Journal of Agricultural and Applied Economics, 38,2(August 2006):299–316

Two Dimensions of the Spatial Distribution of Housing: Dependency and Heterogeneity across Tennessee's Six Metropolitan Statistical Areas

Seong-Hoon Cho, Christopher D. Clark, and William M. Park

A two-stage multinomial logit selection model is used to model the relationship between demographic characteristics and housing density across Tennessee's six metropolitan statistical areas. The study finds that there is both spatial correlation and heterogeneity in the spatial distribution of housing both within and across the six areas. For example, Memphis, the most densely populated area, has the least amount of spatial correlation among housing density at the neighborhood level, while Johnson City, which has the lowest overall housing density, has the highest degree of spatial correlation.

Key Words: community, housing density, spatial dependency, spatial heterogeneity

JEL Classifications: C31, R21

Urban Americans are increasingly living in areas with lower housing densities. One-third more land per person was consumed by urban use in the 1990s than in the 1970s (Daniels and Bowers). As housing densities decline, the pace of urban/suburban development increases, producing what is commonly referred to as sprawl. From 1992 to 1997, an average of 890,000 hectares of land was developed in the United States each year, which is more than 1.5 times the average for the previous 10-year period (U.S. Department of Agriculture [USDA]). What is truly remarkable is that the growth in urbanized land is occurring even in metropolitan areas that are losing population. For example, the Pittsburgh, PA, metropolitan area lost 8% of its population between 1982 and 1997, yet its urbanized land area grew by close to 43% during this period (Fulton et al.). A variety of explanations have been suggested for this trend, including the increasing number of new households due to aging of the population, increased divorce rates, and/or higher incomes (Katz).

While the trend toward lower housing density may offer affordable private spaces set back from streets and commercial areas, it also takes a toll on open space and environmental amenities. A recent report by the American Farmland Trust revealed that every year in the United States, over 400,000 hectares of productive farmland and open space are developed for other uses (American Farmland Trust). Similarly, the Sierra Club estimates that approximately 44,500 hectares of wetlands are lost each year to sprawl (Sierra Club). As the pace of sprawl increases, so do concerns over the social and financial burdens associated with an expanding infrastructure.

Seong-Hoon Cho, Christopher D. Clark, and William M. Park are, respectively, assistant professor, assistant professor, and professor at Department of Agricultural Economics, the University of Tennessee.

For example, studies suggest that gasoline consumption could be 20% to 30% lower in cities like Houston and Phoenix if their urban structure more closely resembled that of Boston or Washington (Newman and Kenworthy 1989, 1999). As a result, many local governments have turned to impact fees to address the higher infrastructure costs, such as water and sewer hookups, associated with outlying developments (Snyder and Bird). Another response to this trend has been the promotion of "smart growth" policies emphasizing the preservation of open space through higher housing densities and more clustered housing (e.g., U.S. EPA). These policies rely not only on modifying the land use pattern in new developments but also on increasing the housing supply on land served by existing infrastructure, such as existing neighborhoods (Smart Growth Online).

However, to develop and implement policies that effectively address this trend, it is desirable, if not necessary, to have a better understanding of two different aspects of the spatial pattern of development—dependency and heterogeneity. Spatial dependency, also known as spatial autocorrelation, is inherent in all geographic data, as expressed by Tobler's First Law of Geography, which states that "everything is related to everything else, but near things are more related than distant things" (Tobler). Thus, examining spatial dependency allows us to understand whether and how development in one area is correlated with development in other areas. Spatial heterogeneity is, in general terms, variation in some condition or measure from one geographic area to another. A more specific definition relating to residential location distinguishes between homogeneity and heterogeneity within and among neighborhoods:

> A *homogeneous neighborhood* is defined as a spatially contiguous region of the city with sufficient population to be considered a neighborhood in which all resident households and housing units have the same characteristics. A *heterogeneous neighborhood* is one in which such characteristics vary in one or more dimensions . . . Using the above definitions, *heterogeneity among neighborhoods* describes a metropolitan area with many neighborhoods that may be relatively homogeneous internally but that differ substantially among themselves (i.e., segregated neighborhoods). *Homogeneity among neighborhoods* includes neighborhoods that may be quite heterogeneous internally but are alike in their heterogeneity (i.e., integrated or mixed neighborhoods). (Vandell, p. 105).

For this paper, increasing spatial heterogeneity implies increasing divergence among different neighborhoods.

Thus, this paper explores ways to analyze or characterize the spatial distribution of housing and the relationship of this distribution to demographic characteristics in terms of both spatial dependency and heterogeneity. The development of tests for spatial dependence in linear regression models, as well as efficient and consistent estimators for these types of models, has been an important part of the spatial econometric literature for the last few decades (e.g., Anselin; Can 1990, 1992; Cliff and Ord; Dubin 1992, 1998; LeSage; Leung, Mei, and Zhang; McMillen 1992, 2003; Tse). Various localized modeling techniques have been proposed to capture spatial heterogeneity (Casetti; Fotheringham and Brunsdon; Getis and Ord). Some models incorporate both spatial dependency and spatial heterogeneity (Anselin; Can 1992).

In this study, we posit that housing location decisions are not conducted in isolation but start with the selection of a community, based in part on household demographic characteristics. To model this choice, we examine the effects of demographic characteristics on choice of metropolitan statistical area (MSA).[1]

---

[1] The general concept of a metropolitan statistical area (MSA) is one of a large population nucleus, together with adjacent communities that have a high degree of economic and social integration with that nucleus. Each MSA must contain either a place with a minimum population of 50,000 or a Census Bureau–defined urbanized area and a total MSA population of at least 100,000. An MSA may also include one or more outlying counties that have close economic and social relationships with the central county. An outlying county must have a specified level of commuting to the central counties and also must meet certain standards regarding metropolitan character, such as population density, urban population, and population growth.

Once an MSA is selected, the selection of a specific neighborhood is dependent, in part, upon the housing density of that and surrounding neighborhoods. Thus, given the choice of MSA, the relationship between demographics and density is examined at the census-block group level. The spatial dependencies of these relationships are examined both within and across MSAs. Finally, more general measures of the degree of residential clustering within each MSA are provided.

The remainder of this paper is organized as follows. In the next section, we describe the empirical model used to examine the spatial dependency and heterogeneity of housing density. The subsequent section presents a description of the study area and data. The estimation results from the empirical models follow, and we close with a discussion of the results and their possible interpretation.

Empirical Model

Since households make the broader choice of community before they make the more site-specific choice of neighborhood, the housing density of a neighborhood needs to be modeled within the context of the choice of community. In this study, MSAs represent the broader communities, while census-block groups (CBGs) represent individual neighborhoods. The MSAs serve well as communities because households within an MSA have a high degree of economic and social integration. The MSAs also cover much broader areas than CBGs. As for neighborhoods, Meen and Andrew provide two conditions for defining a group of neighbors as a neighborhood. First, income and other variables must either be growing at the same rate in each location or exhibit a common stochastic trend. Second, the structure of the housing market must be the same over the space occupied by the group of neighbors. Previous studies have shown that CBG data meet these two criteria and that specifying neighborhood variables and aggregating housing data at the CBG level can lead to robust hedonic price estimations (Cao and Cory; Geoghegan, Waigner, and Nancy; Goodman).

Because the pattern, size, and shape of CBGs vary within and between geographic areas,[2] the number of residences within a CBG is not a sufficient indicator of density. Instead, we calculate housing density for each CBG by dividing the number of residences by the number of acres within the CBG. While housing density for a CBG is a continuous variable, the choice of an MSA is a discrete choice. Thus, the housing density for a CBG is estimated within the constraint of the household's choice of an MSA using a two-stage multinomial logit selection model (Greene 1995, pp. 656–661).

