

**Community Choices and Housing Decisions:
A Spatial Analysis of the Southern Appalachian Highlands**

Seong-Hoon Cho, David H. Newman and David N. Wear

Seong-Hoon Cho is post-doctoral associate, Warnell School of Forest Resources, University of Georgia; David H. Newman is professor, Warnell School of Forest Resources, University of Georgia; and David N. Wear is project leader, Southern Research Station, U.S. Department of Agriculture, Forest Service.

Corresponding author:
Seong-Hoon Cho
Warnell School of Forest Resources
University of Georgia
Athens, GA 30602
Phone: 706-542-9699
Fax: 706-542-8356
E-mail: shc4982@owl.forestry.uga.edu

*Selected Paper prepared for presentation at the Southern Agricultural Economics Association
Annual Meeting, Alabama, February 2003*

*Copyright 2003 by Seong-Hoon Cho, David H. Newman and David N. Wear. All rights reserved.
Readers may make verbatim copies of this document for non-commercial purposes by any means,
provided that this copyright notice appears on all such copies.*

**Community Choices and Housing Decisions:
A Spatial Analysis of the Southern Appalachian Highlands**

Abstract

This paper examines land development using an integrated approach that combines residential decisions about choices of community in the Southern Appalachian region with the application of the GIS (Geographical Information System). The empirical model infers a distinctive heterogeneity in the characteristics of community choices. The results also indicate that socioeconomic motives strongly affect urban housing decisions while environmental amenities affect those of rural housing.

Key Words: community choices, housing decisions, spatial econometrics.

Community Choices and Housing Decisions: A Spatial Analysis of the Southern Appalachian Highlands

Land development has drawn increasingly more attention in the last few decades, partly because of changes in land use patterns. For example, the amount of urban land per person is increasing faster than the population: one-third more land per person was consumed by urban uses in 1990 than in 1970 (Daniels and Bowers 1997). Residential development, driven by residential preferences within the constraints of land use regulations, is the dominant force in overall development. Understanding residential choices is the key to understanding much about land development.

In the standard model of a monocentric city (e.g., Alonso 1964; Mills 1981; Muth 1969), residential development is modeled as the choice of location that provides the best tradeoff between land costs and transportation costs. The standard model has been extended in a number of ways, including consideration of urban growth dynamics (e.g., Fujita 1982; Anas 1978), environmental amenities (e.g., Wu 2001; Bruecker et al. 1999; Polinsky and Shavell 1976), and multiple income groups and employment centers (e.g., McMillen and McDonald 1989). Recent empirical analyses of this type have been improved through the incorporation of spatial statistics with the Geographical Information System (GIS) (e.g., Ding 2001; Lake et al. 2000; Geoghegan et al. 1997). GIS and spatial statistics allow for spatially explicit analysis by providing flexibility in specifying models and measuring variables.

Economic models of land use have been applied to both broad units and fine units, based on the spatial scale of land use. Models of broad units examine patterns of land use from a macro viewpoint. These models generally use counties or county groupings as units to highlight how socioeconomic factors and physical landscape features influence land use allocations (Alig

1986; Hardie and Parks 1997; Miller and Plantinga 1999; Plantinga 1996; and Hardie, et al. 2000).

Models of fine units, on the other hand, provide analyses of spatially explicit land use decisions. These models estimate the direct influence of site-specific factors because they are applied at a fine resolution. For example, the road construction and access influences on land development (e.g., Chomitz and Gray 1995; Nelson and Hellerstein 1997; Dale et al. 1993) and the influences of location, topography, and ownership (Turner et al. 1996; Spies et al. 1994) are analyzed in this framework.

Even though each type of model independently serves a valuable function, they both have limitations. Macro-scale analyses do not capture information in a spatially explicit framework, while micro-scale analyses may miss out on broader physical and social phenomena. Wear and Bolstad (1998) explain the limits of land use models for different units. They point out that land use models of spatially broad units may not provide direct insights into the fine-scale socio-economic and physical consequences of land-use changes. They also discuss the limitations of fine-scale units, including the resolution of the definition of land use. For example, residential presence in the satellite images of forest cover (e.g., Wear and Flamm 1993; Turner et al. 1996) may not capture site-specific land uses. One type of model could be complemented by the other type of model, yet there has been no attempt to link models of different scales in the previous literature.

This paper examines land development using an integrated approach that combines residential decisions about choices of community (broad units) with site-specific information regarding development using US Census blocks of the Southern Appalachian region (fine-scale units). We do this with the application of GIS and econometric tools. Residential development

plays an increasingly important role in the Southern Appalachian region's land development. Because institutional factors such as land use regulations have only a minor influence on the area's development, the Southern Appalachians provide a less complicated study site for testing our methodology.

The Empirical Model

Residential decisions are modeled in two stages in order to link community choice with a site-specific census block. In the first stage, we model the choice of a community type in broad units. The community types are classified as urban-dominant, urban-moderate, rural-moderate and rural-dominant communities according to the types of housing. A multinomial logit framework is used to examine heterogeneity in the characteristics of different community choices. In the second stage, residential decisions based on site-specific census blocks are modeled using housing density equations. The housing density found in the 1990 U.S. Census at the block level is used to examine site-specific residential decisions. The estimates of the community choice models are then incorporated into the housing density equations as a form of self-selection variable. We do this to check if a self-selection bias arises in the formation of the community-type choice. The spatial variables in the housing density equations, a combination of distance and location attributes, are incorporated through the application of GIS.

