The Review of Regional Studies, Vol. 35, No. 3, 2005, pp. 311– 335 *ISSN 1553-0892*

The Review of Regional Studies
Oklahoma State University

A Spatial Analysis of Housing Growth

Seong-Hoon Cho

Department of Agricultural Economics, University of Tennessee, Knoxville, TN 37996-4518, e-mail: scho9@utk.edu

Burton C. English

Department of Agricultural Economics, University of Tennessee Knoxville, TN 37996-4518, e-mail: benglish@utk.edu

Roland K. Roberts

Department of Agricultural Economics, University of Tennessee Knoxville, TN 37996-4518, e-mail: rrobert3@utk.edu

Abstract

This study accounts for spatial spillover and spatial heterogeneity to estimate housing growth between 1990 and 2000 using Tennessee census-block group data. A deterministic estimation method, inverse distance weighted averaging is used to create neighborhood variables that can capture more accurate measurement of spatial spillover while reducing multicollinearity. The block-group specific local effects on housing growth are computed using locally weighted regression, and they are mapped using a geographical information system.

Keywords: Spatial spillover; Spatial heterogeneity; Locally weighted regression; Multicollinearity; Housing growth

JEL classification: C51; R11; R12

The authors thank the editor and two anonymous referees for helpful comments and suggestions.

1. INTRODUCTION

Understanding housing growth is important to planners and policy makers because housing growth can impact natural resources, forest fragmentation, stormwater runoff, degradation or loss of wildlife habitat, loss of open spaces necessary for human health and recreation, and others. The success of resource management in a prosperous, growing economy depends largely on the ability to reconcile residential growth patterns with the natural landscape. Housing growth resulting from residential development is the dominant force driving economic growth in Tennessee, and any systematic study of economic growth should evaluate factors affecting changes in the housing stock.

Housing increased 20 percent in Tennessee from 2.0 million units in 1990 to 2.4 million in 2000, placing Tennessee 12th among the 50 states in housing growth. This growth was greater than for the nation, which saw an increase of 13 percent from 102 million to 115 million housing units during the same period. Tennessee's growth in housing was closely related to the 16 percent growth in population (U.S. Census Bureau 2005).

The spatial pattern of residential development in the 1990s tended to be low-density development commonly referred to as "sprawl." Tennessee's metropolitan areas of Nashville and Knoxville have consistently been ranked as two of the nation's most sprawling metropolitan areas. Tennessee's increase in developed area between 1992 and 1997 ranked 7th among the 50 states.[1]

Tennessee housing growth has given rise to concerns about declining environmental quality. The state's metropolitan areas have been declared ozone non-attainment areas, and the impairment of its waterways has been directly linked to residential and urban development. Rapid growth has put pressure on public services such as sewage treatment, health care, and road maintenance. Despite these rising concerns, a systematic study of housing growth as it relates to socioeconomic and location factors has not been done.

Most studies of housing growth used structural models of housing demand and supply (Follain 1979; Poterba 1984; Topel and Rosen 1988; Crone and Mills 1991; DiPasquale and Wheaton 1994; Blackley 1999). Several studies evaluated fluctuations in the U.S. housing stock using residential building permits (Jaffe and Rosen 1979; Thom 1985; Puri and Van Lierop 1988; McGinnis 1994; Hancock and Wilcox 1997; Rahman and Muhammad 1997; Chan 1999). County-level models of regional housing construction were developed by Conway and Howard (1980), Clark and McGibany (1990), McGibany (1991), Skidmore and Peddle (1998), and McDonald and McMillen (2000). These county-level models typically linked additions to the housing stock to population, household income, and other economic variables; but they did not examine whether changes in the housing stock were related to changes in demographic characteristics and spatial attributes.

[1] Urban and built-up areas are defined by the 1997 National Resources Inventory.

Many factors affecting housing growth are spatially explicit. Location characteristics are frequently paramount in determining real estate value. A significant advantage of a spatially explicit model is the ability to incorporate substantial spatial detail, allowing evaluation of the impacts of selected location factors on housing growth. Location factors influence housing growth in two ways: 1) spatial variation in location factors that directly relate to housing growth, and 2) externalities associated with the location of housing growth. These externalities are called "adjacency effects" because they capture spatial spillover on a given area from neighboring areas. Both relationships are important in predicting housing growth.

Understanding the impacts of spatially varying relationships and spatial spillover within the context of efficient and consistent linear regression has been the subject of numerous studies (e.g., Cliff and Ord 1973; Anselin 1988; McMillen 1992; LeSage 1997; Leung, Mei, and Zhang 2000; Tse 2002; Biles 2003; McMillen 2003a). Dubin (1992, 1998) and Can (1990, 1992) provided estimates of regression parameters in the presence of spatial spillover. Switzer, Kowalik, and Lyon (1982) estimated neighborhood effects of spatial spillover using a prior probability method. While these models captured spatial spillover, spatial heterogeneity was not accommodated.

Spatial heterogeneity exists because every location has an intrinsic degree of uniqueness in the context of the rest of the spatial system. Various localized modeling techniques have been proposed and used to capture spatial heterogeneity (Casetti 1972; Cleveland and Devlin 1988; Getis and Ord 1992; Fotheringham and Brunsdon 1999). The existence of spatial heterogeneity suggests that the estimated parameters of a spatial model are inadequate descriptors of the process being modeled at any given location due to parameter drift across space (Anselin, Dodson, and Hudak 1993; Fotheringham and Rogerson 1993; Fotheringham, Charlton, and Brunsdon 1996, 1997).

One approach that has gained popularity in dealing with spatial heterogeneity is locally weighted regression. Locally weighted regression using geographic coordinates as weights was first proposed by Cleveland and Devlin (1988). Their regression model was estimated through a multivariate smoothing procedure, "fitting a function." Meese and Wallace (1991) and McMillen (1996) also adopted this technique, and others have applied it to identify spatial variations in relationships with spatial heterogeneity (e.g., McMillen and McDonald 1997; Fu and Somerville 2001; Fotheringham, Brunsdon, and Charlton 2002; and McMillen 2003b). While such models of locally weighted regression have focused on spatial heterogeneity, they have not incorporated the impacts of spatial spillover.

The objective of this research was to develop a housing-growth model to capture both spatial spillover and spatial heterogeneity. The model was then applied to Tennessee housing growth as an illustrative example.

2. EMPIRICAL MODEL

We specified and estimated a reduced-form equation for housing growth (growth in housing stock) driven by equilibrium between the supply and demand for new housing structures. This reduced-form equation was derived from a two-equation recursive system in which current-period growth in housing stock was assumed to be a function of exogenous and predetermined variables, namely location and neighborhood variables and current-period growth in housing value. Another structural equation assumed that current-period growth in housing value was a function of exogenous and predetermined variables, namely the same location and neighborhood variables and growth in housing value during the previous period. Thus, housing growth in reduced form was a function of location and neighborhood variables and growth in housing value during the previous period (i.e., Glaeser and Gyourko 2001). We arrived at the final reduced-form equation by specifying three equations with increasing complexity. The reduced-form equation with location and neighborhood variables was first specified using simple averages of the variables from the adjacent neighbors, and the second was specified using inverse distance weighted averages of the variables from all neighbors. The third equation was specified to accommodate spatial heterogeneity and neighborhood effects through inverse distance weighted averaging similar to the second equation.

Applying the prior probability model proposed by Switzer, Kowalik, and Lyon (1982), housing growth within a census-block group accommodating the spatial spillover effects from adjacent neighboring locations was specified as:

$$(1) \qquad Q = Xb + W_1 Xb' + \varepsilon_1$$

where Q was a vector of the differences between numbers of housing units in 2000 and 1990 divided by the number of housing units in 1990 for each location, X was a matrix of location variables that explain housing growth in each location, ε_1 was an error term, and b was a vector of parameters to be estimated. W_1 was an $n \times n$ "contiguity matrix" with diagonal elements of zero and off-diagonal elements of $1/n_i$ for all census-block groups that are contiguous to census-block group i; and n_i is the number of census-block groups that are contiguous to census-block group i. Thus, the neighborhood variables (W_1X) are simple averages of the variables within X for the census-block groups contiguous to census-block group i. The coefficients b' capture the spatial spillover effects of spatially lagged variables (location variables for adjacent neighbors).

