, including five variables for housing characteristics and conditions, four variables for amenities, four variables for accessibility, three variables for existing land uses, ten variables for socio-economic characteristics, and four variables for comprehensive planning and zoning. This model uses the same logarithmic and Box-Cox transformations as the individual neighborhood models. The value of the $R^2$ in this integrated model (0.54) is higher than the $R^2$ of any individual model.

Table 6.3 presents the results of the estimation of the multi-neighborhood model when using the Box-Cox transformation with two parameters ($\lambda$ for the dependent variable and $\mu$ for all the independent variables, except the dummy variables). An iterative OLS grid estimation is performed to find the maximum log-likelihood model, with values of $\lambda$ and $\mu$ ranging from -1 to +4, with a step of 0.01. The same thirty variables are used. The best model has a higher R-square (0.56) than the model with the logarithmic form, and the optimal transformation parameters are $\lambda = 0.30$ for the dependent variable (house price), and $\mu = 0.02$ for all independent variables.

To test whether the logarithmic model is significantly different from the optimal Box-Cox model, the log-likelihood ratio test is used. If the two forms are equivalent, the statistics $2(LK_{\text{max}} - LK_0)$ is approximately distributed as a $\chi^2$ with 3,143 degree of freedom, where $LK_{\text{max}}$ is the optimal log-likelihood and $LK_0$ is the logarithmic model’s log-likelihood. Because $2(LK_{\text{max}} - LK_0) = 400.82 > \chi^2_{d.f=3143}(0.05) = 77.92$, the null
| Variable | Label | Parameter Estimate | t Value | Pr > |t| |
|----------|-------|--------------------|---------|------|---|
| Intercept|       | -33.84428          | -4.19   | <.0001*** |

### Housing Characteristics and Conditions

| Variable | Label | Parameter Estimate | t Value | Pr > |t| |
|----------|-------|--------------------|---------|------|---|
| BASE     | BASEMENT | 0.12400          | 3.25   | 0.0012*** |
| LAGE     | AGE OF HOUSE | 0.14213          | 17.84  | <.0001*** |
| LFLOR    | FLOOR AREA | 0.30031          | 7.66   | <.0001*** |
| BXFIRE   | FIREPLACES | 0.00086914      | 3.79   | 0.0002*** |
| BXGARA   | GARAGE CAPACITY | 0.00111        | 2.93   | 0.0034*** |

### Amenities

| Variable | Label | Parameter Estimate | t Value | Pr > |t| |
|----------|-------|--------------------|---------|------|---|
| LDIHO    | DIST.HOSPITAL | 0.22461          | 6.34   | <.0001*** |
| LDIPO    | DIST.REG.SHOP.MALL | -0.10696       | -3.99  | <.0001*** |
| LDIRI    | DIST.RIVER | 0.03816          | 3.51   | 0.0005*** |
| LDITO    | DIST.TOWN HALL | -0.02511       | -1.66  | 0.0972* |

### Accessibilities

| Variable | Label | Parameter Estimate | t Value | Pr > |t| |
|----------|-------|--------------------|---------|------|---|
| LASFI    | ACCESS.FIRE STATIONS | 0.43167        | 5.39   | <.0001*** |
| LASIN    | ACCESS.INDUSTRIALS | 0.97503        | 5.60   | <.0001*** |
| LASBA    | ACCESS.BANKS | -0.72531        | -5.76  | <.0001*** |
| LASPO    | ACCESS.POST OFFICES | -0.03349      | -2.23  | 0.0257** |

### Socio-economic Characteristics

| Variable | Label | Parameter Estimate | t Value | Pr > |t| |
|----------|-------|--------------------|---------|------|---|
| LPNOW_3  | POP.WHITE OF OWN.B3 | 0.15708        | 1.79   | 0.0743* |
| BXPNOW_3 | POP.ASIAN OF OWN.B3 | 0.00105        | 2.76   | 0.0059*** |
| BXPBOW_2 | POP.BLACK OF OWN.B2 | 0.00173        | 5.50   | <.0001*** |
| LWONG_3  | WORK NON AGRIC.B3 | 3.62713         | 2.32   | 0.0207*** |
| LDCAR_3  | DRIVING CAR.B3 | 1.21677         | 2.69   | 0.0073*** |
| LDENS_1  | NET POP.DENSITY.B1 | 0.17305        | 9.46   | <.0001*** |
| LAVPR_1  | AVE.PRICE NEIGH.HOU.B1 | 0.66440      | 18.49  | <.0001*** |
| LWOOU_1  | WORK OUTSIDE CO.B1 | 0.29366        | 4.43   | <.0001*** |
| LTXS     | PROPERTY TAXS | -0.17730       | -1.82  | 0.0695* |
| LINCB_1  | MIDDLE INCOME.B1 | 0.10632        | 3.30   | 0.0010*** |

### Existing Land Use

| Variable | Label | Parameter Estimate | t Value | Pr > |t| |
|----------|-------|--------------------|---------|------|---|
| BXPAGR_1 | AGRICULTURAL LAND USE.B1 | -0.00164    | -2.38  | 0.0175** |
| LPRES_1  | RESIDENTIAL LAND USE.B1 | 0.13853     | 4.57   | <.0001*** |
| BXPCOM_3 | COMMERCIAL LAND USE.B3 | 0.00247     | 5.60   | <.0001*** |

(continued)
Table 6.2: Continued

| Variable | Label                                   | Parameter Estimate | t Value | Pr > |t| |
|----------|-----------------------------------------|--------------------|---------|------|---|
| Zoning and Planning                  |                    |                    |         |      |   |
| PLAN    | COMPREHENSIVE PLAN                      | 0.06096            | 2.69     | 0.0071*** |
| BXPZAG_1| ZON.AGRI.AND OPEN SP.B1                 | -0.00085049        | -2.13    | 0.0335** |
| BXPZRE_1| ZON.LOW AND MED.RESID.B1               | 0.00050749         | 2.00     | 0.0454** |
| BXPZNUR_1| ZON.COM,IND,AND OTHER.B1               | 0.00046148         | 1.97     | 0.0492** |

R-Square= 0.54
Number of observations= 3144

Note:

(***) = Significant at 1% level.
(**) = Significant at 5% level.
(*) = Significant at 10% level.
| Variable | Label                          | Parameter | t Value | Pr > |t| |
|----------|--------------------------------|-----------|---------|------|---|
| Intercept|                                | -1038.13570 | -3.77   | 0.0002*** |

**Housing Characteristics and Conditions**

| Variable | Label                          | Estimate | t Value | Pr > |t| |
|----------|--------------------------------|----------|---------|------|---|
| BASE     | BASEMENT                       | 2.91942  | 2.14    | 0.0325** |
| BOXAGE   | AGE OF HOUSE                   | 4.78642  | 17.37   | <.0001*** |
| BOXFLOR  | FLOOR AREA                     | 11.91090 | 9.88    | <.0001*** |
| BOXFIRE  | FIREPLACES                     | 0.04798  | 2.93    | 0.0034*** |
| BOXGARA  | GARAGE CAPACITY                | 0.05729  | 2.12    | 0.0338** |

**Amenities**

| Variable | Label                          | Estimate | t Value | Pr > |t| |
|----------|--------------------------------|----------|---------|------|---|
| BOXDIHO  | DIST.HOSPITAL                  | 6.87795  | 5.49    | <.0001*** |
| BOXDIPO  | DIST.REG.SHOP.MALL             | -4.02527 | -4.27   | <.0001*** |
| BOXDIRI  | DIST.RIVER                     | 1.48147  | 3.75    | 0.0002*** |
| BOXDITO  | DIST.TOWN HALL                 | -1.10393 | -2.04   | 0.0416** |

**Accessibilities**

| Variable | Label                          | Estimate | t Value | Pr > |t| |
|----------|--------------------------------|----------|---------|------|---|
| BOXASFI  | ACCESS.FIRE STATIONS           | 11.46434 | 4.94    | <.0001*** |
| BOXASIN  | ACCESS.INDUSTRIALS             | 22.71566 | 4.71    | <.0001*** |
| BOXASBA  | ACCESS.BANKS                   | -21.01597 | -5.66   | <.0001*** |
| BOXASPO  | ACCESS.POST OFFICES            | -1.20498 | -2.20   | 0.0279** |

**Socio-economic Characteristics**

| Variable | Label                          | Estimate | t Value | Pr > |t| |
|----------|--------------------------------|----------|---------|------|---|
| BOXPWOW_3| POP.WHITE OF OWN.B3            | 2.35014  | 0.81    | 0.4161 |
| BOXPAOW_3| POP.ASIAN OF OWN.B3            | 0.07185  | 2.76    | 0.0058*** |
| BOXPBOW_2| POP.BLACK OF OWN.B2            | 0.10769  | 5.00    | <.0001*** |
| BOXWONG_3| WORK NON AGRIC.B3              | 63.25422 | 1.23    | 0.2170 |
| BOXDCAR_3| DRIVING CAR.B3                 | 30.35982 | 2.05    | 0.0405** |
| BOXDENS_1| NET POP.DENSITY.B1             | 5.22928  | 8.08    | <.0001*** |
| BOXAVPR_1| AVE.PRICE NEIGH.HOU.B1         | 20.65093 | 20.51   | <.0001*** |
| BOXWOOU_1| WORK OUTSIDE CO.B1             | 8.05510  | 3.68    | 0.0002*** |
| BOXTAXS  | PROPERTY TAXS                  | -3.51216 | -1.09   | 0.2776 |
| BOXINC_1 | MIDDLE INCOME.B1               | 2.00027  | 1.87    | 0.0621* |

**Existing Land Use**

| Variable | Label                          | Estimate | t Value | Pr > |t| |
|----------|--------------------------------|----------|---------|------|---|
| BOXPAGR_1| AGRIC. LAND USE.B1             | -0.07941 | -1.65   | 0.0991* |
| BOXPRES_1| RESID. LAND USE.B1             | 3.34041  | 3.31    | 0.0009*** |
| BOXPCOM_3| COMMER. LAND USE.B3            | 0.15909  | 5.04    | <.0001*** |

(continued)
### Table 6.3: Continued

| Variable  | Label                                  | Parameter Estimate | t Value | Pr > |t| |
|-----------|----------------------------------------|--------------------|---------|------|---|
| PLAN      | COMPREHENSIVE PLAN                     | 2.44321            | 3.02    | 0.0026*** |
| BOXPZAG_1 | ZON.AGRI.& OPEN SP.B1                  | -0.06486           | -2.33   | 0.0198**  |
| BOXPZRE_1 | ZON.LOW & MED.RES.B1                   | 0.03776            | 2.15    | 0.0313**  |
| BOXPZRNR_1| ZON.COM,IND,& OTH.B1                   | 0.03356            | 2.04    | 0.0413**  |

\[ \lambda = 0.30 \]
\[ \mu = 0.02 \]
Maximum Log-Likelihood= -35375.62
R-square= 0.56
Number of observations= 3144

Note:

(***)= significant at 1% level;
(**)= significant at 5% level;
(* )= significant at 1% level.
hypothesis of functional form equivalence is to be rejected at the $\alpha = 0.05$ level of significance.

The results show that all the variables on housing characteristics and conditions, including availability of basement (BASE), floor area (BOXFLOR), fire places (BOXFIRE), age of house (BOXAGE), and garage capacity (BOXGARA), are significant with positive impacts on housing values.

In the amenity category, distances to hospitals (BOXDIHO), regional shopping malls (BOXDIPO), rivers (BOXDIRI), and town halls (BOXDITO) do impact housing values. The greater the distance to a hospital, the higher the price, as does the distance to rivers. In contrast, the shorter the distance to regional shopping malls, the higher the housing values. The explanatory arguments have already been discussed in the cases of the individual neighborhood models.

Accessibilities to fire stations and industrial sites have significant and positive impacts on housing values, while accessibilities to banks and post offices have significant and negative impacts. The possible reasons for these effects have been discussed earlier.

Among the socio-economic characteristics, the proportion of people driving cars to work, the proportion of people who work outside the county, the net population density, and the average price of housing in the neighborhood, have all positive impacts on housing values. The significance of the proportion of people driving cars to work and the proportion of people who work outside the county emphasize that commuting is the common lifestyle of the suburbs.

The significance of middle-income variable confirms the willingness-to-pay of suburban homeowners to live close to wealthy people. The significance of this variable is
also related to the high level of the average price of single-family houses sold in 2000 in Delaware County ($226,935). In contrast, the average price for single-family housing affordable to low-income people is estimated by Green (1999) at only $ 75,000. Rising income increases the purchasing power of middle-income households. As argued by Tiebout (1956), these people will ‘vote with their feet’, since their ability and demand have made it possible for them to move to better jurisdictions, particularly in the suburbs.

All the variables used as proxies for existing land uses are significant, in the same way as in the individual models, and for the same reasons. All the variables used as proxies for zoning and comprehensive planning are also significant, and with impacts similar to those observed in the individual models.

The pattern of existing land and land zoned for residential and commercial uses, concentrated at the centers of municipalities and townships, may localize development and prevent scattered development, leapfrog, and sprawl development. However, a high willingness to pay for both existing land and land zoned for residential and commercial usages may lead to excessive demand for this particular land, which may induce scattered development, leapfrog, and sprawl in the future. In some areas, existing residential and commercial land uses and land zoned for such uses are located along main roads, such as Route 23, which will induce more ribbon/sprawl development. These patterns of spatial development have been characterized by scholars as inefficient development, because of high infrastructure and transportation costs.

Zoning can be used as a tool for controlling and directing spatial development, but the empirical results imply that it can induce high housing prices. In this case, zoning
may reduce the problem of sprawl/leapfrog development, but it may increase other problems, such as spillover effects, housing exclusiveness, and unaffordable housing.

### 6.3 Elasticity of Variables

In order to compare the effects of each independent variable on housing values, the elasticity of the non-dummy variables is computed. The elasticity represents the percentage change in the dependent variable (housing value) resulting from a one percentage change in an independent variable. The elasticity of variables in a Box-Cox model is computed as follows:

\[
\varepsilon_k = a_k \cdot \frac{(X_k)^\mu}{P^{(\lambda)}}
\]

where:

\[
P^{(\lambda)} = P^{\lambda} - 1
\]

\[
P^{(\lambda)} = a_0 + a_1 (X_1)^\mu + \ldots + a_k(X_k)^\mu
\]

- \(P^{(\lambda)}\) = housing value;
- \(X_k = \) the value of the independent variable k;
- \(\lambda = \) parameter of the dependent variable;
- \(\mu = \) parameter of the independent variables.

