Projecting Spatial Pattern of Housing Growth in Tennessee

Seong-Hoon Cho, Christopher D. Clark, William M. Park[1]

Department of Agricultural Economics, University of Tennessee

Series title and number: "Projecting the Pattern and Density of Development at the Rural-Urban Interface" 137017

Date of publication: May 12, 2005

Postal and e-mail address for the authors: Seong-Hoon Cho, 314 Morgan Hall, Department of Agricultural Economics, University of Tennessee, Knoxville, TN 37996-4518 scho@mail.ag.utk.edu ; Christopher D. Clark, 302 Morgan Hall, Department of Agricultural Economics, University of Tennessee, Knoxville, TN 37996-4518 cdclark@utk.edu ; William M. Park, 302 Morgan Hall, Department of Agricultural Economics, University of Tennessee Knoxville, TN 37996-4518 wpark@utk.edu

Abstract: Housing growth in Tennessee that incorporates spatial spillover and spatial heterogeneity at the level of census-block group is projected. A deterministic interpolation technique is adopted to create alternative neighborhood variables that captures spatial spillover of neighborhood effects on housing growth without multicollinearity. The maps drawn using the geographically weighted regression parameter estimates revealed that the local marginal effect of the housing price increases on housing growth gradually increases as one moves eastward. The population growth in the adjacent neighborhood-block group has about 10% of marginal effect of population growth in its own block group. The marginal effect of population growth is relatively higher in the approximate area of Cumberland Plateau while the local marginal effect of spatial spillover of population growth in adjacent neighbor is relatively higher in the East Tennessee and also the area South of Nashville.

Key words: housing growth, multicollinearity, spatial spillover, spatial heterogeneity

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Providence, Rhode Island, July 24-27, 2005

[1] Assistant Professor, Assistant Professor, Professor, Department of Agricultural Economics, University of Tennessee.

Projecting Spatial Pattern of Housing Growth in Tennessee

1. Introduction

Between 1990 and 2000, the housing in the State of Tennessee went from 2.0 million to 2.4 million houses, an increase of 20% which ranks Tennessee 12th in housing growth. This increase is much greater than the change in national housing which saw an increase from 102 million to 115 million houses, an increase of 13% during the same period of time. This large increase of the number of houses in Tennessee is closely tied to a population increase from 4.9 million to 5.7 million or increase of 16% (U.S. Census Bureau). The spatial pattern of the rapid residential development has tended to be low density development commonly referred to as "sprawl." Tennessee's metropolitan areas, particularly Nashville and Knoxville have consistently been ranked as two of the nation's most sprawling metropolitan areas. In fact, Tennessee's increase of developed area[2] was the 7th largest among all 50 states in the period of 1992-1997.

This relatively rapid growth in the State of Tennessee has given rise to concerns over declining environmental quality. The metropolitan areas across the state have been declared ozone non-attainment areas, and the impairment of waterways has been directly linked to residential and urban development. The rapid growth also puts pressure on public services, such as sewage treatment, health service, and road maintenance. Despite the rise to concerns over declining environmental quality and pressure on public services by the rapid growth, there has been a lack of systematic studies on understanding the propensity of growth as it relates to socioeconomic and location factors in the State of Tennessee. Since housing growth as a result

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

of residential development is the dominant driving force of growth in the State of Tennessee, any systematic study of the growth needs to incorporate modeling housing stock.

Understanding housing growth is important to planners and policy makers because housing growth has many impacts on natural resources – among them forest fragmentation, increased stormwater runoff, degradation or loss of wildlife habitat, and loss of open space necessary for human health and recreation. Many of these consequences of residential growth are serious issues. The success of resource management in a prosperous, growing land depends largely on the ability to reconcile residential growth patterns with the natural landscape.

Most of previous studies of housing growth examine structural models of housing demand and supply (Blackley, 1999; DiPasquale and Wheaton, 1994; Follain, 1979; Poterba, 1984; Topel and Rosen, 1988; Crone and Mills, 1991). Many used residential building permits to examine fluctuations of housing stock (Chan, 1999; Hancock and Wilcox, 1997; Jaffee and Rosen, 1979; Puri and Van Lierop, 1988; Rahman and Muhammad, 1997; and Thom, 1985; McGinnis, 1994). These studies examined the residential development at the national level. Fewer studies have examined residential building at the state or local level. Models for regional housing construction at county level were developed (Conway and Howard, 1980; Clark and McGibany, 1990; McGibany, 1991; Skidmore and Peddle, 1998; McDonald and McMillen, 2000). The models link additions to the housing stock to the levels of population, income per household, and other variables. While the studies of the regional housing stock focused on influences of population, income, and other economic variables on changes in housing stock, they did not examine whether changes in the housing stock were related to changes in demographic characteristics and spatial attributes.

Many of the factors that affect housing growth are spatially explicit. It is frequently said of real estate that what matters most is location. A significant advantage of a spatially explicit model is that it can readily incorporate substantial spatial detail, allowing analysis of how various locational factors influence housing growth. The role of locational factors in housing growth can be examined in two interrelated methods. One form of geographic influence involves externalities associated with the location of the house growth. These types of externalities are called adjacency effects as they capture the spatial spillover on a given area by a neighboring area. In addition to spillovers from neighboring area, spatially varying relationships also enter into housing growth. These kinds of relationships may be called spatial heterogeneity of housing growth.

The understanding of spatial spillover and spatially varying relationships in linear regression models as well as the development of efficient and consistent estimators for these types of models has been an important part of researches over the last few decades (e.g., McMillen, 2003; Tse, 2002; Leung et al., 2000; LeSage, 1997; McMillen, 1992; Anselin, 1988; Cliff and Ord, 1973). A number of literature estimates regression parameters in the presence of spatial spillover (e.g., Dubin, 1992 and 1998; Can, 1990 and 1992). Various localized modeling techniques were proposed to capture spatial heterogeneity (Casetti, 1972; Getis and Ord, 1992; Fotheringham and Brunsdon, 1999). Since every location has an intrinsic degree of uniqueness due to its situation with respect to the rest of the spatial system, spatial heterogeneity occurs. Because of the spatial heterogeneity, the estimated parameters of a spatial model are inadequate descriptors of the process at any given location due to parameter drift across space (Anselin, et al. 1993; Fotheringham, et al. 1996, 1997; Fotheringham and Rogerson 1993). Some models incorporate both spatial spillover and spatial heterogeneity (Anselin, 1988; Can, 1992). While such models of spatial spillover and heterogeneity have focused on modeling stochastic processes of interrelation of point locations, to our knowledge, the socio-demographic and location factors of housing growth have not been similarly modeled within a spatially-explicit context.

In this paper, the housing growth for a group of neighbors related to changes in demographic variables that incorporate spatial spillover and spatial heterogeneity is analyzed. The neighborhood effects of spatial spillover are captured in the form of neighborhood variables that are associated with the nearest neighbors' spatial dependency on housing growth. Because multicollinearity between the variables and the neighborhood variables associated with the nearest neighbors is problematic, a deterministic interpolation technique, a Geographical Information System (GIS) tool, is adopted to create alternative neighborhood variables that can capture spatial spillover without multicollearity problems. In addition to the spatial spillover, a geographically weighted regression (GWR) model proposed in Fotheringham, et al. (2002) is adopted to capture spatial heterogeneity of the relationship between housing growth and socio-demographic and location factors.

2. The Empirical Model

The empirical model used in this analysis follows the model originally developed by Conway and Howard (1980) and later adopted by McGibany (1991), Skidmore and Peddle (1998), and McDonald and McMillen (2000). They estimated a reduced-form equation of demand and supply of new housing structures under equilibrium conditions in order to learn about propensity of the change in the housing stock. Following the reduced form, the residential development in our model is defined as an equilibrium point of demand and supply of new houses. The change in housing numbers through additions to the housing stock over time in a given neighbor is expressed as the rate of residential development,

$$Q = \beta_0 + \sum_k \beta_k X_k + \varepsilon, \tag{1}$$

where Q is a change in number of houses between 1990 and 2000 with respect to the base year, 1990 at the block-group level, i.e., $Q = [E(Q_{2000}) - E(Q_{1990})]/E(Q_{1990})$, X_k is a vector of independent variables for k=1, 2, …, m, ε is an error term β_0 and β_k are parameters to be estimated. Skidmore and Peddle (1998) suggest that the change in housing numbers in a neighbor over time is functional form of a vector of neighbor attribute, the average assessed value of property, and a vector of variables representing the regions. Following the general guidance of the literature, the change in housing numbers in a neighbor over time is specified as a function of a vector of independent variables X_k, which consist of neighbor attributes including demographic characteristics, the average assessed value of property, a vector of variables representing regions including distance variables.