Variables for each location included in X were growth in housing value during the previous period, population growth, and other characteristics of the location. Previous research suggested that population growth is a good indicator of housing growth within an area (i.e., McDonald and McMillen 2000). Three spatial characteristics indicating the desirability of housing within an area were chosen: the ratio of area under water to total area, distance to the nearest interstate highway, and distance to the center of the nearest metropolitan statistical area (MSA). The location variables used to portray characteristics of an area included growth in white population between 1990 and 2000, growth in married residents between 1990 and 2000, growth in college graduates between 1990 and 2000, the vacancy ratio in 1990, and the median age of the housing stock in 1990.

Our model used census-block groups as geographic locations, which was done previously by others (Goodman 1977; Cao and Cory 1981; Geoghegan, Waigner, and Bockstael 1997). In their studies, the linear aggregations of housing data at the census-block group level yielded robust empirical estimations. Two conditions must be met for linear aggregation of the number of houses across geographic boundaries (Meen and Andrew 1998). Those conditions are: (1) income and other variables must grow at the same rate in each geographic area or exhibit a common stochastic trend, and (2) the structure of the housing markets must be the same over the geographic areas within the aggregate boundaries. The unit of observation in our model was the census-block group within the boundaries of the 2000 Census.

Inclusion of X and W_1X introduced the possibility of multicollinearity, which can seriously degrade the standard errors of the estimates and render hypothesis testing inconclusive. If the correlation coefficient between two regressors is greater than 0.8 or 0.9, multicollinearity may be a serious problem (Judge et al. 1982, p. 620). Multicollinearity can also be detected by variance inflation factors (Maddala 1992). Variance inflation factors (vif) are a scaled version of the multiple correlation coefficients between variable k and the rest of the independent variables. Specifically, $\text{vif}_k = 1/\left(1 - R_k^2\right)$, where R_k is the multiple correlation coefficient. Multicollinearity occurs when two (or more) variables are related. Rather than removing one or more variables that are highly correlated, an alternative would be to create neighborhood variables that are less collinear with X.

One method of reducing multicollinearity is inverse distance weighted averaging (IDWA), a deterministic estimation method where values at unsampled points are determined by a linear combination of values at known sampled points. A neighborhood about the interpolated point is identified, and a weighted average is taken of the observations within the neighborhood. The weights are a decreasing function of distance (Shepard 1968). They change according to the linear distance of the samples from the unsampled point. Not only can spatially lagged neighborhood variables created by IDWA reduce multicollinearity, they more accurately measure spatial spillover effects because they account for distance to neighboring areas.

A modification of the housing growth model in equation (1) that accounts for spatial spillover using IDWA is:

$$(2) \qquad Q = Xb + W_2Xb'' + \varepsilon_2$$

where $W_2 X = \left[\sum_{j=1}^{n} \left(x_{jk} \omega_{ij} \right) \Big/ \sum_{j=1}^{n} \omega_{ij} \right]_{ik}$, x_{ji} is the k^{th} variable for the j^{th} census-block group, ω_{ij} is derived from the inverse of the distance between census-block group i and the j^{th} census-block group ($i \neq j$) (Keckler 1995; Song and DePinto 1995), coefficients b'' capture spatial spillover effects, and ε_2 is an error term. In this model, neighborhood variables include the effects of all census-block groups within the study area, not just those that are contiguous to census-block group i. Parameter estimates from equation (2) capture the effects on housing growth in census-block group i from its own explanatory variables and from the inverse distance weighted explanatory variables of all other census-block groups within the study area.

An implicit assumption made in equations (1) and (2) is that relationships between variables measured at different locations are constant over space. If structural variations in housing growth exist, constant variables measured at different locations would result in model misspecification. For instance, the impact that the variable measuring distance to the center of the nearest MSA has on housing growth in different locations may vary across a study region. If such variations in relationships exist over space (spatial heterogeneity), housing growth models based on equations (1) and (2) are clearly misspecified because they assume these relationships are constant.

The locally weighted regression model extends the traditional regression framework by allowing parameters to be estimated locally to accommodate spatial heterogeneity. Thus, equation (2) can be rewritten as equation (3) for locally weighted regression:

$$(3) \qquad Q = \left(\beta \otimes X \right) 1 + \left(\beta'' \otimes W_2 X \right) 1 + \varepsilon_3$$

where $\otimes$ is a logical multiplication operator in which each element of β is multiplied by the corresponding element of X, each element of β'' is multiplied by the corresponding element of $W_2 X$, and ε_3 is error term. If equation (3) has n data points and k explanatory variables, β, X, β'', and $W_2 X$ have dimensions $n \times (k + 1)$; and 1 is a conformable vector of 1's. The matrix β now consists of n sets of local parameters. It has the following structure.

$$(4) \qquad \beta = \begin{bmatrix} \beta_0(u_1, v_1) & \beta_1(u_1, v_1) & \cdots & \beta_k(u_1, v_1) \\ \beta_0(u_2, v_2) & \beta_1(u_2, v_2) & \cdots & \beta_k(u_2, v_2) \\ \cdots & \cdots & \cdots & \\ \beta_0(u_n, v_n) & \beta_1(u_n, v_n) & \cdots & \beta_k(u_n, v_n) \end{bmatrix}$$

where (u_i, v_i) denotes the spatial coordinates of census-block group i. The matrix β'' also contains n sets of local parameters and has the same structure as β.

To calibrate the model, a modified weighted least squares approach is taken so that the data are weighted according to their proximity to census-block group i. Thus, the weighting of any point is not constant but varies with census-block group i. In this local housing growth model, observations closer to census-block group i are weighed more than those farther away. This weighting gives census-block groups nearer to i more influence in the estimation of $\beta(u_i, v_i)$ and $\beta''(u_i, v_i)$ than census-block groups farther away. That is:

$$(5) \qquad \hat{\beta}(u_i, v_i) = \left(Z^T W(u_i, v_i) Z\right)^{-1} Z^T W(u_i, v_i) Q,$$

where $\hat{\beta}$ includes a combined matrix of β and β'' with dimension $n \times 2(k + 1)$; Z is a matrix of explanatory variables representing a combination of X and $W_2 X$ with the same $n \times 2(k + 1)$ dimension; and $W(u_i, v_i)$ is an $n \times n$ diagonal matrix with diagonal elements (w_{ij}) denoting the weight given to data point j in the calibration of the model for census-block group i. The diagonal elements of the weight matrix are:

$$(6) \qquad w_{ij} = \left[1 - \left(d_{ij} / B\right)^2\right]^2 \text{ if } d_{ij} < B$$

$$= 0 \text{ otherwise}$$

where d_{ij} is the Euclidean distance between points i and j and B is a selected bandwidth. At regression point i, the weight of the data point is unity and falls to zero when the distance between i and j is greater than or equal to B.

As B approaches infinity, w_{ij} approaches 1 regardless of d_{ij},, in which case the parameter estimates become uniform; thus, local weighted regression is equivalent to OLS. Conversely, as B becomes smaller, the parameter estimates increasingly depend on neighboring observations in close proximity to location i; hence, the parameter estimates have increasing variances. Cleveland (1979) suggested using a cross-validation (CV) approach for selection of the optimal bandwidth. This approach takes the form:[2]

$$(7) \qquad CV = \sum_{i=1}^{n} \left[q_i - \hat{q}_{\neq i}(B)\right]^2,$$

where $\hat{q}_{\neq i}(B)$ is the fitted value for the housing growth variable and q_i is observations for point i omitted from the fitting process. The bandwidth is chosen to minimize CV. Thus, in the local weighted regression model, only census-block groups up to the optimal level of B are assigned non-zero weights for the nearest neighbors of census-block group i. The weights for these nearest neighbors decrease with their distance from census-block group i.