In the first stage, the probability of a household's choice of an MSA is estimated as a function of the household's characteristics. For this study, the household's choice is assumed to be limited to Tennessee's six MSAs—Chattanooga, Jackson, Johnson City, Knoxville, Memphis, and Nashville. The household's choice among the six communities is modeled in a multinomial logit framework. In the second stage, we estimate the housing density per acre for a CBG, conditional upon the choice of an MSA, using a sample selection model to correct for the selection bias that would occur from the estimation, due to systematic differences in the characteristics of the MSAs.

Because individual data on the more than one million households in these six MSAs are not available, the model is framed so as to use the attributes of the 2,783 CBGs that are contained within the MSAs. In fact, the CBG is the smallest level of geography in which detailed household characteristics are publicly available. Following Lee, the model for the likelihood of a household choosing to locate

[2] "Although most people intuitively think of census-block groups as being rectangular or square, of about the same size, and occurring at regular intervals, as in many large U.S. cities, census block configurations are actually quite different. The pattern, size, and shape of census blocks vary within and between areas. Factors that influence the overall configuration of census blocks 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).

Table 1. Definition of Variables

Variable	Definition
Housing density	Number of housing units per acre
Income	Median family income in 1999 (dollars)
Born in Northeast	Proportion of CBG residents born in the Northeast
Born in Midwest	Proportion of CBG residents born in the Midwest
Born in South	Proportion of CBG residents born in the South
Born in foreign country	Proportion of CBG residents born in foreign country
Female householders	Ratio of female householders in nonfamily[1] households to total households
Bachelor's degree	Ratio of residents 25 years or older who have received a bachelor's degree to the total CBG population 25 years or older
Master's degree	Ratio of residents 25 years or older who have received a master's degree to the total CBG population 25 years or older
Professional degree	Ratio of residents 25 years or older who have received a professional degree to the total CBG population 25 years or older
Doctorate degree	Ratio of residents 25 years or older who have received a doctorate degree to the total CBG population 25 years or older
Senior citizens	Proportion of CBG residents 65 years or older
5-year tenure	Proportion of CBG residents who have been residents of same house for 5 years or more
Number of kids	Ratio of the number of children in family households to total households

[1] Nonfamily as defined by the U.S. Census Bureau is a householder living alone or with nonrelatives only.

in one of the six MSAs is estimated using the following form of the multinomial logit model,

$$(1)\qquad \text{Prob}(J = j) = \frac{\exp(Z' \cdot \gamma_j)}{\sum_{i=1}^{J} \exp(Z' \cdot \gamma_i)},$$

where j = 1, 2, 3, 4, 5, 6 represents the MSAs of Jackson, Johnson City, Memphis, Chattanooga, Knoxville, and Nashville, respectively. We assume that the characteristics of households at a CBG, Z, are represented by the median, mean, and ratio that measure central tendency.

To avoid indeterminacy, the parameter vector of the Jackson MSA, γ_i, is normalized to zero. This normalization renders the estimated parameters (γ_j) uninterpretable. We can however draw inferences from the computed "marginal effects" of the elements of Z relative to sample averages. The marginal effects in the model are partial derivatives of the probabilities with respect to the determi nants:

$$(2)\qquad m_j = \frac{\partial P_j}{\partial Z} = P_j\left(\gamma_j - \sum_{j=1}^{J} P_j\gamma_j\right) = P_j(\gamma_j - \bar{\gamma}).$$

The statistical significances of these effects are estimated by the asymptotic covariance matrix of m_j (Greene 1997, p. 916–17). While the parameter vector γ_i is normalized to zero, the vector of marginal effects δ_j is constrained to sum to zero. This normalization means that any one marginal effect δ_j can be interpreted as the net effect of an increase in the value of the corresponding determinant Z on the decision to live in a particular MSA. The estimated marginal effects and the statistical significance of these effects are shown in Table 3 (see Table 1 for definitions of all the variables used in the analysis and Table 2 for the mean values and standard deviations of these variables, individually for each MSA and aggregated across all six MSAs).

The multinomial logit model of Equation (1) is estimated by maximum likelihood, retaining the coefficients, estimated asymptotic covariance matrix, and the full set of predicted probabilities to compute lambda, λ_j, to be used for the second stage estimation of the sample selection model,

$$(3) \quad \hat{\lambda}_j = \phi[\Phi^{-1}(\hat{P}_j)]/\hat{P}_j$$

In the second stage, the housing density per acre for a CBG is estimated using the sample selection model (Lee), which is expressed as

$$(4) \quad y_{ij} = \beta' Z - \theta_j \lambda_j + \eta_j,$$

where y_{ij} is the housing density of CBG i conditional upon the choice of community j in the first stage, Z is a vector of household characteristics influencing housing density, and η_j is a residual capturing errors. We call Equation (4) the "global model" as opposed to the "local model," which is derived below.

When we estimate Equation (4), we forego the implicit assumption that relationships between variables measured at different locations are constant over the MSA. Otherwise, the existence of structural variations between the demographic characteristics and housing density within the MSAs would represent a misspecification of the data. Thus, a geographically weighted regression (GWR) model (Fotheringham, Brunsdon, and Charlton) is adopted to identify spatial variations in relationships at the neighborhood level under the sample selection model. As a result, we estimate the following local housing density equations for the six MSAs of Tennessee (note that the notation j denoting a particular MSA is omitted for simplicity):

$$(5) \quad y_i = \sum_k [\beta_k(u_i, v_i) Z_{ki}] - \beta_\lambda(u_i, v_i)\lambda_i + \eta_i,$$

where (u_i, v_i) denotes the coordinates of the centroid for the ith CBG and $\beta_k(u_i, v_i)$ is a parameter estimate for the ith CBG, derived from the continuous function $\beta_k(u, v)$. Thus, GWR allows for a continuous surface of parameter values, with measurements of the surface taken at certain points to denote the spatial variability of the surface (Fotheringham, Brunsdon, and Charlton).

In this local housing density model, an observation is weighted in accordance with its proximity to point i in order to account for the fact that an observation near point i has 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,

$$(6) \quad \hat{\beta}(u_i, v_i) = [Z_k^T \mathbf{W}(u_i, v_i) Z_k]^{-1} Z_k^T \mathbf{W}(u_i, v_i) y,$$

where $\mathbf{W}(u_i, v_i)$ is an $n \times n$ matrix whose diagonal elements w_{il} denote the geographical weighting of each of the n variables for the ith CBG and the centroid of the lth CBG, and the off-diagonal elements are zero. Z_k is a vector of explanatory variables. The diagonal elements of the weight matrix, w_{il}, are equal to

$$(7) \quad w_{il} = \exp\left[-\frac{1}{2(d_{il}/b)^2}\right]$$

where d_{il} is the Euclidean distance[3] between the centroids of CBGs i and l, and b is a smoothing parameter or bandwidth. The bandwidth is chosen by minimizing the function

$$(8) \quad \Delta(b) = \sum_{i=1}^{n} [y_i - \hat{y}_i(b)]^2,$$

where $\hat{y}_i(b)$ is the fitted value of y_i, with the observation at location i omitted from the fitting process. This process is known as the cross-validation procedure (Fotheringham, Brunsdon, and Charlton). As a result, the weight matrix is such that if i and l are the same CBG, then $w_{il} = 1$, and the weighting of the observations from other CBG will decrease pursuant to a Gaussian curve as the distance between the centroids of i and l increases.