The model is estimated with normalized 1990 Census Long Form Data in 2000 boundaries and 2000 long form data at a census-block group level for a group of neighbors. A census-block group is used as a group of neighbors in our study because the census-block groups meet two conditions for linear aggregation of number of houses for a group of neighbors (Meen and Andrew, 1998). The two conditions are: (1) Income and other variables are to be growing at the same rate in each location of neighbors or exhibit a common stochastic trend. (2) The structure of the housing markets is to be the same over the space within the boundary of neighbors. A few previous studies have specified the census-block groups as groups of neighbors (Goodman 1977; Cao and Cory 1981; Geoghegan et al. 1997). In their studies, the linear aggregations of housing data at the census block-group level yield robust empirical estimations.

In order to estimate neighborhood effects of spatial spillover in the model, a prior probability method proposed by Switzer et al. (1982) is applied. The equation for the housing growth that encountered spatial spillover is expressed as:

(2) $$Q = \beta_0 + \sum_k \beta_k X_k + \sum_k \beta_k' X_k' + \varepsilon$$

For simplification, let X_k' take the value of X_k that are associated with the nearest neighbors' spatial spillover. The values of the X_k' are calculated such a way that a mean value of a variable of neighbor block groups, which are adjacent to a block group you are measuring spatial spillover of is assigned to the block group. The mean value of the variable of neighbor block groups count only the values of the variable of the surrounding block groups not the value of the block group itself. Because the multicollinearity between the variables and the variables associated with the nearest neighbors may lead to unstable estimates, the model themselves would be improper, let alone their estimates.

Although there have been many suggestions about how to detect multicollinearity, there are no certain guidelines. A commonly used rule of thumb is that if the correlation coefficient between the values of two regressors is greater than 0.8 or 0.9, then multicollinearity is a serious problem (Judge et. al. 1982, p. 620). The best solution is to understand the cause of multicollinearity and remove it. The multicollinearity occurs because two (or more) variables are related, removing one of the two or more may require decision makings of which variables to remove. The decision for which variables to remove does not have a firm guideline. Rather than removing the variables that are highly correlated, an alternative way to combine the variables and the variables associated with the nearest neighbors is suggested in this research.

We suggest to combine the variables that cause muticollinearity by adopting a GIS tool of deterministic interpolation technique called Inverse Distance Weighted Averaging (IDWA). A neighborhood about the interpolated point is identified and a weighted average is taken of the observation values within the neighborhood. The interpolation is based on the idea that points that are close to one another in space have more similar characteristics than ones further away. The weights are a decreasing function of distance. The neighborhood size determines how many points are included in the inverse distance weighting. The neighborhood size can be specified in terms of its radius or the number of points. A number of block-group in our model specifies the number of points. The various sizes and shapes of the block-groups make difficult to specify the neighborhood in terms of its radius. There is no clear rule choosing how many points for the neighborhood size. The IDWA for the variables that are highly correlated, $\tilde{X}_k$ is calculated using the following equation

(3)
$$\tilde{X}_k = \frac{\sum_{i=1}^{n} \frac{\tilde{X}_i}{d_{ij}^{w}}}{\sum_{i=1}^{n} \frac{1}{d_{ij}^{w}}},$$

n represents the total number of sample data values, d_{ij} denotes the separation distance between interpolated value and the data value, and w denotes the weighting power (Keckler, 1995; Song and Depinto, 1995). The equation for the housing growth that encountered spatial dependency without the multicollinearity problem using the IDWA is expressed as

(4)
$$Q = \beta_0 + \sum_k \beta_k \hat{X}_k + \sum_k \beta_k X_k'' + \sum_k \tilde{\beta}_k \tilde{X}_k + \varepsilon ,$$

where $\hat{X}_k$ and X_k'' are the original variables and the neighborhood variables that are not highly correlated respectively.

An implicit assumption made throughout the equations (1), (2), and (4) are that relationships between variables measured at different locations are constant over space. If there are essential structural variations of housing growth in the study area, then constant variables measured at different locations would represent a misspecification of the data. For instance, a value of distance to Center of Business District (CBD) as a determinant of housing growth in different parts of a study region may vary. If such variations in relationships exist over space, then the housing growth models of the equations (1), (2), and (4) clearly are misspecifications of reality because it assumes these relationships to be constant. A geographically weighted regression (GWR) model proposed by Fotheringham et al. (2002) is adopted to identify these variations in relationships in space or spatial heterogeneity.

The GWR model extends traditional regression framework by allowing parameters to be estimated locally so that the models in equations (1), (2), and (4) are rewritten as

(5) $$Q_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) X_{ki} + \varepsilon_i ,$$

(6) $$Q_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) X_{ki} + \sum_k \beta_k(u_i, v_i)' X_{ki}' + \varepsilon_i , \text{ and}$$

(7) $$Q_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) \hat{X}_{ki} + \sum_k \beta_k(u_i, v_i) X_{ki}'' + \sum_k \tilde{\beta}_k(u_i, v_i) \tilde{X}_{ki} + \varepsilon_i$$

respectively where, (u_i, v_i) denotes the coordinates of the i th point in space and $\beta_k(u_i, v_i)$ is a realization of the continuous function $\beta_k(u, v)$ at point *i*. That is, there to be a continuous surface of parameter values, and measurements of this surface are taken at certain points to denote the spatial variability of the surface (Fotheringham et al. 2002).

In GWR models an observation is weighted in accordance with its proximity to point *i* in order to account the fact that an observation near to point *i* have more of an influence in the estimation of the $\beta_k(u_i, v_i)$ s than do observations located farther from *i*. That is,

(8) $$\hat{\beta}(u_i, v_i) = (Z_k{}^T W(u_i, v_i) Z_k)^{-1} Z_k{}^T W(u_i, v_i) Q \text{ ,}$$

where $W(u_i, v_i)$ is an $n \times n$ matrix whose diagonal elements w_{ij} denotes the geographical weighting of each of the *n* observed data for regression point *i* and a specific point *j* in space at which data are observed, and the off-diagonal elements are zero. Z_k is a vector of explanatory variables. The diagonal elements of the weight matrix, w_{ij} is obtained as

(9) $$w_{ij} = \exp[-1/2(d_{ij}/b)^2]$$

where d_{ij} is the Euclidean distance[3] between point *i* and *j*, *b* is a bandwidth. The bandwidth can be determined using the cross-validation procedure. The parameter takes the form:

(10) $$\Delta(b) = \sum_{i=1}^{n}(Q_i - \hat{Q}_i(b))^2$$

where $\hat{Q}_i(b)$ is the fitted value Q_i with the observation at location *i* omitted from the fitting process. Choose b_0 as a desirable value of the bandwidth such that $\Delta(b_0) = \min \Delta(b)$. If *i* and *j* coincide, the $w_{ij} = 1$ and the weighting of other data will decrease according to a normal distribution curve as the distance between *i* and *j* increases.

To evaluate the prediction accuracy of the housing growth model, out-of-sample prediction is computed using the estimated coefficients from the models. For each out-of-sample prediction, absolute difference between the estimated value of the predicted and the actual

[3] Distance between objects or values that is computed as a straight line.

housing growth is calculated. The percentage of differences within 50% of the observed housing growth is estimated.

3. Data

Normalized 1990 Census Long Form Data in 2000 boundaries and 2000 long form data at census-block group level in spatial form are used for the estimation of the empirical models. The normalized data was created by a private data provider, GeoLytics. The 1990 data had to be normalized in 2000 boundaries because of changes in geographic definitions of the boundaries. This enables comparisons between 1990 and 2000 Long Form data to be made in standard 2000 geographies. Appendix 1 discusses how GeoLytics normalized the 1990 Long Form census data to various 2000 geographies.

The records of 4,014 census-block groups[4] in the form of polygons within the boundary of the State of Tennessee are used for this study. Each polygon represents one census-block group. GIS is utilized to generate distance variables. The distance variables are created using 'Near' the ArcToobox, GIS tool for geoprocessing. The closest distances to six Metropolitan Statistical Areas (MSAs) from each census-block group are computed using the distance from centroid of census-block group to the nearest points of CBD 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 is used to calculate the coordinates of centroid:

[4] Although most people intuitively think of 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 group 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. The census blocks in remote areas may be large and irregular and may contain many square miles (U.S. Census Bureau).