[2] This process is almost identical as choosing B using a "least squares" criterion except for the fact that the observation for point i is omitted.

Equations (1) – (3) were estimated using GWR 3.0 (Charlton, Fotheringham, and Brunsdon 2002). Equation (1) with equally weighted neighborhood variables was estimated for comparison with equation (2) with neighborhood variables created with IDWA. These equations were estimated with OLS (e.g., Haining 1990, pp. 344-350).[3] Equation (3) was estimated with locally weighted regression using the CV approach to determine optimal bandwidth. Sensitivity analysis was conducted for bandwidths of plus and minus 50 percent of the *B* selected by the CV approach.

3. DATA

Normalized 1990 Census long form data in 2000 boundaries and 2000 long form data at the census-block group level in spatial form were used to estimate equations (1) – (3). The 1990 data were normalized to 2000 boundaries because the geographic definitions of boundaries changed between 1990 and 2000. This normalization enabled comparisons between 1990 and 2000 long form data in standard 2000 geographies. The normalized data were created by a private data provider, GeoLytics. Their normalization procedures are presented in Appendix 1. Normalized 1980 data were not available to create housing value growth in the previous period. Thus, the 1980 census-block group level was overlaid on the 2000 boundaries using the geoprocessing option of ArcGIS 9.1 (ESRI, 2005). Housing values for the 462 larger census-block groups in 1980 were assigned to the most closely matching geographies of the 4,014 census-block groups in 2000.

The records of 4,014 census-block groups in the form of polygons within the boundary of the state of Tennessee were used for this study.[4] Each polygon represented one census-block group. GIS was utilized to generate the distance variables. The distance variables were created using "Near," the ArcToolbox GIS tool for geoprocessing. The closest distances to six MSAs from each census-block group were computed using the distance from the centroid of a census-block group to the nearest points of the central business districts for the six MSAs. The centroid of an area is similar to the center of mass of a body. The calculation of a centroid involves only the geometric shape of the area. The following formula was used to calculate the coordinates of the centroid $(\bar{x}, \bar{y})$ for each census-block group:

---

[3] Application of the locally weighted regression to equations (1) could be easily done. However, estimation of the model would add little new information regarding locally weighted regression and housing growth.

[4] Although most people intuitively think of a census-block group as being rectangular or square, of about the same size, and occurring at regular intervals, as in many large cities of the United States, census-block group configurations actually are quite different. The pattern, size, and shape of census-block groups vary within and between areas. Factors that influence the overall configuration of census-block groups include topography; the size and spacing of water features; the land survey system; and the extent, age, type, and density of urban and rural development. Census blocks in remote areas may be large and irregular and may contain many square miles (U.S. Census Bureau 2005).

$$(8) \qquad \bar{x} = \frac{\sum_{I=1}^{n} x_I a_I}{\sum_{I=1}^{n} a_I}, \bar{y} = \frac{\sum_{I=1}^{n} y_I a_I}{\sum_{I=1}^{n} a_I}$$

where x_I and y_I are, respectively, the distances from the *x*-axis and *y*-axis to differential area center for individual segment *I* a series of *n* regular shapes (such as triangles or squares where the centroid is easily calculated), a_I is differential area for individual segment *I*. The closest distance to the interstate highway from each census-block group was calculated using the same tool.

The definitions and descriptive statistics of the variables used for the empirical estimates are shown in Tables 1 and 2, respectively. During the 1990s, the average number of housing units for the census-block groups increased by 22 percent. Over the same time period, the average population per acre for the census-block groups increased by 18 percent. The higher rate of housing growth relative to population growth in Tennessee may have been due to a decline in average household size from 2.56 to 2.48 during the 1990s (U.S. Census Bureau 2005). The difference may also reflect second home development on the Cumberland Plateau and near Smoky Mountain National Park, the most visited national park in the country. These second homes are owned by residents both inside and beyond Tennessee borders. During the 1980s, the median housing value increased by 60 percent.

TABLE 1

Definitions of Variables

Variable	Definition
Housing growth	Difference between numbers of housing units in 2000 and 1990 divided by the number of housing units in 1990
Lagged housing value growth	Difference between median housing values in 1990 and 1980 divided by the median housing value in 1980
Population growth	Difference between Tennessee populations in 2000 and 1990 divided by population in 1990
Age of housing stock	Median age of housing stock in 1990
Vacancy rate	Vacant housing units divided by total housing units in 1990
Water ratio	Area of water divided by total area
White ratio change	Difference between numbers of white residents in 1990 and 2000 divided by the number of white residents in 1990
Marriage ratio change	Difference between numbers of married residents in 1990 and 2000 divided by the number of married residents in 1990
College ratio change	Difference between numbers of college graduates in 1990 and 2000 divided by the number of college graduate in 1990
Distance to MSA	Distance to nearest metropolitan statistical area in miles
Distance to interstate highway	Distance to nearest interstate highway in miles

TABLE 2

Sample Statistics of Variables

Variable	Mean	Std Dev	Min	Max
Dependent variable				
Housing growth	0.222	0.750	-0.969	23.936
Location variable				
Lagged housing value growth	0.601	0.190	-0.118	1.900
Population growth	0.184	0.808	-0.955	29.771
Age of housing stock (100 year)	0.383	0.104	0.000	0.720
Vacancy rate	0.083	0.063	0.000	0.762
Water ratio	0.016	0.054	0.000	0.798
White ratio change	-0.050	0.114	-0.735	0.644
Marriage ratio change	-0.036	0.073	-1.000	0.546
College ratio change	0.020	0.051	-0.241	1.000
Distance to MSA (100 miles)	0.231	0.197	0.000	0.856
Distance to interstate highway (100 miles)	0.061	0.079	0.000	0.052
Neighborhood variable using spatially lagged value				
Lagged housing value growth†	0.629	0.244	0.185	1.127
Population growth†	0.190	0.898	-0.043	6.04
Age of housing stock (100 year) †	0.390	0.119	0.144	0.690
Vacancy rate†	0.079	0.035	0.020	0.487
Water ratio†	0.017	0.032	0.000	0.315
White ratio change†	-0.050	0.078	-0.577	0.105
Marriage ratio change†	-0.036	0.031	-0.162	0.077
College ratio change†	0.020	0.020	-0.040	0.101
Distance to MSA (100 miles)†	0.230	0.195	0.000	0.841
Distance to interstate highway (100 miles)†	0.061	0.076	0.000	0.517
Neighborhood variable using inverse distance weighted average				
Lagged housing value growth‡	0.623	0.230	0.174	1.121
Population growth‡	0.186	0.311	-0.282	3.605
Age of housing stock (100 year)‡	0.389	0.116	0.149	0.701
Vacancy rate‡	0.083	0.036	0.023	0.491
Water ratio‡	0.017	0.032	0.000	0.297
White ratio change‡	-0.050	0.079	-0.578	0.111
Marriage ratio change‡	-0.036	0.030	-0.163	0.050
College ratio change‡	0.020	0.019	-0.042	0.102
Distance to MSA (100 miles)‡	0.225	0.174	0.010	0.831
Distance to interstate highway (100 miles)‡	0.065	0.074	0.003	0.509

Sample size = 3,993. Dagger (†) indicates neighborhood variable using spatially lagged value, and ‡ indicates neighborhood variable using inverse distance weighted average.

The white population ratio declined by 5 percent during the 1990s. Slight gains occurred in the black population, 16 percent to 16.4 percent; and more significant gains occurred in the Hispanic/Latino ethnic groups, 0.7 percent to 2.2 percent during the same period. During the 1990s, the ratio of married households decreased by 4 percent. This decline may have resulted from an increase in the divorce rate. Tennessee had the 4th highest divorce rate among the 50 states in 1994 (U.S. Census Bureau 2005).

4. EMPIRICAL ESTIMATES

Estimates of the housing growth models in equations (1) – (2) are reported in Table 3, and summary estimates for equation (3) are reported in Table 4. Sensitivity analysis of the bandwidth selected with the CV method is reported in Appendix Tables A-1 and A-2.