The Choice of Community Types

Suppose a household tries to choose a community from among four possible types of communities. The types of communities are based on degree of urbanization. Let u_j^* be the

household's expected utility from choosing a type of community j . The community j is indexed as 1, 2, 3 and 4 for urban-dominant, urban-moderate, rural-moderate and rural-dominant communities, respectively:

$$(1) \quad u_j^* = Z' \gamma_j + e_j$$

where Z is a vector of community characteristics influencing the choice of the community and e_j is a residual capturing errors in perception and optimization by the household. The household's utility in choosing an alternative community is not observable, but their choice of a community is. Let J be a polychotomous index denoting the household's type of community.

$$(2) \quad J = j \text{ if and only if}$$

$$u_j^* = \max(u_1^*, u_2^*, u_3^*, u_4^*).$$

Maddala (1983) shows that if the residuals e_j are independently distributed with an extreme value distribution, then the choice of the type of community can be represented by a multinomial logit model (Maddala 1983, pp. 60). Following McFadden (1973), disturbances are assumed to be independent and identically distributed with a Weibull distribution. This implies that the probability of choosing a type of community j by the household can be expressed as

$$(3) \quad P_j \equiv \Pr(J = j) = \frac{\exp(Z' \cdot \gamma_j)}{\sum_{i=1}^J \exp(Z' \cdot \gamma_i)}$$

$$j = 1, 2, 3, 4.$$

The multinomial logit model is estimated using the urban-dominant community as a base of community choice (see the discussion in the *Estimates of the Community Model*). Previous studies (e.g., Nechyba and Strauss 1998; Rapaport 1997) suggest that individual community choices are specified as a function of household characteristics and community attributes. Here

we consider the influence of individual-specific characteristics (the household characteristics of education level and political view) and choice-specific attributes (the community attributes of population density, crime level, stability, and level of air pollution).

In the first stage, the multinomial logit model in equation (3) is estimated. We also estimate the marginal effects of explanatory variables on the choice of alternative communities as

$$(4) \quad m_j = \frac{\partial P_j}{\partial Z} = P_j \left(\gamma_j - \sum_{j=1}^J P_j \gamma_j \right).$$

These marginal effects depend on the sign and magnitude of many coefficients. The statistical significances of these effects are estimated by the asymptotic covariance matrix of m_j ¹.

Housing Density Equations

The residential decisions are directly reflected by the housing counts of a given area. The housing count per km² of the 1990 U.S. Census block is defined as housing density. The housing market is assumed to be in equilibrium; this requires that households optimize their residential choices. Community choices are assumed to be made prior to residential choices. With these assumptions, the housing density can be described as a function of the socioeconomic and environmental characteristics of the block, in addition to a self-selection variable in the formation of community choice. The following housing density equation is estimated in the second stage.

$$(5) \quad h_{ij} = x' \beta_j - \theta_j \hat{\lambda}_j + e_{ij},$$

¹ A detailed description of marginal effects and their asymptotic covariance of multinomial logit can be found at pp. 916-17 in Greene (1997).

$$j = 1, 2, 3, 4$$

where h_{ij} is the housing density of a block i at community j ; x is a vector of socioeconomic variables and environmental variables; $\hat{\lambda}_j$ is a self-selection variable for community j ; and e_{ij} is a residual capturing errors in perception and optimization by the household's choice of a site-specific block and a community. The self-selection variable is estimated using the following equation (Lee 1983).

$$(6) \quad \hat{\lambda}_j \equiv \phi[\Phi^{-1}(\hat{P}_j)] / \hat{P}_j,$$

where $\hat{P}_j = \frac{\exp(Z' \cdot \gamma_j)}{\sum_{i=1}^J \exp(Z' \cdot \gamma_i)}$ from the estimates of the first stage. The form of self-selection

variable incorporates residential decisions about choices of communities into the residential decisions concerning blocks. We consider explanatory variables x to include socioeconomic variables describing housing value, income, population density, crime rate, stability, education, political view, travel time to work, distance to any city, distance to major city, distance to major roads and a road index. The environmental variables of distance to major open spaces, distance to lakes, air pollution, elevation, the stream index, and the open space index are considered (see the discussion of data in the next section).

The housing density equations are estimated using cross-sectional data. Because the block size and characteristics of residential decisions are different across the blocks, heteroscedasticity is likely to be present. The null hypothesis of no heteroscedasticity is tested using the Lagrange multiplier (LM) test suggested by Greene (1997, pp. 653-58). The null hypothesis is rejected at the 1% significance level for each equation. Heteroscedasticity is corrected using the technique suggested by Kmenta (1986, pp. 270-76). The transformed

equation system is then estimated using the SUR estimator. A test for selectivity bias is a test for $\theta_j = 0$, $j = 1, 2, 3, 4$. If the null hypothesis of $\theta_j = 0$ is rejected, there is self-selection in choosing a type of community, j , and estimation without the self-selection variable will be biased.