(11) $$C_x = \frac{\sum_n A_n C_{xn}}{\sum A_n}\ ,\ C_y = \frac{\sum_n A_n C_{y_n}}{\sum A_n}$$

where C_x is the distance from the y-axis to the centroid, C_y is the distance from the x-axis to the centroid, $A = \int f(x)dx$, and (C_x, C_y) is the coordinates of the centroid. The closest distance to the interstate highway from each census-block group is 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 per acre for the block groups increased by 22%. Over the same time period, the average population per acre for the block groups increased by 18%. The higher rate of housing growth relative to population growth may be due to a decline of average household size of Tennessee, 2.56 to 2.48 during the 1990s (U.S. Census Bureau). The difference may also be a reflection of the importance of second home development in the Cumberland Plateau, the Smokey Mountains, and the most-visited National Park in the country on its doorstep of East Tennessee. A lot of the second homes are owned by residents outside of the State of Tennessee.

The ratio of white population was declined by 5% during the 1990s. The decline was occurred with slight gains occurring in the black population, 16% to 16.4%, but more significant gains in the Hispanic/Latino racial groups, 0.7% to 2.2% during the same period of time. The ratio of married households was declined by 3.6% during the 1990s. The decline of the married households may be due to the increase of divorce rate. The Tennessee is reported as the State with 4th highest divorce rate out of the 50 States in 1994 (U.S. Census Bureau).

4. Empirical Estimates

The models in the equations (1) and (5) are calibrated using Ordinary Least Square (OLS) and GWR respectively to estimate parameters reported in Table 3. The OLS is referred as a global model as opposed to a local model for the GWR. The adjusted R^2 value of the global model is 0.7907 indicating a reasonable explanatory performance but it still leaves 20.93% of the variance in housing growth unexplained. Some of this unexplained variance may result from assuming the relationships in the model to be constant over space. By minimizing (10), *b* is set to 1.593 degree of latitude which is approximately 110 miles. The improved adjusted R^2 value of 0.7921 for the local model and the reduction in the Akaike Information Criterion (AIC)[5] (Akaike 1973; Sakamoto et al. 1986) from 3017.2 in the global model to 3004.5 in the local model suggest that the local model is better fit than the global model. The analysis of variance between groups (ANOVA) shows that the local model also reduces residual sum of square error from 494.4 in the global model to 489.1, an improvement of 5.3. The F value of 2.87 from the ANOVA, which is greater than the critical value of 2.04, confirms that the local model is a statistically significant improvement of the global model at the level of 1%.

The significance of the spatial variability in the local parameter is estimated by conducting Monte Carlo test. The test is based on the Monte Carlo significance test procedure developed by Hope (1968). The results of the Monte Carlo test on the local estimates in the Table 3 shows that the spatial variation of the variables of distances to Knoxville CBD, Johnson City CBD, and the closest interstate highway are statistically significant at the level of 5%.

Based on the Table 3, a discussion of the instances in which variables are statistically significant at least the level of 5% follows. An increase in growth rate of housing price decreases the housing growth significantly at the 5% level. This reflects the economy of housing

[5] An AIC is computed for each of a number of competing models fitted to a given data set, and the model with the smallest AIC is deemed to be the best fit to the data.

market in action. When there is an increase in housing price housing demand increases, housing supply slowly adjust to meet housing demand, housing price increases, transaction volume increases, and finally demand decreases which slows down housing growth. The global model predicts that an increase of 1% growth of housing price decreases 0.018% housing growth and median estimates of the local model shows slightly higher effect of 0.023%. As expected, population growth is shown to be the most significant single determinant of the housing growth. The global model predicts that an increase of 1% population growth increases housing growth rate by 0.837%. A decrease of distance to Knoxville CBD by 100 mile increases the housing growth by 0.054%.

For the demographic variables, an increase of white ratio of a block group by 1% increases the housing growth by 0.139%. Although non-white ratio increased across the region during the 1990s, the number of houses grew more at the block groups where white ratio is higher. The local parameter estimates shows that although 50% of local marginal effects are within the range between 0.050 of lower quartile and 0.144 of upper quartile, some regions have negative marginal effect. An increase of marriage ratio is shown to increase the housing growth significantly at the level of 1%. This positive effect of marriage ratio on housing growth may be due to the fact that married households are likely to choose to move from renting to the home ownerships. The positive and statistically significant effect of the changes of ratio of college degree on housing growth may reflect the fact that more households with college degrees are likely to have more home ownerships.

The estimates of the housing growth model with neighborhood effects before resolving multicollearity problems of the equation (2) for the global regression and the equation (6) for the local regression are shown in Table 4. The ANOVA shows that the sum of square error is reduced from 488.6 in the global model to 479.3 in the local model, an improvement of 9.3. The F value of the test, 2.55, which is greater than the critical value of 1.70, confirms the fact that the local model is a statistically significant improvement of the global model at the level of 1%. The inclusions of the neighborhood variables for the measure of spatial spillover did not change the significances and signs of the original variables estimated in the equations (1) and (5) except the distance to Knoxville CBD. Among the neighborhood variables, the neighborhood effect of population growth on housing growth is positive and significant at the 1% level. An increase of population growth in the neighbor block groups by 1% increases the housing growth by 0.077% which is about 10% of marginal effect of population growth in their own block groups, 0.827.

The estimates of the spatial spillover after resolving multicollinearity problems using the equation (3) for the global model and the equation (7) for the local model are shown in the Table 5. The ANOVA shows that the sum of square error is reduced from 489.1 in the global model to 480.1 in the local model, an improvement of 9.0. The F-value of the test, 2.96, which is greater than the critical value of 1.88, confirms that the local model is a statistically significant improvement of the global model at the level of 1%. The Monte Carlo significance test shows that spatial variations of the interpolated values for the distances to Johnson City, Jackson, and Knoxville CBDs are statistically significant at the level of 0.1%.

Figure 1 shows the relationship between the housing growth predicted by the model with neighborhood effect after resolving multicollearity and observed housing growth. Out-of-sample prediction revealed that 2,217 of the 4,014 block groups, 55.2%, have predicted housing growth within 50% of the observed housing growth. Figure 2 shows the spatial distribution of observed housing growth. It shows that Memphis, Nashville, and Knoxville are the areas with concentration of high housing growth. It also shows that the housing growth is more intense at the suburban areas rather than centers of the MSAs.

The spatial distributions of local marginal effects of the variables that are statistically significant at least level of 5% (growth of housing price, population growth, marriage ratio, college degree ratio, neighborhood of population growth, and interpolated value of distance to Knoxville) in the model with neighborhood effects after resolving multicollinearity problems are mapped in the Figures 3-8. Although the interpolated value of distance to Knoxville is only variable out of the six significant variables found to be also statistically significant in the Monte Carlo test for spatial heterogeneity, the other five variables are mapped to see if there is any pattern of changes worth to recognize.

The Figure 3 shows that the local marginal effect of growth of housing price on housing growth gradually increases toward the east from the west. This increasing marginal effect toward the east may be due to the surprisingly strong gain of housing price in the Nashville and Knoxville during the 1990s with the greater surprise in Knoxville than Nashville. The greater surprise of the housing price increase is, the greater its impact on slow down of the housing growth is.

Figure 4 shows that the local marginal effect of population growth is relatively higher in the area east of the Middle Tennessee, approximate location of Cumberland Plateau. The Cumberland Plateau is among the fastest growing rural areas in Tennessee. The area has been attracting second-home development for the increasing recreational uses. This may be the reason why marginal effect of population growth on housing growth, which excludes an increase of second-home owners as population growth but counts an increase number of houses for housing growth, is relatively higher than the rest of region.

Figure 5 shows that the local marginal effect of change of marriage ratio gradually increases toward the east. Figure 6 shows that the local marginal effect of change of college degree ratio gradually increases toward the west from the east. Figure 7 shows that the local marginal effect of neighborhood population growth is relatively higher in the East Tennessee including the areas of Knoxville and Johnson City and also the area South of Nashville. Figure 8 shows that the local marginal effect of interpolation value of distances to Knoxville is higher in the East Tennessee relative to the Middle Tennessee. The constant marginal effect throughout the Middle Tennessee is higher than the constant marginal effect throughout the East Tennessee. Surprisingly the local marginal effect is higher in the area around Memphis which is around 400 miles from Knoxville. There is no good explanation for this unexpected high marginal effect of the distance to CBD on housing growth of the area quite far away from the CBD.