TABLE 3

Parameter Estimate of Housing Growth Model

Variable	Equation 1		Equation 2	
	Location Variables	Neighborhood Variables	Location Variables	Neighborhood Variables
Intercept	0.213***		0.173***	
	(0.036)		(0.036)	
Lagged housing value growth	-0.008	-0.018	-0.033**	-0.011
	(0.006)	(0.033)	(0.014)	(0.037)
Population growth	0.856***	-0.013**	0.845***	0.074***
	(0.006)	(0.006)	(0.006)	(0.019)
Age of housing stock (100 years)	-0.136**	-0.144***	-0.107**	-0.044
	(0.055)	(0.046)	(0.050)	(0.057)
Vacancy rate	-0.662***	0.725***	-0.724***	0.784***
	(0.087)	(0.163)	(0.096)	(0.179)
Water ratio	0.053	-0.209	0.082	-0.240
	(0.111)	(0.186)	(0.123)	(0.210)
White ratio change	0.062	0.157	0.254**	0.048
	(0.058)	(0.105)	(0.112)	(0.062)
Marriage ratio change	0.163**	0.297	0.136*	0.307
	(0.068)	(0.209)	(0.072)	(0.232)
College ratio change	0.419***	-0.196	0.430***	-0.413
	(0.092)	(0.280)	(0.096)	(0.305)
Distance to MSA (100 miles)	0.414	0.027	0.322	-0.357
	(0.398)	(0.467)	(0.380)	(0.386)
Distance to interstate highway (100 miles)	-0.016	-0.432	0.082	-0.074
	(0.437)	(0.404)	(0.480)	(0.512)
Adjusted R^2	0.857		0.857	
Mean variance inflation factor	41.23		8.02	

Numbers in parentheses are standard error. Sample size = 3,993. ***, **, and * denote statistically significant at 1%, 5%, and 10% level.

TABLE 4

Parameter Estimate Summary of Local Housing Growth Model Using Inverse Distance Weighted Average Variables (Equation 3, Bandwidth: 110 Miles)

	Location Variables					Neighborhood Variables				
Variable	Min	Low Quartile	Median	Upper Quartile	Max	Min	Low Quartile	Median	Upper Quartile	Max
Intercept	0.168	0.174	0.174	0.174	0.174					
Lagged housing value growth	-0.035	-0.035	-0.035	-0.035	-0.003	-0.035	-0.010	-0.010	-0.009	-0.009
Population growth	0.845	0.845	0.845	0.845	0.845	0.074	0.074	0.074	0.074	0.076
Age of housing stock (100 year)	-0.117	-0.107	-0.107	-0.107	-0.107	-0.046	-0.046	-0.046	-0.046	-0.029
Vacancy rate	-0.726	-0.725	-0.724	-0.723	-0.678	0.783	0.783	0.784	0.785	0.806
Water ratio	0.080	0.080	0.081	0.082	0.129	-0.354	-0.238	-0.237	-0.235	-0.234
White ratio change	0.253	0.254	0.255	0.255	0.286	0.038	0.048	0.048	0.048	0.048
Marriage ratio change	0.135	0.136	0.136	0.137	0.153	0.270	0.304	0.306	0.307	0.308
College ratio change	0.429	0.430	0.430	0.430	0.432	-0.567	-0.409	-0.408	-0.406	-0.405
Distance to MSA (100 miles)	0.313	0.314	0.315	0.316	0.405	-0.441	-0.351	-0.350	-0.349	-0.348
Distance to interstate highway (100 miles)	0.078	0.079	0.081	0.083	0.182	-0.172	-0.075	-0.073	-0.071	-0.069
Adjusted R^2	0.857									

Sample size = 3,993.

The lower mean vif for equation (2) (8.02) than for equation (1) (41.23) suggests that using IDWA to create the neighborhood variables reduced collinearity between the location and neighborhood variables. The higher level of collinearity in equation (1) is manifest through serious degradation of the standard errors of the coefficients for lagged housing value growth and white ratio change, which were statistically significant in equation (2) but not in equation (1). Further evidence of reduced collinearity was obtained by estimating an equation without neighborhood variables (not reported in a table). The mean vif for that equation was 1.82. Although the mean vif for equation (2) is somewhat higher than for the equation without neighborhood variables, the same location variables were statistically significant in both equations, suggesting that multicollinearity did not seriously degrade the standard errors of the location-variable coefficients in equation (2).

The coefficient for the neighborhood variable, age of housing stock, was statistically significant in equation (1) but not in equation (2), which initially suggested that multicollinearity may still have been a problem in equation (2). Further examination, however, provided evidence to the contrary. The vif of 2.67 for the coefficient of the neighborhood variable, age of housing stock, indicated that another explanation may exist for its lack of statistical significance. Perhaps IDWA did not capture the neighborhood effect of age of housing stock as well as the simple average of adjacent neighbors used in estimating equation (1). This explanation is intuitively more convincing because the age of housing stock may be a variable whose neighborhood effect exists only in the adjacent neighbors and not beyond. IDWA may dilute the effects of the adjacent neighbors by giving inappropriate weight to age of housing stock in census-block groups that are spatially distant and less weight to the adjacent neighbors. Thus, a simple average of the adjacent neighbors may be a more appropriate tool for capturing the spatial spillover effect for the age of housing stock. Although outside the scope of this research, mixing neighborhood variables created with IDWA and simple averages of adjacent neighbors needs further investigation.

Since collinearity diagnostics indicated a clear multicollinearity problem, a locally weighted regression estimated with neighborhood variables that ignores distance to neighboring areas may not be reliable and was not estimated. No statistical tests are available for the parameters of locally weighted regression. Hence, to save space, results for the locally weighted regression model in equation (3) are discussed for variables that are statistically significant at the 5 percent level in equation (2). Discussion of results summarizes the marginal effect of an explanatory variable by presenting the marginal effect for the census-block group with the median marginal effect (median housing growth) (Table 4).

Figure 1 shows the relationship between housing growth predicted by the model in equation (3) and observed housing growth. Housing growth, when using the estimated model and assuming a correct prediction occurred when the predicted value was within 50 percent of actual, predicted correctly 62 percent of the time. An increase of 10 percent in the value of housing growth during the previous decade decreased median housing growth by 0.35 percent with the neighborhood effect not significant. Lagged housing value growth is inversely related to housing growth. The marginal effect of lagged housing value growth in absolute value, at the 110-mile bandwidth, is lowest in east Tennessee (0.03 percent) and increases westwardly with the highest estimate in west Tennessee (0.35 percent) (Figure 2).

Increasing neighborhood population growth by 10 percent would increase median housing growth by 0.74 percent, while the same increase of the location variable would increase the median housing growth by 8.45 percent. Thus, these estimates indicate that the neighborhood effect of population growth on housing growth in a census-block group is about one-tenth the effect of population growth within the census-block group itself. The local marginal effects of the population growth location variable is lowest around Knoxville and Chattanooga and increases slightly as one moves away from southeast Tennessee (Figure 3). The local marginal effect of neighborhood population growth is highest in west Tennessee and decreases slightly as one moves toward east Tennessee (Figure 4). Despite slight differences in the local marginal effects, both location and neighborhood effects of population growth on housing growth are relatively homogeneous throughout the state.