In addition, we calculated a number of different univariate statistics to provide a more general view of how the spatial pattern of housing density varies from one MSA to another. These include both spatial and nonspatial statistics. The nonspatial statistics include the typical frequency statistics used to summarize a distribution (i.e., mean, median, and quartiles) and the Gini index (G), which is borrowed from the literature on income in-

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

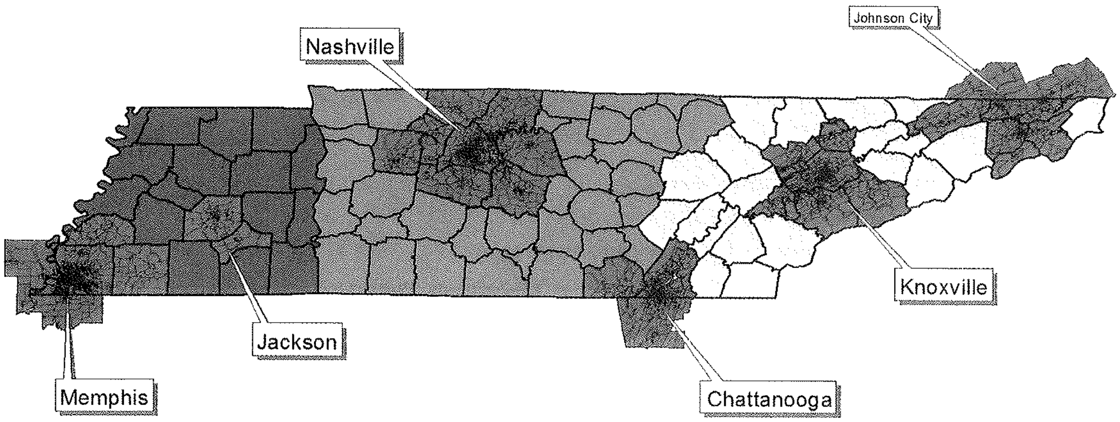

Figure 1. Map of the six MSAs of the Three Grand Divisions of Tennessee at the County and CBG Level

equality. The value of the Gini index is equal to the mean of the difference between every possible pair of block groups divided by the mean μ,

$$G = \frac{\sum_{i=1}^{n}\sum_{l=1}^{n}|y_i - y_l|}{2n^2\mu}, \tag{9}$$

where y_i and y_l are the aggregate housing density per acre for census-block groups i and l and n is the number of observations (Damgaard and Weiner; Dixon et al.). Higher values of the Gini index imply a more unequal distribution of housing density across an MSA.

Finally, Moran's Index (Moran) is used to measure the extent of residential clustering within each MSA. The formula to calculate Moran's index (I) is as follows:

$$I = \frac{N\sum_i\sum_l \mathbf{W}_{il}(y_i - \bar{y})(y_l - \bar{y})}{\left(\sum_i\sum_l \mathbf{W}_{il}\right)\sum_i (y_i - \bar{y})^2}, \tag{10}$$

where N is the number of CBGs, y_i is the aggregate housing density per acre for CBG i, $\bar{y}$ is the mean of the housing density across all CBGs within the MSA, and $\mathbf{W}_{ij}$ is the binary weight matrix of the general cross-product statistic, such that $\mathbf{W}_{ij} = 1$ if CBGs i and j are adjacent and zero for all CBGs that are not adjacent. The I value is similar to a correlation coefficient, varying between -1 and 1. When spatial correlation is high, the coefficient is correspondingly high and an I value greater than zero indicates positive spatial correlation (densely populated CBGs tend to be located near other densely populated CBGs), while an I value below zero indicates negative spatial correlation (densely populated CBGs tend to be located near sparsely populated CBGs).

Study Area and Data

The CBG data used were from the 2000 Long Form Census for Tennessee's six MSAs—Chattanooga, Jackson, Johnson City, Knoxville, Memphis, and Nashville. These MSAs encompass areas, not only in Tennessee, but also in Virginia, Georgia, Mississippi, and Arkansas, since Johnson City, Memphis, and Chattanooga are located along the state boundary (see Figure 1 for a map depicting the location of the six MSAs).

Tennessee and its six MSAs are an interesting subject of study for a number of reasons. First, according to the 1997 National Resources Conservation Service Inventory (NRCS), Tennessee has the fourth highest rate of land development among all U.S. states. Between 1982 and 1997, Tennessee's developed area increased from 607,000 hectares to 959,000 hectares, or 58%, which greatly exceeded the national average of 34% for this time period. Of the 352,000 hectares of land developed during this time period, about

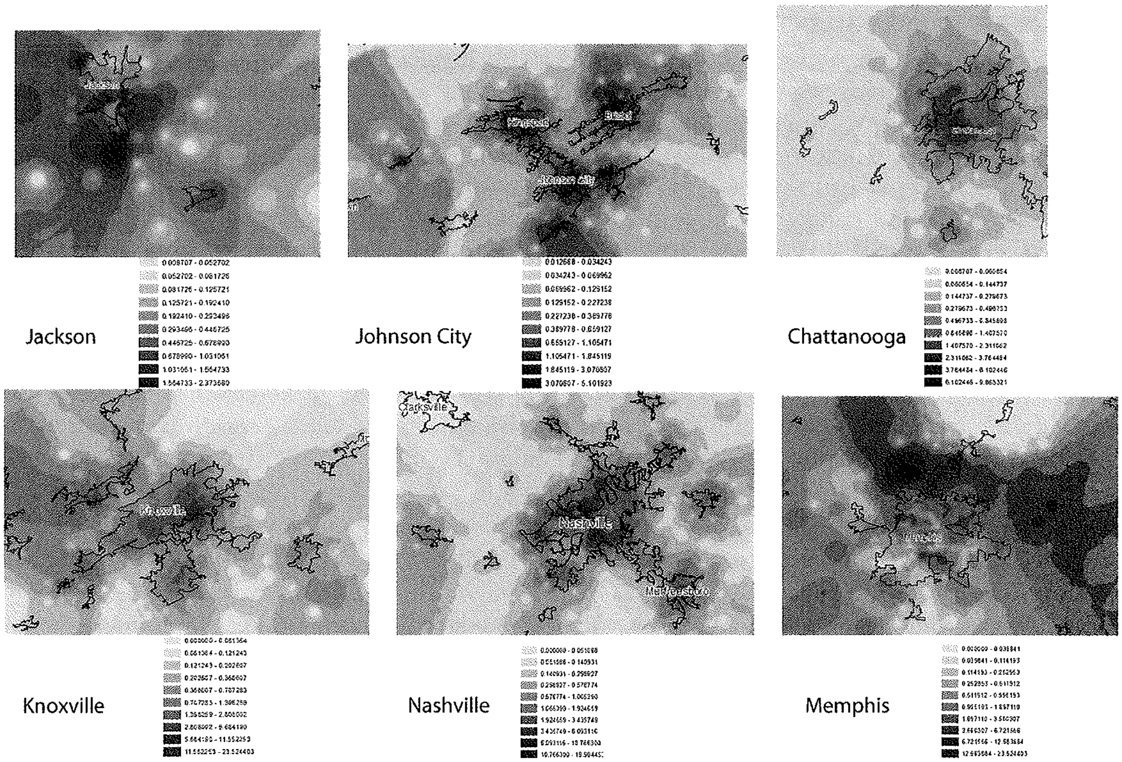

Figure 2. Housing Densities of the six MSAs

137,000 hectares or 39% was prime farmland. Further, the pace of development in Tennessee seems to be increasing. From 1982 to 1992, an average of 18,600 hectares of farmland, forests, and other open space were developed each year, while this rate increased to 32,400 hectares per year during the period from 1992 to 1997 (NRCS). Similarly, the state's population growth rate increased from 29,000 people per year during the 1980s to 81,000 people per year during the 1990s (U.S. Census Bureau). Further, this growth is disproportionately occurring in Tennessee's metropolitan areas, since the majority of the state's population growth between 1990 and 2000 was concentrated in the 21 Tennessee counties comprising the Knoxville, Nashville, and Memphis MSAs (out of a total of 95 counties). The area also provides a relatively uncomplicated study site for testing our methodology because institutional factors such as land use regulations have only a minor influence on the area's housing choices.

