Air Qual Atmos Health (2010) 3:41–51
DOI 10.1007/s11869-009-0049-8

Measuring the value of air quality: application of the spatial hedonic model

Seung Gyu Kim · Seong-Hoon Cho · Dayton M. Lambert · Roland K. Roberts

Received: 22 January 2009 / Accepted: 11 August 2009 / Published online: 29 September 2009

Abstract This study applies a hedonic model to assess the economic benefits of air quality improvement following the 1990 Clean Air Act Amendment at the county level in the lower 48 United States. An instrumental variable approach that combines geographically weighted regression and spatial autoregression methods (GWR-SEM) is adopted to simultaneously account for spatial heterogeneity and spatial autocorrelation. SEM mitigates spatial dependency while GWR addresses spatial heterogeneity by allowing response coefficients to vary across observations. Positive amenity values of improved air quality are found in four major clusters: (1) in East Kentucky and most of Georgia around the Southern Appalachian area; (2) in a few counties in Illinois; (3) on the border of Oklahoma and Kansas, on the border of Kansas and Nebraska, and in east Texas; and (4) in a few counties in Montana. Clusters of significant positive amenity values may exist because of a combination of intense air pollution and consumer awareness of diminishing air quality.

Keywords Air quality · Hedonic model · Spatial autocorrelation · Spatial heterogeneity · Total suspended particulates

S. G. Kim (✉) · S.-H. Cho · D. M. Lambert · R. K. Roberts
Department of Agricultural Economics,
The University of Tennessee,
Knoxville, TN 37996-4518, USA
e-mail: sgkim@utk.edu

S.-H. Cho
e-mail: scho9@utk.edu

D. M. Lambert
e-mail: dmlambert@tennessee.edu

R. K. Roberts
e-mail: rrobert3@utk.edu

Introduction

Since the inception of the Clean Air Act (CAA), aggregate emissions of the six principal pollutants, ozone, particulate matter, carbon monoxide, sulfur dioxide, nitrogen dioxide, and lead, were reduced by 25% in the United States, even as U.S. gross domestic product (GDP) increased by 161%, energy consumption increased by 42%, and vehicle miles traveled increased by 149% during the 1970–2002 period (U.S. Environmental Protection Agency (EPA) 2008a). Despite obvious improvements in air quality, the amount of amenity value created by the reduction in pollutants is still controversial (Chay and Greeonstone 2003; EPA 1997; Goklany 1999; Greenstone 2004; Henderson 1996).

Two methods have been used to measure these kinds of amenity values. One is to use a survey-based method such as travel cost or contingent valuation (Alberini et al. 1997; Carson and Mitchell 1993; Krupnick et al. 2002). Hedonic pricing is another approach. Smith and Huang (1995) used a meta-analysis of 37 cross-sectional hedonic studies and found that a decrease in total suspended particulates (TSP) of 1 μg/m^3 results in a 0.05–0.07% increase in property values. From this small increase, many researchers conclude that individuals either place a small value on air quality or the hedonic approach cannot produce reliable estimates of the marginal willingness to pay for air quality improvement (Chay and Greeonstone 2005). For example, Berry (1976) showed that marginal willingness to pay for reduction in TSP was $1.38 (1982–84 dollars) in Chicago in 1968. The estimations of Palmquist (1982, 1983) ranged between −$89.5 and $108.9 in 1982 and between −$76.1 and $98.5 in 1983 (1982–84 dollars) in multiple cities, i.e., Minneapolis, Houston, Chicago, Los Angeles, Philadelphia, and Detroit. These results may be associated with potential bias and loss of efficiency that can result when spatial effects such as spatial autocorrelation and spatial heterogeneity are ignored in the hedonic estimation process (Anselin and Lozano 2007). None of the 37 studies used by Smith and Huang (1995) and the more recent work by Chay and Greeonstone (2005) explicitly modeled spatial effects when estimating marginal willingness to pay for reduced TSP.

In contrast, spatial econometric methods have become commonplace in empirical studies of housing and real estate, leading to spatial hedonic modeling. Reviews of the basic specifications and estimation methods are provided in Anselin (1998), Basu and Thibodeau (1998), Pace et al. (1998), Dubin et al. (1999), Gillen et al. (2001), and Pace and LeSage (2004), among others. The most recent hedonic price studies assumed that housing price at a given location is simultaneously determined by prices of neighboring houses. These studies typically used a spatial process model (Whittle 1954), in which an endogenous variable specifies spatial interactions between prices observed at transaction points plus exogenous variables relating housing attributes and possibly other geographic or demographic variables to sale prices and a disturbance term.