It is a challenge to incorporate all the independent variables for the housing density equations because there may be multicollinearity among them. Although there have been many suggestions about how to detect multicollinearity, there are no certain guidelines. A commonly used rule 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. The correlation coefficients are reported on Table A-1 in the Appendix. Few of the correlation coefficients are shown to be close to 0.8 (e.g., correlation between housing values and education level, income and education level, housing values and income, and road index and population density). The seriousness of the multicollinearity is examined by deletion of the regressors involved with high correlation coefficients. We did not detect serious fluctuations in the coefficients, nor serious changes of statistical significance resulting from the deletion of the regressors with high correlation coefficients (see Table A-2 in the Appendix). Thus, the suspected multicollinearity is not a serious problem in the housing density equation.

Study Area and Data

The area of our study is the Blue Ridge province of the Southern Appalachian Highlands; it includes all of the mountainous portions of western North Carolina, northern Georgia, southeastern South Carolina, eastern Tennessee, southwestern Virginia and southeastern West Virginia. Within this region, 3,687 blocks of the 1990 U.S. Census are used (see Figure 1). The

eastern portion of the region is dominated by the Blue Ridge Mountains, which rise abruptly from the Piedmont province, forming a rugged and diverse landscape. Regionwide, the area of developed land has increased considerably over the past 20 years. Much of this development has been at the expense of cropland and pasture. Though the region has the greatest concentration of federally-owned land in the eastern United States, the vast majority of the region's land is privately owned. The population of the region increased by 27.8 percent between 1970 and 1990. Despite this growth, the population density in the study area remains below the average for the six states that contain the study area (U.S. Forest Service 1996).

Two principal data sources were used in this study: Applied Geographic Solutions, Thousands Oaks, California, which collects demographic, housing, crime risk and pollution data from the U.S. Census, the FBI and the EPA; and Geography Network, a web service which provides geographic data from the Environmental System Research Institute (ESRI), Redlands, California. The ArcView, computer software was employed to generate the database, using the data from the two principal sources. Distance calculations were made using a raster system where all data were arranged in grid cells. Distances were measured as the Euclidean distance from the centroid of the census block to the nearest edge of a feature. The sum of length and the sum of area were calculated using ArcScripts downloaded from ESRI. The census blocks are bounded on all sides by visible features, such as streets, roads, streams, and railroad tracks, and by invisible boundaries, such as cities, towns, townships, and county limits, property lines, and short, imaginary extensions of streets and roads. The census blocks in remote areas may be large and irregular and may contain many square miles (U.S. Census Bureau 1990).

The dependent variable of the community choice model is a community index. We constructed an index to classify each block into urban-dominant, urban-moderate, rural-moderate

and rural-dominant communities. The classification is based on information about housing types from the U.S. Census. The U.S. Census divides housing types into urban core, urban non-core, rural farm, and rural non-farm, based on the population of each block. Specifically, we calculated the ratio of housing types of urban core and urban non-core to all housing types for each block. A block is identified as an urban-dominant community if all the housing types of each block are urban core or urban non-core. 554 blocks of 3,687 blocks or 1 % of the total study area are identified as urban-dominant communities. A block is identified as an urban-moderate community if the percent of urban core and urban non-core housing is greater than or equal to 50 % and less than 100 %. A total of 1,027 blocks or 6 % of the total area are identified as urban moderate communities. A block is identified as a rural-moderate community if the percent of rural farm and rural non-farm housing is greater than 0 % or less than 50 %. 495 blocks or 10 % of the total area are identified as rural-moderate communities. A block is identified as a rural-dominant community if all the housing types of each block are rural farm or rural non-farm. A total of 1,611 blocks or 83 % of the total area are identified as rural-dominant communities.

The dependent variable for the housing density equation is the housing density of each block. The housing density is the number of houses per km^2 of area. It is the ratio of the total number of houses of the urban core, urban non-core, rural farm, and rural non-farm types to the area of each block in km^2 . The dependent variables, explanatory variables and their definitions are shown in Table 1. Descriptive statistics for the variables are given in Table 2.

Estimation Results

Estimates of the Community Model

Parameter estimates for the multinomial logit model are presented in Table 3. The multinomial logit models the probabilities of households in an urban-dominated community relocating to other communities because the base of community choice is set to be an urban-dominated community, in our estimation. The marginal effects of independent variables on the choices between staying in an urban-dominant community and relocating to other communities are shown in Table 4. Sixteen of eighteen marginal effects are significant at the 1 % level, indicating that the model fits the data well.

The results show that community choice is significantly affected by the household characteristics of education level and political view. Education level is positively correlated with a choice of urban-moderate community, but it is negatively correlated with choices of rural communities. More educated households in an urban-dominated community are more likely to relocate to an urban-moderate community, but they are less likely to relocate to rural communities. Political view is correlated with choices of non-urban dominated communities (urban-moderate, rural-moderate and rural-dominant communities). The more conservative households in the urban-dominated community are more likely to relocate to other communities. These results indicate that more educated households choose to move toward urban communities, while conservative households choose to move away from urban-dominated communities.

The results show that community choice is significantly affected by the community attributes of population density, crime level, stability, and pollution. Population density and crime rate are all negatively correlated with the choices of non-urban dominated communities. Households that are currently located in urban-dominated communities are less likely to relocate

to other communities experiencing increases of population density and/or crime rate.