In addition, the six MSAs span Tennessee's three grand divisions—East, Middle, and West Tennessee (see Figure 1 for a map depicting the three grand divisions). These divisions reflect significant geographic and socioeconomic differences (Ballard). East Tennessee is mountainous and historically isolated, and its three MSAs—Knoxville, Chattanooga, and Johnson City—are all located along the western edge of the Great Smoky Mountains. The highlands of Middle Tennessee present a rolling to steeply sloping landscape that has come to be largely dominated by Tennessee's most populous city and the area's only MSA—Nashville. Finally, Memphis and Jackson are located in the lowlands of West Tennessee, which are largely flat and rural. See Figure 2 for maps illustrating the spatial distribution of housing density across the six MSAs.

Estimation Results

Table 2 presents descriptive statistics of the variables used in the model estimations. Memphis has the highest housing density and the highest average number of children per household. Nashville has the highest ratios of people with a bachelor, master, or professional degree and, not surprisingly, the highest average household income, while the two least populous MSAs—Jackson and Johnson City—rank at the bottom of the list for all four of these categories. Knoxville, which has Tennessee Valley Authority, the University of Tennessee, and Oak Ridge National Laboratory, has the highest percentage of population with a doctorate degree. Nashville has the highest percentage of households headed by women. Johnson City has the highest proportion of residents with a tenure of five years or more, while the largest MSA—Memphis—has the lowest. Nashville has the highest ratio of population from the Midwestern and Northeastern United States and from foreign countries.

The marginal effects of the explanatory variables on the choice of community from the multinomial logit model are presented in Table 3. To conserve space, the coefficient estimates used to calculate the marginal effects are not reported. The Pseudo R^2 of 31%, log-likelihood estimates of $-3{,}120$, and correct prediction of 57% show that the model fits the data reasonably well. A discussion of the instances in which variables are statistically significant at the 5% level or above follows.

Three of the demographic characteristics are statistically significant at the 5% level or above for all six MSAs, while only one is not significant at the 5% level for any of the MSAs. In all, seven of the variables are significant at the 5% level or better for at least 4 of the 6 MSAs. In general, the marginal effects for the demographic characteristics correspond to what one would expect from the community averages displayed in Table 2, i.e., the marginal effect is negative for those MSAs with an average below the aggregate average and positive for those with averages above the aggregate. Instances in which the variable is significant but does not have the anticipated sign (i.e., proportion of residents born in Northeast in the Johnson City MSA) possibly suggest correlation between the variable and one or more other variables.

More generally, if one were to consider the statistically significant marginal effects at the 10% level or above, then the proportion of those effects with the anticipated sign decreases as the number of significant marginal effects across communities decreases. Thus, for the three characteristics with significant marginal effects for all six communities—density, residents born in the Midwest, and residents born in the South—all have the anticipated sign (18/18); when the marginal effects are significant for 5 of the 6 communities, 83% (15/18) have the anticipated sign; when the marginal effects are significant for three communities, 39% (7/18) have the anticipated sign.

Parameter estimates for the global model symbolized by Equation (4) are presented in Table 4. The adjusted R^2 statistics for the global model regressions for the individual MSAs range from a low of 0.22 to a high of 0.43. While this result might be considered a reasonable explanatory performance, it does leave 57% to 78% of the variance in housing density unexplained. Some of this unexplained variance may result from the global model's implicit assumption that the relationship between housing density and the explanatory variables are constant across the individual MSAs (but, of course, vary from one MSA to another). This assumption is relaxed for the local model, which is discussed below.

For the global model, the lambda values, which indicate self-selection or sorting of CBGs among the MSAs, are statistically significant at either the 1% or 5% level for all of the MSAs except Jackson. The self-selection bias variables detect the fact that households' community choices would not have the same effects on the households' site-specific housing demands. This implies a distinctive heterogeneity in the demographic characteristics of the CBGs from one MSA to another. The coefficients of the lambda variable are negative in Johnson City and Chattanooga, while they are positive in Knoxville, Memphis, and Nashville. The coefficients are negative for those communities with a number of CBGs below the aggregate average for all MSAs and positive for those communities with a number of CBGs above the aggregate average.

Two of the 13 demographic variables in the global model—median family income and ratio of residents born in a foreign country—are statistically significant at the 1% or 5% levels and have the anticipated sign (negative for income and positive for foreign born) for all six MSAs. The coefficient for the ratio of female heads of household is positive for all six MSAs and significant at the 1% level for five and at the 10% level for the remaining MSA. Of the remaining 10 demographic variables, seven are statistically significant at either the 1% or 5% levels for at least two different MSAs. Of these seven variables, four have consistent signs across the MSAs for which the variable is significant. The exceptions are residents born in the South, which has a positive sign for Johnson City and Chattanooga but a negative sign for Memphis, and ratio of residents with doctorate degree, which is positive for Knoxville and Nashville but negative for Johnson City. In general, the relationships between housing density and the demographic variables are relatively homogeneous across the different communities. Neighborhoods with a higher proportion of residents born in the Northeast or a foreign country or that have a higher percentage of female-headed households or residents with a bachelor or professional degree are more likely to be more densely populated. Conversely, neighborhoods with a higher median income or a higher proportion of residents who are at least 65 years old or who have lived there for five or more years tend to be less densely populated.

The relative sizes of the marginal effects reported in Table 4 also provide information as to the nature of the relationship between the demographic characteristics and housing density. For example, a 10% increase in a CBG's proportion of residents born in a foreign country is associated with an increase of approximately one and a half residences per acre in Johnson City, but only half of a residence per acre in Nashville. Similarly, a $1,000 increase in median household income is associated with decreases of 0.05 and 0.02 residences per acre in Jackson and Johnson City, respectively.

The median values of the parameter estimates for the local model symbolized by Equation (5) are presented in Table 5. By minimizing Equation (8), bandwidths are set to be 0.45, 0.11, 0.07, 0.77, 0.16, and 0.08 degrees of latitude for Jackson, Johnson City, Chattanooga, Knoxville, Nashville, and Memphis, which correspond to approximately 50, 12, 78, 86, 18, and 9 km, respectively. The local regression model produces an increase in the adjusted R^2 of 0.19 for Chattanooga, 0.1 for Memphis, and 0.07 for Nashville over the adjusted R^2 statistics for the global model for these MSAs (the adjusted R^2 statistics for the remaining MSAs remain unchanged). The analysis of variance between groups shows reductions in the sum of squared errors of 125 in Chattanooga, 296 in Nashville, and 854 in Memphis. Thus, for at least the Chattanooga, Memphis, and Nashville regressions, the local model appears to represent a significant improvement over the global model.