Table 6.4 presents the results of computing the elasticities of the variables over the sample, showing their variabilities. The floor area variable has the highest mean elasticity among housing characteristics, with 0.3294, while the availability of fireplaces has the smallest one (0.0009). The mean floor area elasticity is consistent with the results of other studies. For example, Singell and Lillydahl (1990) find an elasticity of 0.34.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELAS_AGE</td>
<td>AGE OF HOUSE</td>
<td>0.1188</td>
<td>0.0873</td>
<td>0.2616</td>
</tr>
<tr>
<td>ELAS_FLOR</td>
<td>FLOOR AREA</td>
<td>0.3294</td>
<td>0.2519</td>
<td>0.7381</td>
</tr>
<tr>
<td>ELAS_FIRE</td>
<td>FIREPLACES</td>
<td>0.0009</td>
<td>0</td>
<td>0.0021</td>
</tr>
<tr>
<td>ELAS_GARA</td>
<td>GARAGE CAPACITY</td>
<td>0.0012</td>
<td>0</td>
<td>0.0026</td>
</tr>
<tr>
<td>ELAS_DIHO</td>
<td>DIST.HOSPITAL</td>
<td>0.1702</td>
<td>0.1268</td>
<td>0.3723</td>
</tr>
<tr>
<td>ELAS_DIPO</td>
<td>DIST.REG.SHOP.MALL</td>
<td>-0.0988</td>
<td>-0.2278</td>
<td>-0.0734</td>
</tr>
<tr>
<td>ELAS_DIRI</td>
<td>DIST.RIVER</td>
<td>0.0351</td>
<td>0.0259</td>
<td>0.0787</td>
</tr>
<tr>
<td>ELAS_DITO</td>
<td>DIST.TOWN HALL</td>
<td>-0.0264</td>
<td>-0.0604</td>
<td>-0.0193</td>
</tr>
<tr>
<td>ELAS_ASFI</td>
<td>ACCESS.FIRE STATIONS</td>
<td>0.3366</td>
<td>0.2518</td>
<td>0.7692</td>
</tr>
<tr>
<td>ELAS_ASIN</td>
<td>ACCESS.INDUSTRIALS</td>
<td>0.6974</td>
<td>0.5202</td>
<td>1.5996</td>
</tr>
<tr>
<td>ELAS_ASBA</td>
<td>ACCESS.BANKS</td>
<td>-0.6032</td>
<td>-1.3727</td>
<td>-0.4506</td>
</tr>
<tr>
<td>ELAS_ASPO</td>
<td>ACCESS.POST OFFICES</td>
<td>-0.0277</td>
<td>-0.0650</td>
<td>-0.0207</td>
</tr>
<tr>
<td>ELAS_PWOW_3</td>
<td>POP.WHITE OF OWN.B3</td>
<td>0.0608</td>
<td>0.0455</td>
<td>0.1374</td>
</tr>
<tr>
<td>ELAS_PAOW_3</td>
<td>POP.ASIAN OF OWN.B3</td>
<td>0.0018</td>
<td>0</td>
<td>0.0041</td>
</tr>
<tr>
<td>ELAS_PBOW_2</td>
<td>POP.BLACK OF OWN.B2</td>
<td>0.0025</td>
<td>0</td>
<td>0.0062</td>
</tr>
<tr>
<td>ELAS_WONG_3</td>
<td>WORK NON AGRIC.B3</td>
<td>1.6424</td>
<td>1.2248</td>
<td>3.7397</td>
</tr>
<tr>
<td>ELAS_DCAR_3</td>
<td>DRIVING CAR.B3</td>
<td>0.7871</td>
<td>0.5872</td>
<td>1.7886</td>
</tr>
<tr>
<td>ELAS_DENS_1</td>
<td>NET POP.DENSITY.B1</td>
<td>0.1272</td>
<td>0.0925</td>
<td>0.2901</td>
</tr>
<tr>
<td>ELAS_AVPR_1</td>
<td>AVE.PRICE NEIGH.HOU.B1</td>
<td>0.6267</td>
<td>0.4820</td>
<td>1.3897</td>
</tr>
<tr>
<td>ELAS_WOOU_1</td>
<td>WORK OUTSIDE CO.B1</td>
<td>0.2074</td>
<td>0.1550</td>
<td>0.4620</td>
</tr>
<tr>
<td>ELAS_TAXS</td>
<td>PROPERTY TAXS</td>
<td>-0.0898</td>
<td>-0.2044</td>
<td>-0.0669</td>
</tr>
<tr>
<td>ELAS_INCB_1</td>
<td>MIDDLE INCOME.B1</td>
<td>0.0510</td>
<td>0.0378</td>
<td>0.1153</td>
</tr>
<tr>
<td>ELAS_PAGR_1</td>
<td>AGRIC. LAND USE.B1</td>
<td>-0.0017</td>
<td>-0.0045</td>
<td>0</td>
</tr>
<tr>
<td>ELAS_PRES_1</td>
<td>RESID. LAND USE.B1</td>
<td>0.0850</td>
<td>0.0633</td>
<td>0.1956</td>
</tr>
<tr>
<td>ELAS_PCOM_3</td>
<td>COMMER. LAND USE.B3</td>
<td>0.0037</td>
<td>0</td>
<td>0.0090</td>
</tr>
<tr>
<td>ELAS_PZAG_1</td>
<td>ZON.AGRI.&amp; OPEN SP.B1</td>
<td>-0.0015</td>
<td>-0.0036</td>
<td>0</td>
</tr>
<tr>
<td>ELAS_PZRE_1</td>
<td>ZON.LOW &amp; MED.RES.B1</td>
<td>0.0008</td>
<td>0</td>
<td>0.0017</td>
</tr>
<tr>
<td>ELAS_PZNR_1</td>
<td>ZON.COM,IND,&amp; OTH.B1</td>
<td>0.0007</td>
<td>0</td>
<td>0.0019</td>
</tr>
</tbody>
</table>

Table 6.4 Elasticity of Variables
Hughes and Turnbull (1996) and McDonald (1980) report elasticities of 0.32 and 0.30, respectively. Increasing floor area does not lead to a proportionate increase in price, probably because of the high share of fixed costs in total building costs.

The age of a house is significant and has a positive impact on housing values. This finding is consistent with Michaels and Smith (1990) and Hughes and Turnbull (1996). The mean elasticity of this variable is 0.12. A homeowner’s willingness-to-pay for an older house is higher due to the generally established character of the area, and the quality of older houses. Note that the average age of houses in Delaware County is 11 years, which suggests the availability of nice neighborhoods without dilapidated houses. The maximum elasticity of this variable is 0.26.

The average price of housing in the neighborhood has a mean elasticity of 0.63, and a maximum elasticity of 1.39. Grieson and White (1989) find a mean elasticity of 0.57 for this variable.

Among the amenity variables, distances to hospitals and regional shopping malls have the highest mean elasticities. Residential land use has the highest mean elasticity (0.085) among existing land uses. The mean elasticity of existing agricultural land use (-0.0017) is close to the mean elasticity of agricultural zoning (-0.0015). Among the socio-economic characteristics, the mean elasticities for the variables related to working in non-agriculture activities, average price of housing in the neighborhood, driving cars to works, and working outside the county, are higher than those of the other variables.

The percentage of people who work in non-agriculture activities (LWONG) has a high mean elasticity value (1.64). The high elasticity value of this variable and the impacts of the land-use variables (negative impact of agricultural land-use, negative
impact of land zoned for agricultural activities, the positive impact of residential land use, and the positive impact of land zoned for residential activities), all imply that homeowners in Delaware County prefer to reside in urbanized neighborhoods. Agricultural land use and agricultural zoning are not valued by homeowners as “green value”. Hart (1991) describes this phenomenon as “the perimetropolitan bow-wave phenomenon”, or the new metropolitan frontier in the urban-rural fringe. He argues that the fringe area for one generation becomes a suburban zone for the next one, and agricultural activities are pushed farther out to a new fringe. Hart describes what might happen to four generations on a single farm as the bow-wave gradually engulfs it. The first generation, like its immediate forebears, practices dairy farming, and resents the encroaching urbanization. The second generation decides to intensify farming by growing vegetables. The third generation begins a nursery operation and may even build a greenhouse. The fourth generation sells the land and retires to Florida. This description shows that the pressure of development in fringe areas may lead to the “impermanence syndrome”, that encourages premature disinvestment in agriculture.
CHAPTER 7

CONCLUSIONS

This research differs from past hedonic price studies and contributes to this area in the following way:

1. **Assessing neighborhood impacts by using circular buffers**

   A new set of variables has been considered to assess the impacts of housing and neighborhood characteristics on single-family suburban housing values. Neighborhoods are defined in a novel way as circular buffers surrounding each individual housing unit. Different neighborhood (buffer) sizes are considered (100 meters, 400 meters, 1 mile radii) to account for differential effects of price determinants. Neighborhood data, including parcels data, master field data, and Census data at the block and block groups levels, have been converted into these buffers by using GIS tools and a complex allocation process. GIS tools have also been used to measure shortest distances to amenities and to compute accessibilities, using parcel centroids. These data sources and conversions provide a contrast to past studies, which have generally used grids and binary variables for capturing amenities and accessibilities. In addition, the use of detailed land-use and zoning shares in each township and municipality also contrasts this research from past studies.
Land-use and zoning data have also been converted into the circular neighborhood buffers. Models have first been estimated for each individual neighborhood, using fifty variables representing housing and neighborhood characteristics. A multi-neighborhood model is then estimated, that combines the significant variables of the three neighborhood models. The final version of the multi-neighborhood model is estimated using the Box-Cox transformation. The maximum log-likelihood model has the following Box-Cox parameters: $\lambda = 0.30$ for the dependent variable (price) and $\mu = 0.02$ for all independent variables, with $R^2 = 0.56$.

2. **Understanding the interactions between different suburbanization theories**

The floor area, fireplace, age of house, basement, and garage capacity variables, all proxies for housing characteristics and conditions, are significant and with positive impacts on housing values. The existing land uses that surround a house have also significant impacts on housing values, particularly residential and commercial land uses. Proximity to the regional shopping mall has also a significant and positive impact on housing values. Given the significances of such variables as property taxes, zoning, and the shares of people with middle income, driving cars to works, working in non-agricultural activities, and working outside the county, the empirical results support both the “natural evolution theory” and “flight from blight theory”.

3. **Understanding the role of zoning and comprehensive planning**

Zoning and comprehensive planning are found to significantly impact housing values. They provide certainty and perfect information on future land
uses. In terms of elasticity, lands zoned for residential and commercial usages have about the same positive impact on housing values. However, the impact of agricultural zoning is negative, probably because it is perceived as constraining future land development. Zoning and comprehensive planning are important to control growth in the suburbs, but these results suggest that they induce high housing prices, which may create other problems, such as spillover effects and unaffordable housing for low-income people.

The public policies needed to control the fast-growing development in Delaware County are as follows:

1. **Spatial development policy**

   Since Delaware County has been one of the fastest growing counties in the U.S., local policies are needed to control this growth and sprawl development. One important policy is to encourage infill development in existing residential and commercial centers, particularly in the southern part of the county and in all municipality centers. Such infill policy encourages more compact patterns of development. It is likely that the high demand for new housing and commercial land will make it difficult to control development solely by zoning. Activation of impact fees for new developments and limiting building permits in particular areas may be additional options to prevent excessive sprawl development in this county.
2. **Housing Development Policy**

Since a high level of housing price negatively impacts low-income people, the county needs a policy that prevents total exclusion of low-income people in this area. One policy to accommodate them is to build single-family housing at affordable prices, in the northern part of the county, where land prices are lower.

3. **Zoning coordination and the provision of Comprehensive Planning**

Because zoning ordinances are mainly prepared by municipalities/townships, with a limited involvement of the county government, it would be useful for adjacent municipalities and townships to coordinate their zoning. Such coordination and cooperation would harmonize and synchronize zoning, particularly in boundary areas between adjacent municipalities. Although zoning and comprehensive planning have significant impacts on housing values, only ten of the townships/municipalities have formulated a Comprehensive Plan. Since this document complements zoning, especially in regulating density and maximum parcels size, all municipalities/townships should have this kind of document in place. Like zoning coordination, Comprehensive Planning should be coordinated among adjacent municipalities/townships.
Future research should focus on the following:

1. **Capturing data for adjacent counties and expanding housing markets**

   One limitation of this research is the unavailability of data for adjacent counties. In order to analyze the complete set of neighborhood characteristics, it is recommended that future research capture data for adjacent counties. This research has shown that zoning and existing land uses are significant variables in explaining housing values in suburban areas, while, however, focusing on a single housing market. To capture a broader market, future research could include multi-county/regional housing markets, which would make it possible to compare the price determinants of different housing markets.

2. **Expanding the neighborhood concept**

   This research has considered three different neighborhoods, defined as circular buffers with radii of 100 meters (immediate neighborhood), 400 meters (medium neighborhood), and 1 mile (large neighborhood). Other types and sizes of neighborhoods could be explored in future research, in order to further test neighborhood effects. Also, future research could focus on econometric specifications different from the Box-Cox transformation.

3. **Expanding the composition of variables**

   Future research could expand the focus on the interaction between existing land uses and zoning, for example accounting for the difference between (1) existing residential land use zoned for residential use, and (2) existing agricultural land use zoned as residential. Such combinations would allow for broader testing of the impacts of zoning on housing values.
BIBLIOGRAPHY


88. Orfield, Myron, and Thomas Luce. 2002. *Ohio Metro Patterns: A Regional Agenda for Community and Stability*. Ameregis, Minneapolis, MN.


APPENDIX A:

PARCELS DATA
### Appendix A  Parcels Data

<table>
<thead>
<tr>
<th>ITEM</th>
<th>ITEM DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARCEL-NO</td>
<td>Parcel Identification Number</td>
</tr>
<tr>
<td>TAXDIST</td>
<td>Tax District</td>
</tr>
<tr>
<td>TAXID</td>
<td>Old Tax ID</td>
</tr>
<tr>
<td>OWNER1</td>
<td>Owner's last &amp; first name</td>
</tr>
<tr>
<td>OWNER2</td>
<td>Co-Owner's first name</td>
</tr>
<tr>
<td>ADDR1</td>
<td>Street number and name</td>
</tr>
<tr>
<td>ADDR2</td>
<td>City, State and zip code</td>
</tr>
<tr>
<td>LEGAL</td>
<td>Legal Description</td>
</tr>
<tr>
<td>CLASS</td>
<td>Property Class</td>
</tr>
<tr>
<td>ACRES</td>
<td>Property Acreage</td>
</tr>
<tr>
<td>MARKET_LAND</td>
<td>Land's Market Value</td>
</tr>
<tr>
<td>MARKET_IMPR</td>
<td>Improvement's Market Value</td>
</tr>
<tr>
<td>MARKET_TOTAL</td>
<td>Total Market Value</td>
</tr>
<tr>
<td>TAXABLE_LAND</td>
<td>Land's Taxable Value</td>
</tr>
<tr>
<td>TAXABLE_IMPR</td>
<td>Improvement's Taxable Value</td>
</tr>
<tr>
<td>TAXABEL_TOTAL</td>
<td>Total Taxable Value</td>
</tr>
<tr>
<td>CAUV</td>
<td>Current Agriculture Use Valuation</td>
</tr>
<tr>
<td>SALE1_DATE</td>
<td>Most recent sales date</td>
</tr>
<tr>
<td>SALE2_DATE</td>
<td>Second recent sales date</td>
</tr>
<tr>
<td>SALE1_AMNT</td>
<td>Most recent sales value</td>
</tr>
<tr>
<td>SALE2_AMNT</td>
<td>Second recent sales value</td>
</tr>
<tr>
<td>SALE1_NUM</td>
<td>Most recent sales number</td>
</tr>
<tr>
<td>SALE2_NUM</td>
<td>Second recent sales number</td>
</tr>
<tr>
<td>SALE1_CONVEY</td>
<td>Most recent conveyance number</td>
</tr>
<tr>
<td>SALE2_CONVEY</td>
<td>Second recent conveyance number</td>
</tr>
<tr>
<td>NUM_L-UNITS</td>
<td>Number of living units</td>
</tr>
<tr>
<td>STORYHGT</td>
<td>Story Height</td>
</tr>
<tr>
<td>BASEMENT</td>
<td>Basement</td>
</tr>
<tr>
<td>CRAWLSP</td>
<td>Crawl Space</td>
</tr>
<tr>
<td>YRBUILT</td>
<td>Year Built</td>
</tr>
<tr>
<td>YRREMOD</td>
<td>Year Remodeled</td>
</tr>
<tr>
<td>TOT_FIN_L</td>
<td>Total Finished Living Area</td>
</tr>
<tr>
<td>ROOMS_TOT</td>
<td>Total Rooms</td>
</tr>
<tr>
<td>BDROOMS</td>
<td>Number of Bedrooms</td>
</tr>
<tr>
<td>FAMROOMS</td>
<td>Family-Rooms</td>
</tr>
<tr>
<td>DINROOMS</td>
<td>Dining-Rooms</td>
</tr>
<tr>
<td>HAFBATHS</td>
<td>Half Bathrooms</td>
</tr>
<tr>
<td>FULBATHS</td>
<td>Full Bathrooms</td>
</tr>
<tr>
<td>ATTIC</td>
<td>Attic</td>
</tr>
<tr>
<td>GARAGE_TYPE</td>
<td>Garage Type</td>
</tr>
<tr>
<td>GARAGE_CAP</td>
<td>Number of Garages</td>
</tr>
</tbody>
</table>