Households of an urban-dominated community are more likely to relocate to other communities with a greater stability. Urban-dominated households are more likely to relocate to rural communities with lower levels of air pollution. However, the relocation of households from urban-dominated communities to urban-moderate communities is not significantly affected by pollution level, reflecting little difference in the air pollution level between urban-dominated and urban-moderate communities. These results indicate that households choose to live in less-urbanized communities for safety, less crowding, more stability and a less air-polluted environment.

Estimates of the Housing Density Model

The results of the housing density models of the four different types of communities are shown in the Table 5. Of the seventy-six housing density coefficients (nineteen variables in each of the four equations), thirty-seven are significant at the 5 % level. The system weighted R^2 is between 0.84 and 0.91.

The self-selection variables are taken from the multinomial logit model. There is substantial evidence that self-selection occurred in the households' choices of communities. The coefficient of the self-selection variable is statistically significant at the 1 % level in the housing equations for rural communities. It is also statistically significant at the 5 % level in urban housing equations. These results suggest that the community choices have different effects on the communities themselves. This implies a distinctive heterogeneity in the characteristics found in the community types observed in the region.

The parameter estimates of the housing density equations for different communities show that variables affecting housing density vary across the communities. Housing densities are affected more by socioeconomic variables in urban communities, while they are affected more by environmental variables in rural communities. Of the twenty-four socioeconomic coefficients (twelve variables in each of the dominant and moderate equations), sixteen in the urban communities and nine in the rural communities are statistically significant at the 5 % level. Of the twelve environmental coefficients (six variables in each of the dominant and moderate equations), no variables in the urban communities and eight in the rural communities are statistically significant at the 5 % level.

The effects of socioeconomic variables on housing densities across urban and rural communities also vary, even though the difference in socioeconomic effects is not as drastic as the difference in environmental effects. Population density, crime rate, education, political view, travel time to work, and road index commonly affect housing density in both urban- and rural-dominated communities. A higher population density requires more housing. The marginal effects of population density on the urban communities are higher than those of the rural communities. This suggests that an equal increase in population density increases housing density more in the urban communities than it does in the rural communities. This finding provides evidence that housing developments in urban communities are more responsive to increased population than housing developments in rural communities. A lower crime rate and higher levels of education attract more housing, either in urban-dominated communities or rural-dominated communities. The marginal effects of these two variables in the urban communities are higher than those in the rural communities. They indicate that safety and the education level of the community are common concerns of urban and rural households, but the degree of the

concern is greater in the urban communities. A less conservative political viewpoint is correlated with more housing. An increase in the travel time to work increases with housing density. This suggests that people of the region are indifferent to driving longer distances to meet their other housing requirements. The coefficient for the road index is positive and statistically significant at the 1 % level in both urban- and rural-dominated communities. This suggests that road accessibility is important to houses in any type of community.

Housing value, income, stability, and the distance to major roads have significant effects on housing density in urban communities, but they are not significant in rural communities. Housing density is negatively associated with housing value in urban communities. This is evidence that supports the notion that the more sparsely-developed houses in an urban area are more highly-priced. Housing density is positively associated with income in urban communities and is negatively associated with the stability of households in urban communities. This indicates that households in stable urban communities prefer to not be located in densely developed housing. Housing density is higher in urban-dominated communities, where the houses are closer to a major road.

Four of six environmental variables are statistically significant at the 1 % level in the rural-dominated communities. Rural dominated households are more likely to locate in the blocks that are closer to lakes, at higher elevations, and with greater access to streams and open space. Environmental variables did not have a substantial impact on the housing densities of urban communities. Clear differences in the effects of environment factors on housing densities between urban and rural communities imply heterogeneity in the characteristics found in the community choices observed in the region; this confirms significant self-selection.

All coefficients for the distance to a lake are negative across the urban and rural communities, although the coefficients of only the rural communities are significant at the 5 % level. This shows that both urban and rural households enjoy the environmental amenities of lakes but the attractions are only substantial to rural households. Elevation and access to streams are statistically significant at the 1 % level in both rural-moderate and rural-dominated communities. This indicates that the environmental amenities of higher elevation and a greater access to streams draw a substantial number of households to rural communities. The coefficient for the open space index is positive and statistically significant at the 1 % level only in rural-dominated communities. This suggests that access to open space is significantly important only to rural-dominated households.

Distance to the closest city is not a significant factor across the communities, and distance to the closest major city is not a significant factor in urban communities. This result may be explained by the relatively smaller and fewer cities observed in the region. The impact of distance to the closest major city is positive and significant at the 1 % level only in rural-dominant communities. This implies that rural-dominated households enjoy remoteness more than the positive utilities of being close to major cities. Air pollution is not a significant factor in housing decisions across the communities, perhaps reflecting that air quality under each community choice of the region is relatively homogeneous. Thus, the air quality is not a significant factor of housing choice within each community, even though it is a significant factor of alternative community choices, as shown in the estimates of the community model.

Concluding Remarks

This paper makes the first attempt to develop a spatial econometric model that combines broad units and fine-scale units with the application of GIS. The importance of our findings lies in their ability to present a coherent multi-scale model of housing decisions in the Southern Appalachian region.