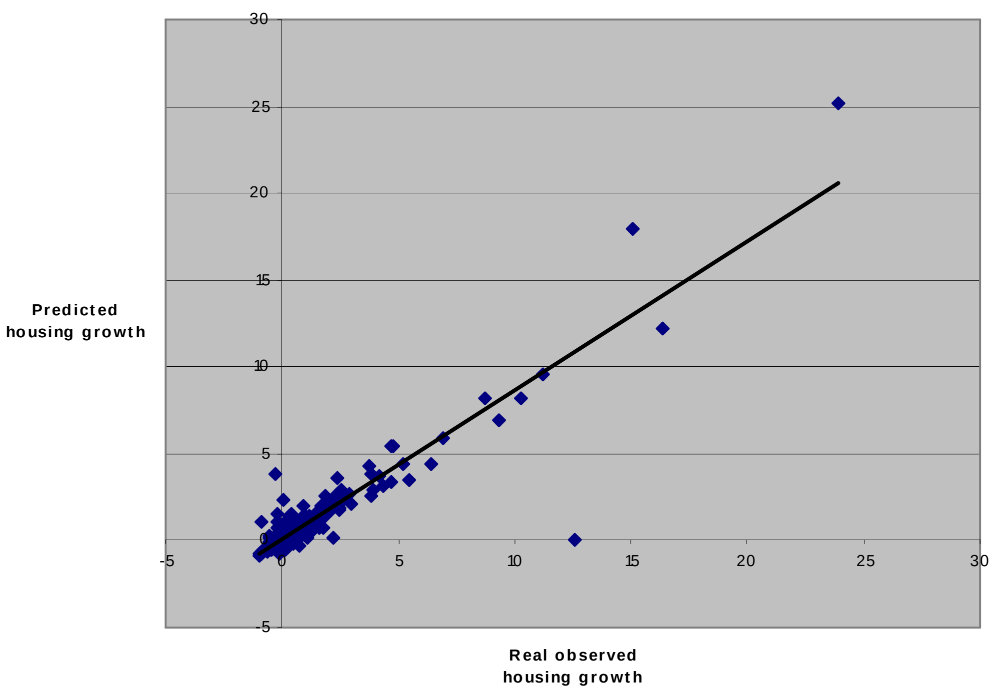

FIGURE 1. Observed and Predicted Housing Growth Using Inverse Distance Weighted Average Variables

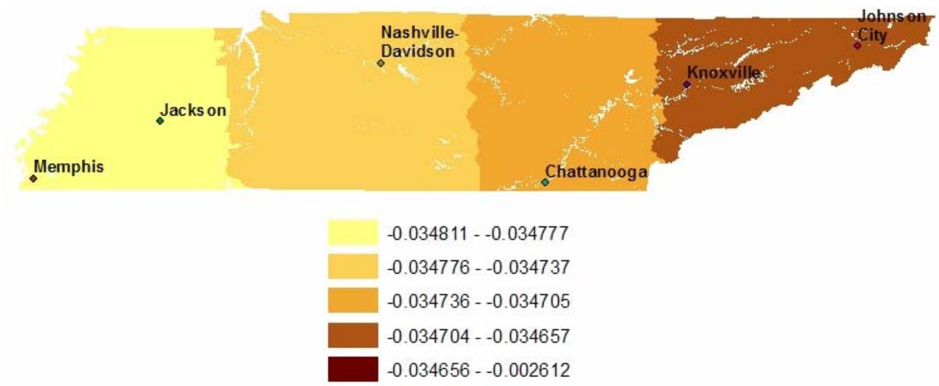

FIGURE 2. Local Marginal Effect of
Lagged Housing Value Growth

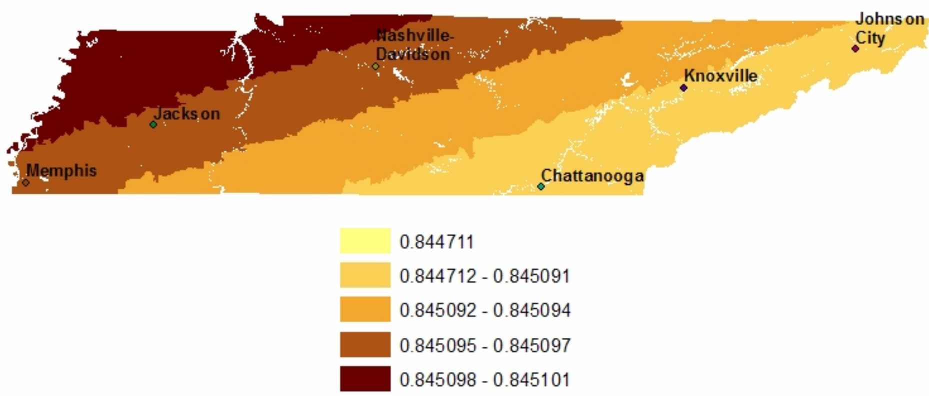

FIGURE 3. Local Marginal Effect of Population Growth

For both the white ratio and the college graduate ratio, the neighborhood effects are insignificant. An increase in the white ratio by 10 percent increases median housing growth by 2.55 percent. Local marginal effects increase as one moves from the west toward the east, although the difference is small (Figure 5). An increase in the college graduate ratio by 10 percent increases median housing growth by 4.30 percent, while the neighborhood effect is not significant. Local marginal effects increase from the east toward the west, though the effects are not substantially different (Figure 6).

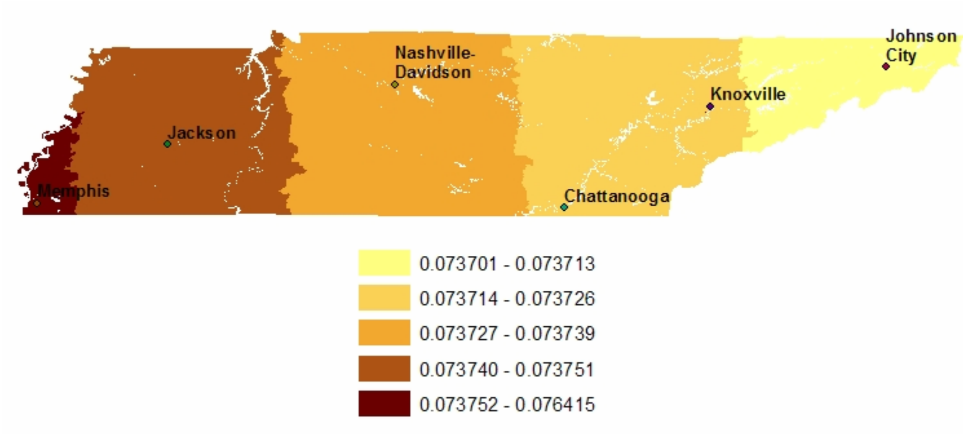

FIGURE 4. Local Marginal Effect of Neighborhood Population Growth

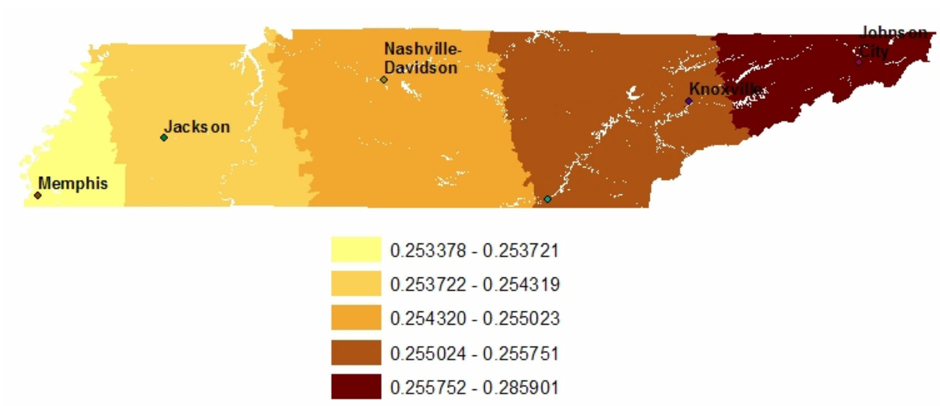

FIGURE 5. Local Marginal Effect of White Ratio Change

An increase in the neighborhood vacancy rate by 10 percent increases median housing growth by 7.84 percent, while the same percentage increase in the local vacancy rate decreases median housing growth by 7.24 percent. The opposite effects of the location and neighborhood variables is worthy of note. The own effect gradually increases as one moves from the east toward the west, while the neighborhood effect gradually increases as one moves from the west toward the east. Local marginal effects of the location variable varies in absolute value between 6.78 percent around Johnson City and 7.26 percent around Memphis (Figure 7), while local marginal effects of the neighborhood variable vary between 7.83 percent around Memphis and 8.06 percent around Johnson City (Figure 8).