(continued)
### Appendix A : Continued

<table>
<thead>
<tr>
<th>ITEM</th>
<th>ITEM DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>RECROOM</td>
<td>Recreation Room</td>
</tr>
<tr>
<td>AIRCON</td>
<td>Air-conditioning</td>
</tr>
<tr>
<td>FIREPL_STACKS</td>
<td>Number of Fireplaces</td>
</tr>
<tr>
<td>HEATING</td>
<td>Heating Type</td>
</tr>
<tr>
<td>GRADE</td>
<td>Grade</td>
</tr>
<tr>
<td>GRADE_ADJSIGN</td>
<td>Grade Adjustment</td>
</tr>
<tr>
<td>GRADE_ADJFACT</td>
<td>Grade Adjustment Factor</td>
</tr>
<tr>
<td>MAILNAME1</td>
<td>Owner's last and first name for mailing purposes</td>
</tr>
<tr>
<td>MAILNAME2</td>
<td>CO Owner's first name for mailing purposes</td>
</tr>
<tr>
<td>MAILADDR1</td>
<td>Street number and name for mailing purposes</td>
</tr>
<tr>
<td>MAILADDR2</td>
<td>City State and Zip Code for mailing purposes</td>
</tr>
<tr>
<td>SCHOOL</td>
<td>School District Name</td>
</tr>
<tr>
<td>TWP_NAME</td>
<td>Township Name</td>
</tr>
<tr>
<td>MUN_NAME</td>
<td>Municipality Name</td>
</tr>
<tr>
<td>SUB_ID</td>
<td>A Subdivision Location Number</td>
</tr>
<tr>
<td>SUB_NAME</td>
<td>Subdivision Name as it appears on the plat</td>
</tr>
<tr>
<td>MAPSHEET</td>
<td>Geographic Location of the parcel</td>
</tr>
</tbody>
</table>
## Appendix B  Master Data of Delaware County

<table>
<thead>
<tr>
<th>ITEM</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>PACEL_NO</td>
<td>Parcel Identification Number</td>
</tr>
<tr>
<td>IMAGENAME</td>
<td>Image Identifier with .jpg extension</td>
</tr>
<tr>
<td>X_COORD</td>
<td>Structure's X coordinate</td>
</tr>
<tr>
<td>YCOORD</td>
<td>Structure's Y coordinate</td>
</tr>
<tr>
<td>CLASS_CODE</td>
<td>Land Use Code (see table below)</td>
</tr>
<tr>
<td>LANDUSE</td>
<td>Land Use Description</td>
</tr>
<tr>
<td>STREETNUMB</td>
<td>Structure's Number</td>
</tr>
<tr>
<td>UNIT_TYPE</td>
<td>Unit Type (i.e APT, STE, etc.)</td>
</tr>
<tr>
<td>UNITNUMBER</td>
<td>Multy family Structure Unit Numbers</td>
</tr>
<tr>
<td>PREDIR</td>
<td>Prefix Direction</td>
</tr>
<tr>
<td>PRETYPE</td>
<td>Road Pre-types (i.e. State Routes, US Highway)</td>
</tr>
<tr>
<td>NAME</td>
<td>Road Name</td>
</tr>
<tr>
<td>TYPE</td>
<td>Road Type</td>
</tr>
<tr>
<td>SUFDIR</td>
<td>Suffix Direction</td>
</tr>
<tr>
<td>CITY</td>
<td>City's Names</td>
</tr>
<tr>
<td>STATE</td>
<td>State's Designation</td>
</tr>
<tr>
<td>ZIP_CODE</td>
<td>5-Digit Zip Code</td>
</tr>
<tr>
<td>ZIP+4</td>
<td>4 digit zip +4 code (i.e.7905)</td>
</tr>
<tr>
<td>PRIMARYIMA</td>
<td>Primary Image Designation (True or False)</td>
</tr>
<tr>
<td>VISIBLE</td>
<td>Structure's Visibility from Road ROW (True or False)</td>
</tr>
<tr>
<td>RETAKE</td>
<td>Necessity for Retake of the Photo (True or False)</td>
</tr>
<tr>
<td>MSAG_CODE</td>
<td>E-911’s Street Segment Codes</td>
</tr>
<tr>
<td>SOURCE</td>
<td>Source of data (See table below)</td>
</tr>
<tr>
<td>REG_NO</td>
<td>Mobile Home's Registration Number</td>
</tr>
<tr>
<td>PARK_CODE</td>
<td>Mobile Home Parks Number</td>
</tr>
<tr>
<td>MH_PIN</td>
<td>Mobile Homes Parcel Number</td>
</tr>
<tr>
<td>ALIAS</td>
<td>Name of Establishment</td>
</tr>
<tr>
<td>OCC_STAT</td>
<td>Occupancy Status Designation</td>
</tr>
<tr>
<td>ADDRESS</td>
<td>Address of Property</td>
</tr>
</tbody>
</table>

(continued)
Appendix B: Continued

<table>
<thead>
<tr>
<th>BUILDING USE</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential (Single family)</td>
<td>1</td>
</tr>
<tr>
<td>Residential (Duplex)</td>
<td>2</td>
</tr>
<tr>
<td>Residential (Triplex)</td>
<td>3</td>
</tr>
<tr>
<td>Residential (Condominium)</td>
<td>4</td>
</tr>
<tr>
<td>Residential (Rental Apartments)</td>
<td>5</td>
</tr>
<tr>
<td>Residential (Mobile Homes)</td>
<td>6</td>
</tr>
<tr>
<td>Residential (Camp Ground, Other)</td>
<td>7</td>
</tr>
<tr>
<td>Commercial (Retail)</td>
<td>8</td>
</tr>
<tr>
<td>Commercial (Neighborhood/Regional Shopping Center)</td>
<td>9</td>
</tr>
<tr>
<td>Commercial (Office)</td>
<td>10</td>
</tr>
<tr>
<td>Commercial (Other)</td>
<td>11</td>
</tr>
<tr>
<td>Industrial (Heavy)</td>
<td>12</td>
</tr>
<tr>
<td>Industrial (Light)</td>
<td>13</td>
</tr>
<tr>
<td>Agriculture</td>
<td>14</td>
</tr>
<tr>
<td>Exempt, Utility</td>
<td>15</td>
</tr>
<tr>
<td>Vacant</td>
<td>16</td>
</tr>
<tr>
<td>Commercial (condominium)</td>
<td>17</td>
</tr>
</tbody>
</table>
APPENDIX C:

SAS PROGRAMS
SAS PROGRAM FOR DERIVING HOUSING CHARACTERISTICS AND CONDITIONS DATA BASED ON PARCELS DATA

PROC IMPORT DATAFILE = 'c:\variable_a_b\SFH_2000_UID.xls' out = sfh;
DATA sfh2000;
SET sfh;
PARCEL_N2 = PUT(PARCEL_NO, 14.);

DATA _NULL_; SET sfh2000; FILE OUT DSD DELIMITER =','; FILENAME OUT 'c:\variable_a_b\SFH_2000_UIDNEW.DA1';
PUT PARCEL_N2 $ 1-14 TAXDIST 15-16 CLASS 17-19 ACRES 20-28 3 SALE1_DATE 29-37 SALE1_AMNT 38-46 STORYHGT 47-49 1 BASEMENT $ 50-54 YRBUILT 55-59 YRREMOD 60-64 TOT_FIN_L_ 65-70 ROOMS_TOT 71-72 BDROOMS 73-74 FULBATHS 75-76 HAFBATHS 77-78 GARAGE_CAP 79-81 1 AIRCOND $ 82-89 FIREPL_STA 90-91 HEATING $ 92-96 SCHOOL $ 97-117 TWP_NAME $ 118-130 SHAPE_AREA 131-147 4 SOURCEUID 148-152;

RUN;
SAS PROGRAM FOR COMPUTING ALLOCATION FACTORS OF BLOCKS
BASED ON INTERSECTION OF RESIDENTIAL, BLOCK, AND BUFFER’S SHAPE FILES
---------------------------------------------------------------------;

PROC IMPORT DATAFILE = 'c:\RESBLOBU_400NEW\JOINALF_400APR.dbf' out = s
REPLACE;
DATA s;
SET s;
IF Sum_AREA = 0 THEN DO; ALF_AREA = 0; POP_AREA = 0; END;
IF Sum_AREA > 0 THEN DO; ALF_AREA = AREA/Sum_AREA; POP_AREA = ALF_AREA*
TOTAL; END;
IF AREA_1 = 0 THEN DO; BED_BU = 0; POP_BED = 0; END;
IF SHAPE_AR_1 = 0 THEN DO; BED_BU = 0; POP_BED = 0; END;
IF MULTI_ID > 0 THEN BED_BUM = MAX(EXISTING_U, TOTAL_UNIT) * 2 *
AREA/AREA_1; ELSE BED_BUM = 0;
IF MULTI_ID = 0 THEN BED_BUS = BDROOMS_1 * AREA/SHAPE_AR_1; ELSE
BED_BU = BED_BUM + BED_BUS;
IF Sum_BED_BL = 0 THEN DO; ALF_BED = 0; POP_BED = 0; END;
IF Sum_BED_BL > 0 THEN DO; ALF_BED = BED_BU/Sum_BED_BL;
POP_BED = ALF_BED * TOTAL; END;
FILENAME OUT 'C:\BUF400\ALFFID_APR_BUF400.dat';
DATA _NULL_; FILE OUT;
PUT PARCEL_N_2 1-14 STFID 15-30 ALF_AREA 31-41 5 ALF_BED 42-52 5
POP_AREA 53-63 3 POP_BED 64-74 3 FID_BUF_AP 75-79;

DATA jan;
INFILE 'c:\BUF400\ALFFID_APR_BUF400.dat' TRUNCOVER;
INPUT PARCEL_N_2 1-14 STFID 15-30 ALF_AREA 31-41 5 ALF_BED 42-52 5
POP_AREA 53-63 3 POP_BED 64-74 3 FID_BUF_AP 75-79;

DATA fid;
INFILE 'c:\sucahyono7\july1\apr2000_fid.dat';
INPUT FID_BUF_AP 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
PROC SORT DATA = fid;
BY FID_BUF_AP;

* MERGE data sets by FID;
DATA janfid;
MERGE jan(IN=A) fid(IN=B); BY FID_BUF_AP; IF A;
PROC SORT DATA = janfid;
BY PARCEL_N2;
FILENAME OUT 'C:\BUF400\ALFFID_APR_ALL_BUF400.dat';
DATA _NULL_; FILE OUT;
PUT PARCEL_N2 1-14 STFID 15-30 ALF_AREA 31-41 5 ALF_BED 42-52 5
POP_AREA 53-63 3 POP_BED 64-74 3;
RUN;
*------------------------------------------------------------
SAS PROGRAM FOR COMPUTING ALLOCATION FACTORS OF BLOCK GROUPS 
BASED ON ALLOCATION FACTORS OF BLOCKS 
-------------------------------------------------------------;

*READ ALLOCATION FACTORS OF BLOCKS 
------------------------------------;
DATA jan;
INFILE 'C:\BUF100_SEP\ALF_BUF100_N2.dat' TRUNCOVER;
INPUT PARCEL_N2 1-14 STFID 15-30 ALF_AREA 31-41 5 ALF_BED 42-52 5 
POP_AREA 53-63 3 POP_BED 64-74 3;
PROC SORT DATA = jan;
   BY STFID; FORMAT STFID 15.;
DATA block;
INFILE 'c:\sucahyono7\sumblock_areabed2.dat';
INPUT STFID 1-15 SUM_AREA 16-28 3 SUM_BED_BL 29-41 3;
PROC SORT DATA = block;
   BY STFID; FORMAT STFID 15.;
* MERGE data sets by STFID;
DATA janblock;
   MERGE jan(IN=A) block(IN=B); BY STFID; IF A;
      AREA = ALF_AREA*SUM_AREA;
      BEDS = ALF_BED*SUM_BED_BL;
PROC SORT DATA = janblock; BY STFID; FORMAT STFID 15. PARCEL_N2 14. 
AREA 12.3 BEDS 12.3;

* BLOCK GROUPS DATA 
---------------------;
DATA bg;
INFILE 'c:\BGSTFID_STFID.dat';
INPUT STFID 1-15 BGSTFID 16-28;
PROC SORT DATA = bg;
   BY STFID; FORMAT STFID 15.;
* MERGE data sets by STFID;
DATA janblockbg;
   MERGE janblock(IN=A) bg(IN=B); BY STFID; IF A;
PROC SORT DATA = janblockbg; BY PARCEL_N2 BGSTFID; FORMAT PARCEL_N2 14. 
BGSTFID 12.;
PROC MEANS NOPRINT SUM; VAR AREA BEDS; BY PARCEL_N2 BGSTFID;
   OUTPUT OUT = BGJAN SUM= AREA2 BEDS2;
PROC SORT DATA = BGJAN; BY BGSTFID; FORMAT BGSTFID 12.;
DATA sumbg;
  INFILE 'c:\sucahyono7\SUMBG_AREABED2.dat';
  INPUT BGSTFID 1-12 SUM_AREABG 13-25 3 SUM_BEDBG 26-38 3;

PROC SORT DATA = sumbg;
  BY BGSTFID; FORMAT BGSTFID 12.;

* ALLOCATION FACTORS OF BLOCK GROUPS
-------------------------------------;

DATA calcalf;
  MERGE BGJAN(IN=A) sumbg(IN=B); BY BGSTFID; IF A;

    ALF_AREABG = AREA2/SUM_AREABG;
    ALF_BEDBG = BEDS2/SUM_BEDBG;

PROC SORT DATA= calcalf; BY PARCEL_N2; FORMAT BGSTFID 12. PARCEL_N2 14.;

FILENAME OUT 'c:\BUF100_SF3\ALFBG_SF3_BUF100.dat';
DATA _NULL_;
  SET CALCALF; FILE OUT;
  PUT PARCEL_N2 1-14 BGSTFID 15-27 AREA2 28-40 3 BEDS2 41-53 3 SUM_AREABG
      54-66 3 SUM_BEDBG 67-79 3 ALF_AREABG 80-88 5 ALF_BEDBG 89-97 5;
  FORMAT BGSTFID 12. PARCEL_N2 14.;
RUN;
* SAS PROGRAM FOR CONVERTING DATA FROM BLOCKS INTO BUFFER'S DATA
BASED ON BLOCK'S ALLOCATION FACTORS AND STF1 DATA
---------------------------------------------------------------;

* DERIVING STF1 DATA FOR DELAWARE COUNTY FROM CENSUS 2000
---------------------------------------------------------;

FILENAME IN1 'c:\SF1.DA0' LRECL=16886;
DATA SF1; INFILE IN1 DSD TRUNCOVER DELIMITER = ',';
   LENGTH  KEY  $ 26 COUNTY $ 3 TRACT $ 6 BG $ 1 BLK $ 3;
   INPUT
     KEY $;
     COUNTY=SUBSTR(KEY,3,3);
     TRACT=SUBSTR(KEY,16,6);
     BG=SUBSTR(KEY,22,1);
     BLK=SUBSTR(KEY,24,3);
     DROP KEY;
   IF COUNTY = '041';
   STFID = 39*10000000000000 + COUNTY*10000000000 + TRACT*10000 + BG*1000 + BLK;
   DROP COUNTY TRACT BG BLK;
   proc sort data = SF1; BY STFID;
DATA _NULL_; SET SF1; FILE OUT DSD DELIMITER = ','; FILENAME OUT 'c:\SF1\DELA_STFID_X595.DA1' LRECL=16886;
   PUT STFID 1-15 X1-X595;
RUN;

* CONVERSION OF BLOCK DATA INTO BUFFER 400 M
---------------------------------------------;

FILENAME IN1 'c:\SF1\DELA_STFID_X595.DA1' LRECL=16886;
FILENAME IN2 'c:\BUF400\ALFFID_JAN_ALL_BUF400.dat';
FILENAME OUT 'c:\BUF400\DELA_X595_JAN_BUF400.DA1' LRECL=64464;