5. Summary and Conclusion

This study takes account of spatial spillover and spatial heterogeneity to estimate housing growth at the level of census-block group. A GIS tool of deterministic interpolation technique called Inverse Distance Weighted Averaging is used to create neighborhood variables that capture spatial spillover without multicollearity problems. The GWR approach is adopted to account spatial heterogeneity. The global and local models with and without neighborhood effect and before and after resolving the multicollearity problems are presented. The conclusion of the study is summarized in the following.

The ANOVA confirms that the local model using GWR is a significant improvement of the global model in the estimation of housing growth at the level of census-block group within Tennessee in the period of 1990-2000. An increase in housing price decreases the housing growth across the region while the local marginal effect gradually increases toward the east from the west. The marginal effect of population growth is relatively higher in the approximate area of Cumberland Plateau. While the marginal effect of change of marriage ratio gradually increases toward the east, the marginal effect of change of college degree ratio gradually increases toward the west. The local marginal effect of spatial spillover of population growth in adjacent neighbor is relatively higher in the East Tennessee and also the area South of Nashville. The population growth in the adjacent neighborhood-block group has about 10% of marginal effect of population growth in its own block group. The constant marginal effect of interpolation value of distances to Knoxville throughout the Middle Tennessee is higher than the constant local marginal effect throughout the East Tennessee.

Based on the local model estimates of our study, policy makers could build programs which encourage or discourage housing growth at the local level depending on the various marginal effects of the variables on housing growth. The estimates could be used to predict future housing growth at the local level given projected local demographic characteristics. For example, our study found that local marginal effect of housing price on housing growth is negative across the State and it is greater in the East Tennessee relative to the West Tennessee. Policy makers could anticipate more housing growth in the East Tennessee relative to the West Tennessee with the same amount of lower housing price in projection. If the policy makers intend to encourage more housing growth using a local policy that lowers housing price, they would expect greater responses of housing growth in the East Tennessee than in the West Tennessee. Based on the higher local marginal effect of population growth on housing growth in the Cumberland plateau, local policymakers in the Cumberland plateau would need to reserve a greater budget for infrastructure expansion need for new housings relative to the rest of the regions with the same anticipated increase in population growth.

Based on the results of our study, growth drivers play out in distinctive ways at local level. These distinctively different growth drivers imply that growth of an area has to be managed differently 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 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 next logical step of the analysis would be to apply the GWR in hedonic housing price model. Learning of varying magnitude of the different determinants of housing value at local level would help us understand dynamics of housing market of the area at local level. Understanding the dynamics of housing markets at local level would help us recognizing the structures of submarket of the area. Housing demand at an individual level could be used for a better analysis of more fine scale units if the 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 the GIS database that could be 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.

References

Akaike H. 1973. Information theory and an extension of the maximum likelihood principle. In Petrob B., Csaki, F. (eds) 2nd Symposium on information theory. Budapest, Akademiai Kiado, 267-281.

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

Anselin, L., Dodson, R., Hudak, S. 1993. "Linking GIS and spatial data analysis in practice." *Geographical Systems* 1, 3-23.

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

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

Cao, T., Cory, D. 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.

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

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

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

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression: the analysis of spatially varying relationships. John Wiley & Sons Ltd., West Sussex, England.

Fotheringham, A., Brunsdon, C. 1999. "Local forms of spatial analysis." *Geographical Analysis* 31, 341-358.

Fotheringham, A., Charlton, M., Brundson, C. 1997. "Two techniques for exploring non-stationary in geographical data." *Geographical Systems* 4, 59-82.

Fotheringham, A., Charlton, M., Brundson, C. 1996. "The geography of parameter space: An investigation of spatial non-stationarity." *International Journal of Geographical Information Systems* 4, 59-82.

Fotheringham, A., Rogerson, P. 1993. "GIS and spatial analytical problems." *International Journal of Geographical Information Systems* 7, 3-19.

Gao, X., Asami, Y., Chung, C., 2002. "An Empirical Evaluation of Hedonic Regression Models." Symposium on Geospatial Theory, Processing, and Applications.

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

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

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

Hope, A. 1968. "A simplified Monte Carlo significance test procedure." *Journal of the Royal Statistical Society Series B* 30, 582-598.

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

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

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

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

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

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

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

Melser, D. 2004. "The Hedonic Regression Time–Dummy Method and the Monotonicity Axiom." working paper.

Sakamoto, Y., Ishiguro, M., Kitagawa, G. 1986. Akaike Information Criterion Statistics. Kluwer, Dordrecht.

Song, J., DePinto, J. 1995. "A GIS-based Data Query System." Presented at the International Association for Great Lakes Research (IAGLR) Conference, Windsor, Ontario.

Switzer, P., Kowalik, W., Lyon, R., 1982. "A prior probability method for smoothing discriminant analysis classification maps." *Mathematical Geology*. 14, 433-444.

Tse, R.C. 2002. Estimating Neighborhood Effects in House Prices: Towards a New Hedonic Model Approach. *Urban Studies* 39, 1165-1180.

U.S. Census Bureau http://www.census.gov/

Table 1. Definition of Variables

Variable	Definition
Housing growth	Difference between number of housing units in 2000 and 1990 over number of housing units in 1990
Growth of housing price	Difference between housing price in 2000 and 1990 over housing price in 1990
Growth of population	Difference between population in 2000 and 1990 over population in 1990
Water ratio	Area of water over area of total census-block group
Distance to Knoxville CBD (mile)	Distance between the centroid of each census-block group and Knoxville CBD
Distance to Johnson City CBD (mile)	Distance between the centroid of each census-block group and Johnson City CBD
Distance to Chattanooga CBD (mile)	Distance between the centroid of each census-block group and Chattanooga CBD
Distance to Jackson CBD (mile)	Distance between the centroid of each census-block group and Jackson CBD
Distance to Nashville (mile)	Distance between the centroid of each census-block group and Nashville CBD
Distance to Memphis CBD (mile)	Distance between the centroid of each census-block group and Memphis CBD
Distance to the closest Interstate (mile)	Distance between the centroid of each census-block group and the closest interstate highway
Change of white ratio	Change in ratio of white residents between year 2000 and 1990
Change of marriage ratio	Change in ratio of married household between years 2000 and 1990
Change of college ratio	Change in ratio of college graduate between years 2000 and 1990

Table 2. Descriptive Statistics

	Mean	Std Dev	Min	Max
Housing growth	0.22468	0.77007	-1.0	23.93548
Growth of housing price	0.62146	0.54349	-1.0	16.88491
Growth of population	0.18444	0.81633	-1.0	29.77108
Water ratio	0.01628	0.05385	0	0.79791
Distance to Knoxville CBD (mile)	163.7	112.7	0	358.1
Distance to Johnson City CBD (mile)	236.1	133.1	0	449.1
Distance to Chattanooga CBD (mile)	142.5	78.4	0	279.9
Distance to Jackson CBD (mile)	162.9	100.6	0	401.9
Distance to Nashville (mile)	121.5	72.8	0	282.8
Distance to Memphis CBD (mile)	207.7	133.3	0	474.9
Distance to the closest Interstate (mile)	6.1	7.9	0	52.4
Change of white ratio	-0.04982	0.11445	-0.73497	0.76383
Change of marriage ratio	-0.03572	0.07258	-1.0	0.54545
Change of college ratio	0.02046	0.05342	-0.24054	1.0
Neighborhood of growth of housing price	0.61162	0.29156	-0.51755	4.77005
Neighborhood of growth of population	0.18802	0.38975	-0.49648	6.04490
Neighborhood of water ratio	0.01505	0.03160	0	0.38874
Neighborhood of distance to Knoxville CBD (mile)	163.7	112.6	0.5	354.4
Neighborhood of distance to Johnson City CBD (mile)	236.1	133.1	0.9	445.4
Neighborhood of distance to Chattanooga CBD (mile)	142.5	78.3	0.9	275.9
Neighborhood of distance to Jackson CBD (mile)	162.9	100.4	0.8	399.2
Neighborhood of distance to Nashville CBD (mile)	121.5	72.7	0.5	280.5
Neighborhood of Memphis CBD (mile)	207.7	133.1	0.6	471.9
Neighborhood of distance to the closest interstate (mile)	6.0	7.6	0.1	51.7
Neighborhood of change of white ratio	-0.04929	0.08744	-0.64342	0.19665
Neighborhood of change of marriage ratio	-0.03565	0.03833	-0.23149	0.07939
Neighborhood of change of college ratio	-0.00890	0.01664	-0.30359	0.07369
Interpolation value of distance to Knoxville CBD (mile)	170.3	111.7	0	358.1
Interpolation value of distance to Johnson City CBD (mile)	241.8	134.1	0	449.1
Interpolation value of distance to Chattanooga CBD (mile)	146.4	78.6	0	279.9
Interpolation value of distance to Jackson CBD (mile)	158.7	101.0	0	401.9