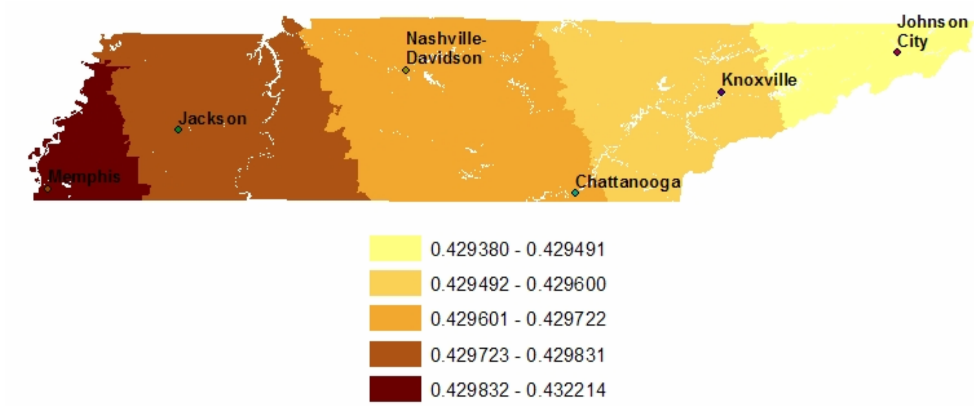

FIGURE 6. Local Marginal Effect of College Graduate Ratio Change

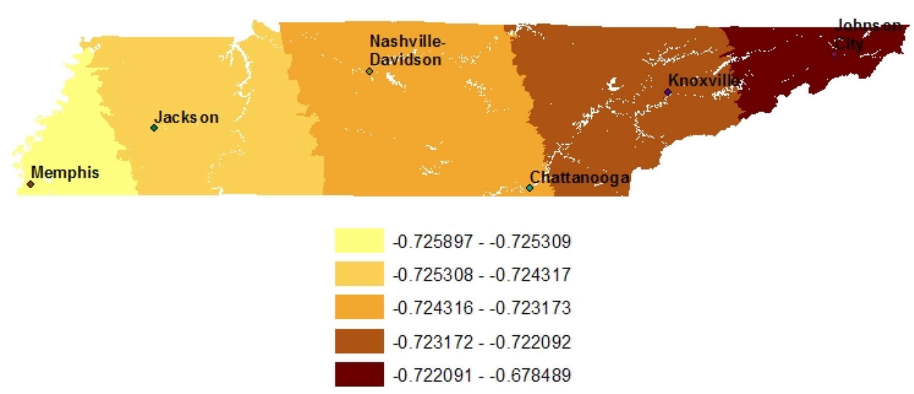

FIGURE 7. Local Marginal Effect of Vacancy Ratio

An increase in age of housing stock by 10 years decreases median housing growth by 1 percent, while no significant neighborhood effect exists. Despite slight differences of the local marginal effects, the own effects of age of housing stock on housing growth are relatively homogeneous throughout the state (Figure 9).

To examine the volatility of local regression estimates, the housing growth model with IDWA is estimated using a bandwidth that is 50 percent larger and 50 percent smaller than the bandwidth found using the CV approach in estimating equation (3). Estimates using these larger and smaller bandwidths are reported in Appendix Tables A-1 and A-2, respectively. The median value of the local marginal effects using both 165-

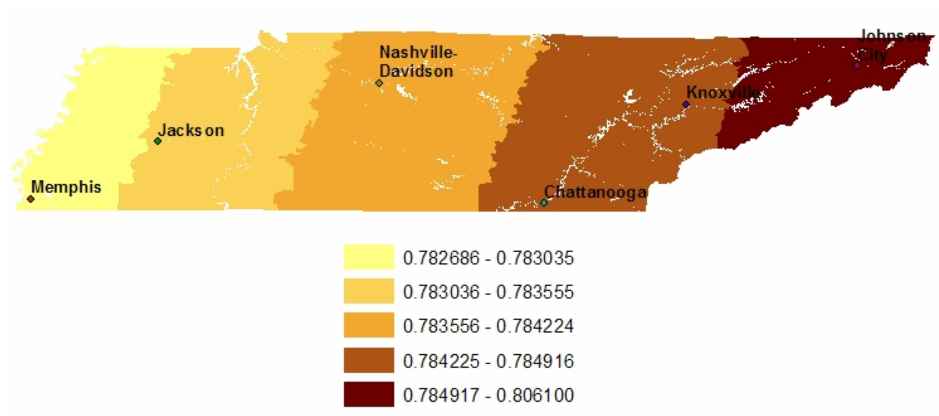

FIGURE 8. Local Marginal Effect of Neighborhood Vacancy Ratio

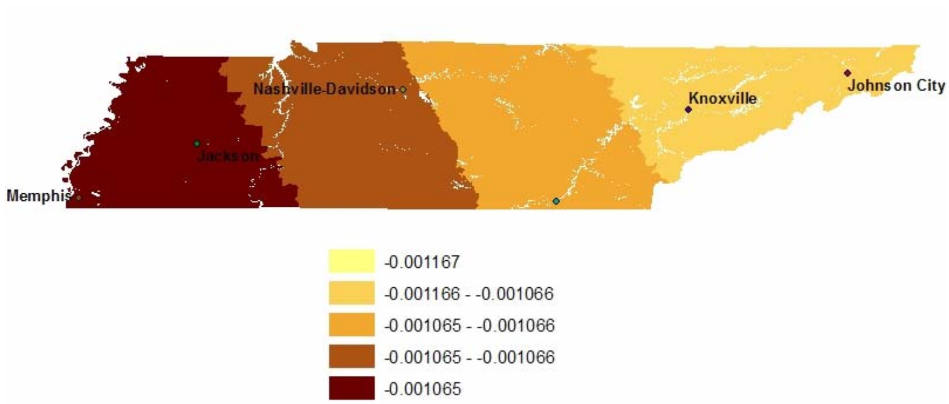

FIGURE 9. Local Marginal Effect of Age of Housing Stock

mile and 55-mile bandwidths are fairly close to the median estimates using the CV approach that identified an optimal bandwidth of 110 miles. However, with a bandwidth of 165 miles, almost no variation across the state exists in the local marginal effects. As the bandwidth widens to 165 miles, the spatial heterogeneity captured by locally weighted regression using the CV approach is not captured and the β estimates are close to those estimated by OLS in equation (2). This sensitivity analysis emphasizes the trade-off between a smaller bandwidth, which retains the spatial heterogeneity inherent in the variables, and the need to produce estimates that vary smoothly over the spatial regions of the study area (larger bandwidth).

5. SUMMARY AND CONCLUSION

This research accounted for spatial spillover and spatial heterogeneity in developing a predictive housing growth model using census-block group data from 1980, 1990, and 2000. A deterministic estimation method, IDWA, was used to create neighborhood variables that captured spatial spillover while overcoming problems with multicollinearity.

The block-group specific local effects on housing growth were computed using locally weighted regression. The local effects of the variables that are statistically significant at the 5 percent level in equation (2) were mapped using GIS. Mapping the locally weighted regression estimates revealed spatial patterns of their local marginal effects on housing growth. Based on our local-model estimates, planners could develop programs that encourage or discourage housing growth in particular locations depending on how variables affect housing growth in those locations.

The estimates could be used to predict future housing growth at the local level given projected local characteristics and neighborhood characteristics. For example, the effect of a vacancy ratio change within a block-group on its own housing growth was found to be negative across the state, and local marginal effects were greater in west Tennessee than in east Tennessee. Planners could anticipate greater housing growth within block-groups in west Tennessee than in east Tennessee with similar reductions in their own vacancy ratios. A positive neighborhood effect of a change in the vacancy ratio on housing growth was found throughout the state. A positive neighborhood effect implies that planners within a block-group could anticipate greater housing growth if vacancy rates increased among its neighbors. The neighborhood effect of population growth on housing growth within a block-group was estimated at about one-tenth of its own effect. This information can be useful for planners who try to project spillover effects of population growth on housing growth.