The first-stage analysis yields estimates of the marginal effects of household characteristics and community attributes in community choices. We found that people who choose to live in less-urbanized communities value safety, less crowding, more stability, and a cleaner environment. The second-stage analysis yields the marginal effects of the socioeconomic and environmental characteristics in the residential choices for different communities. There is a distinctive heterogeneity of the characteristics found in the community choices observed in the region. The socioeconomic motives of urban communities and the environmental motives of rural communities are more weighted in their housing decisions. Specifically, housing development in urban communities is more responsive to increased population density than housing development in rural communities. Safety and the education level of the community are a greater concern to urban households. More sparsely developed houses in urban communities are more highly-priced. The higher income in urban communities attracts more housing. Households in stable urban communities dislike being located in densely-developed housing. Houses are more likely to be closer to a major road in urban-dominated communities. On the other hand, the environmental amenities of proximity to a lake, higher elevation, greater access to streams, and greater access to open spaces draw a substantial number of households to rural communities.

Based on the results of our study, growth drivers play out in distinctive ways in different community types. These distinctively different growth drivers imply that growth of an area has to be managed differently according to community type. These findings indicate that as development proceeds, shifts between community types will bring changes in their social structures. These changes will likely give rise to conflict as development proceeds and will have implications for how subsequent development might be organized across a landscape.

One of the weaknesses of the study is in the resolution of the block level in the site-specific housing choice model. Housing choices 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 can be built using a database of individual houses from county tax assessors' offices, the census dataset of block levels, and the GIS database that can be created using information about individual houses. While collecting a dataset from the 98 counties of the Southern Appalachian region would be extremely expensive, a sample study for some selected counties in which all the types of communities are contained might be feasible.

The next step to this research might be to develop predictive models of land use choice that incorporate socio-economic and environmental influences at the micro level. Another direction for further research would be to address the conflict between old settlers and newcomers to the region. This region is increasingly divided into social structures of old settlers and newcomers who move to this area mainly in pursuit of retirement, vacation homes and second homes. The interests of these two groups conflict in many ways, including in the area of housing decisions. The models we used in this study can be modified to investigate the heterogeneity of these two groups in the area.

References

- Alig R. "Econometric Analysis of the Factors Influencing Acreage Trends in the Southeast." *Forest Science* 32(1986):119-34.
- Alonso, W. *Location and Land Use*. Cambridge: Harvard University Press, 1964.
- Anas, A. "Dynamics of Urban Residential Growth." *Journal of Urban Economics* 5(January 1978):66-87.
- Brueckner, J., J. Thisse, and Y. Zenou. "Why Is Central Paris Rich and Downtown Detroit Poor? An Amenity-Based Theory." *European Economic Review* (January 1999):91-107.
- Chomitz K., and D. Gray. "Roads, Land, Markets and Deforestation: A Spatial Model of Land Use in Belitz." *World Bank Economic Review* 10(1995):487-512.
- Dale, V., R. O'Neill, M. Pedlowski, and F. Southworth. "Causes and Effects of Land-Use Change in Central Rondonia, Brazil" *Photogrammetric Engineering and Remote Sensing* 59(1993):997-1005.
- Daniels, T. and D. Bowers. *Holding Our Ground*. Island Press, Washington, D.C. (1997).
- Ding C. "An Empirical Model of Urban Spatial Development." *Review of Urban and Regional Development Studies* 13(November 2001):173-86.
- Fujita, M. "Spatial Patterns of Residential Development." *Journal of Urban Economics* 12(July 1982):22-52.
- Geoghegan, J., L. Weinger, and N. Bockstael. "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS." *Ecological Economics* 23(December 1997):251-64.
- Greene, W. *Econometric Analysis*. Upper Saddle River NJ: Prentice Hall, 1997.
- Hardie, I., and P. Parks. "Land Use with Heterogeneous Land Quality: An Application of an Area Base Model." *American Journal of Agricultural Economics* 79(May 1997):299-310.
- Hardie, I., P. Parks, P. Gottlieb, and D. Wear. "Responsiveness of Rural and Urban Land Uses to Land Rent Determinants in the U.S. South." *Land Economics* 76(November 2000):659-73.

Kmenta, J. *Elements of Econometrics*. New York: Macmillan, 1986.

Lake, I., I. Bateman, B. Day, and A. Lovett. "Improving Land Compensation Procedures via GIS and Hedonic Pricing." *Environment and Planning C: Government and Policy* 18(2000): 681-96.

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

Maddala, G. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.

McFadden, D. "Conditional Logit Analysis of Qualitative Choice Behavior." in P. Zarembka, ed., *Frontier of Econometrics*. New York, Academic Press, 1973.

McMillen, D., and J. McDonald. "Selectivity Bias in Urban Land Value Functions." *Land Economics* 65(November 1989):341-51.

Miller, D., and A. Plantinga. "Modeling Land Use Decisions with Aggregate Data." *American Journal of Agricultural Economics* 81(1999):180-94.

Mills, D. "Growth, Speculation and Sprawl in a Monocentric City." *Journal of Urban Economics* 10(September 1981):201-26.

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

Nechyba, T., and R. Strauss. "Community Choice and Local Public Services: A Discrete Choice Approach." *Regional Science and Urban Economics* 28(January 1998):51-73.

Nelson, G., and D. Hellerstein. "Do Roads Cause Deforestation? Using Satellite Images in Econometric Analysis of Land Use." *American Journal of Agricultural Economics* 79(1997):80-8.