*SPLIT DATA
-----------;

DATA SF1; INFILE IN1 DSD TRUNCOVER DELIMITER = ',';
   INPUT STFID 1-15
     X1-X595;
DATA SF1A; SET SF1; KEEP STFID X1-X100;
PROC SORT DATA= SF1A; BY STFID;
DATA SF1B; SET SF1; KEEP STFID X101-X200;
PROC SORT DATA=SF1B; BY STFID;
DATA SF1C; SET SF1; KEEP STFID X201-X300;
PROC SORT DATA=SF1C; BY STFID;
DATA SF1D; SET SF1; KEEP STFID X301-X400;
PROC SORT DATA= SF1D; BY STFID;
DATA SF1E; SET SF1; KEEP STFID X401-X500;
PROC SORT DATA= SF1E; BY STFID;
DATA SF1F; SET SF1; KEEP STFID X501-X595;
PROC SORT DATA= SF1F; BY STFID;

*READ ALLOCATION FACTORS OF BLOCKS
------------------------------------;

DATA JAN; INFILE IN2;
INPUT PARCEL_N2 1-14 STFID 15-30 ALF_AREA 31-41 5 ALF_BED 42-52 5
POP_AREA 53-63 3 POP_BED 64-74 3;
PROC SORT DATA= JAN; BY STFID;

* COMPUTATION OF BUFFER DATA GRADUALLY
---------------------------------------;
DATA SF1;
   MERGE SF1A(IN=A) JAN(IN=B); BY STFID; IF B;
      ARRAY XX(100) X1-X100;
      ARRAY XXA(100) XA1-XA100;
      DO I = 1 TO 100;
         XXA(I)=XX(I)*ALF_BED;
      END;
      DROP X1-X100 I;
PROC SORT DATA= SF1; BY PARCEL_N2; FORMAT PARCEL_N2 14.;
PROC MEANS DATA=SF1 NOPRINT SUM; VAR XA1-XA100; BY PARCEL_N2;
   OUTPUT OUT= topa SUM=XAA1-XAA100;

DATA SF1;
   MERGE SF1B(IN=A) JAN(IN=B); BY STFID; IF B;
      ARRAY XX(100) X101-X200;
      ARRAY XXA(100) XA101-XA200;
      DO I = 1 TO 100;
         XXA(I)=XX(I)*ALF_BED;
      END;
      DROP X101-X200 I;
PROC SORT DATA= SF1; BY PARCEL_N2; FORMAT PARCEL_N2 14.;
PROC MEANS DATA=SF1 NOPRINT SUM; VAR XA101-XA200; BY PARCEL_N2;
   OUTPUT OUT= topb SUM=XAA101-XAA200;

DATA SF1;
   MERGE SF1C(IN=A) JAN(IN=B); BY STFID; IF B;
      ARRAY XX(100) X201-X300;
      ARRAY XXA(100) XA201-XA300;
      DO I = 1 TO 100;
         XXA(I)=XX(I)*ALF_BED;
      END;
      DROP X201-X300 I;
PROC SORT DATA= SF1; BY PARCEL_N2; FORMAT PARCEL_N2 14.;
PROC MEANS DATA=SF1 NOPRINT SUM; VAR XA201-XA300; BY PARCEL_N2;
   OUTPUT OUT= topc SUM=XAA201-XAA300;

DATA SF1;
   MERGE SF1D(IN=A) JAN(IN=B); BY STFID; IF B;
ARRAY XX(100) X301-X400;
ARRAY XXA(100) XA301-XA400;
DO I = 1 TO 100;
  XXA(I)=XX(I)*ALF_BED;
END;
DROP X301-X400 I;
PROC SORT DATA= SF1; BY PARCEL N2; FORMAT PARCEL N2 14.;
PROC MEANS DATA=SF1 NOPRINT SUM; VAR XA301-XA400; BY PARCEL N2;
  OUTPUT OUT= topd SUM=XAA301-XAA400;

DATA SF1;
  MERGE SF1E(IN=A) JAN(IN=B); BY STFID; IF B;
  ARRAY XX(100) X401-X500;
  ARRAY XXA(100) XA401-XA500;
  DO I = 1 TO 100;
    XXA(I)=XX(I)*ALF_BED;
  END;
  DROP X401-X500 I;
PROC SORT DATA= SF1; BY PARCEL N2; FORMAT PARCEL N2 14.;
PROC MEANS DATA=SF1 NOPRINT SUM; VAR XA401-XA500; BY PARCEL N2;
  OUTPUT OUT= tope SUM=XAA401-XAA500;

DATA SF1;
  MERGE SF1F(IN=A) JAN(IN=B); BY STFID; IF B;
  ARRAY XX(95) X501-X595;
  ARRAY XXA(95) XA501-XA595;
  DO I = 1 TO 95;
    XXA(I)=XX(I)*ALF_BED;
  END;
  DROP X501-X595 I;
PROC SORT DATA= SF1; BY PARCEL N2; FORMAT PARCEL N2 14.;
PROC MEANS DATA=SF1 NOPRINT SUM; VAR XA501-XA595; BY PARCEL N2;
  OUTPUT OUT= topf SUM=XAA501-XAA595;

DATA SF1; MERGE topa topb topc topd tope topf; BY PARCEL N2; FILE OUT
DSD DELIMITER =','; FILENAME OUT 'c:\BUF400\DELA_X595_JAN_BUF400.DA1'
LRECL=64464;
PUT PARCEL N2 1-14
  XAA1-XAA595; FORMAT PARCEL N2 14.;
run;

* MERGING STF1 DATA FROM X1 TO X595 FOR ALL BUF400
--------------------------------------------------;
FILENAME IN10 'c:\BUF400\DELA_X595_OCT_BUF400.DA1' LRECL=64464;
FILENAME IN11 'c:\BUF400\DELA_X595_NOV_BUF400.DA1' LRECL=64464;
FILENAME IN12 'c:\BUF400\DELA_X595_DEC_BUF400.DA1' LRECL=64464;
FILENAME OUT 'c:\BUF400\DELA_SF1_X595_JANDEC_BUF400.DA1' LRECL=64464;

DATA SF1A; INFILE IN1 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1A; BY PARCEL N2;

DATA SF1B; INFILE IN2 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1B; BY PARCEL N2;

DATA SF1C; INFILE IN3 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1C; BY PARCEL N2;

DATA SF1D; INFILE IN4 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1D; BY PARCEL N2;

DATA SF1E; INFILE IN5 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1E; BY PARCEL N2;

DATA SF1F; INFILE IN6 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1F; BY PARCEL N2;

DATA SF1G; INFILE IN7 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1G; BY PARCEL N2;

DATA SF1H; INFILE IN8 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1H; BY PARCEL N2;

DATA SF1I; INFILE IN9 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1I; BY PARCEL N2;

DATA SF1J; INFILE IN10 DSD TRUNCOVER DELIMITER = ',
   INPUT PARCEL N2 1-14 XAA1-XAA595; FORMAT PARCEL N2 14.;
PROC SORT DATA= SF1J; BY PARCEL_N2;

DATA SF1K; INFILE IN11 DSD TRUNCOVER DELIMITER = ',';
   INPUT PARCEL_N2 1-14
       XAA1-XAA595; FORMAT PARCEL_N2 14.;
PROC SORT DATA= SF1K; BY PARCEL_N2;

DATA SF1L; INFILE IN12 DSD TRUNCOVER DELIMITER = ',';
   INPUT PARCEL_N2 1-14
       XAA1-XAA595; FORMAT PARCEL_N2 14.;
PROC SORT DATA= SF1L; BY PARCEL_N2;

DATA SF1ALL;
   MERGE SF1A SF1B SF1C SF1D SF1E SF1F SF1G SF1H SF1I SF1J SF1K SF1L;
   BY PARCEL_N2; FILE OUT DSD DELIMITER =','; FILENAME OUT 'c:\BUF400\DELA_SF1_X595_JANDEC_BUF400.DA1' LRECL=64464;
   PUT PARCEL_N2 1-14
       XAA1-XAA595; FORMAT PARCEL_N2 14.;
run;
*SAS PROGRAM FOR CONVERTING DATA FROM BLOCK GROUPS INTO BUFFER’S DATA BASED ON BLOCK GROUP’S ALLOCATION FACTORS AND STF3 DATA*

FILENAME IN1 'c:\SF3\SF3.DA3' LRECL=16886;
FILENAME IN2 'c:\BUF400_SF3\ALFFID_BGAPR_ALL_BUF400.dat' LRECL=97;
FILENAME OUT 'c:\BUF400_SF3\ALFFID_SF3DAB_BGAPR_BUF400.DA1'
LRECL=64464;

*READ STF3 DATA FOR DELAWARE COUNTY
------------------------------------;
DATA SF3; INFILE IN1 DSD TRUNCOVER DELIMITER = ',';
LENGTH COUNTY $ 3
TRACT $ 6
BLKGRP $ 1;
INPUT COUNTY $;
TRACT $;
BLKGRP $;
X1-X5453; /*THERE ARE 5453 SF3 VARIABLES IN SF3.DA3*/
IF COUNTY = '041';
BGSTFID = 39*10000000000 + COUNTY*10000000 + TRACT*10 + BLKGRP;
PROC SORT DATA=SF3; BY BGSTFID;

*READ ALLOCATION FACTORS OF BLOCK GROUPS
----------------------------------------;
DATA ALF_BGAPR;
INFILE IN2;
INPUT PARCEL_N2 1-14 BGSTFID 15-27 AREA2 28-40 3 BEDS2 41-53 3 SUM_AREABG 54-66 3 SUM_BEDBG 67-79 3 ALF_AREABG 80-88 5 ALF_BEDBG 89-97 5;
FORMAT BGSTFID 12. PARCEL_N2 14.;
PROC SORT DATA = ALF_BGAPR; BY BGSTFID;

DATA MER;
MERGE SF3(IN=A) ALF_BGAPR(IN=B); BY BGSTFID; IF B;
ARRAY XX(5453) X1-X5453;
ARRAY XXA(5453)XA1-XA5453;

* CONVERSION DATA FROM STF3 OF BLOCK GROUPS INTO BUFFERS DATA
-------------------------------------------------------------;
DO I = 1 TO 5453;
XXA(I)=XX(I)*ALF_BEDBG;
END;
DROP X1-X5453 I;
PROC SORT DATA= MER; BY PARCEL_N2; FORMAT PARCEL_N2 14.;
PROC MEANS NOPRINT SUM; VAR XA1-XA5453; BY PARCEL_N2;
OUTPUT OUT= top SUM=XAA1-XAA5453;
DATA _NULL_; SET top; FILE OUT DSD DELIMITER =','; FILENAME OUT 'c:\BUF400_SF3\ALFFID_SF3DAB_BGAPR_BUF400.DA1' LRECL=64464;
PUT PARCEL_N2 $ XAA1-XAA5453; FORMAT PARCEL_N2 14.;
RUN;
SAS PROGRAM FOR COMPUTING AVERAGE PRICE OF HOUSING IN NEIGHBORHOOD (BUFFER 1 MILE) BASED ON INTERSECTION OF BUFFER 1 MILE AND PARCEL’S SHAPE FILES

* MONTH: JANUARY;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS\BUF1M_PARCELS_JAN.dbf' replace
  out = jan;
DATA jan;
SET jan;
IF CLASS_1 >= 510 AND CLASS_1 <= 515;
KEEP FID_BUF_JA MARKET_T_1;
proc sort data=jan; by FID_BUF_JA;
PROC MEANS NOPRINT DATA=jan; VAR MARKET_T_1; BY FID_BUF_JA;
   OUTPUT OUT= topjan MEAN(MARKET_T_1)= AVPR3;
data topjan;
set topjan;
proc sort data=topjan; BY FID_BUF_JA;

DATA a1; INFILE 'c:\sucahyono7\SFH_JANFID_NEW.dat' DSD DELIMITER = ','; INPUT FID_BUFFER 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
FID_BUF_JA = FID_BUFFER;
proc sort data = a1; BY FID_BUF_JA;
data janal; merge topjan(IN=A) a1(IN=B); IF A;BY FID_BUF_JA;
proc sort data = janal; BY PARCEL_N2;

* MONTH: FEBRUARY;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS\BUF1M_PARCELS_FEB.dbf' replace
  out = feb;
DATA feb;
SET feb;
IF CLASS_1 >= 510 AND CLASS_1 <= 515;
KEEP FID_BUF_FE MARKET_T_1;
proc sort data=feb; by FID_BUF_FE;
PROC MEANS NOPRINT DATA=feb; VAR MARKET_T_1; BY FID_BUF_FE;
   OUTPUT OUT= topfeb MEAN(MARKET_T_1)= AVPR3;
data topfeb;
set topfeb;
proc sort data=topfeb; BY FID_BUF_FE;

DATA a2; INFILE 'e:\sucahyono7\july1\feb2000_fid.dat' DSD DELIMITER = ','; INPUT FID_BUF_FE 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a2; BY FID_BUF_FE;
data feba2; merge topfeb(IN=A) a2(IN=B); IF A;BY FID_BUF_FE;
proc sort data = feba2; BY PARCEL_N2;
* MONTH: MARCH;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS2\buf1m_parmar.dbf' replace out = mar;
DATA mar;
SET mar;
IF CLASS >= 510 AND CLASS <= 515;
KEEP FID_BUF_MA MARKET_TOT;
proc sort data=mar; by _FID_BUF_MA_

PROC MEANS NOPRINT DATA=mar; VAR MARKET_TOT; BY FID_BUF_MA;
OUTPUT OUT= topmar MEAN(MARKET_TOT)= AVPR3;
data topmar;
set topmar;
proc sort data=topmar; BY FID_BUF_MA;

DATA a3; INFILE 'e:\sucahyono7\july1\mar2000_fid.dat' DSD DELIMITER = ',',;
INPUT FID_BUF_MA 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a3; BY FID_BUF_MA;

data mara3; merge topmar(IN=A) a3(IN=B); IF A;BY FID_BUF_MA;
proc sort data = mara3; BY PARCEL_N2;

* MONTH: APRIL;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS2\buf1m_par_apr.dbf' replace out = apr;
DATA apr;
SET apr;
IF CLASS >= 510 AND CLASS <= 515;
KEEP FID_BUF_AP MARKET_TOT;
proc sort data=apr; by FID_BUF_AP;

PROC MEANS NOPRINT DATA=apr; VAR MARKET_TOT; BY FID_BUF_AP;
OUTPUT OUT= topapr MEAN(MARKET_TOT)= AVPR3;
data topapr;
set topapr;
proc sort data=topapr; BY FID_BUF_AP;

DATA a4; INFILE 'e:\sucahyono7\july1\apr2000_fid.dat' DSD DELIMITER = ',',;
INPUT FID_BUF_AP 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a4; BY FID_BUF_AP;

data apra4; merge topapr(IN=A) a4(IN=B); IF A;BY FID_BUF_AP;
proc sort data = apra4; BY PARCEL_N2;

* MONTH: MAY;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS2\bulmay.dbf' replace out = may;
DATA may;
SET may;
IF CLASS >= 510 AND CLASS <= 515;
KEEP FID_BUF_MA MARKET_TOT;
proc sort data=may; by FID_BUF_MA;

PROC MEANS NOPRINT DATA=may; VAR MARKET_TOT; BY FID_BUF_MA;
   OUTPUT OUT= topmay MEAN(MARKET_TOT)= AVPR3;
data topmay;
set topmay;
proc sort data=topmay; by FID_BUF_MA;

DATA a5; INFILE 'e:\sucahyono7\july1\may2000_fid.dat' DSD DELIMITER = ',';
INPUT FID_BUF_MA 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a5; by FID_BUF_MA;

data maya5; merge topmay(IN=A) a5(IN=B); IF A;BY FID_BUF_MA;
proc sort data = maya5; by PARCEL_N2;

* MONTH: JUNE;
PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS2\buf1m_parjun.dbf' replace out = jun;
DATA jun;
SET jun;
IF CLASS >= 510 AND CLASS <= 515;
KEEP PARCEL_NO FID_BUF_JU;
proc sort data=jun; by PARCEL_NO;