We use Monte Carlo simulation to test the statistical significance of the spatial variation of the local model coefficients across each MSA. For a given number of times, say, n, the geographical coordinates of the observations are randomly permuted against variables. Note that since we are only questioning the geographical variability of the observations, we do not permute the independent variables against the dependent variable. Thus, we have n values of the variance of the coefficient of interest, which we use in an experimental distribution. We compare the actual value of the variance against this list to obtain an experimental significance level. This test shows that the spatial variability of only one coefficient of one MSA is statistically significant at the 5% level, while four other coefficients are statistically significant at the 10% level. Out of these five coefficients for which spatial variabilities are statistically significant, only one is also statistically significant in the global regression—the proportion of residents in the Johnson City CBGs born in the South. The rest of the four coefficients are the proportion of residents with master's degrees in the Jackson CBGs and Johnson City CBGs, and the proportion of residents in the Knoxville CBGs born in Northeast and Midwest. Figure 3 is a

Table 2. Descriptive Statistics for the Six MSAs

	Jackson	Johnson City	Chattanooga	Knoxville	Nashville	Memphis	Total
Income (US$)	41,009	38,492	44,284	44,363	51,447	44,824	45,679
	(16,540)	(12,852)	(16,015)	(19,431)	(26,280)	(25,663)	(22,841)
Housing density	0.9458	0.7392	1.0654	1.0747	1.5886	2.4310	1.5783
	(1.1129)	(0.9657)	(1.0993)	(1.5813)	(1.8957)	(2.4776)	(1.9743)
Born in Northeast	0.0145	0.0299	0.0265	0.0393	0.0402	0.0205	0.0306
	(0.0137)	(0.0248)	(0.0223)	(0.0309)	(0.0350)	(0.0254)	(0.0300)
Born In Midwest	0.0864	0.0493	0.0563	0.0835	0.0954	0.0643	0.0740
	(0.0438)	(0.0279)	(0.0356)	(0.0427)	(0.0565)	(0.0416)	(0.0480)
Born in South	0.1123	0.2563	0.2775	0.1582	0.1595	0.2397	0.2068
	(0.0507)	(0.1138)	(0.1455)	(0.0625)	(0.0672)	(0.1132)	(0.1106)
Born in foreign country	0.0166	0.1193	0.0252	0.0201	0.0468	0.0310	0.0303
	(0.0284)	(0.0154)	(0.0327)	(0.0311)	(0.0742)	(0.0505)	(0.0524)
Female householders	0.1473	0.1652	0.1629	0.1696	0.1700	0.1544	0.1632
	(0.0754)	(0.0802)	(0.0811)	(0.0964)	(0.1036)	(0.1060)	(0.0977)
Bachelor's degree	0.1125	0.1043	0.1216	0.1324	0.1628	0.1252	0.1335
	(0.0898)	(0.0739)	(0.0952)	(0.0959)	(0.1215)	(0.1090)	(0.1068)
Master's degree	0.0391	0.0343	0.0413	0.0510	0.0550	0.0463	0.0473
	(0.0385)	(0.0300)	(0.0398)	(0.0472)	(0.065)	(0.0590)	(0.0542)
Professional degree	0.0128	0.0123	0.0149	0.0160	0.0197	0.0170	0.0167
	(0.0197)	(0.0164)	(0.0210)	(0.0239)	(0.0331)	(0.0299)	(0.0276)
Doctorate degree	0.0066	0.0081	0.0064	0.0118	0.0098	0.0068	0.0086
	(0.0087)	(0.0154)	(0.0101)	(0.0187)	(0.0173)	(0.0132)	(0.0153)
Senior citizens	0.1340	0.1619	0.1445	0.1438	0.1191	0.1115	0.1292
	(0.0713)	(0.0725)	(0.0638)	(0.0668)	(0.0779)	(0.0769)	(0.0754)
5-year tenure	0.9318	0.9457	0.9404	0.9419	0.9355	0.9252	0.9351
	(0.0331)	(0.0208)	(0.0248)	(0.0265)	(0.0331)	(0.0403)	(0.0333)
Number of kids	0.7059	0.5611	0.6420	0.5886	0.6710	0.8140	0.6852
	(0.1706)	(0.1453)	(0.1952)	(0.2031)	(0.2700)	(0.3034)	(0.2655)
Number of CBGs	82	314	335	469	770	839	2,809
Hectares	*422,498*	*1,008,094*	*683,574*	*874,673*	*1,443,335*	*1,260,447*	*5,692,624*
Total Population	*134,645*	*533,966*	*518,303*	*743,676*	*1,302,136*	*1,188,772*	*4,421,498*

Notes: Numbers in parentheses are standard deviations. Note that *Hectares* and *Total Population* are not used in the model estimations but added for reference.

Table 3. Marginal Effects of Multinomial Logit Model

	Jackson	Johnson City	Chattanooga	Knoxville	Nashville	Memphis
Constant	0.0549	0.0757	−0.4868	0.7909***	.6044	−1.0390***
	(0.0372)	(0.1322)	(0.3321)	(0.2983)	(0.3737)	(0.4471)
Income (1,000 USD)	−0.0002	−0.0023***	−0.0021***	−0.0033***	0.0043***	0.0035***
	(0.0001)	(0.0004)	(0.0008)	(0.0007)	(0.0009)	(0.0011)
Housing density	−0.0024**	−0.0404***	−0.0615***	−0.0382***	0.0217***	0.1208***
	(0.0011)	(0.0050)	(0.0081)	(0.0068)	(0.0070)	(0.0092)
Born in Northeast	−0.3075***	0.5809***	0.4823	2.5497***	3.2610***	−6.5664***
	(0.0970)	(0.1537)	(0.4319)	(0.3356)	(0.4756)	(0.6839)
Born in Midwest	0.1198***	−1.0700***	−2.2900***	0.9157***	3.1340***	−0.8098**
	(0.0376)	(0.1448)	(0.2730)	(0.2028)	(0.2787)	(0.3690)
Born in South	−0.2035***	0.3465***	1.5738***	−1.6971***	−3.6664***	3.6467***
	(0.0421)	(0.0557)	(0.1183)	(0.1289)	(0.1769)	(0.1929)
Born in foreign country	−0.0617	−1.0411***	0.4049*	−1.7024***	1.1468***	1.2535***
	(0.0394)	(0.1775)	(0.2223)	(0.2756)	(0.2131)	(0.3020)
Female householders	−0.0228	−0.1707***	−0.1097	−0.3924***	0.2509*	0.4443**
	(0.0192)	(0.0524)	(0.1263)	(0.1105)	(0.1470)	(0.2039)
Bachelor's degree	0.0450**	0.1808***	0.1802	0.1256	0.2725	−0.8042***
	(0.0213)	(0.0628)	(0.1583)	(0.1329)	(0.1746)	(0.2391)
Master's degree	0.0205	−0.3300**	−0.5075*	0.0909	−0.5345	1.2606***
	(0.0204)	(0.1319)	(0.3014)	(0.1961)	(0.3113)	(0.3642)
Professional degree	0.0330	0.13091	0.4289	−0.3114	−1.7669***	1.4854**
	(0.0631)	(0.2066)	(0.4861)	(0.4166)	(0.5223)	(0.6507)
Doctorate degree	0.1505	1.0780***	−1.2387	3.4977***	−0.2996	−3.1880***
	(0.0971)	(0.2866)	(0.8391)	(0.5612)	(0.8452)	(1.1059)
Senior citizens	0.0298	0.1426***	0.0837	0.2279*	−0.1364	−0.3476
	(0.0195)	(0.0523)	(0.1451)	(0.1216)	(0.1757)	(0.2293)
5-year tenure	−0.0322	0.1779	0.5857*	−0.0860	−0.3477	−0.2977
	(0.0376)	(0.1358)	(0.3396)	(0.3074)	(0.3873)	(0.4604)
Number of kids	−0.0083	−0.2004***	−0.1477***	−0.4668***	−0.1487***	0.9719***
	(0.0071)	(0.0288)	(0.0552)	(0.0517)	(0.0612)	(0.0768)

Notes: Numbers in parentheses are standard errors. Pseudo R^2 is 0.31; Correct prediction is 57%.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Table 4. Parameter Estimates for Global Regression