Interpolation value of distance to Nashville CBD (mile)	120.9	74.5	0	282.8
Interpolation value of distance to Memphis CBD (mile)	202.0	134.4	0	474.9
Interpolation value of distance to the closest Interstate (mile)	5.7	6.9	0	30.8

Table 3. Parameter Estimates without Neighborhood Effects

	Global Model	Local Model					
	Estimates	Minimum	Lower Quartile	Median	Upper Quartile	Maximum	P-value Monte Carlo test
Intercept	0.12255 (0.20208)	-0.63933	0.05575	0.07189	0.11355	0.28653	0.08
Growth of housing price	-0.01839* (0.01081)	-0.04282	-0.03821	-0.02314	-0.00595	-0.00026	0.31
Growth of population	0.83729** (0.00687)	0.76387	0.82842	0.85671	0.87273	0.88501	0.74
Water ratio	0.05162 (0.10477)	-0.64313	-0.09036	0.14667	0.15991	0.17251	0.14
Distance to Knoxville CBD	-0.00054* (0.00026)	-0.00556	-0.00061	-0.00053	-0.00050	0.00098	0.00***
Distance to Johnson City CBD	0.00007 (0.00051)	-0.00141	0.00013	0.00020	0.00025	0.00567	0.00***
Distance to Chattanooga CBD	0.00023 (0.00019)	-0.00014	0.00016	0.00020	0.00025	0.00039	0.87
Distance to Jackson CBD	-0.00038 (0.00033)	-0.00290	-0.00048	-0.00042	-0.00032	0.00094	0.11
Distance to Nashville	0.00030 (0.00018)	0.00031	0.00035	0.00037	0.00039	0.00110	0.76
Distance to Memphis CBD	0.00016 (0.00047)	-0.00104	0.00012	0.00034	0.00039	0.00189	0.12
Distance to the closest Interstate	0.00130 (0.00079)	0.00038	0.00071	0.00125	0.00231	0.00263	0.04*
Change of white ratio	0.13853** (0.05610)	-0.09232	0.05003	0.10712	0.14415	0.17517	0.28
Change of marriage ratio	0.25020** (0.08424)	0.00374	0.19621	0.31599	0.32358	0.45856	0.37
Change of college ratio	0.48152**	0.33834	0.37714	0.41199	0.54758	0.68649	0.37

(0.10844)	
Number of observations: 4014	Number of observations: 4014
Adjusted R-square: 0.7907	Adjusted R-square: 0.7921
Akaike Information Criterion: 3017.2	Akaike Information Criterion: 3004.5
Residual sum of squares: 494.4	Residual sum of squares: 489.1
	Bandwidth: 110 miles

Numbers in parentheses are standard error. *** indicates statistical significance at the 0.1% level; ** indicates statistical significance at the 1% level; * indicates statistical significance at the 5% level.

Table 4. Parameter Estimates with Neighborhood Effects before Resolving Multicollearity Problems

	Global Model	Local Model					
	Estimates	Minimum	Lower Quartile	Median	Upper Quartile	Maximum	P-value Monte Carlo test
Intercept	-0.03721 (0.29405)	-0.95377	-0.23791	-0.20311	-0.11846	0.03122	0.42
Growth of housing price	-0.02534** (0.01090)	-0.06256	-0.05275	-0.03033	-0.00732	-0.00085	0.26
Growth of population	0.82687** (0.00710)	0.71151	0.82102	0.83122	0.85612	0.86631	0.61
Water ratio	0.16219 (0.12950)	-0.67038	-0.06214	0.28632	0.32504	0.36540	0.19
Distance to Knoxville CBD	0.00998 (0.00814)	0.00663	0.00763	0.00820	0.00907	0.01931	0.70
Distance to Johnson City CBD	-0.00196 (0.00465)	-0.00122	-0.00052	-0.00026	0.00029	0.00409	0.92
Distance to Chattanooga CBD	-0.01223 (0.00835)	-0.03345	-0.01470	-0.00995	-0.00945	-0.00786	0.16
Distance to Jackson CBD	-0.01164 (0.01248)	-0.07492	-0.02181	-0.00942	-0.00754	-0.00617	0.01**
Distance to Nashville CBD	0.00709 (0.00728)	0.00501	0.00548	0.00872	0.01530	0.04607	0.00***
Distance to Memphis CBD	0.00972 (0.00874)	0.00501	0.00633	0.00912	0.01577	0.04052	0.09
Distance to the closest Interstate	-0.00428 (0.00512)	-0.01400	-0.00907	-0.00517	-0.00071	0.00335	0.05*
Change of white ratio	0.09658 (0.07359)	-0.21781	-0.04713	0.04125	0.12974	0.17297	0.32
Change of marriage ratio	0.21067** (0.08424)	-0.06969	0.14267	0.29251	0.29721	0.40259	0.20
Change of college ratio	0.43548**	0.34653	0.36700	0.38425	0.48435	0.59235	0.56

	(0.10863)						
Neighborhood of growth of housing price	0.03414 (0.02460)	-0.03811	-0.01835	-0.01281	0.12343	0.22908	0.29
Neighborhood of growth of population	0.07659** (0.01509)	0.06455	0.07263	0.08462	0.08683	0.11665	0.83
Neighborhood of water ratio	-0.28944 (0.22242)	-0.56784	-0.48788	-0.25939	-0.04780	0.11704	0.38
Neighborhood of distance to Knoxville CBD	-0.01054 (0.00817)	-0.02659	-0.00963	-0.00874	-0.00820	-0.00729	0.44
Neighborhood of distance to Johnson City CBD	0.00225 (0.00467)	-0.00428	0.00081	0.00114	0.00161	0.00588	0.82
Neighborhood of distance to Chattanooga CBD	0.012495 (0.00836)	0.00812	0.00972	0.01029	0.01489	0.03425	0.16
Neighborhood of distance to Jackson CBD	0.01153 (0.01252)	0.00593	0.00727	0.00932	0.02187	0.07358	0.01**
Neighborhood of distance to Nashville CBD	-0.00685 (0.00729)	-0.04497	-0.01505	-0.00845	-0.00518	-0.00470	0.00***
Neighborhood of Memphis CBD	-0.00954 (0.00871)	-0.04043	-0.01508	-0.00867	-0.00562	-0.00442	0.08
Neighborhood of distance to the closest interstate	0.00541 (0.00536)	-0.00465	0.00190	0.00691	0.01008	0.01535	0.07
Neighborhood of change of white ratio	0.03406 (0.15848)	-0.32548	-0.26227	-0.16675	0.03019	0.03964	0.54
Neighborhood of change of marriage ratio	-0.01823 (0.18643)	-0.38972	-0.34290	-0.05168	0.29039	0.46994	0.15
Neighborhood of change of college ratio	0.09302 (0.80206)	-0.33829	-0.15158	1.28641	1.80812	2.08321	0.15
Number of observations: 4014 Adjusted R-square: 0.7924 Akaike Information Criterion: 2997.9 Residual sum of squares: 488.6		Number of observations: 4014 Adjusted R-square: 0.7948 Akaike Information Criterion: 2982.4 Residual sum of squares: 479.3 Bandwidth: 110 miles					

Numbers in parentheses are standard error. *** indicates statistical significance at the 0.1% level; ** indicates statistical significance at the 1% level; * indicates statistical significance at the 5% level.