PROC IMPORT DATAFILE =
'E:\BUF1M_PARCEL_PRICE\parcel_64283_short_new.xls' replace out = par;
DATA par;
SET par;
KEEP PARCEL_NO MARKET_TOT;
proc sort data=par; by PARCEL_NO;

data junpar; MERGE jun(IN=A) par(IN=B); IF A;BY PARCEL_NO;
Proc sort data=junpar; by PARCEL_NO;

PROC MEANS NOPRINT DATA=junpar; VAR MARKET_TOT; BY FID_BUF_JU;
   OUTPUT OUT= topjun MEAN(MARKET_TOT)= AVPR3;
data topjun;
set topjun;
proc sort data=topjun; by FID_BUF_JU;

DATA a6; INFILE 'e:\sucahyono7\july1\jun2000_fid.dat' DSD DELIMITER = ',';
INPUT FID_BUF_JU 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a6; by FID_BUF_JU;

data juna6; merge topjun(IN=A) a6(IN=B); IF A;BY FID_BUF_JU;
proc sort data = juna6; by PARCEL_N2;
* MONTH: JULY;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS2\buf1m_parjul.dbf' replace out = jul;
DATA jul;
SET jul;
IF CLASS >= 510 AND CLASS <= 515;
KEEP PARCEL_NO FID_RESBLO;
proc sort data=jul; by PARCEL_NO;

PROC IMPORT DATAFILE =
'E:\BUF1M_PARCEL_PRICE\parcel_64283_short_new.xls' replace out = par;
DATA par;
SET par;
KEEP PARCEL_NO MARKET_TOT;
proc sort data=par; by PARCEL_NO;

Data julpar; MERGE jul(IN=A) par(IN=B); IF A;BY PARCEL_NO;
Proc sort data=julpar; BY FID_RESBLO;

PROC MEANS NOPRINT DATA=julpar; VAR MARKET_TOT; BY FID_RESBLO;
    OUTPUT OUT= topjul MEAN(MARKET_TOT)= AVPR3;
data topjul;
set topjul;
proc sort data=topjul; BY FID_RESBLO;

DATA a7; INFILE 'e:\sucahyono7\july1\jul2000_fid.dat' DSD DELIMITER = ',,'
    INPUT FID_RESBLO 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a7; BY FID_RESBLO;

data jula7; merge topjul(IN=A) a7(IN=B); IF A;BY FID_RESBLO;
proc sort data = jula7; BY PARCEL_N2;

* MONTH: AUGUST;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS3\buf1m_paraug.dbf' replace out = aug;
DATA aug;
SET aug;
IF CLASS >= 510 AND CLASS <= 515;
KEEP PARCEL_NO FID_BUF_AU;
proc sort data=aug; by PARCEL_NO;

PROC IMPORT DATAFILE =
'E:\BUF1M_PARCEL_PRICE\parcel_64283_short_new.xls' replace out = par;
DATA par;
SET par;
KEEP PARCEL_NO MARKET_TOT;
proc sort data=par; by PARCEL_NO;

Data augpar; MERGE aug(IN=A) par(IN=B); IF A;BY PARCEL_NO;
Proc sort data=augpar; BY FID_BUF_AU;

PROC MEANS NOPRINT DATA=augpar; VAR MARKET_TOT; BY FID_BUF_AU;
PROC MEANS NOPRINT DATA=seppar; VAR MARKET_TOT; BY FID_BUF_SE;
   OUTPUT OUT= topsep MEAN(MARKET_TOT)= AVPR3;
data topsep;
set topsep;
proc sort data=topsep; BY FID_BUF_SE;

DATA a9; INFILE 'e:sucahyono7\july1\sep2000_fid.dat' DSD DELIMITER = ', '; INPUT FID_BUF_SE 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a9; BY FID_BUF_SE;

data sepa9; merge topsep(IN=A) a9(IN=B); IF A;BY FID_BUF_SE;
proc sort data = sepa9; BY PARCEL_N2;

* MONTH: OCTOBER;

PROC IMPORT DATAFILE = 'E:BULF1M_PARCELS3\buf1m_paroit.dbf' replace out = oct;
DATA oct;
SET oct;
IF CLASS >= 510 AND CLASS <= 515;
KEEP PARCEL_NO FID_BUF_OC;
PROC MEANS NOPRINT DATA=seppar; VAR MARKET_TOT; BY FID_BUF_OC;
   OUTPUT OUT= topsep MEAN(MARKET_TOT)= AVPR3;
data topsep;
set topsep;
proc sort data=topsep; BY FID_BUF_OC;
proc sort data=oct; by PARCEL_NO;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCEL_PRICE\parcel_64283_short_new.xls' replace out = par;
DATA par;
SET par;
KEEP PARCEL_NO MARKET_TOT;
proc sort data=par; by PARCEL_NO;

Data octpar; MERGE oct(IN=A) par(IN=B); IF A;BY PARCEL_NO;
Proc sort data=octpar; BY FID_BUF_OC;

PROC MEANS NOPRINT DATA=octpar; VAR MARKET_TOT; BY FID_BUF_OC;
OUTPUT OUT= topoct MEAN(MARKET_TOT)= AVPR3;
data topoct;
set topoct;
proc sort data=topoct; BY FID_BUF_OC;

DATA a10; INFILE 'e:\sucahyono7\july1\oct2000_fid.dat' DSD DELIMITER = ',';
INPUT FID_BUF_OC 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a10; BY FID_BUF_OC;

data octa10; merge topoct(IN=A) a10(IN=B); IF A;BY FID_BUF_OC;
proc sort data = octa10; BY PARCEL_N2;

* MONTH: NOVEMBER;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS3\buf1m_parnov.dbf' replace out = nov;
DATA nov;
SET nov;
IF CLASS >= 510 AND CLASS <= 515;
KEEP PARCEL_NO FID_BUF_NO;
proc sort data=nov; by PARCEL_NO;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCEL_PRICE\parcel_64283_short_new.xls' replace out = par;
DATA par;
SET par;
KEEP PARCEL_NO MARKET_TOT;
proc sort data=par; by PARCEL_NO;

Data novpar; MERGE nov(IN=A) par(IN=B); IF A;BY PARCEL_NO;
Proc sort data=novpar; BY FID_BUF_NO;

PROC MEANS NOPRINT DATA=novpar; VAR MARKET_TOT; BY FID_BUF_NO;
OUTPUT OUT= topnov MEAN(MARKET_TOT)= AVPR3;
data topnov;
set topnov;
proc sort data=topnov; BY FID_BUF_NO;

DATA a11; INFILE 'e:\sucahyono7\july1\nov2000_fid.dat' DSD DELIMITER = ',';
INPUT FID_BUF_NO 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = all; BY FID_BUF_NO;
data novall; merge topnov(IN=A) all(IN=B); IF A;BY FID_BUF_NO;
proc sort data = novall; BY PARCEL_N2;

* MONTH: DECEMBER;

PROC IMPORT DATAFILE = 'E:\BUF1M_PARCELS3\buf1m_pardec.dbf' replace out = dec;
DATA dec;
SET dec;
IF CLASS >= 510 AND CLASS <= 515;
KEEP PARCEL_NO FID_BUF_DE;
proc sort data=dec; by PARCEL_NO;

PROC IMPORT DATAFILE =
'E:\BUF1M_PARCEL_PRICE\parcel_64283_short_new.xls' replace out = par;
DATA par;
SET par;
KEEP PARCEL_NO MARKET_TOT;
proc sort data=par; by PARCEL_NO;

Data decpar; MERGE dec(IN=A) par(IN=B); IF A;BY PARCEL_NO;
Proc sort data=decpar; BY FID_BUF_DE;

PROC MEANS NOPRINT DATA=decpar; VAR MARKET_TOT; BY FID_BUF_DE;
OUTPUT OUT= topdec MEAN(MARKET_TOT)= AVPR3;
data topdec;
set topdec;
proc sort data=topdec; BY FID_BUF_DE;

DATA a12; INFILE 'e:\sucathyono7\july1\dec2000_fid.dat' DSD DELIMITER = ',';
INPUT FID_BUF_DE 1-4 PARCEL_N2 5-19; FORMAT PARCEL_N2 14.;
proc sort data = a12; BY FID_BUF_DE;

data deca12; merge topdec(IN=A) a12(IN=B); IF A;BY FID_BUF_DE;
proc sort data = deca12; BY PARCEL_N2;

*MERGING ALL MONTHS;

DATA merall; MERGE jan1 feba2 mara3 apra4 maya5 juna6 jula7 auga8
sepa9 octa10 novall deca12; BY PARCEL_N2; FORMAT PARCEL_N2 14.;
PROC SORT DATA=merall; BY PARCEL_N2;

DATA _NULL_; SET merall; FILE OUT DSD DELIMITER =','; FILENAME OUT
'E:\BUF1M_PARCEL_PRICE\AVPR3_BUF1M_ALL.DA1';
PUT PARCEL_N2 1-14 AVPR3 15-24 2;
FORMAT PARCEL_N2 14.;
run;
*----------------------------------------------------------
SAS PROGRAM FOR COMPUTING ACCESSIBILITY INDICES
OF EMPLOYMENT CENTERS BASED ON CTPP DATA AND DISTANCE DATA
-----------------------------------------------------------;

*READ CTPP/REGIONAL EMPLOYMENT’S DATA
----------------------------------------;
FILENAME IN1 'C:\BG_JOB_COUNTY\DELA_CTPP_B_S41.DA1';
DATA a1; INFILE IN1 DSD DELIMITER = ','; INPUT S4_1 1-5 BGSTFID 6-18;
FORMAT BGSTFID 12.;
FILENAME IN2 'C:\BG_JOB_COUNTY\FAIR_CTPP_B_S41.DA1';
DATA a2; INFILE IN2 DSD DELIMITER = ','; INPUT S4_1 1-5 BGSTFID 6-18;
FORMAT BGSTFID 12.;
FILENAME IN3 'C:\BG_JOB_COUNTY\FRA_CTPP_B_S41.DA1';
DATA a3; INFILE IN3 DSD DELIMITER = ','; INPUT S4_1 1-5 BGSTFID 6-18;
FORMAT BGSTFID 12.;
FILENAME IN4 'C:\BG_JOB_COUNTY\LICK_CTPP_B_S41.DA1';
DATA a4; INFILE IN4 DSD DELIMITER = ','; INPUT S4_1 1-5 BGSTFID 6-18;
FORMAT BGSTFID 12.;
FILENAME IN5 'C:\BG_JOB_COUNTY\MADI_CTPP_B_S41.DA1';
DATA a5; INFILE IN5 DSD DELIMITER = ','; INPUT S4_1 1-5 BGSTFID 6-18;
FORMAT BGSTFID 12.;
FILENAME IN6 'C:\BG_JOB_COUNTY\PICK_CTPP_B_S41.DA1';
DATA a6; INFILE IN6 DSD DELIMITER = ','; INPUT S4_1 1-5 BGSTFID 6-18;
FORMAT BGSTFID 12.;
DATA EMPLOYMENT; SET a1 a2 a3 a4 a5 a6; BY BGSTFID;
PROC MEANS DATA = EMPLOYMENT; VAR S4_1;

*READ DISTANCE DATA
---------------------;
PROC IMPORT DATAPASDATA =
'C:\BG_JOB_DIST_COUNTY\dist_del_del2_stfid_mile_exc.xls' REPLACE out =
z1;
DATA z1;
SET z1;
BGSTFID = TARGETUID;
BGSTFID_DELA = SOURCEUID;
Distance_m = disdel;
proc sort data=z1; BY BGSTFID;
PROC IMPORT DATAPASDATA =
'C:\BG_JOB_DIST_COUNTY\dist_del_fair_stfid_mile_exc.xls' REPLACE out =
z2;
DATA z2;
SET z2;
BGSTFID = TARGETUID;
BGSTFID_DELA= SOURCEUID;
Distance_m = del_fair;
proc sort data=z2; BY BGSTFID;

PROC IMPORT DATAFILE =
'C:\BG_JOB_DIST_COUNTY\dist_del_fra_stfid_mile_exc.xls' REPLACE out = z3;
DATA z3;
SET z3;
BGSTFID = TARGETUID;
BGSTFID_DELA= SOURCEUID;
Distance_m = del_fra;
proc sort data=z3; BY BGSTFID;

PROC IMPORT DATAFILE =
'C:\BG_JOB_DIST_COUNTY\dist_del_lick_stfid_mile_exc.xls' REPLACE out = z4;
DATA z4;
SET z4;
BGSTFID = TARGETUID;
BGSTFID_DELA= SOURCEUID;
Distance_m = del_lick;
proc sort data=z4; BY BGSTFID;

PROC IMPORT DATAFILE =
'C:\BG_JOB_DIST_COUNTY\dist_del_madi_stfid_mile_exc.xls' REPLACE out = z5;
DATA z5;
SET z5;
BGSTFID = TARGETUID;
BGSTFID_DELA= SOURCEUID;
Distance_m = del_madi;
proc sort data=z5; BY BGSTFID;

PROC IMPORT DATAFILE =
'C:\BG_JOB_DIST_COUNTY\dist_del_pick_stfid_mile_exc.xls' REPLACE out = z6;
DATA z6;
SET z6;
BGSTFID = TARGETUID;
BGSTFID_DELA= SOURCEUID;
Distance_m = del_pick;
proc sort data=z6; BY BGSTFID;

*COMPUTING ACCESSIBILITY INDICES
----------------------------------;

DATA Distance; SET z1 z2 z3 z4 z5 z6; BY BGSTFID;
DATA EMPDIS; MERGE EMPLOYMENT DISTANCE; BY BGSTFID;
IF distance_m > 0 THEN DO;
Index1 = S4_1/(distance_m) ** 0.5;
index2 = S4_1/(distance_m);
index3 = S4_1/(distance_m) ** 1.5;
index4 = S4_1/(distance_m) ** 2;
index5 = S4_1/(distance_m) ** 2.5;
END;
proc sort data = EMPDIS; BY BGSTFID_DELA;
PROC MEANS NOPRINT SUM; VAR index1 index2 Index3 Index4 Index5; BY BGSTFID_DELA;
     OUTPUT OUT= all SUM= employ_index1 employ_index2 employ_index3 employ_index4 employ_index5;
proc means data=all; VAR employ_index1 employ_index2 employ_index3 employ_index4 employ_index5;
run;
SAS PROGRAM FOR COMPUTING THE PROPORTION OF ZONING IN EACH BUFFER
BASED ON INTERSECTION OF BUFFERS AND ZONING
OF EACH TOWNSHIP/MUNICIPALITY

* ZONING IN BUFFERS 1 MILE
------------------------;

PROC IMPORT DATAFILE =
' e:\BUFLM_ZON_DELACITY\buflm_APR_zon_delacity_delcizon.dbf' out = a
REPLACE;
DATA a;
SET a;
IF ZONING = "A-1" THEN ag = AREA;
IF ZONING = "R-1" OR ZONING = "R-2" OR ZONING = "R-3" THEN lo = AREA;
IF ZONING = "R-4" OR ZONING = "R-6" OR ZONING = "R-7" THEN hi = AREA;
IF ZONING = "M-1" OR ZONING = "M-2" THEN in = AREA;
IF ZONING = "B-1" OR ZONING = "B-2" OR ZONING = "B-3" OR ZONING = "B-4"
OR ZONING = "B-6" OR ZONING = "PO/I" THEN co = AREA;

PROC SORT DATA= a; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo hi in co; BY FID_BUF_AP;
OUTPUT OUT= top1 SUM= agric lores hires indus commer;

PROC IMPORT DATAFILE = 'e:\BUFLM_ZONING_APR\buflm_apr_zon_ashl.dbf' out = b REPLACE;
DATA b;
SET b;
IF ZONING = "R-1" OR ZONING = "R-2" OR ZONING = "R-3" THEN lo = AREA;
IF ZONING = "GS" OR ZONING = "ROAD" THEN op = AREA;
IF ZONING = "R-4" OR ZONING = "R-5" THEN hi = AREA;
IF ZONING = "M-1" THEN in = AREA;
IF ZONING = "C-1" OR ZONING = "C-2" THEN co = AREA;