	Jackson	Johnson City	Chattanooga	Knoxville	Nashville	Memphis
Intercept	−0.7097	4.7938*	5.6378**	3.2847	4.1288***	8.7699***
	(4.1534)	(2.4821)	(2.4697)	(1.3966)	(1.3900)	(1.6710)
Income (1,000 USD)	−0.0490***	−0.0210***	−0.0220***	−0.0400***	−0.0210***	−0.0430***
	(0.0120)	(0.0046)	(0.0060)	(0.0064)	(0.0040)	(0.0062)
Born in Northeast	3.0522	10.4526**	12.7288***	0.1569	−0.6122	−3.4211
	(10.4658)	(4.3244)	(4.8820)	(2.4449)	(1.9105)	(3.7736)
Born In Midwest	−2.1823	−2.2976	−13.4618***	−1.9584	−0.5748	−2.7607
	(3.1342)	(3.3854)	(3.4048)	(1.5118)	(1.1768)	(2.2363)
Born in South	3.6807	8.6601***	10.1047***	−2.0355*	0.8444	−1.7898**
	(3.7472)	(3.3482)	(3.3010)	(1.1645)	(1.1761)	(0.7108)
Born in foreign country	9.5718**	16.0429***	6.7687***	8.3987***	4.8404***	9.5434***
	(4.2439)	(3.1086)	(1.9746)	(2.5405)	(0.8581)	(1.6077)
Female householders	3.4186*	5.0824***	3.9862***	3.7510***	7.5295***	2.8163***
	(2.0590)	(0.8125)	(0.9212)	(0.7492)	(0.8230)	(0.9256)
Bachelor's degree	2.1880	0.8928	0.8081	4.0556***	−0.7313	5.9885***
	(2.1353)	(0.9902)	(1.2019)	(1.1412)	(0.7860)	(1.4280)
Master's degree	4.7807	−0.7432	1.5080	0.5758	−1.6464	−3.1446*
	(3.9578)	(2.3125)	(2.1793)	(2.2594)	(1.1603)	(1.7162)
Professional degree	7.0005	−4.6472	−2.6486	10.9844***	10.2722***	9.6750**
	(6.7630)	(3.1624)	(3.5286)	(4.0501)	(2.7513)	(3.8010)
Doctorate degree	5.2973	−8.4275***	−3.0842	11.5496**	14.8174***	10.6580
	(13.9186)	(4.1088)	(6.1430)	(4.8986)	(4.1075)	(7.0656)
Senior citizens	1.1389	1.7010*	−1.0596	−2.5877**	−2.3445***	−3.7552***
	(2.0844)	(0.9202)	(1.0666)	(1.0609)	(0.9367)	(1.2326)
5-year tenure	3.5245	−4.5795*	−2.8477	−0.8859	−3.6047***	−4.6296***
	(4.1710)	(2.6601)	(2.6238)	(1.4173)	(1.4353)	(1.7186)
Number of kids	0.9687	0.0017	−0.0654	−0.7338*	0.7204***	−0.7237*
	(0.8217)	(0.4374)	(0.3937)	(0.4076)	(0.3129)	(0.3703)
Lambda	−1.2221	−0.9070**	−1.3655***	0.0027**	0.0031***	0.0071***
	(0.9219)	(0.3875)	(0.4068)	(0.0012)	(0.0013)	(0.0017)
Adjusted R^2	0.32	0.43	0.25	0.34	0.28	0.22
Residual sum of squares	55.5	159.3	289.1	748.2	1,943.5	3,923.7

Notes: Numbers in parentheses are standard errors.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Table 5. Median Value of Parameter Estimates for Local Regression

	Jackson	Johnson City	Chattanooga	Knoxville	Nashville	Memphis
Intercept	−0.8098	2.5247	5.8010	3.5082	4.1710	11.2617
	(0.8600)	(0.8000)	(0.8200)	(0.1800)	(0.9800)	(0.9600)
Income (1,000 USD)	−0.0500	−0.0280	−0.0140	−0.0400	−0.0190	−0.0280
	(0.9900)	(0.4600)	(0.9900)	(0.8900)	(0.9200)	(0.7100)
Born in Northeast	4.3605	17.2352	8.4181	0.2468**	−1.2580	2.8967
	(0.6100)	(0.3800)	(0.9900)	(0.0200)	(0.9900)	(0.9800)
Born in Midwest	−2.1829	−5.6300	−11.4474	−1.9902*	0.7013	1.2248
	(0.8300)	(0.1700)	(0.9900)	(0.1000)	(0.8300)	(0.9900)
Born in South	3.9226	14.1014*	6.2221	−2.0944	0.3035	−3.2774
	(0.1400)	(0.0900)	(0.9900)	(0.1900)	(0.7700)	(0.9000)
Born in foreign country	9.4869	13.8225	1.4765	8.4234	3.3070	6.1556
	(0.7200)	(0.5000)	(0.8800)	(0.9700)	(0.9300)	(0.9800)
Female householders	3.3877	4.1250	2.3108	3.6522	6.8444	4.0537
	(0.3300)	(0.9800)	(0.9900)	(0.9100)	(0.9900)	(0.9800)
Bachelor's degree	2.1831	−0.4048	−1.3836	3.9737	−2.5853	2.1708
	(0.9700)	(0.8900)	(0.9900)	(0.8400)	(0.2500)	(0.9300)
Master's degree	4.4374*	−3.9220*	3.5230	0.5052	−2.0459	−5.3688
	(0.0600)	(0.1000)	(0.9900)	(0.8300)	(0.9900)	(0.9900)
Professional degree	7.1185	−0.6897	0.0298	11.1098	10.9030	7.1136
	(0.5100)	(0.9900)	(0.9700)	(0.9800)	(0.5400)	(0.9900)
Doctorate degree	5.7140	−8.4656	−9.0680	11.7042	13.6522	6.2348
	(0.7200)	(0.8400)	(0.9900)	(0.4100)	(0.7800)	(0.9900)
Senior citizens	1.2722	0.6556	−2.5190	−2.6168	−3.1548	−4.2429
	(0.5400)	(0.8500)	(0.9900)	(0.9600)	(0.7200)	(0.7000)
5-year tenure	3.8274	−0.8359	−2.5317	−1.0328	−2.8190	−6.6331
	(0.6400)	(0.7600)	(0.9400)	(0.3800)	(0.9900)	(0.9900)
Number of kids	0.9868	0.2451	−0.5398	−0.7875	0.7786	−1.3619
	(0.7500)	(0.8900)	(0.9600)	(0.5100)	(0.7500)	(0.9600)
Lambda	−1.3347	−1.4315	−1.0190	0.0027	0.0032	0.0062
	(0.2500)	(0.1600)	(0.9900)	(0.8100)	(0.9900)	(0.9600)
Adjusted R^2	0.32	0.43	0.44	0.34	0.35	0.32
Residual sum of squares	54.5	86.5	163.8	742.1	1,647.5	3,069.8

Notes: Numbers in parentheses are p values for the Monte Carlo significance test.
* Significant at the 10% level.
** Significant at the 5% level.