Table 5. Parameter Estimates with Neighborhood Effects after Resolving Multicollearity Problems

	Global Model	Local Model					
	Estimates	Minimum	Lower Quartile	Median	Upper Quartile	Maximum	P-value Monte Carlo test
Intercept	0.00802 (0.30516)	-0.71887	-0.20705	-0.15087	-0.08705	0.13563	0.62
Growth of housing price	-0.02561** (0.01089)	-0.06248	-0.05322	-0.03073	-0.00776	-0.00129	0.26
Growth of population	0.82681** (0.00709)	0.70898	0.82084	0.83119	0.85628	0.86640	0.59
Water ratio	0.15316 (0.12862)	-0.70851	-0.08771	0.27792	0.32573	0.36679	0.14
Change of white ratio	0.09559 (0.07354)	-0.21472	-0.04767	0.04098	0.12680	0.16747	0.35
Change of marriage ratio	0.21143** (0.08416)	-0.05895	0.14811	0.29111	0.29537	0.39807	0.25
Change of college ratio	0.43011** (0.10844)	0.33935	0.36199	0.37891	0.47981	0.58597	0.56
Neighborhood of growth of housing price	0.03483 (0.02456)	-0.03699	-0.01706	-0.01244	0.12400	0.22916	0.29
Neighborhood of growth of population	0.07667** (0.01508)	0.06453	0.07266	0.08442	0.08662	0.11708	0.83
Neighborhood of change of white ratio	0.03730 (0.15834)	-0.34474	-0.27256	-0.16581	0.03743	0.04676	0.49
Neighborhood of change of marriage ratio	-0.01915 (0.18623)	-0.39056	-0.33460	-0.05123	0.28767	0.46670	0.14
Neighborhood of change of college ratio	0.08645 (0.80157)	-0.36415	-0.17246	1.29577	1.89361	2.19939	0.10
Neighborhood of water ratio	-0.29976 (0.22160)	-0.57764	-0.50156	-0.26017	-0.04301	0.13679	0.32

Interpolation value of distance to Knoxville CBD	-0.00052* (0.00027)	-0.00545	-0.00062	-0.00054	-0.00052	0.00050	0.00***
Interpolation value of distance to Johnson City CBD	0.00018 (0.0007)	-0.00074	0.00042	0.00051	0.00060	0.00571	0.00***
Interpolation value of distance to Chattanooga CBD	0.00023 (0.00019)	0.00005	0.00020	0.00024	0.00030	0.00047	0.87
Interpolation value of distance to Jackson CBD	-0.00018 (0.00034)	-0.00304	-0.00035	-0.00022	-0.00014	0.00132	0.00***
Interpolation value of distance to Nashville CBD	0.00028 (0.00018)	0.00028	0.00032	0.00035	0.00036	0.00128	0.69
Interpolation value of distance to Memphis CBD	0.00013 (0.00066)	-0.00142	0.00026	0.00056	0.00065	0.00181	0.11
Interpolation value of distance to the closest Interstate	0.00088 (0.00084)	-0.00111	0.00070	0.00082	0.00142	0.00175	0.61
Number of observations: 4014 Adjusted R-square: 0.7926 Akaike Information Criterion: 2987.7 Residual sum of squares: 489.1		Number of observations: 4014 Adjusted R-square: 0.7951 Akaike Information Criterion: 2964.4 Residual sum of squares: 480.1 Bandwidth: 110 miles					

Numbers in parentheses are standard error. *** indicates statistical significance at the 0.1% level; ** indicates statistical significance at the 1% level; * indicates statistical significance at the 5% level

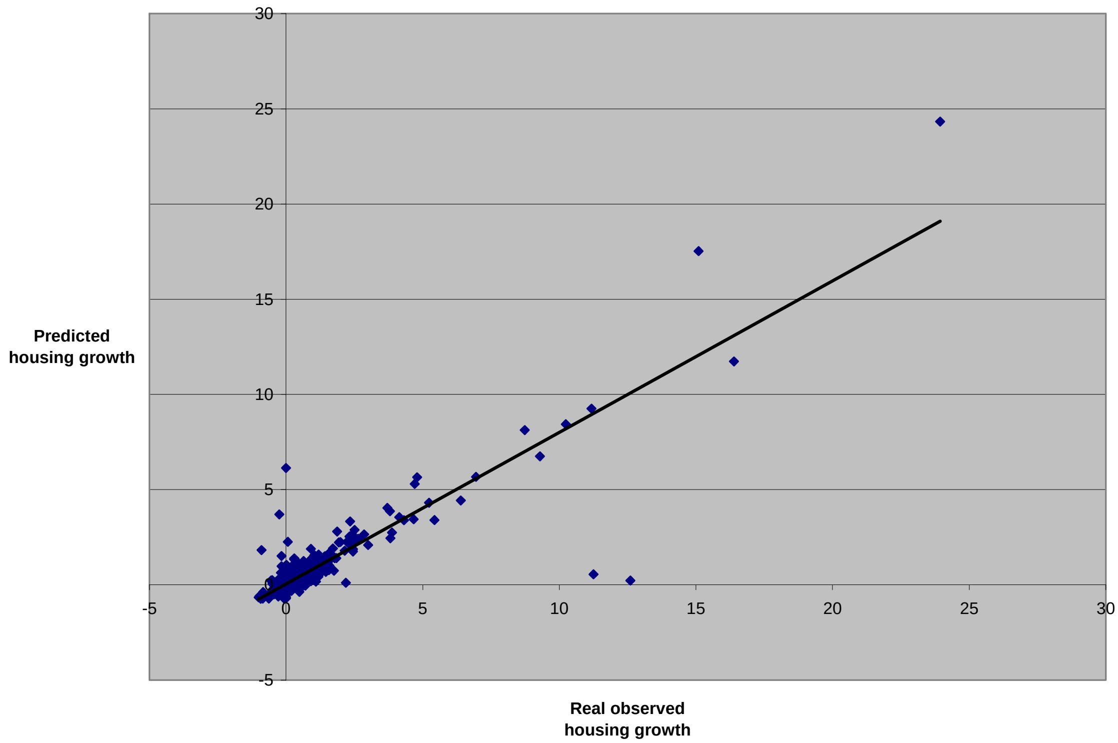

Figure 1. Observed vs. Predicted Housing Growth using the Model with Neighborhood Effects after Resolving Multicollearity Problems