Based on the results of our study, growth drivers play out in distinctive ways at the local level with spatial spillover effects. These different growth drivers imply that growth of an area has to be managed in a different way according to the variations of the relationships. These findings indicate that as development proceeds, regional shifts will bring changes in their social structures differently at the local level. These changes will likely give rise to conflict as development proceeds and will have implications for how subsequent development might be organized across a region.

The logical next step for analysis of housing growth is to apply locally weighted regression to a hedonic housing price model with neighborhood effects using IDWA. Such a model would assist in understanding the dynamics of local housing markets and help in recognizing the structures of local submarkets. Analyzing the impact of the different determinants on the value of housing at the local level would help us understand housing demand at a census-block level and could be used for a better analysis at finer scale units if individual housing data were readily available. This data set could be built using a database of individual houses from county tax assessors' offices, the census dataset of block levels, and a GIS database created using information about individual houses. While collecting a dataset from the 95 counties of the entire state of Tennessee would be extremely expensive, a sample study for some selected counties in which all the characteristics of growth are contained might be feasible.

Another extension of this research is to investigate the characteristics of neighborhood effects and the neighborhood variables used to capture them. For example, although overall performance was better for the model with neighborhood variables created using IDWA than for the model using the simple average of adjacent neighbors, the neighborhood variable for age of housing stock created using IDWA did not perform as well as its adjacent-neighbor counterpart. Because the neighborhood effect of a particular variable may be confined to the adjacent neighbors and neighborhood effect of another variable may be influential far beyond the adjacent neighbors, one method of creating neighborhood variables may not be desirable. Further research should investigate the ranges of influence different neighborhood effects have on housing growth. The most appropriate methods for creating neighborhood variables should correlate well with the ranges of their associated neighborhood effects. By investigating the ranges of influence of important neighborhood effects, more appropriate methods for capturing neighborhood effects can be developed.

REFERENCES

Anselin, L., 1988. *Spatial Econometrics: Methods and Model.* Kluwer Academic: Dordrecht.

Anselin, L., R. Dodson, and S. Hudak, 1993. "Linking GIS and Spatial Data Analysis in Practice," *Geographical Systems* 1, 3-23.

Biles, J., 2003. "Using Spatial Econometric Techniques to Estimate Spatial Multipliers: An Assessment of Regional Economic Policy in Yucatan, Mexico," *The Review of Regional Studies* 33, 121-141.

Blackley, D., 1999. "The Long-Run Elasticity of New Housing Supply in the United States: Empirical Evidence for 1950 to 1994," *Journal of Real Estate Finance Economics* 18, 25-42.

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

_____, 1992. "Specification and Estimation of Hedonic Housing Price Models," *Regional Science and Urban Economics* 22, 453-474.

Cao, T. and D. Cory, 1981. "Mixed Land Uses, Land-Use Externalities, and Residential Property Values: A Reevaluation," *Annals of Regional Science* 16, 1-24.

Casetti, E., 1972. "Generating Models by the Expansion Method: Applications to Geographic Research," *Geographic Analysis* 4, 81-91.

Chan, T., 1999. "Residential Construction and Credit Market Imperfection," *Journal of Real Estate Finance Economics* 18, 125-139.

Charlton, M., S. Fotheringham, and C. Brunsdon, 2002. *User's Manual and Installation Guide, Geographically Weighted Regression: User's Manual and Installation Guide.*

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

Clark, D. and J. McGibany, 1990. *A Pooled Countywide Model of Housing Construction: 1978-1989*, Working Paper 90-103, College of Business Administration, Marquette University.

Cleveland, W.S., 1979. "Robust Locally Weighted Regression and Smoothing Scatter-plots," *Journal of the American Statistical Association* 74, 829-836.

Cleveland, W.S. and S.J. Devlin, 1988. "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," *Journal of the American Statistical Association* 83, 596-610.

Conway, R. and C. Howard, 1980. "A Forecasting Model for Regional Housing Construction," *Journal of Regional Science* 20, 1-10.

Crone, T. and L. Mills, 1991. "Forecasting Trends in the Housing Stock Using Age-Specific Demographic Projections," *Journal of Housing Research* 2, 1-20.

DiPasquale, D. and W. Wheaton, 1994. "Housing Market Dynamics and the Future of Housing Prices," *Journal of Urban Economics* 35, 1-27.

Dubin, R., 1992. "Spatial Autocorrelation and Neighborhood Quality," *Regional Science and Urban Economics* 22, 433-452.

_____, 1998. "Spatial Autocorrelation: A Primer," *Journal of Housing Economics* 7, 304-327.

ESRI, Environmental Systems Research Institute, 2005. *ArcGIS 9.1*.

Follain, J., 1979. "The Price Elasticity of the Long-Run Supply of New Housing Construction," *Land Economics* 55, 190-199.

Fotheringham, A. and C. Brunsdon, 1999. "Local Forms of Spatial Analysis," *Geographical Analysis* 31, 341-358.

Fotheringham, A., C. Brunsdon, and M. Charlton, 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. John Wiley & Sons Ltd.: West Sussex, England.

Fotheringham, A., M. Charlton, and C. Brunsdon, 1996. "The Geography of Parameter Space: An Investigation of Spatial Non-Stationarity," *International Journal of Geographical Information Systems* 10, 605-627.

_____, 1997. "Two Techniques for Exploring Non-Stationary in Geographical Data," *Geographical Systems* 4, 59-82.

Fotheringham, A. and P. Rogerson, 1993. "GIS and Spatial Analytical Problems," *International Journal of Geographical Information Systems* 7, 3-19.

Fu, Y. and C. Somerville, 2001. "Site Density Restrictions: Measurement and Empirical Analysis," *Journal of Urban Economics* 49, 404-423.

Geoghegan, J., L. Waigner, and N. Bockstael, 1997. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis using GIS," *Ecological Economics* 23, 251-264.

Getis, A. and K. Ord, 1992. "The Analysis of Spatial Association by Use of Distance Statistics," *Geographical Analysis* 24, 189-206.

Glaeser, E. and J. Gyourko, 2001. "Urban Decline and Durable Housing," Working Paper 8598, National Bureau of Economic Research.

Goodman, A., 1977. "A Comparison of Block Group and Census Tract Data in a Hedonic Housing Price Model," *Land Economics* 53, 483-487.

Haining, R., 1990. *Spatial Data Analysis in the Social and Environment Sciences*. Cambridge University Press: Cambridge.

Hancock, D. and J. Wilcox, 1997. "Bank Capital, Nonbank Finance, and Real Estate Activity," *Journal of Housing Research* 8, 75-105.

Jaffe, D. and K. Rosen, 1979. "Mortgage Credit Availability and Residential Construction," *Brookings Papers on Economic Activity* 2, 333-376.

Judge, G., C. Hill, W. Griffiths, H. Lütkepohl, and T. Lee, 1982. *Introduction to the Theory and Practice of Econometrics*. John Wiley & Sons: New York.

Keckler, D., 1995. *The Surfer Manual*. Golden Software, Inc.: Golden, CO.

LeSage, J., 1997. "Regression Analysis of Spatial Data," *Journal of Regional Analysis and Policy* 27, 83-94.

Leung, Y., C. Mei, and W. Zhang, 2000. "Testing for Spatial Autocorrelation Among the Residuals of the Geographically Weighted Regression," *Environmental and Planning A* 32, 871-890.

Maddala, G., 1992. *Introduction to Econometrics*, 2nd Edition. MacMillan: New York.

McDonald, J. and D. McMillen, 2000. "Residential Building Permits in Urban Counties: 1990-1997," *Journal of Housing Economics* 9, 175-186.

McGibany, J., 1991. "The Effect of Property Tax Rate Differentials on Single-Family Housing Starts in Wisconsin, 1978-1989," *Journal of Regional Science* 31, 347-359.

McGinnis, H., 1994. "Determining the Impact of Economic Factors on Local Government Growth Policy: Using Time-Series Analysis and Transfer Function Models," *Urban Studies* 31, 233-246.