This study uses a spatial hedonic model to estimate the effects of TSP reductions on changes in median housing values at the county level in the lower 48 United States. An instrumental variables approach that combines geographically weighted regression and spatial autoregression methods (GWR-SEM) is adopted to simultaneously account for spatial heterogeneity and spatial autocorrelation. The spatial autoregression method (SEM) corrects for spatial dependency. At the same time, spatial heterogeneity can be addressed by allowing response coefficients to vary across observations using geographically weighted regression.

Empirical model

Under the assumption that the housing market is in equilibrium, a household chooses to reside in a location that maximizes utility as follows:

$$Max\ u[h_i(a_i, g_i), y - p_i(a_i, g_i)], \tag{1}$$

where $u(\cdot)$ is a continuous twice-differentiable utility function with $u'>0$ and $u''<0$; h_i is the flow of housing services from location i, which is a function of a_i (air quality) and g_i (vector of other housing attributes), y is household income; and p_i is the price of house i, also a function of a_i and g_i (Bruecker 1990; Capozza and Helsley 1989). The difference between household income and housing price is total expenditures on commodities other than housing (i.e., a composite numéraire commodity). The utility gained from consuming housing services subject to a budget constraint generates a quasilinear function with respect to all other goods.

The solution to the consumer's utility maximization problem allows the testing of hypotheses about consumer behavior using the hedonic price model. Although Equation 1 is typically applied to individual housing data, interest here is in aggregate housing demand for a group of neighbors. The county is used as the unit of observation in identifying groups of neighbors because the CAA amendments are federally imposed county-level regulations and attainment status is monitored at the county level.

Previous research suggests that air quality measures are potentially endogenous (e.g., Anselin and Lozano 2007; Anselin and Le Gallo 2006; Chattopadhyay 1999; Chay and Greeonstone 2005). Local air quality measures are likely correlated with unobserved local economic factors that also affect housing prices. Therefore, ordinary least square (OLS) estimates may produce biased and inconsistent parameter estimates. TSP attainment status is a reliable instrument because it is correlated with the TSP change and not correlated with the error term in the price equation (Bayer et al. 2006; Chay and Greeonstone, 2005). Following Chay and Greeonstone (2005), a first–difference model using TSP attainment status as an instrument is specified:

$$\Delta P_i = \Delta X_i^{'}\beta + \theta\Delta \mathrm{TSP}_i + \delta R_i + \alpha R_i{*}\Delta \mathrm{TSP}_i + e_i, \tag{2}$$

$$\Delta \mathrm{TSP}_i = \Delta X_i^{'}\mathit{\Pi}_X + Z_{95i}\mathit{\Pi}_Z + \upsilon_i, \tag{3}$$

where ΔP_i is the change in the logarithm of housing price, ΔX_i is the change in a vector of observed characteristics, $\Delta \mathrm{TSP}_i$ is the TSP change (ton/km^2) between 1990 and 2000, R_i is a vector of regional dummy variables to control for region-specific heterogeneity, $R_i{*}\Delta \mathrm{TSP}_i$ is a vector of their interactions with the TSP change, Z_{95i} is the mid-decade TSP attainment status in county i for 1995, and e_i and υ_i are the unobserved determinants of changes in housing price and TSP, respectively.

Mid-decade TSP attainment status (Z_{95i}) is used as an instrument in Eq. 3 to account for potential endogeneity of the TSP change in Eq. 2. The mid-decade attainment status is a better candidate instrument than the attainment designation at the beginning of the decade because a smaller time window is available for general equilibrium responses to affect the composition of households and houses. Because mid-decade attainment status is correlated with large reductions in air pollution and increases in county-level housing prices, its use as an instrument minimizes the bias from omitted variables on the housing price–air pollution gradient. This instrument is also uncorrelated with most observable determinants of housing prices, including economic shocks.

Anselin and Lozano (2007) raised another important issue regarding the spatial structure of housing values in the hedonic model. They suggested that the hedonic model should account for the effects of neighboring housing values with a spatial lag model. Without accounting for spatial lag effects, inference may be compromised because spatial lag dependence yields inconsistent and biased estimates (Anselin 1998). In an attempt to correct potential problems caused by spatial lag dependence, the spatial lag of housing prices is included in Eq. 2:

$$\Delta P_i = \rho W \Delta P_j + \Delta X_i' \beta + \theta \Delta \mathrm{TSP}_i + \delta R_i + \alpha R_i * \Delta \mathrm{TSP}_i + \xi_i \tag{4}$$

where W is the minimum Arc distance weight matrix ensuring that every observation has at least one neighbor; and ρ is the spatial autoregressive parameter explaining the spatial lag dependence between housing prices ($W\Delta P_i$). An instrumental variables approach (or two-stage least squares regression) is used to estimate Eq. 4 (Anselin 1998). The first stage entails regressing all exogenous variables and their first and second spatial lags on the spatial lag of the differenced housing prices (Anselin 1998; Kelejian and Prucha 1999):

$$W\Delta P_i = \delta_1 \Delta \widetilde{X}_i + \delta_2 W \Delta \widetilde{X}_i + \delta_3 W^2 \Delta \widetilde{X}_i + \mu_i, \tag{5}$$

where $\Delta \widetilde{X}_i$ is a matrix containing ΔX_i, $\Delta \mathrm{TSP}_i$, R_i and $R_i * \Delta \mathrm{TSP}_i$. In the second stage, Eq. 4 is estimated using the predicted values of $W\Delta P_i$ from Eq. 5.