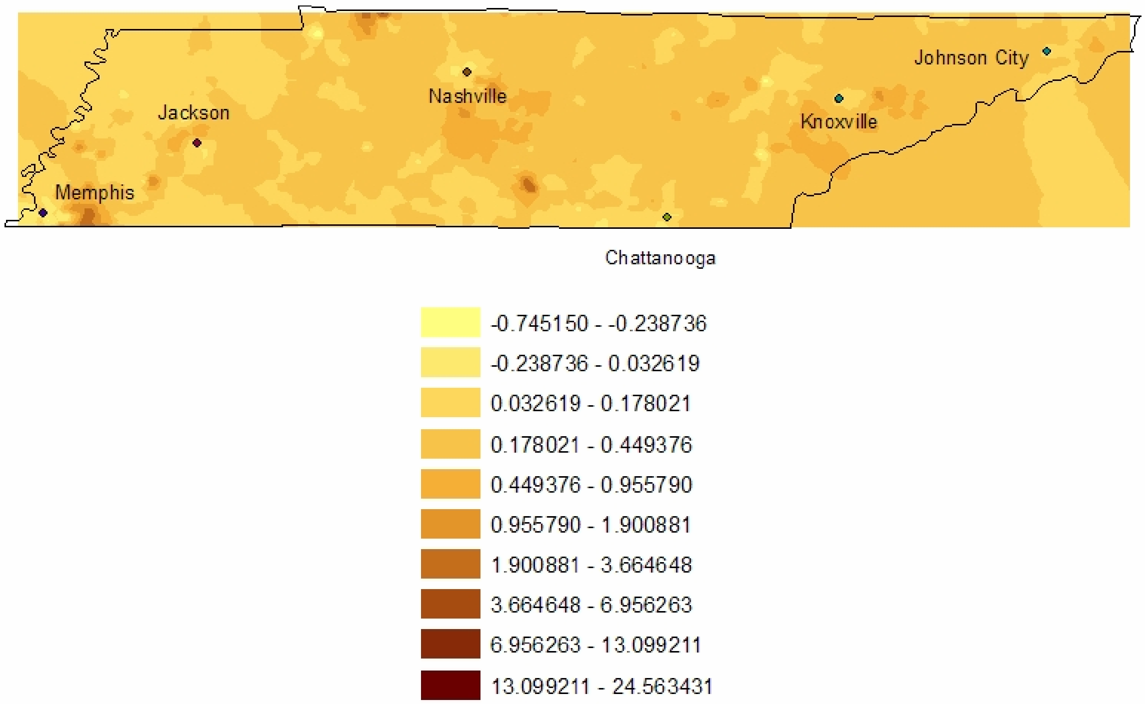

Figure 2. Housing Growth

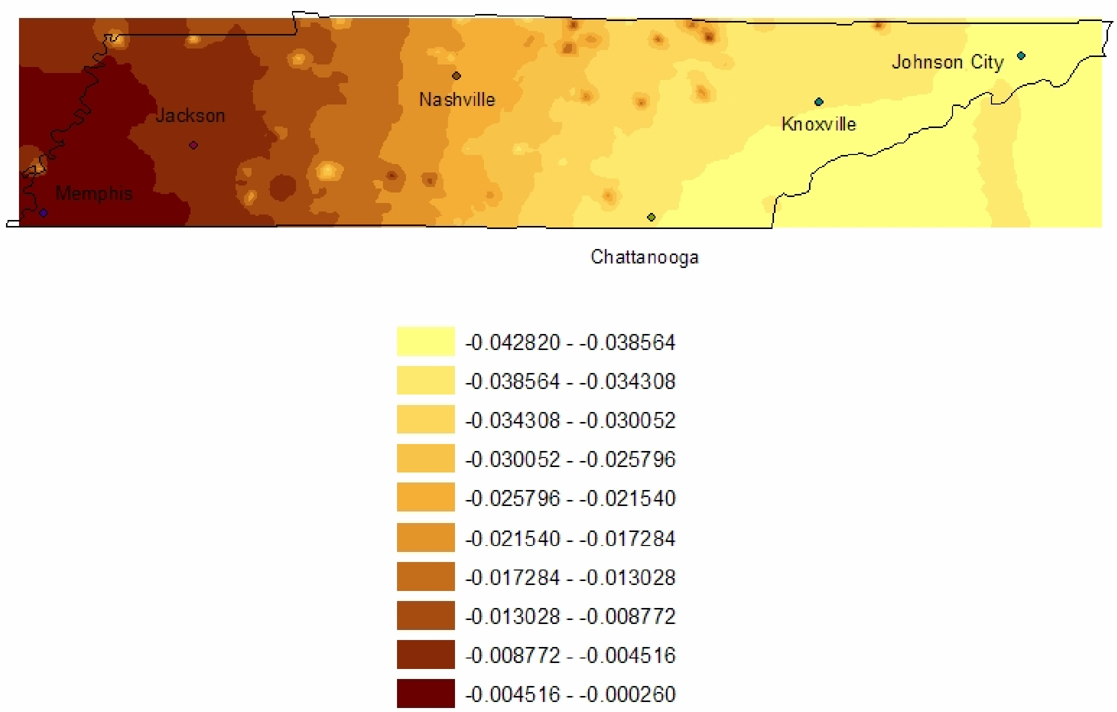

Figure 3. Local Marginal Effect of Housing Price on Housing Growth

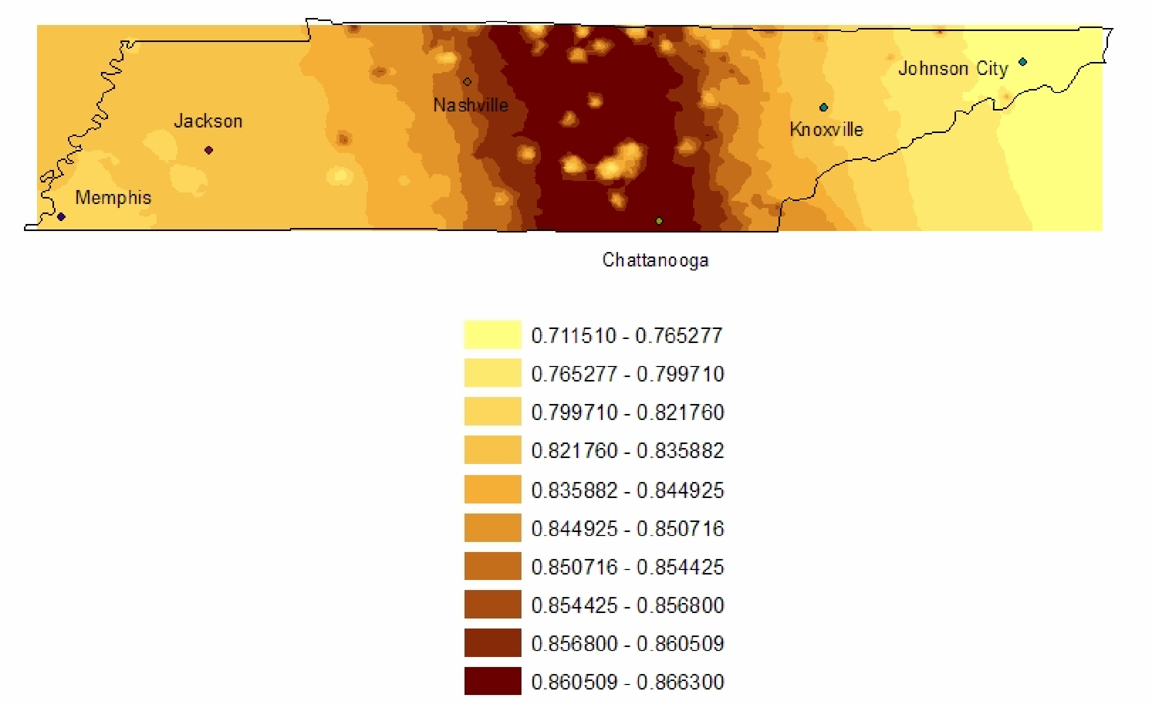

Figure 4. Local Marginal Effect of Population on Housing Growth

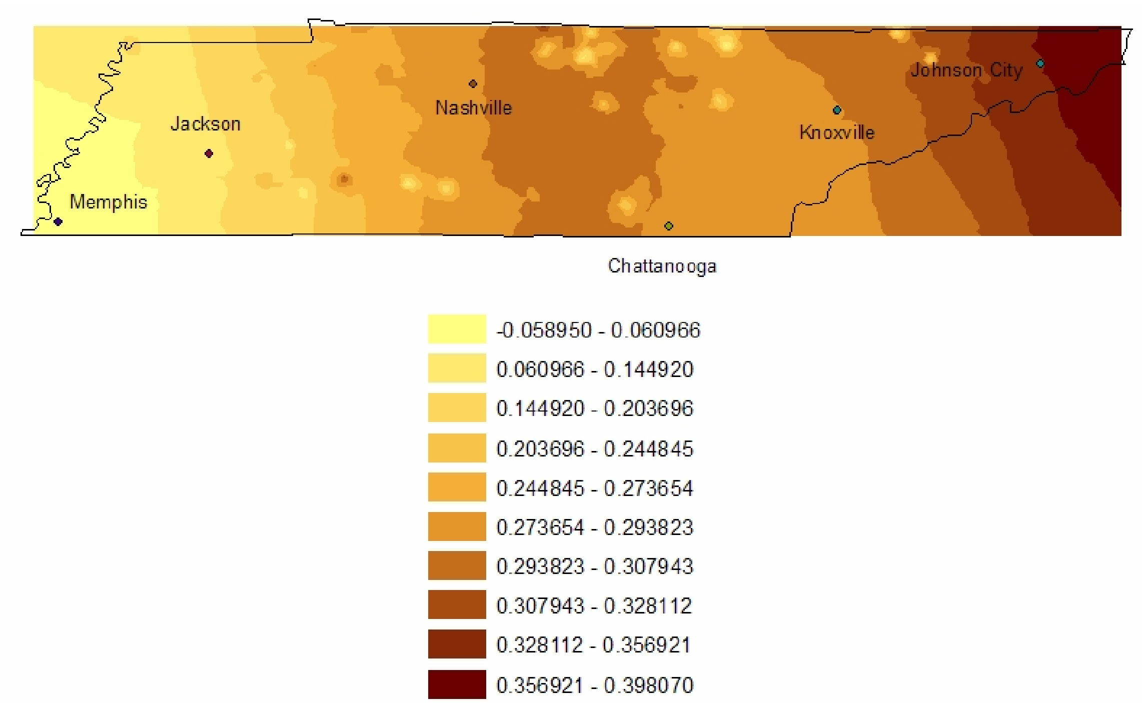

Figure 5. Local Marginal Effect of Marriage Ratio on Housing Growth

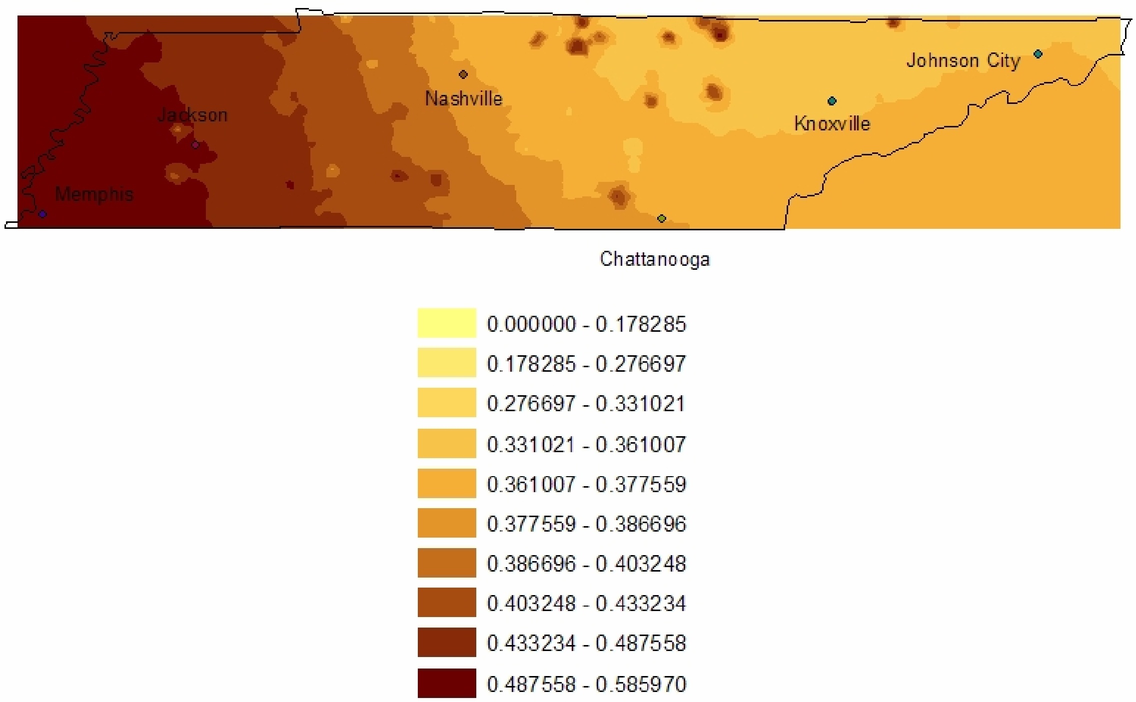

Figure 6. Local Marginal Effect of College Ratio on Housing Growth

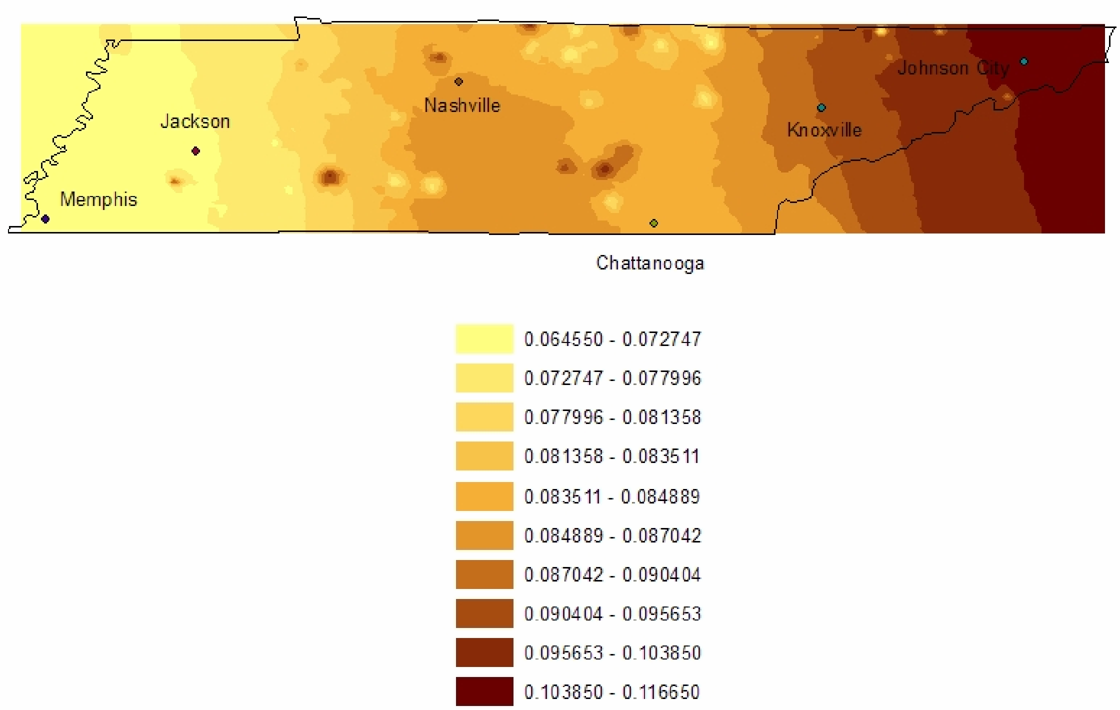

Figure 7. Local Marginal Effect of Neighborhood Population on Housing Growth