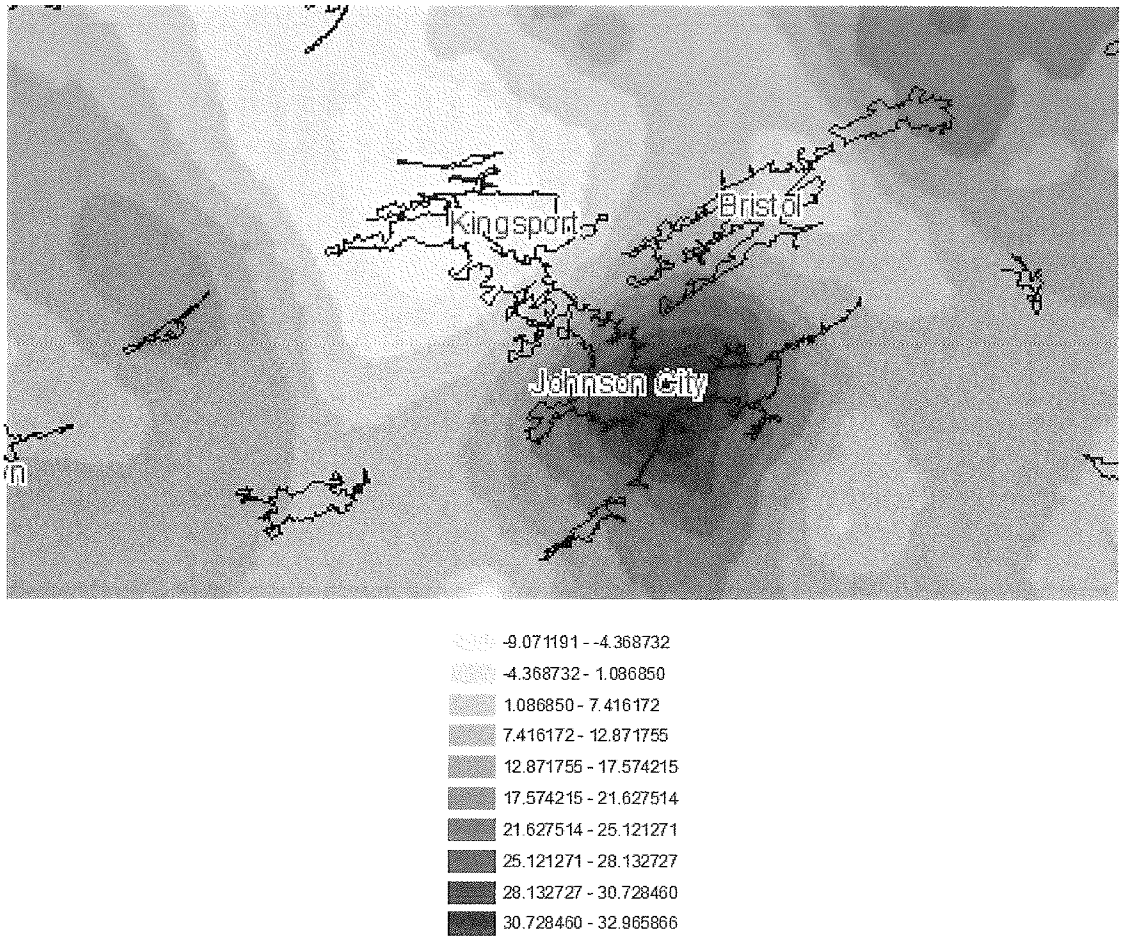

Figure 3. Spatial Distribution of Marginal Effect of the Ratio of Residents Born in the South on Housing Density at Johnson City MSA

map of the spatial variability of the marginal effect of this variable for the Johnson City MSA, drawn using a deterministic interpolation technique—inverse distance weighting.[4] The map shows an interesting pattern over the three cities that dominate the MSA (this area is known as the Tri-Cities because of the presence of the three similarly sized cities located in close proximity to each other). The marginal effect is generally positive throughout the MSA, but it is strongest in the center of Johnson City and gradually weakens as one moves away from the city center. The two areas where the effect is weakest, if not actually negative, are within the cities of Kingsport and Bristol. What this distinctive spatial heterogeneity implies is that there are other factors in the Johnson City MSA that exert a noticeable influence over housing density and that are somehow related to the proportion of residents born in the South.

A more general examination of the pattern

[4] Inverse distance weighting is based on the idea that points that are close to one another in space have more similar characteristics than ones further away. Thus, the weights are a decreasing function of distance. The weights are constructed by identifying an area around the interpolated point and a weighted average is taken of the observed values within this area. The size of the area determines how many points are included in the weighting. The size can be specified in terms of its radius or the number of points it includes. For this analysis, we chose to include a 24.14 kilometer radius from the center of MSAs since the various sizes and shapes of the CBGs make it difficult to specify a radius (Shepard).

Table 6. Spatial Dependency

MSA	Mean Housing Density	Housing Density Percentiles 25	50 (median)	75	Gini Coefficient	Moran's Index for Housing Density per Acre
Jackson	0.9458	0.0523	0.4628	1.4953	0.67	0.06
Johnson City	0.7392	0.1010	0.2929	1.0201	0.65	0.27
Chattanooga	1.0654	0.2108	0.7386	1.6367	0.56	0.25
Knoxville	1.0747	0.1679	0.5556	1.5192	0.61	0.20
Nashville	1.5886	0.2151	1.0461	2.4295	0.58	0.18
Memphis	2.4310	0.8157	1.9688	3.3507	0.49	0.03

of residential density across the six MSAs is presented in Table 6 and Figure 4. Figure 4 is a box plot of the distribution of CBG housing density across each MSA, with each box representing the middle half of the distribution and the vertical line in each box the median value. The mean and quartiles of each distribution are reproduced in Table 6. Taken together, they illustrate the greater density and dispersion associated with the larger cities. For the most part, the distributions are heavily skewed toward low-density neighborhoods. The only exception is Memphis, which is much more evenly distributed and much more dispersed. The Gini coefficient is a measure of how evenly housing density is spread across the different CBGs within an MSA, with higher values representing a more uneven distribution. With this statistic, it appears that, in general, the more populous MSAs have the most even distributions, possibly indicating that the less populous MSAs include some rural areas.

Finally, the Moran's Index indicates the degree of residential clustering across the CBGs for each MSA. The Moran's index for all of the MSAs is positive, indicating that on average, more densely populated CBGs tend to be located near other densely populated CBGs. Memphis, which has a mean housing density

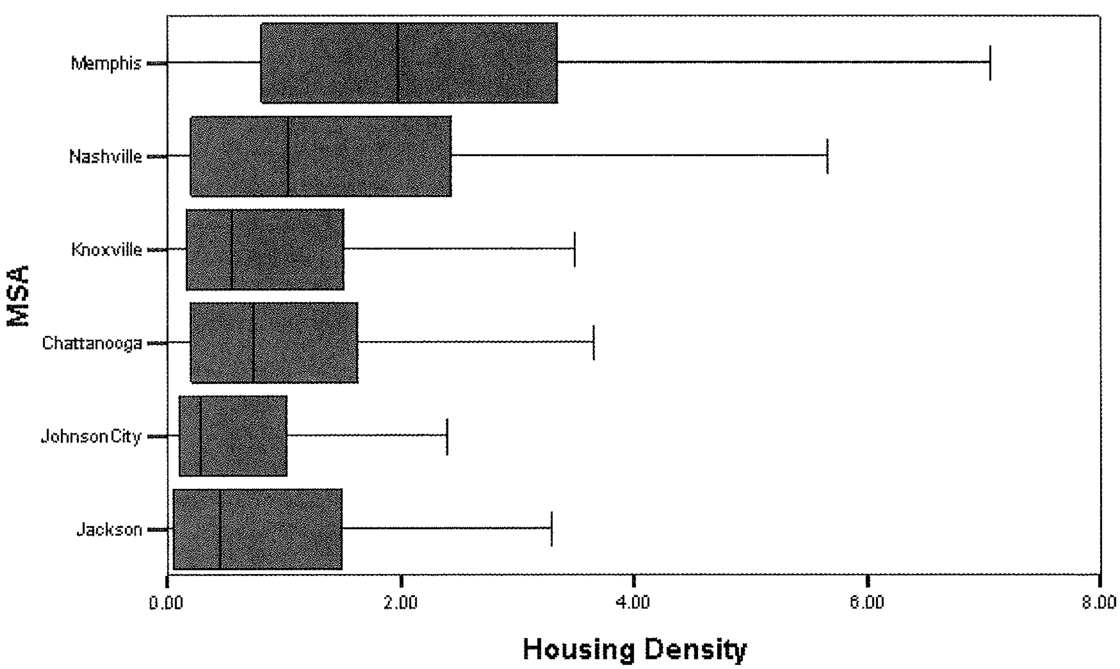

Figure 4. Distribution of Housing Density for the CBGs of the six different MSAs.

of over 2.43 residences per acre and a relatively dispersed distribution of density over the CBGs, has the lowest Moran's Index score (0.03). Thus, while Memphis' neighborhoods are, on average, relatively densely populated, there is a relatively weak spatial relationship between the densities of its neighborhoods. On the other hand, Johnson City, which is the least densely populated MSA, with a mean of 0.74 and a median of 0.29 houses per acre, has the highest Moran's Index (0.27). Thus, the relatively sparsely populated Johnson City has the greatest degree of clustering among the CBGs with the relatively higher housing densities, implying the inclusion of distinct rural areas into the MSA. As a result, it is clear that for at least these six MSAs, no single statistic can capture a representative picture of the spatial pattern of housing within the MSA.