PROC SORT DATA= b; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR lo hi in co op; BY FID_BUF_AP;
OUTPUT OUT= top2 SUM= lores hires indus commer open;

PROC IMPORT DATAFILE =
' e:\BUFLM_ZON_DELACITY\buflm_APR_zon_delatwp_delzon.dbf' out = c
REPLACE;
DATA c;
SET c;
IF ZONING = "FR-1" THEN ag=AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "PRD" THEN me = AREA;
IF ZONING = "I" OR ZONING = "PI" THEN in = AREA;
IF ZONING = "PCD" THEN co = AREA;

PROC SORT DATA= c; BY FID_BUF_AP;

PROC MEANS NOPRINT SUM; VAR ag me in co op; BY FID_BUF_AP;
   OUTPUT OUT= top3 SUM= agric meres indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_berk.dbf' out = d REPLACE;
   DATA d;
   SET d;
   IF ZONING = "FR-1" THEN ag=AREA;
   IF ZONING = "A-1" THEN ag=AREA;
   IF ZONING = "River" OR ZONING = "Road" OR ZONING = "PRCD" THEN op = AREA;
   IF ZONING = "PRD" THEN me = AREA;
   IF ZONING = "PID" THEN in = AREA;
   IF ZONING = "PCD" OR ZONING = "PIND" THEN co = AREA;

PROC SORT DATA= d; BY FID_BUF_AP;

PROC MEANS NOPRINT SUM; VAR ag me in co op; BY FID_BUF_AP;
   OUTPUT OUT= top4 SUM= agric meres indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_berl.dbf' out = e REPLACE;
   DATA e;
   SET e;
   IF ZONING = "FR-1" OR ZONING = "A-1" THEN ag=AREA;
   IF ZONING = "River" OR ZONING = "Road" OR ZONING = "US Land" THEN op = AREA;
   IF ZONING = "PRD" THEN me = AREA;
   IF ZONING = "PID" THEN in = AREA;
   IF ZONING = "PCD" OR ZONING = "PIND" THEN co = AREA;

PROC SORT DATA= e; BY FID_BUF_AP;

PROC MEANS NOPRINT SUM; VAR ag lo me in co op; BY FID_BUF_AP;
   OUTPUT OUT= top5 SUM= agric lores meres indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_brow.dbf' out = f REPLACE;
   DATA f;
   SET f;
   IF ZONING = "FR-1" THEN ag=AREA;
   IF ZONING = "C-2" OR ZONING = "PC" THEN co = AREA;
   IF ZONING = "R-2" THEN lo = AREA;
   IF ZONING = "PID" THEN in = AREA;
   IF ZONING = "River" OR ZONING = "ROW" THEN op = AREA;
PROC SORT DATA= f; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo co op in; BY FID_BUF_AP;
    OUTPUT OUT= top6 SUM= agric lores commer open indus;
PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_conc.dbf' out = g REPLACE;
DATA g;
SET g;
IF ZONING = "FR-1" THEN ag=AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "R-2" THEN lo = AREA;
IF ZONING = "PRD" THEN me = AREA;
IF ZONING = "R-6" THEN hi = AREA;
IF ZONING = "M-1" THEN in = AREA;
IF ZONING = "B-1" OR ZONING = "B-2" OR ZONING = "B-3" OR ZONING = "B-4"
    OR ZONING = "PCD" THEN co = AREA;
PROC SORT DATA= g; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo me hi in co op; BY FID_BUF_AP;
    OUTPUT OUT= top7 SUM= agric lores meres hires indus commer open;
PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_gale.dbf' out = h REPLACE;
DATA h;
SET h;
IF ZONING = "FR-1" OR ZONING = "A-1" THEN ag=AREA;
IF ZONING = "R-2" THEN lo=AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF DSC_ZONE = "Planned Residence District" THEN me = AREA;
IF ZONING = "PC" THEN co = AREA;
PROC SORT DATA= h; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo me in co op; BY FID_BUF_AP;
    OUTPUT OUT= top8 SUM= agric lores meres indus commer open;
PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_geno.dbf' out = i REPLACE;
DATA i;
SET i;
IF ZONING = "RR" OR ZONING = "SR" THEN lo = AREA;
IF ZONING = "PD-1" THEN me = AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "PD-3" THEN in = AREA;
IF ZONING = "CB" OR ZONING = "CF" OR ZONING = "PD-2" THEN co = AREA;
PROC SORT DATA= i; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR lo me in co op; BY FID_BUF_AP;
PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_harl.dbf' out = j REPLACE;
DATA j;
SET j;
IF ZONING = "AR-1" THEN ag=AREA;
IF ZONING = "R-1" OR ZONING = "R-2" THEN lo=AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "PRD" THEN me = AREA;
IF ZONING = "PID" THEN in = AREA;
IF ZONING = "C-2" OR ZONING = "PCD" THEN co = AREA;

PROC SORT DATA= j; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo me co op in; BY FID_BUF_AP;
OUTPUT OUT= top9 SUM= lores meres indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_king.dbf' out = k REPLACE;
DATA k;
SET k;
IF ZONING = "FR-1" THEN ag=AREA;
IF ZONING = "Road" OR ZONING = "REC" THEN op = AREA;
IF ZONING = "R-3" THEN lo= AREA;
IF ZONING = "C-2" THEN co = AREA;

PROC SORT DATA= k; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo co; BY FID_BUF_AP;
OUTPUT OUT= top10 SUM= agric lores commer open indus;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_libe.dbf' out = l REPLACE;
DATA l;
SET l;
IF ZONING = "FR-1" THEN ag=AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "PR" OR ZONING = "PERRC" THEN me = AREA;
IF ZONING = "R-3" THEN lo = AREA;
IF ZONING = "PI" OR ZONING = "I" THEN in = AREA;
IF ZONING = "C-2" OR ZONING = "PC" THEN co = AREA;

PROC SORT DATA= l; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo me in co op; BY FID_BUF_AP;
OUTPUT OUT= top11 SUM= agric lores meres commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_oran.dbf' out = m REPLACE;
PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_oran.dbf' out = m REPLACE;
DATA m;
SET m;
IF ZONING = "FR-1" THEN ag = AREA;
IF ZONING = "Alum Creek" OR ZONING = "High Banks" OR ZONING = "River"
OR ZONING = "Road" THEN op = AREA;
IF ZONING = "SF-PRD" THEN me = AREA;
IF ZONING = "MF-PRD" THEN hi = AREA;
IF ZONING = "PI" THEN in = AREA;
IF ZONING = "C-2" OR ZONING = "PC" THEN co = AREA;
PROC SORT DATA= m; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag me hi in co op; BY FID_BUF_AP;
OUTPUT OUT= top12b SUM= agric meres hires indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_oxfo.dbf' out = n REPLACE;
DATA n;
SET n;
IF ZONING = "FR-1" THEN ag = AREA;
IF ZONING = "Road" THEN op = AREA;
IF ZONING = "I" THEN in = AREA;
PROC SORT DATA= n; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag in op; BY FID_BUF_AP;
OUTPUT OUT= top14 SUM= agric indus open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_port.dbf' out = o REPLACE;
DATA o;
SET o;
IF Zoning = "A-1" THEN ag = AREA;
IF zoning = "Road" THEN op = AREA;
IF zoning = "R-1" THEN lo = AREA;
PROC SORT DATA= o; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo op; BY FID_BUF_AP;
OUTPUT OUT= top15 SUM= agric lores open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_powe.dbf' out = p REPLACE;
DATA p;
SET p;
IF ZONID = 2505 THEN me = AREA;
IF ZONID = 2501 OR ZONID = 2502 THEN lo = AREA;
IF ZONID = 1 THEN op = AREA;
IF ZONID = 2503 OR ZONID = 2504 OR ZONID = 2506 OR ZONID = 2507 THEN co
= AREA;
IF ZONID = 2508 THEN in = AREA;
PROC SORT DATA= p; BY FID_BUF_AP;

PROC MEANS NOPRINT SUM; VAR lo me co in op; BY FID_BUF_AP;
OUTPUT OUT= top16 SUM= lores meres commer indus open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_radn.dbf' out = q REPLACE;
DATA q;
SET q;
IF ZONING = "FR-1" THEN ag=AREA;
IF ZONING = "Road" OR ZONING = "River" OR ZONING = "REC" THEN op = AREA;
IF ZONING = "PI" OR ZONING = "I" THEN in = AREA;
IF ZONING = "PINS" OR ZONING = "C-2" OR ZONING = "PC" THEN co = AREA;
IF ZONING = "R-2" THEN lo=AREA;
PROC SORT DATA= q; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo in co op; BY FID_BUF_AP;
OUTPUT OUT= top17 SUM= agric lores indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_scio.dbf' out = r REPLACE;
DATA r;
SET r;
IF ZONING = "FR-1" THEN ag=AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "I" OR ZONING = "PI" OR ZONING = "QD" THEN in = AREA;
IF ZONING = "C-2" OR ZONING = "PC" THEN co = AREA;
PROC SORT DATA= r; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag in co op; BY FID_BUF_AP;
OUTPUT OUT= top18 SUM= agric indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_shaw.dbf' out = s REPLACE;
DATA s;
SET s;
IF ZONID = 1 THEN op=AREA;
IF ZONID = 1009 THEN ag = AREA;
IF ZONID = 1001 THEN lo = AREA;
IF ZONID = 1004 OR ZONID = 1005 OR ZONID = 1006 OR ZONID = 1008 THEN co = AREA;
PROC SORT DATA= s; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo co op; BY FID_BUF_AP;
OUTPUT OUT= top19 SUM= agric lores commer open;
PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_sunb.dbf'
out = t REPLACE;
DATA t;
SET t;
IF ZONING = "A-1" THEN ag = AREA;
IF ZONING = "Road" THEN op = AREA;
IF ZONING = "R-1" OR ZONING = "R-2" OR ZONING = "R-3" THEN lo = AREA;
IF ZONING = "PRD" THEN me = AREA;
IF ZONING = "R-4" OR ZONING = "R-5" THEN hi = AREA;
IF ZONING = "I" OR ZONING = "PID" THEN in = AREA;
IF ZONING = "C-1" OR ZONING = "C-2" OR ZONING = "C-3" OR ZONING = "PCD"
THEN co = AREA;

PROC SORT DATA = t; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo me hi in co op; BY FID_BUF_AP;
OUTPUT OUT = top20 SUM = agric lores meres hires indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_tren.dbf'
out = u REPLACE;
DATA u;
SET u;
IF ZONING = "FR" THEN ag = AREA;
IF ZONING = "RR" THEN lo = AREA;
IF ZONING = "RIVER" OR ZONING = "ROAD" THEN op = AREA;
IF ZONING = "LI" THEN in = AREA;
IF ZONING = "CB" THEN co = AREA;

PROC SORT DATA = u; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag lo in co op; BY FID_BUF_AP;
OUTPUT OUT = top21 SUM = agric lores indus commer open;

PROC IMPORT DATAFILE = 'e:\BUF1M_ZONING_APR\buf1m_apr_zon_troy.dbf'
out = v REPLACE;
DATA v;
SET v;
IF ZONING = "FR-1" THEN ag = AREA;
IF ZONING = "River" OR ZONING = "Road" THEN op = AREA;
IF ZONING = "PCD" THEN co = AREA;
IF ZONING = "PID" THEN in = AREA;

PROC SORT DATA = v; BY FID_BUF_AP;
PROC MEANS NOPRINT SUM; VAR ag co op in; BY FID_BUF_AP;
OUTPUT OUT = top22 SUM = agric commer open indus;

Data mergetop;
MERGE top1 top2 top3 top4 top5 top6 top7 top8 top9 top10 top11 top12
top12b top14 top15 top16 top17 top18 top19 top20 top21 top22; BY
FID_BUF_AP;
proc means noprint SUM; VAR agric lores meres hires indus commer open;
BY FID_BUF_AP;
output out = tot SUM= tagric tlores tmeres thires tindus tcommer topen;
data tot;
set tot;

all = tagric + tlores + tmeres + thires + tindus + tcommer + topen;

FILE OUT DSD DELIMITER =','; FILENAME OUT 'c:\BUF1M_ZON_APR.DA1';
   PUT FID_BUF_AP 1-4 tagric 5-16 2 tlores 17-28 2 tmeres 29-40 2
   thires 41-52 2 tindus 53-64 2 tcommer 65-76 2 topen 77-88 2 all 89-100 2;
run;
SAS PROGRAM FOR THE REGRESSION OF INDIVIDUAL NEIGHBORHOOD (FOR BUFFER 1 MILE)

*READ DATA

DATA WORK; INFILE 'c:\VAR_DEC2005\VAR_ALLBUFF.DA1' DSD DELIMITER = '',';
INPUT PARCEL_N2 1-14 SALE 15-23 STOR 24-26 1 BASE 27-28 YEAR 29-33 FLOR
34-39 ROOM 40-41 FBAT 42-43 HBAT 44-45 GARA 46-48 1 AIRC 49-50 FIRE 51-52
LOTS 53-69 4 TAXS 70-76 4 DIAR 77-89 3 DIAM 90-102 3
DICE 103-115 3 DIEL 116-128 3 DIFL 129-141 3 DIGA 142-154 3 DGAF 155-167 3
dgof 168-180 3 DIHO 181-193 3 DIMH 194-206 3 DIPO 207-219 3 DIRA
220-232 3 DIRI 233-245 3 DISC 246-258 3 DITO 259-271 3
DIDW 272-278 3 DENS 279-282 2 PAGR 283-287 2 PCOM 288-292 2 PIND
293-297 2 POSP 298-302 2 PRES 303-307 2 AVPR 308-317 2
ASEM 318-325 2 ASEM2 326-333 4 ASEM3 334-341 5 ASEM4 342-352 9 ASBA1
353-365 3 ASBA2 366-378 3 ASBA3 379-391 3 ASBA4 392-404 3 ASBR 405-417
3 ASBR2 418-430 3 ASBR3 431-443 3 ASBR4 444-456 3 ASFI1 457-469 3 ASFI2
470-482 3 ASFI3 483-495 3 ASFI4 496-508 3 ASSG 509-521 3 ASSU 522-534
3 ASSU3 535-547 3 ASSU4 548-560 3
ASIN 561-573 3 ASIN2 574-586 3 ASIN3 587-599 3 ASIN4 600-612 3 ASGO1
613-625 3 ASGO2 626-638 3 ASGO3 639-651 3 ASGO4 652-664 3 ASPO 665-677
3 ASPO2 678-690 3 ASPO3 691-703 3 ASPO4 704-716 3
NUHO 717-722 2
OCHO 723-727 2
VAHO 728-732 2
OWHO 733-737 2
REHO 738-742 2
WHHO 743-747 2
BHKO 748-752 2
ASHO 753-757 2
HIHO 758-762 2
POOW 763-767 2
PORE 768-772 2
PWOW 773-777 2
PWRE 778-782 2
PBOW 783-787 2
PBRE 788-792 2
PAOW 793-797 2
PARE 798-802 2
PHOW 803-807 2
PHRE 808-812 2
SIZA 813-817 2
SIZB 818-822 2
SIZC 823-827 2
SIZD 828-832 2
SIZE 833-837 2
OWNA 838-842 2
OWNB 843-847 2
OWNC 848-852 2
RENA 853-857 2
RENB 858-862 2