Plantinga, A. "The Effects of Agricultural Policies on Land Use and Environmental Quality." *American Journal of Agricultural Economics* 78(1996):1082-91.

Polinsky, M., and S. Shavell. "Amenities and Property Values in a Model of an Urban Area." *Journal of Public Economics* 5(January-February 1976):119-29.

Rapaport, C. "Housing Demand and Community Choice: An Empirical Analysis." *Journal of Urban Economics* 42(September 1997):243-61.

Spies, T., W. Ripple, and G. Bradshaw. "Dynamics and Pattern of a Managed Coniferous Forest Landscape in Oregon." *Ecological Applications* 4(1994):555-68.

Turner, M., D. Wear, and R. Flamm. "Land Ownership and Land-Cover Change in the Southern Appalachian Highlands and the Olympic Peninsula." *Ecological Applications* 6(1996):1150-72.

U.S. Census Bureau. "Technical Documentation for Census of Population and Housing 1990." Washington, D.C. (1990).

U.S. Forest Service. "The Southern Appalachian Assessment Online Reports" coordinated through Southern Appalachian Man and Biosphere Program (1996).

Wear, D., and P. Bolstad. "Land-Use Changes in Southern Appalachian Landscapes: Spatial Analysis and Forecast Evaluation." *Ecosystems* 1(November 1998):575-94.

Wear, D., and R. Flamm. "Public and Private Disturbance Regimes in the Southern Appalachians." *Natural Resource Modeling* 7(1993):379-97.

Wu, J. "Environmental Amenities and the Spatial Pattern of Urban Sprawl." *American Journal of Agricultural Economics* 83(August 2001):691-97.

Table 1. Definition of Variables

Variable	Definition
Dependent Variables	
Community index	Index for a type of community of urban-dominant, urban-moderate, rural-moderate, rural-dominant
Housing density	Number of houses within 1 km ² of area
Socioeconomic Variables	
Housing value	Median value of owner-occupied houses in \$1,000
Income	Per capita income in \$1,000
Population density	Population within 1 km ² of area
Crime rate	Number of reported crimes, from vehicle theft to murder
Stability	Ratio of occupancies with 5 years or more to total occupancies
Education	Median school years
Political view	Ratio of population with political outlook very conservative and somewhat conservative to total population
Travel time to work	Travel time to work per employee in minutes
Distance to any city	Distance from a center of each block to the nearest city, town or village in km
Distance to major city	Distance from a center of each block to the nearest city with more than 50,000 population in km
Distance to major road	Distance from a center of each block to the nearest primary highway with limited access, interstate highways and toll highways, in km
Road index	Total distance of all roads in km within 1 km ² of area
Environmental Variables	
Distance to major open spaces	Distance from a center of each block to the nearest major open space including national park service land, national forest or other federal land, state or local parks or forests in km
Distance to lakes	Distance from a center of each block to the nearest major lake or reservoir in km
Pollution	NO ₂ level
Elevation	Mean elevation of each block in km
Stream index	Total distance of streams and rivers of each block in km within 1 km ² of area
Open space index	Ratio of total area of major open space to total area of each block

Table 2. Descriptive Statistics of Variables

	Initial	Mean	Min	Max	Std. Dev.
Dependent Variables					
Community index	CI	2.14	1	4	1.14
Housing density (per km ²)	HD	0.16	0	6.15	0.26
Socioeconomic Variables					
Housing value (\$1,000)	HV	59.79	0	330.77	24.66
Income (\$1,000)	IC	12.14	0	64	4.65
Population density (per km ²)	PD	0.36	0.0007	20.04	0.64
Crime rate	CR	78.23	1	558	73.66
Stability (%)	ST	0.59	0	1	0.13
Education (year)	ED	11.54	7.55	16.4	0.87
Political view (%)	PV	0.42	0.22	0.52	0.06
Travel time to work (min)	TW	19.40	2.5	62	5
Distance to any city (km)	DA	5.10	0.05	28.07	3.86
Distance to major city (km)	DM	51.28	0.27	165.69	37.42
Distance to major road (km)	DR	12.00	0	79.82	14.14
Road index (km)	RI	0.019	0	0.08	0.01
Environmental Variables					
Distance to major open space (km)	DO	16.63	0	60.55	10.81
Distance to lake (km)	DL	6.74	0	51.97	6.38
Pollution	PL	88.50	55	129	12.26
Elevation (km)	EL	0.45	0.20	1.34	0.22
Stream index (km)	SI	0.004	0	0.12	0.0028
Open space index (%)	OS	0.004	0	0.97	0.04

Table 3. Parameter Estimates for the Multinomial Logit Model of Community Choices

	Urban-Moderate	Rural-Moderate	Rural-Dominant
Constant	-1.9565**	3.3455**	7.2635**
Education	0.1083**	-0.2896**	-0.6181**
Political view	1.3222*	1.9548**	2.0088**
Population density	-3.9594**	-61.7434**	-166.507**
Crime rate	-0.0027**	-0.0146**	-0.0134**
Stability	1.7098**	2.0769**	4.2697**
Pollution	0.0069	-0.0141*	-0.0195**

Note: Log likelihood, -2938.72; ** indicates statistical significance at the 1 % level; * indicates statistical significance at the 5 % level