Summary and Conclusions

The relationship between demographics and housing density is modeled in a 2-stage multinomial logit sample selection model. The sample selection model was used to account for any selection bias that would result from individuals choosing a community or MSA prior to choosing a particular neighborhood or CBG. The regression estimates that there was some sorting of CBGs at the MSA level. The second stage of the model that regresses demographic characteristics on the housing densities of CBGs was estimated using two different models. The global model assumed that relationships between demographic variables and housing density were constant across a particular MSA, while the local model allowed this relationship to vary across the geographic area of the MSA.

The first stage of the model correctly predicts the location of the CBG within a particular MSA 57% of the time. The lambda values representing self-selection bias in the global model detect the fact that community choice would not have the same effects on housing density. The results of the global model suggest that neighborhoods with a higher proportion of households headed by a female, or residents born in the Northeast or a foreign country, or residents with a bachelor or professional degree are more likely to be more densely populated. Conversely, neighborhoods with a higher median income or a greater proportion of senior citizens or residents with a tenure of five or more years tend to be less densely populated.

The local model produced a better fit for three of the six MSAs. However, only the relationship between the proportion of residents born in the South in the Johnson City MSA and housing density was both statistically significant in the global model and displayed a statistically significant spatial variation over the MSA. Mapping of the marginal effects from the local regression show that the effect of this variable on housing density is positive over most of the MSA (as is the marginal effect in the global regression), but negative in a couple of areas within the MSA. Thus the relationship between the proportion of residents born in the South and housing density shows distinctive spatial heterogeneity within the Johnson City MSA.

Univariate analysis of the spatial distribution of housing within the different MSAs reveals that while Memphis' neighborhoods are on average relatively densely populated, there is a relatively weak spatial relationship between the housing density of these neighborhoods. On the other hand, Johnson City, which has the lowest housing density overall, displays the highest degree of spatial correlation between the housing densities of its neighborhoods, indicating that while Johnson City's neighborhoods are, on average not that densely populated, the neighborhoods with the higher housing densities do tend to be located near each other.

References

American Farmland Trust. *Farming on the Edge.* Washington D.C.: AFT, 2002.

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

Ballard, D. "Three for One: East, Middle, and West Tennessee: Why They're Different and Likely to Stay That Way." *Tennessee Alumnus* (Winter 1996):9–10.

Can, A. "Specification and Estimation of Hedonic Housing Price Models." *Regional Science and Urban Economics* 22(1992):453–74.

———. "The Measurement of Neighborhood Dynamics in Urban House Prices." *Economic Geography* 66(1990):254–72.

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

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

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

Damgaard, C., and J. Weiner. "Describing Inequality in Plant Size or Fecundity." *Ecology* 81(2000):1139–42.

Daniels, T., and D. Bowers *Holding Our Ground.* Washington, D.C.: Island Press, 1997.

Dixon, P.M., J. Weiner, T. Mitchell-Olds, and R. Woodley. "Bootstrapping the Gini Coefficient of Inequality." *Ecology* 68(1987):1548–51.

Dubin, R. "Spatial Autocorrelation and Neighborhood Quality." *Regional Science and Urban Economics* 22(1992):433–52.

———. "Spatial Autocorrelation: A Primer." *Journal of Housing Economics* 7(1998):304–27.

Fotheringham, A.S., and C. Brunsdon. "Local Forms of Spatial Analysis." *Geographical Analysis* 31(1999):341–58.

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

Fulton, W., R. Pendal, M. Nguyen, and A. Harrison. *Who Sprawls the Most: How Growth Patterns Differ across the United States.* Washington, D.C.: Brookings Institute Center on Urban and Metropolitan Policy, 2001.

Geoghegan, J., L. Waigner, and B. Nancy. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS." *Ecological Economics* 23(1997):251–64.

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

Goodman, A. "A Comparison of Block Group and Census Tract Data in a Hedonic Housing Price Model." *Land Economics* 53(1977):483–87.

Greene, W. *LIMDEP Version 7.0 User's Manual.* Econometric Software, Inc., 1995.

———. *Econometric Analysis.* Upper Saddle River, NJ: Prentice Hall, 1997.

Katz, B. *Smart Growth: The Future of the American Metropolis?* Centre for Analysis of Social Exclusion 58. London: London School of Economics, 2002.

Lee, L. "Generalized Econometric Models with Selectivity." *Econometrica* 51(1983):507–12.

Lesage, J.P. "Regression Analysis of Spatial Data." *Journal of Regional Analysis and Policy* 27(1997):83–94.

Leung, Y., C. Mei, and W. Zhang. "Testing for Spatial Autocorrelation among the Residuals of the Geographically Weighted Regression." *Environmental and Planning A* 32(2000):871–90.

McMillen, D. "Probit with Spatial Autocorrelation." *Journal of Regional Science* 3(1992): 335–48.

———. "Spatial Autocorrelation or Model Misspecification?" *International Regional Science Review* 26(2003):208–17.

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

Moran, P.A.P. "The Interpretation of Statistical Maps." *Journal of the Royal Statistical Society* Series B 10(1948):243–51.

Newman, P., and J. Kenworthy. "Gasoline Consumption and Cities: A Comparison of U.S. Cities with a Global Survey." *Journal of American Planners Association* 55(1989):24–37.

———. *Sustainability and Cities: Overcoming Automobile Dependence.* Washington, D.C.: Island Press, 1999.

NRCS. *Summary Report 1997 National Resources Inventory (revised December 2000).* Washington, D.C.: Natural Resources Conservation Service, United States Department of Agriculture, 2000.

Shepard, D. "A Two-Dimensional Interpolation Function for Irregularly-Spaced Data." *Proceedings of the 23rd National Conference ACM,* ACM Press, 1968, pp. 517–24.

Sierra Club. *Solving Sprawl: 1999 Sierra Club Report.* Internet site: http://www.sierraclub.org.

Smart Growth Online. *Principles of Smart Growth.* Internet site: http://www.sustainable.org.

Snyder, K., and L. Bird. *Paying the Costs of Sprawl: Using Fair-Share Costing to Control Sprawl.* Washington, D.C.: U.S. Department of Energy, 1998.

Tobler, W. "Smooth Pycnophylactic Interpolation for Geographic Regions." *Journal of Statistical Association* 74(1979):519–36.

Tse, R.C. "Estimating Neighborhood Effects in House Prices: Towards a New Hedonic Model Approach." *Urban Studies* 39(2002):1165–80.

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

U.S. Department of Agriculture. *Summary Report—1997 National Resources Inventory.* Washington, D.C.: United States Department of Agriculture, 1999.

U.S. EPA. *Protecting Water Resources with Smart Growth.* Washington, D.C.: United States Environmental Protection Agency, 2004.

Vandell, K.D. "Market Factors Affecting Spatial Heterogeneity among Urban Neighborhoods." *Housing Policy Debate* 6(1995):103–37.