171
RENC1 863-867 2
POTO1 868-873 2
POBH1 874-878 2
POBL1 879-883 2
POAS1 884-888 2
POHI1 889-893 2
POMA1 894-898 2
POMB1 899-903 2
POMC1 904-908 2
POFB1 914-918 2
POFC1 919-923 2
MAMA1 924-928 2
MAFE1 929-933 2
CIUS1 934-938 2
FOBO1 939-943 2
WOIN1 944-948 2
WOOU1 949-953 2
DCAR1 954-958 2
TIMA1 959-963 2
TIMB1 964-968 2
TIMC1 969-973 2
TIMD1 974-978 2
EMHS1 979-983 2
EMBA1 984-988 2
EMMA1 989-993 2
EFHS1 994-998 2
EFBA1 999-1003 2
EFMA1 1004-1008 2
WODO1 1009-1013 2
WMAG1 1014-1018 2
WMNG1 1019-1023 2
WFAG1 1024-1028 2
WFNG1 1029-1033 2
INCA1 1034-1038 2
INCB1 1039-1043 2
INCC1 1044-1048 2
PLAN 1049-1050 PZAG1 1051-1056 2 PZLO1 1057-1062 2 PZME1 1063-1068 2
PZH11 1069-1074 2 PZCO1 1075-1080 2 PZIN1 1081-1086 2 PZOP1 1087-1092 2
PZOT1 1093-1098 2
DENS2 1099-1104 2 AVPR2 1105-1114 2 PAGR2 1115-1119 2 PCOM2 1120-1124 2
PIND2 1125-1129 2 POSP2 1130-1134 2 PRES2 1135-1139 2
NUHO2 1140-1145 2
OCHO2 1146-1150 2
VAHO2 1151-1155 2
OWHO2 1156-1160 2
REHO2 1161-1165 2
WHHO2 1166-1170 2
BLHO2 1171-1175 2
ASHO2 1176-1180 2
HISHO2 1181-1185 2
POOW2 1186-1190 2
PORE2 1191-1195 2
FWOW2 1196-1200 2
FWRE2 1201-1205 2
<table>
<thead>
<tr>
<th>Code</th>
<th>Start-Stop</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBOW2</td>
<td>1206-1210</td>
<td>4</td>
</tr>
<tr>
<td>PBRE2</td>
<td>1211-1215</td>
<td>4</td>
</tr>
<tr>
<td>PAOW2</td>
<td>1216-1220</td>
<td>4</td>
</tr>
<tr>
<td>PARE2</td>
<td>1221-1225</td>
<td>4</td>
</tr>
<tr>
<td>PHOW2</td>
<td>1226-1230</td>
<td>4</td>
</tr>
<tr>
<td>PHRE2</td>
<td>1231-1235</td>
<td>4</td>
</tr>
<tr>
<td>SIZA2</td>
<td>1236-1240</td>
<td>4</td>
</tr>
<tr>
<td>SIZB2</td>
<td>1241-1245</td>
<td>4</td>
</tr>
<tr>
<td>SIZC2</td>
<td>1246-1250</td>
<td>4</td>
</tr>
<tr>
<td>SIZD2</td>
<td>1251-1255</td>
<td>4</td>
</tr>
<tr>
<td>SIZE2</td>
<td>1256-1260</td>
<td>4</td>
</tr>
<tr>
<td>OWNA2</td>
<td>1261-1265</td>
<td>4</td>
</tr>
<tr>
<td>OWNB2</td>
<td>1266-1270</td>
<td>4</td>
</tr>
<tr>
<td>OWNC2</td>
<td>1271-1275</td>
<td>4</td>
</tr>
<tr>
<td>RENA2</td>
<td>1276-1280</td>
<td>4</td>
</tr>
<tr>
<td>RENB2</td>
<td>1281-1285</td>
<td>4</td>
</tr>
<tr>
<td>RENC2</td>
<td>1286-1290</td>
<td>4</td>
</tr>
<tr>
<td>POTO2</td>
<td>1291-1296</td>
<td>4</td>
</tr>
<tr>
<td>POWH2</td>
<td>1297-1301</td>
<td>4</td>
</tr>
<tr>
<td>POBL2</td>
<td>1302-1306</td>
<td>4</td>
</tr>
<tr>
<td>POAS2</td>
<td>1307-1311</td>
<td>4</td>
</tr>
<tr>
<td>POHI2</td>
<td>1312-1316</td>
<td>4</td>
</tr>
<tr>
<td>POMA2</td>
<td>1317-1321</td>
<td>4</td>
</tr>
<tr>
<td>POMB2</td>
<td>1322-1326</td>
<td>4</td>
</tr>
<tr>
<td>POMC2</td>
<td>1327-1331</td>
<td>4</td>
</tr>
<tr>
<td>POFA2</td>
<td>1332-1336</td>
<td>4</td>
</tr>
<tr>
<td>POFB2</td>
<td>1337-1341</td>
<td>4</td>
</tr>
<tr>
<td>POFC2</td>
<td>1342-1346</td>
<td>4</td>
</tr>
<tr>
<td>MAMA2</td>
<td>1347-1351</td>
<td>4</td>
</tr>
<tr>
<td>MAFE2</td>
<td>1352-1356</td>
<td>4</td>
</tr>
<tr>
<td>CIUS2</td>
<td>1357-1361</td>
<td>4</td>
</tr>
<tr>
<td>FOBO2</td>
<td>1362-1366</td>
<td>4</td>
</tr>
<tr>
<td>WOIN2</td>
<td>1367-1371</td>
<td>4</td>
</tr>
<tr>
<td>WOOU2</td>
<td>1372-1376</td>
<td>4</td>
</tr>
<tr>
<td>DCAR2</td>
<td>1377-1381</td>
<td>4</td>
</tr>
<tr>
<td>TIMA2</td>
<td>1382-1386</td>
<td>4</td>
</tr>
<tr>
<td>TIMB2</td>
<td>1387-1391</td>
<td>4</td>
</tr>
<tr>
<td>TIMC2</td>
<td>1392-1396</td>
<td>4</td>
</tr>
<tr>
<td>TIMD2</td>
<td>1397-1401</td>
<td>4</td>
</tr>
<tr>
<td>EMHS2</td>
<td>1402-1406</td>
<td>4</td>
</tr>
<tr>
<td>EMBA2</td>
<td>1407-1411</td>
<td>4</td>
</tr>
<tr>
<td>EMMA2</td>
<td>1412-1416</td>
<td>4</td>
</tr>
<tr>
<td>EFHS2</td>
<td>1417-1421</td>
<td>4</td>
</tr>
<tr>
<td>EFBA2</td>
<td>1422-1426</td>
<td>4</td>
</tr>
<tr>
<td>EFMA2</td>
<td>1427-1431</td>
<td>4</td>
</tr>
<tr>
<td>WODO2</td>
<td>1432-1436</td>
<td>4</td>
</tr>
<tr>
<td>WMAG2</td>
<td>1437-1441</td>
<td>4</td>
</tr>
<tr>
<td>WMNG2</td>
<td>1442-1446</td>
<td>4</td>
</tr>
<tr>
<td>WFA2</td>
<td>1447-1451</td>
<td>4</td>
</tr>
<tr>
<td>WFGN2</td>
<td>1452-1456</td>
<td>4</td>
</tr>
<tr>
<td>INCA2</td>
<td>1457-1461</td>
<td>4</td>
</tr>
<tr>
<td>INCB2</td>
<td>1462-1466</td>
<td>4</td>
</tr>
<tr>
<td>INCC2</td>
<td>1467-1471</td>
<td>4</td>
</tr>
<tr>
<td>PZAG2</td>
<td>1472-1477</td>
<td>6</td>
</tr>
<tr>
<td>PZLO2</td>
<td>1478-1483</td>
<td>6</td>
</tr>
<tr>
<td>PZME2</td>
<td>1484-1489</td>
<td>6</td>
</tr>
<tr>
<td>PZHI2</td>
<td>1490-1495</td>
<td>6</td>
</tr>
<tr>
<td>PZCO2</td>
<td>1496-1501</td>
<td>6</td>
</tr>
<tr>
<td>PZIN2</td>
<td>1502-1507</td>
<td>6</td>
</tr>
<tr>
<td>PZOP2</td>
<td>1508-1513</td>
<td>6</td>
</tr>
<tr>
<td>PZOT2</td>
<td>1514-1519</td>
<td>6</td>
</tr>
</tbody>
</table>
EMBA3 1828-1832 2
EMMA3 1833-1837 2
EFHS3 1838-1842 2
EFBA3 1843-1847 2
EFMA3 1848-1852 2
WODO3 1853-1857 2
WMAG3 1858-1862 2
WMNG3 1863-1867 2
WFAG3 1868-1872 2
WFGN3 1873-1877 2
INCA3 1878-1882 2
INCB3 1883-1887 2
INCC3 1888-1892 2
PZAG3 1893-1898 2 PZLO3 1899-1904 2 PZME3 1905-1910 2 PZHI3 1911-1916 2
PZCO3 1917-1922 2 PZIN3 1923-1928 2 PZOP3 1929-1934 2 PZOT3 1935-1940 2
AGES 1941-1944 REMO 1945-1948 BATH 1949-1951;
FORMAT PARCEL_N2 14.;

*TRANSFORMING VARIABLES
------------------------;

LSALE = LOG(SALE);
LFLOR = LOG(FLOR);
LROOM = LOG(ROOM);
LTAXS = LOG(TAXS);
LDICE = LOG(DICE);
LDIEL = LOG(DIEL);
LDIGA = LOG(DIGA);
LDIHO = LOG(DIHO);
LDIMH = LOG(DIMH);
LDIFO = LOG(DIFO);
LDIRA = LOG(DIRA);
LDIRO = LOG(DIRO);
LDISC = LOG(DISC);
LDITO = LOG(DITO);
LDIDW = LOG(DIDW);
PAGR = PAGR3 + POSP3;
LPAGR = LOG(PAGR);
LPRES = LOG(PRES3);
PZAG = PZAG3 + PZOP3;
LPZAG = LOG(PZAG);
PZRE = PZLO3 + PZME3;
PZNR = PZCO3 + PZIN3 + PZOT3;
LDENS = LOG(DENS3);
LAVPR = LOG(AVPR3);
LASEM = LOG(ASEM3);
LASBA = LOG(ASBA1);
LASFI = LOG(ASFI1);
LASSU = LOG(ASSU1);
LASIN = LOG(ASIN1);
LASGO = LOG(ASGO3);
LASPO = LOG(ASPO3);
AGE = AGES + 1;
REM = REMO + 1;
LAGE = LOG(AGE);
LREM = LOG(REM);
SIZA = SIZA3 + SIZB3;
SIZB = SIZC3 + SIZD3 + SIZE3;
EDHS = EMHS3 + EFHS3;
EDBA = EMBA3 + EFBA3;
EDMA = EMMA3 + EFMA3;
WONG = WMNG3 + WFNG3;
LWONG = LOG(WONG3);
LWOOU = LOG(WOOU3);
LDCAR = LOG(DCAR3);
LCIUS = LOG(CIUS3);
LEDHS = LOG(EDHS);
LEDGA = LOG(EDBA);
LINCB = LOG(INCB3);
LINCC = LOG(INCC3);
LSIZB = LOG(SIZB);
LPWOW = LOG(PWOW3);

LAM = 0.01;
BXHBAT = (HBAT**LAM - 1)/LAM;
BFXBAT = (FXBAT**LAM - 1)/LAM;
BFXFIRE = (FIRE**LAM - 1)/LAM;
BFXGARA = (GARA**LAM - 1)/LAM;
BXPWOW = (PWOW3**LAM - 1)/LAM;
BXPBOW = (PBOW3**LAM - 1)/LAM;
BXPBOW = (P40W3**LAM - 1)/LAM;
BXPZAG = (PZAG**LAM - 1)/LAM;
BXPZRE = (PZRE**LAM - 1)/LAM;
BXPZNR = (PZNR**LAM - 1)/LAM;
BXPCOM = (PCOM3**LAM - 1)/LAM;
BXPAGR = (PAGR**LAM - 1)/LAM;
BXPRES = (PRES3**LAM - 1)/LAM;

*REGRESSION MODEL
------------------;

proc reg data = work;
MODEL LSALE = PLAN LPZAG BXPZRE BXPZNR LPAGR LPRES BXPCOM
BASE LAGE LREM LFLOR BXHBAT BXFIRE LROOM BXBAT BXGARA LSTOR
LTAXS LDICE LDIEL LDIGA LDIHO LDIHM LDIRA LDIRI LDICO LDITO LDIDW
LDENS LAVPR
LASEM LASFI LASSU LASSIN LASBA LASGO LASPO
LPWOW BXPBOW BXPBOW
LWONG LWOOU LDCAR
LCIUS LEDHS LEDBA LINCB LINCC LSIZB;

LABEL PLAN = 'COMPREHENSIVE PLAN'
LPZAG = 'ZON.AGRI.AND OPEN SP.'
BXPZRE = 'ZON.LOW AND MED.RESID.'
BXPZNR = 'ZON.COM,IND,AND OTHER'
LPAGR = 'AGRICULTURAL LAND USE'
LPRES = 'RESIDENTIAL LAND USE'
BXFCOM = 'COMMERCIAL LAND USE'
BASE = 'BASEMENT'
LAGE = 'AGES OF HOUSE'
LREM = 'AGES OF REMOD.HOUSE'
LFLOR = 'FLOOR AREA'
BXFBAT = 'FULL BATHROOMS'
BXFIRE = 'FIREPLACES'
BXHBAT = 'HALF BATHROOMS'
LROOM = 'TOTAL ROOMS'
BXGARA = 'GARAGE CAPACITY'
LSTOR = 'NUMBER OF STORIES'
LTAXS = 'PROPERTY TAXS'
LDICE = 'DIST.CEMETERY'
LDIEL = 'DIST.ELECTRIC LINES'
LDIGA = 'DIST.GAS LINES'
LDIHO = 'DIST.HOSPITAL'
LDIMH = 'DIST.MOB.HOMES'
LDIPO = 'DIST.REG.SHOP.MALL'
LDIRA = 'DIST.RAILROADS'
LDIRI = 'DIST.RIVER'
LDISC = 'DIST.SCHOOL'
LDITO = 'DIST.TOWN HALL'
LDIDW = 'DIST.DOWN.COL.'
LDENS = 'NET POP.DENSITY'
LAVPR = 'AVE.PRICE NEIGH.HOU'
LASEM = 'ACCESS.EMPLOYMENTS'
LASFI = 'ACCESS.FIRE STATIONS'
LASSU = 'ACCESS.SUPERMARKETS'
LASIN = 'ACCESS.INDUSTRIAL SITES'
LASBA = 'ACCESS.BANKS'
LASGO = 'ACCESS.GOLF COURSES'
LASPO = 'ACCESS.POST OFFICES'
LPWOW = 'POP.WHITE OF OWN.'
BXPBOW = 'POP.BLACK OF OWN.'
BXPAMG = 'POP.ASIAN OF OWN.'
LWONG = 'WORK NON AGRIC.'
LWOOU = 'WORK OUTSIDE CO.'
LDCAR = 'DRIVING CAR'
LCIUS = 'U.S CITIZEN'
LEDHS = 'EDU.HIGH SCHOOL'
LEDBA = 'EDU.BACHELOR DEG.'
LINCB = 'MIDDLE INCOME'
LINCC = 'HIGH INCOME'
LSIZB = 'HOUSEHOLD SIZE';
run;
SAS PROGRAM FOR THE REGRESSION OF MULTI-NEIGHBORHOOD
WITH LOGARITHMIC TRANSFORMATIONS

*TRANSFORMING VARIABLES

LSALE = LOG(SALE);
LFLOR = LOG(FLOR);
LTAXS = LOG(TAXS);
LDIHO = LOG(DIHO);
LDIPO = LOG(DIPO);
LDIRI = LOG(DIRI);
LDISC = LOG(DISC);
LDITO = LOG(DITO);

WONG_3 = WMNG3 + WFNG3;
LWONG_3 = LOG(WONG_3);
LDCAR_3 = LOG(DCAR3);
LPWOW_3 = LOG(PWOW3);

PAGR_1 = PAGR1 + POSP1;
PZAG_1 = PZAG1 + PZOP1;
LPRES_1 = LOG(PRES1);

PZRE_1 = PZLO1 + PZME1;
PZNR_1 = PZCO1 + PZIN1 + PZOT1;
LDENS_1 = LOG(DENS1);
LAVPR_1 = LOG(AVPR1);