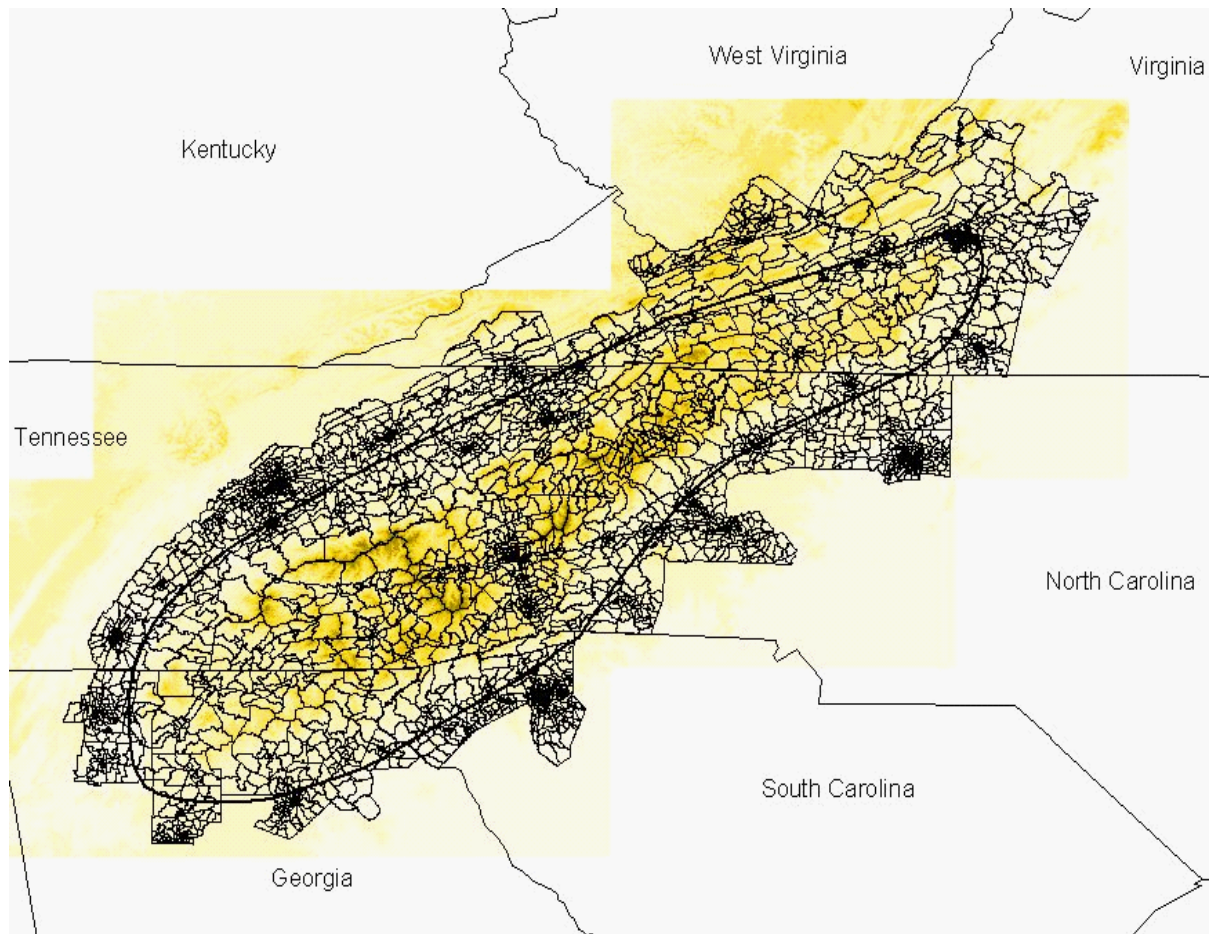
Table 4. Estimated Marginal Effects for Community Choices

	Urban-Moderate	Rural-Moderate	Rural-Dominant
Education	0.8998**	-2.9075**	-3.9944**
Political view	0.3998*	0.7143**	0.4725**
Population density	-1.0263**	-19.3380**	-33.5678**
Crime rate	-0.1521**	-0.9937**	-0.5871**
Stability	0.7263**	1.0661**	1.4107**
Pollution	0.4397	-1.0856**	-0.9664**

Table 5. Parameter Estimates for the Housing Density Equations for Alternative Community Choices