McMillen, D., 1992. "Probit with Spatial Autocorrelation," *Journal of Regional Science* 3, 335-348.

_____, 1996. "One Hundred Fifty Years of Land Values in Chicago: A Non-Parametric Approach," *Journal of Urban Economics* 40, 100-124.

_____, 2003a. "Spatial Autocorrelation or Model Misspecification?" *International Regional Science Review* 26, 208-217.

_____, 2003b. "Neighborhood House Price Indexes in Chicago; A Fourier Repeat Sales Approach," *Journal of Economic Geography* 3, 57-73.

McMillen, D. and McDonald, 1997. "A Nonparametric Analysis of Employment Density in a Polycentric City," *Journal of Regional Science* 37, 591-612.

Meen, G. and M. Andrew, 1998. "On the Aggregate Housing Market Implications of Labour Market Change," *Scottish Journal of Political Economy* 45, 393-419.

Meese, R. and N. Wallace, 1991. "Nonparametric Estimation of Dynamic Hedonic Price Models and the Construction of Residential Housing Price Indices," *Journal of the American Real Estate and Urban Economics Association* 19, 308-332.

Poterba, J., 1984. "Tax Subsidies to Owner-Occupied Housing: An Asset Market Approach," *Quarterly Journal of Economics* 99, 729-752.

Puri, A. and J. Van Lierop, 1988. "Forecasting Housing Starts," *International Journal of Forecasting* 4, 125-134.

Rahman, M. and M. Muhammad, 1997. "Growth in U.S. Housing Starts, Real Consumer Debt, Real GDP, and the Long-Term Real Interest Rate: A Vector Cointegration Analysis," *Applied Economics Letters* 4, 757-759.

Shepard, D., 1968. "A Two-Dimensional Interpolation Function for Irregularity-Spaced Data," Proceedings of the 23rd National Conference Association for Computing Machinery, 517-524, ACM Press.
Song, J. and J. DePinto, 1995. "A GIS-Based Data Query System," presented at the International Association for Great Lakes Research (IAGLR) Conference, Windsor, Ontario.
Skidmore, M. and M. Peddle, 1998. "Do Development Impact Fees Reduce the Rate of Residential Development?" *Growth and Change* 29, 383-400.
Switzer, P., W. Kowalik, and R. Lyon, 1982. "A Prior Probability Method for Smoothing Discriminant Analysis Classification Maps," *Mathematical Geology* 14, 433-444.
Topel, R. and S. Rosen, 1988. "Housing Investment in the United States," *Journal of Political Economy* 96, 718-740.
Thom, R., 1985. "The Relationship Between Housing Starts and Mortgage Availability," *Review of Economics and Statistics* 57, 693-696.
Tse, R., 2002. "Estimating Neighborhood Effects in House Prices: Towards a New Hedonic Model Approach," *Urban Studies* 39, 1165-1180.
U.S. Census Bureau, 2005. Available at: http://www.census.gov/.

APPENDIX

To explain the normalization of 1990 data to 2000 geographies, we start by weighting and converting 1990 block-group data to 2000 areas. 1990 block-group data are used because it is the smallest level of 1990 geography at which the full set of U.S. Census 1990 long form data is available. To facilitate the splitting and merging of 1990 block groups to 2000 areas, Census blocks are used. A Census block is much smaller than a block group. There are approximately 30 to 40 blocks in each block group. And unlike previous censuses, blocks and block groups cover 100 percent of the United States in 1990 and 2000.

The 1990 to 2000 block relations were determined from Tiger/Line 2000, Type 1 and Type 3 records. Eighty-five percent of the blocks had a 1:1 relationship, 10 percent had a 2:1 relationship, and 5 percent had a greater than 2:1 relationship. Block splits between 1990 and 2000 were weighted by an analysis of the 1990 streets. To split a block into parts, the sub-block areas were weighted according to the 1990 streets relating to each 2000 block part. The assumption is that local roads indicate where the population lived. 1990 streets were determined using Tiger/Line 1992. Using Tiger 1992 and Tiger 2000, we created a correspondence between 1990 and 2000 blocks as well as a weighting value. The weighting value was then used to help split block demographics for those blocks that had been split or merged between 1990 and 2000. The file produced by this process is the 1990 to 2000 Block Weighting File (BWF). From this BWF we can roll up the 1990 data to any 2000 geography (tract, zip code, county, etc.).

TABLE A1

Parameter Estimate Summary of Local Housing Growth Model Using Inverse Distance Weighted Average Variables (Equation 3, Bandwidth: 165 Miles)

Variable	Location Variables					Neighborhood Variables				
	Min	Low Quartile	Median	Upper Quartile	Max	Min	Low Quartile	Median	Upper Quartile	Max
Intercept	0.173	0.173	0.173	0.173	0.173					
Lagged housing value growth	-0.033	-0.033	-0.033	-0.033	-0.033	-0.011	-0.010	-0.010	-0.010	-0.010
Population growth	0.845	0.845	0.845	0.845	0.845	0.074	0.074	0.074	0.074	0.074
Age of housing stock (100 year)	-0.045	-0.045	-0.045	-0.045	-0.045	-0.107	-0.107	-0.107	-0.107	-0.107
Vacancy rate	-0.724	-0.724	-0.724	-0.724	-0.722	0.784	0.784	0.784	0.784	0.786
Water ratio	0.082	0.082	0.082	0.082	0.083	-0.243	-0.240	-0.240	-0.240	-0.239
White ratio change	0.047	0.048	0.048	0.048	0.048	0.254	0.254	0.254	0.255	0.256
Marriage ratio change	0.136	0.136	0.136	0.136	0.137	0.304	0.307	0.307	0.307	0.307
College ratio change	0.430	0.430	0.430	0.430	0.430	-0.416	-0.412	-0.412	-0.412	-0.412
Distance to MSA (100 miles)	0.321	0.321	0.321	0.321	0.322	-0.357	-0.356	-0.356	-0.356	-0.356
Distance to interstate highway (100 miles)	0.082	0.082	0.082	0.082	0.086	-0.079	-0.074	-0.074	-0.074	-0.074
Adjusted R^2	0.857									

Sample size = 3,993.

TABLE A2

Parameter Estimate Summary of Local Housing Growth Model Using Inverse Distance Weighted Average Variables (Equation 3, Bandwidth: 55 Miles)

	Location Variables					Neighborhood Variables				
Variable	Min	Low Quartile	Median	Upper Quartile	Max	Min	Low Quartile	Median	Upper Quartile	Max
Intercept	0.174	0.174	0.174	0.174	0.174					
Lagged housing value growth	-0.034	-0.034	-0.034	-0.034	-0.028	-0.016	-0.010	-0.010	-0.010	-0.010
Population growth	0.845	0.845	0.845	0.845	0.845	0.074	0.074	0.074	0.074	0.074
Age of housing stock (100 year)	-0.046	-0.046	-0.046	-0.046	-0.045	-0.109	-0.107	-0.107	-0.107	-0.107
Vacancy rate	-0.725	-0.725	-0.724	-0.724	-0.706	0.784	0.784	0.784	0.785	0.797
Water ratio	0.081	0.081	0.081	0.081	0.095	-0.274	-0.238	-0.237	-0.237	-0.236
White ratio change	0.045	0.048	0.048	0.048	0.048	0.254	0.254	0.254	0.255	0.267
Marriage ratio change	0.136	0.136	0.136	0.136	0.144	0.287	0.305	0.306	0.307	0.307
College ratio change	0.429	0.430	0.430	0.430	0.430	-0.450	-0.409	-0.408	-0.408	-0.407
Distance to MSA (100 miles)	0.316	0.317	0.317	0.317	0.327	-0.364	-0.352	-0.352	-0.351	-0.351
Distance to interstate highway (100 miles)	0.080	0.081	0.081	0.082	0.120	-0.114	-0.075	-0.074	-0.073	-0.072
Adjusted R^2	0.857									
Sample size = 3,993										