LASBA = LOG(ASBA1);
LASFI = LOG(ASFI1);
LASIN = LOG(ASIN1);
LASPO = LOG(ASPO3);
AGE = AGES + 1;
LAGE = LOG(AGE);
LWOOU_1 = LOG(WOOU1);
LINCB_1 = LOG(INCB1);

LAM = 0.01;
BXFIRE = (FIRE**LAM - 1)/LAM;
BXRARA = (GARA**LAM - 1)/LAM;
BXPZAG_1 = (PZAG_1**LAM - 1)/LAM;
BXPZRE_1 = (PZRE_1**LAM - 1)/LAM;
BXPZNR_1 = (PZNR_1**LAM - 1)/LAM;
BXPAGR_1 = (PAGR_1**LAM - 1)/LAM;
BXPBOW_2 = (PBOW2**LAM - 1)/LAM;
BXPZAG_1 = (PZAG1**LAM - 1)/LAM;
BXPZNR_3 = (PZNR1**LAM - 1)/LAM;
BXPCOM_3 = (PCOM3**LAM - 1)/LAM;

*REGRESSION MODEL

proc reg data = work;
MODEL LSALE = PLAN
BASE LAGE LFLOR BXFIRE BXGARA
LDIHO LDIFO LDIRI LDITO
LASFI LASIN LASBA LASPO
LPWOW_3 BXPAOW_3
LWONG_3
BXPZAG_1 BXPZRE_1 BXPZNR_1
LDCAR_3
BXPAGR_1 LPRES_1
LDENS_1 LAVPR_1
LWOOU_1
LINCB_1 BXPCOM_3 LTXS BXPBOW_2;
Run;
*-----------------------------------------------------------
SAS PROGRAM FOR THE REGRESSION OF MULTI-NEIGHBORHOOD
WITH BOX-COX PROCEDURE AND COMPUTING MAXIMUM LIKELIHOOD
BY USING LAMDA AND MU
-----------------------------------------------------------;

* LA AND MU WITH INCREMENTS OF 0.05
--------------------------------------;

LA=0;
START1:
MU=0;
START2:
BOXSALE=(SALE**LA-1)/LA;
BOXAGE = (AGE**MU-1)/MU;
BOXFLOR = (FLOR**MU-1)/MU;
BOXFIRE = (FIRE**MU-1)/MU;
BOXGARA = (GARA**MU-1)/MU;
BOXDIHO = (DIHO**MU-1)/MU;
BOXDIPO = (DIPO**MU-1)/MU;
BOXDIRI = (DIRI**MU-1)/MU;
BOXDITO = (DITO**MU-1)/MU;
BOXASFI = (ASFI**MU-1)/MU;
BOXASIN = (ASIN**MU-1)/MU;
BOXASBA = (ASBA**MU-1)/MU;
BOXASPO = (ASPO**MU-1)/MU;
BOXPWOW_3 = (PWOW**MU-1)/MU;
BOXPAOW_3 = (PAOW**MU-1)/MU;
BOXWONG_3 = (WONG_3**MU-1)/MU;
BOXPZAG_1 = (PZAG_1**MU-1)/MU;
BOXPZRE_1 = (PZRE_1**MU-1)/MU;
BOXPZNR_1 = (PZNR_1**MU-1)/MU;
BOXDCAR_3 = (DCAR3**MU-1)/MU;
BOXPAGR_1 = (PAGR_1**MU-1)/MU;
BOXPRES_1 = (PRES1**MU-1)/MU;
BOXDENS_1 = (DENS1**MU-1)/MU;
BOXAVPR_1 = (AVPR1**MU-1)/MU;
BOXWOOU_1 = (WOOU1**MU-1)/MU;
BOXINCBB_1 = (INCB1**MU-1)/MU;
BOXPZCOM_3 = (PZCOM3**MU-1)/MU;
BOTAXS = (TAXS**MU-1)/MU;
BOXPBOW_2 = (PBOW2**MU-1)/MU;
OUTPUT;

MU=MU+0.05;
IF MU<0.6 THEN GOTO START2;
LA=LA+0.05;
IF LA<0.6 THEN GOTO START1;
PROC SORT; BY LA MU;

PROC REG DATA = WORK OUTEST=EST1 RSQUARE; BY LA MU;
MODEL BOXSALE = PLAN BASE BOXAGE BOXFLOR BOXFIRE BOXGARA
BOXDIHO BOXDIPO BOXDIRI BOXDITO
BOXASFI BOXASIN BOXASBA BOXASPO
PROC PRINT DATA=EST1;
DATA LGLK1; SET EST1;
LK= -1572*LOG(_RMSE_**2)+((LA-1)*(38209.94));
PROC PRINT DATA=LGLK1; VAR LA MU _EDF_ LK _RSQ_; run;

*LA AND MU WITH INCREMENTS OF 0.01  
-----------------------------------;

LA=0.30;
START1:
MU=0;
START2:
BOXSALE=(SALE**LA-1)/LA;
BOXAGE = (AGE**MU-1)/MU;
BOXFLOR = (FLOR**MU-1)/MU;
BOXFIRE = (FIRE**MU-1)/MU;
BOXGARA = (GARA**MU-1)/MU;
BOXDIHO = (DIHO**MU-1)/MU;
BOXDIPO = (DIPO**MU-1)/MU;
BOXDIRI = (DIRI**MU-1)/MU;
BOXDITO = (DITO**MU-1)/MU;
BOXASFI = (ASFI1**MU-1)/MU;
BOXASIN = (ASIN1**MU-1)/MU;
BOXASBA = (ASBA1**MU-1)/MU;
BOXASPO = (ASPO3**MU-1)/MU;
BOXPWOW_3 = (PWOW3**MU-1)/MU;
BOXPAOW_3 = (PAOW3**MU-1)/MU;
BOXWONG_3 = (WONG_3**MU-1)/MU;
BOXPZAG_1 = (PZAG_1**MU-1)/MU;
BOXPZRE_1 = (PZRE_1**MU-1)/MU;
BOXPZNR_1 = (PZNR_1**MU-1)/MU;
BOXDCAR_3 = (DCAR3**MU-1)/MU;
BOXPAGR_1 = (PAGR_1**MU-1)/MU;
BOXPRES_1 = (PRES1**MU-1)/MU;
BOXDENS_1 = (DENS1**MU-1)/MU;
BOXAVPR_1 = (AVPR1**MU-1)/MU;
BOXWOOU_1 = (WOOU1**MU-1)/MU;
BOXINCB_1 = (INCB1**MU-1)/MU;
BOXPCOM_3 = (PCOM3**MU-1)/MU;
BOXTAXS = (TAXS**MU-1)/MU;
BOXPBOW_2 = (PBOW2**MU-1)/MU;
OUTPUT;

MU=MU+0.01;
IF MU<0.06 THEN GOTO START2;
LA=LA+0.01;
IF LA<0.32 THEN GOTO START1;
PROC SORT; BY LA MU;
PROC REG DATA = WORK OUTEST=EST1 RSQUARE; BY LA MU;
MODEL BOXSALE = PLAN BASE BOXAGE BOXFLOR BOXFIRE BOXGARA
BOXDIHO BOXDIPO BOXDIRI BOXDITO
BOXASFI BOXASIN BOXASBA BOXASPO
BOXPWOW_3 BOXPAOW_3
BOXWONG_3
BOXPZAG_1 BOXPZRE_1 BOXPZNR_1
BOXDCAR_3
BOXPAGR_1 BOXPRES_1
BOXDENS_1 BOXAVPR_1
BOXWOOU_1
BOXINC_1 BOXPCOM_3 BOXTAXS BOXPBOW_2;
PROC PRINT DATA=EST1;
DATA LGLK1; SET EST1;
LK= -1572*LOG((_RMSE_**2))+((LA-1)*(38209.94));
PROC PRINT DATA=LGLK1; VAR LA MU _EDF_ LK _RSQ_;
run;
*--------------------------------------------------
SAS PROGRAM FOR COMPUTING ELASTICITY OF VARIABLES
WITH BOX-COX TRANSFORMATIONS
---------------------------------------------------;

WONG_3 = WMNG3 + WFNG3;
PAGR_1 = PAGR1 + POSP1;
PZAG_1 = PZAG1 + PZOP1;
PZRE_1 = PZLO1 + PZME1;
PZNR_1 = PZCO1 + PZIN1 + PZOT1;
AGE = AGES + 1;
REM = REMO + 1;

LA=0.30;
MU=0.02;
BOXAGE = (AGE**MU-1)/MU;
BOXFLOR = (FLOR**MU-1)/MU;
BOXFIRE = (FIRE**MU-1)/MU;
BOXGARA = (GARA**MU-1)/MU;
BOXDIHO = (DIHO**MU-1)/MU;
BOXDIP = (DIPO**MU-1)/MU;
BOXDIRI = (DIRI**MU-1)/MU;
BOXDITO = (DITO**MU-1)/MU;
BOXASFI = (ASFI1**MU-1)/MU;
BOXASIN = (ASIN1**MU-1)/MU;
BOXASBA = (ASBA1**MU-1)/MU;
BOXASPO = (ASPO3**MU-1)/MU;

BOXPWOW_3 = (PWOW3**MU-1)/MU;
BOXPAOW_3 = (PAOW3**MU-1)/MU;
BOXWONG_3 = (WONG_3**MU-1)/MU;
BOXPZAG_1 = (PZAG_1**MU-1)/MU;
BOXPZRE_1 = (PZRE_1**MU-1)/MU;
BOXPZNR_1 = (PZNR_1**MU-1)/MU;
BOXDCAR_3 = (DCAR3**MU-1)/MU;
BOXPAGR_1 = (PAGR_1**MU-1)/MU;
BOXPRES_1 = (PRES1**MU-1)/MU;
BOXDENS_1 = (DENS1**MU-1)/MU;
BOXAVPR_1 = (AVPR1**MU-1)/MU;
BOXWOU_1 = (WOU1**MU-1)/MU;
BOXINCBI = (INCBI**MU-1)/MU;
BOXCOM_3 = (COM3**MU-1)/MU;
BOXTAXS = (TAXS**MU-1)/MU;
BOXPB_2 = (PBOW2**MU-1)/MU;

BOXSALELA = -1038.135 + (4.786*BOXAGE)+ (11.911*
BOXFLOR)+(2.919*BASE)+ (0.048*BOXFIRE)+ (0.057*BOXGARA)+
(6.877*BOXDIHO) - (4.025* BOXDIPO) + (1.481* BOXDIRI) - (1.104*BOXDITO)
+ (11.464*BOXASFI) + (22.715*BOXASIN) - (21.015*BOXASBA) -
(1.205*BOXASPO) + (2.350*BOXPWOW_3) + (0.071*BOXPAOW_3) +
(63.254*BOXWONG_3) +
(2.443*PLAN)-(0.065*BOXPZAG_1) + (0.037*BOXPZRE_1) + (0.033*BOXPZNR_1)
+ (30.359*BOXDCAR_3) -
(0.079*BOXPAGR_1) + (3.340*BOXPRES_1) +
(5.229*BOXDENS_1) + (20.650*BOXAVPR_1) +

183
(8.055*BOXWOOU_1) +
(2.000*BOXINCB_1) + (0.159*BOXPSCOM_3) - (3.512*BOXTAXS) +
(0.107*BOXPBOW_2);
BOXSALELA2 = (BOXSALELA*LA) + 1;
ELAS_AGE= 4.786*(AGE**MU)/(BOXSALELA2);
ELAS_Flor = 11.910*(FLOR**MU)/(BOXSALELA2);
ELAS_FIRE = 0.047*(FIRE**MU)/(BOXSALELA2);
ELAS_Gara = 0.057*(GARA**MU)/(BOXSALELA2);
ELAS_DHO = 6.877*(DIHO**MU)/(BOXSALELA2);
ELAS_DIPO = -4.025*(DIPO**MU)/(BOXSALELA2);
ELAS_DIRI = 1.481*(DIRI**MU)/(BOXSALELA2);
ELAS_DITO = -1.104*(DITO**MU)/(BOXSALELA2);
ELAS_ASFI = 11.464*(ASFI1**MU)/(BOXSALELA2);
ELAS_ASIN = 22.715*(ASIN1**MU)/(BOXSALELA2);
ELAS_ASBA = -21.015*(ASBA1**MU)/(BOXSALELA2);
ELAS_APO = -1.205*(ASPO3**MU)/(BOXSALELA2);
ELAS_PWOW_3 = 2.350*(PWOW3**MU)/(BOXSALELA2);
ELAS_PAOW_3 = 0.071*(PAOW3**MU)/(BOXSALELA2);
ELAS_WONG_3 = 63.254*(WONG_3**MU)/(BOXSALELA2);
ELAS_PZAG_1 = -0.065*(PZAG_1**MU)/(BOXSALELA2);
ELAS_P2RE_1 = 0.037*(P2RE_1**MU)/(BOXSALELA2);
ELAS_P2NR_1 = 0.033*(P2NR_1**MU)/(BOXSALELA2);
ELAS_DCAR_3 = 30.359*(DCAR3**MU)/(BOXSALELA2);
ELAS_PAGR_1 = -0.079*(PAGR_1**MU)/(BOXSALELA2);
ELAS_PMAT_1 = 3.340*(PMAT1**MU)/(BOXSALELA2);
ELAS_DENS_1 = 5.229*(DENS1**MU)/(BOXSALELA2);
ELAS_AVPR_1 = 20.650*(AVPR1**MU)/(BOXSALELA2);
ELAS_WOOU_1 = 8.055*(WOOU1**MU)/(BOXSALELA2);
ELAS_INCB_1 = 2.000*(INCB1**MU)/(BOXSALELA2);
ELAS_PCOM_3 = 0.159*(PCOM3**MU)/(BOXSALELA2);
ELAS_TAXS = -3.512*(TAXS**MU)/(BOXSALELA2);
ELAS_PBOW_2 = 0.107*(PBOW2**MU)/(BOXSALELA2);
KEEP ELAS_AGE
ELAS_Flor
ELAS_FIRE
ELAS_Gara
ELAS_DHO
ELAS_DIPO
ELAS_DIRI
ELAS_DITO
ELAS_ASFI
ELAS_ASIN
ELAS_ASBA
ELAS_APO
ELAS_PWOW_3
ELAS_PAOW_3
ELAS_WONG_3
ELAS_PZAG_1
ELAS_P2RE_1
ELAS_P2NR_1
ELAS_DCAR_3
ELAS_PAGR_1
ELAS_PMAT_1
ELAS_DENS_1
ELAS_AVPR_1
ELAS_WOOU_1
ELAS_INCB_1
ELAS_PCOM_3
ELAS_TAXS
ELAS_PBOW_2;
proc means data=work N MEAN MIN MAX STDDEV;var ELAS_AGE
ELAS_FLOR
ELAS_FIRE
ELAS_GARA
ELAS_DIHO
ELAS_DIPO
ELAS_DIRI
ELAS_DITO
ELAS_ASFI
ELAS_ASIN
ELAS_ASBA
ELAS_ASPO
ELAS_PWOW_3
ELAS_PAOW_3
ELAS_WONG_3
ELAS_PZAG_1
ELAS_PZRE_1
ELAS_PZNR_1
ELAS_DCAR_3
ELAS_PAGR_1
ELAS_PRES_1
ELAS_DENS_1
ELAS_AVPR_1
ELAS_WOOU_1
ELAS_WOOU_1
ELAS_INCB_1
ELAS_PCOM_3
ELAS_TAXS
ELAS_PBOW_2;
PROC PRINT DATA= work;
run;
APPENDIX D:

ZONING OF TOWNSHIPS/MUNICIPALITIES
Figure D.8 Zoning in Concord Township
Figure D.8 Zoning of Galena Municipality
Figure D.9 Zoning of Genoa Township
Figure D.13  Zoning of Marlboro Township.
Figure D.14: Zoning of Orange Township
Figure D.16 Zoning of Porter Township
Figure D.17 Zoning of Powell Municipality
Figure D.18: Zoning of Radnor Township
Figure D.20  Zoning of Shawnee Hills Municipality
Figure D.24  Zoning of Troy Township