	Urban-Dominant	Urban-Moderate	Rural-Moderate	Rural-Dominant
Constant	-0.3082 [*]	0.1192 ^{**}	-0.0342	-0.0856 ^{**}
Socioeconomic Variables				
Housing value	-0.0016 ^{**}	-0.0003 [*]	-0.0001	-0.00003
Income	0.0066 ^{**}	0.0036 ^{**}	0.0001	-0.00004
Population density	3.4983 ^{**}	3.7256 ^{**}	3.0718 ^{**}	3.1338 ^{**}
Crime rate	-0.0001 [*]	0.00004	-0.00003	-0.00001 [*]
Stability	-0.1283 ^{**}	-0.0250	0.0148	-0.0013
Education	0.0216 ^{**}	-0.0155 ^{**}	0.0008	0.0052 ^{**}
Political view	-0.1289 [*]	-0.0008	0.0093	-0.0063 ^{**}
Travel time to work	0.0032 [*]	0.0010 [*]	0.0001	0.0002 ^{**}
Distance to any city	0.0036	0.0010	-0.0002	0.0001
Distance to major city	0.00002	-0.00002	-0.00002	0.00002 ^{**}
Distance to major road	-0.0024 ^{**}	0.0002	0.0001	0.00003
Road index	5.8219 ^{**}	2.5552 ^{**}	1.9362 ^{**}	1.9115 ^{**}
Environmental Variables				
Distance to major open spaces	0.0002	-0.00005	0.0001	0.00005 [*]
Distance to lakes	-0.0005	-0.0004	-0.0002 ^{**}	-0.00007 ^{**}
Air pollution level	-0.0002	-0.0003	0.0001 ^{**}	0.00002 ^{**}
Elevation	-0.0353	0.0200	0.0175 ^{**}	0.0077 ^{**}
Stream index	-1.3668	-0.5716	2.1383 ^{**}	1.7497 ^{**}
Open space index	0.0749	0.0475	0.0326	0.0171 ^{**}
Self-Selection Variable				
λ	0.0440 [*]	0.0033 [*]	-0.0064 ^{**}	0.0002 ^{**}
System weighted R ²	0.90	0.86	0.84	0.91

Figure 1. Study Area



Appendix

Table A-1. Correlation Coefficients of Variables Considered for Housing Density Model

	PD	CR	ST	ED	PV	PL	HV	IC	TW	DA	DM	DR	DO	DL	EL	SI	OS
PD	1																
CR	0.39	1															
ST	-0.37	-0.19	1														
ED	0.19	0.14	-0.37	1													
PV	-0.	-0.07	0.06	-0.07	1												
PL	-0.08	0.18	-0.08	0.10	0.05	1											
HV	0	0.03	-0.27	0.72	-0.01	0.22	1										
IC	-0.05	0.12	-0.11	0.69	-0.03	0.17	0.79	1									
TW	-0.34	-0.30	0.27	-0.27	0.12	-0.02	-0.11	-0.14	1								
DA	-0.33	-0.23	0.24	-0.15	0.06	-0.03	-0.03	-0.09	-0.14	1							
DM	-0.27	-0.33	0.26	-0.29	0.05	-0.44	-0.26	-0.28	0.07	0.14	1						
DR	-0.26	-0.25	0.16	-0.16	0.05	0.15	-0.07	-0.19	0.17	0.25	0.43	1					
DO	-0.05	0.10	0.06	-0.07	0.08	0.22	-0.06	-0.05	0.06	0.04	-0.05	0.10	1				
DL	-0.04	-0.09	0.12	-0.13	0.01	-0.22	-0.16	-0.15	0.12	0.07	0.21	0.04	0.07	1			
EL	-0.17	-0.26	0.18	-0.07	-0.09	-0.29	-0.09	-0.19	0.09	0.21	0.46	0.35	-0.11	0.17	1		
SI	0	0.02	-0.06	-0.04	0.02	0	0.01	-0.01	-0.06	-0.02	0.03	0.03	-0.06	-0.07	-0.03	1	
OS	-0.04	-0.01	0.02	-0.01	0.01	0.01	0	-0.03	0.02	0.10	0.04	0.03	-0.14	-0.08	0.12	0.01	1
RI	0.74	0.53	-0.35	0.18	-0.10	0.02	-0.01	0.05	-0.49	-0.35	-0.33	-0.35	-0.05	-0.07	-0.21	0.04	-0.08

Table A-2. Parameter Estimates for the Housing Density Equations under Alternative Community Choices without Variables (Income, Education Level, and Road Index) to Check Multicollinearity

	Urban Dominant	Urban Moderate	Rural Moderate	Rural Dominant
Constant	-0.3090 [*]	0.1233 ^{**}	-0.0338	-0.0863 ^{**}
Socioeconomic Variables				
Housing value	-0.0018 ^{**}	-0.0005 [*]	-0.0001	-0.00004
Population density	3.5814 ^{**}	3.7269 ^{**}	3.0725 ^{**}	3.1348 ^{**}
Crime rate	-0.0003 [*]	0.00004	-0.00004	-0.00003 [*]
Stability	-0.1292 ^{**}	-0.0252 [*]	0.0151	-0.0022
Political view	-0.1297 [*]	-0.0009	0.0098	-0.0069 ^{**}
Travel time to work	0.0037 [*]	0.0014 [*]	0.0001	0.0002 [*]
Distance to any city	0.0038	0.0010	-0.0002	0.0002
Distance to major city	0.00003	-0.00003	-0.00004	0.00003 ^{**}
Distance to major road	-0.0026 ^{**}	0.0002	0.0001	0.00003
Environmental Variables				
Distance to major open spaces	0.0002	-0.00006	0.0001	0.00005 [*]
Distance to lakes	-0.0005	-0.0005	-0.0003 ^{**}	-0.00007 ^{**}
Air pollution level	-0.0002	-0.0004	0.0001	0.00003
Elevation	-0.0359	0.0214	0.0179 ^{**}	0.0079 ^{**}
Stream index	-1.3674	-0.5721	2.1391 ^{**}	1.7503 ^{**}
Open space index	0.0753	0.0481	0.0332	0.0176 ^{**}
Self-Selection Variable				
λ	0.0444 [*]	0.0038 [*]	-0.0069 ^{**}	0.0003 ^{**}
System weighted R ²	0.86	0.85	0.82	0.90