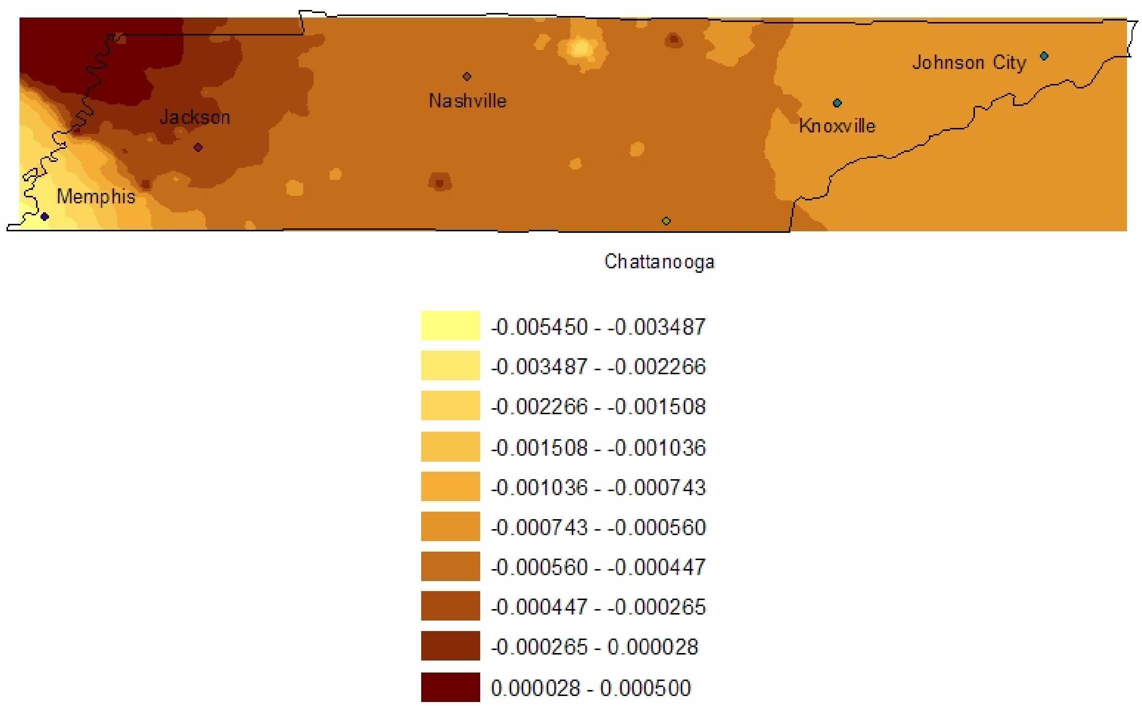

Figure 8. Local Marginal Effect of Interpolation Value of Distance to Knoxville on Housing Growth

Appendix 1

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 is used because it is the smallest level of 1990 geography at which the full set of US 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% of the US in 1990 and 2000.

The 1990 to 2000 Block relations were determined from Tiger/Line 2000, Type 1 and Type 3 records. 85% of the Blocks had a 1:1 relationship, 10% had a 2:1, and 5% had a greater than 2:1. 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.).