GWR is applied to Eqs. 2 and 4 to control for potential spatial heterogeneity at the local level. Equation 2 with GWR is specified as:

$$\Delta P_i = \Delta \widetilde{X}_i' \beta(u_i, v_i) + \varepsilon_i. \tag{6}$$

where ε_i is a random disturbance term; and (u_i, v_i) are location coordinates. Given predicted values of $W\Delta P_i$ from Eq. 5, Eq. 4 can be specified in the GWR framework:

$$\Delta P_i = \rho(u_i, v_i) W \Delta \widehat{P}_i + \Delta \widetilde{X}_i' \beta(u_i, v_i) + \varepsilon_i. \tag{7}$$

The GWR assigns weights to other counties according to their spatial proximity to county i. These weights allow counties in closer proximity to county i to have more influence in the estimation of the local parameters than counties located farther away. An adaptive bi-weight function is used to geographically weigh observations (Fotheringham et al. 2002). The function is "adaptive" in the sense that the trace of the weight matrix is allowed to expand and contract, conditional upon a given location. The bi-weight function for each w_{ij} is:

$$w_{ij} = \left[1 - \left(d_{ij}/d_{\max}\right)^2\right]^2 \text{if } d_{ij} \leq d_{\max}, \text{ otherwise } w_{ij} = 0, \tag{8}$$

where j represents a point in space at which data are observed; i represents any point in space for which parameters are estimated; d_{ij} is the Euclidean distance between point i and j; and $d_{\max}$ is the maximum distance between observation i and its q nearest neighbors (optimal neighborhood bandwidth). The weight attributed to regression point i is one. Weights attributed to the j observations in the neighborhood of i are less than one and become zero when the distance between i and j is greater than $d_{\max}$. Therefore, as d_{ij} increases, the influence of observation j on local regression point i decreases. A cross-validation suggested in Fotheringham et al. (2002; p. 59–62) is used to select the optimal number of neighbors to determine $d_{\max}$ at each location.

To allow for potential correlation between the disturbance terms, the error terms of Eqs. 6 and 7 are assumed to have the structure, $\varepsilon_i = \lambda \sum_{j=1, j\neq i}^{n} w_{ij}\varepsilon_j + \xi_i$, where $\xi_i \sim iid(0, \sigma^2)$; w_{ij} is an element of an n by n row-standardized spatial weight matrix; and λ is a spatial error autoregressive parameter. The GWR residuals are tested for spatial error autocorrelation using a Lagrange multiplier (LM) test (Anselin 1998). In this analysis, a spatial contiguity weight matrix based on Thiessen polygons created by MATLAB was applied to construct the test statistic (Anselin 1998). The statistic is distributed as a χ^2 variate with 1 degree of freedom. The null hypothesis is λ=0.

To test how spatial heterogeneity and spatial autocorrelation can be mitigated by adopting a spatial process model, models with and without the specifics of the spatial process are estimated and compared, i.e., OLS controlling for regional fixed effects (OLS model), GWR to account for spatial heterogeneity (GWR model), and GWR corrected for spatial autocorrelation (GWR-SEM model). The GWR 3.0 (Fotheringham et al. 2002) software was used to calibrate the GWR model. The SAS 9.1(SAS) system was used to estimate the GWR model with spatial lag and error effects because the GWR 3.0 package cannot accommodate specialized error structures.

One of the local measures of spatial associations, Getis-Ord Gi statistics (Gi statistics) is used to identify the clusters of high- and low-marginal effects of the TSP change (Ord and Getis 1995). The analysis is implemented by looking at each county within the context of neighboring counties. The distance for identifying neighboring counties is calculated from Geoda 0.9.5-i as a default minimum threshold distance (Anselin 2005). The local sum of the marginal effects of the TSP change for a county and its neighbors is compared proportionally to the sum of all counties. When the local sum is different from the expected local sum, and that difference is too large to be the result of random chance, those counties sharing similar values are identified as clusters with high (or low) marginal effects of the TSP change (Environmental Systems Research Institute 2007). Once the clusters are mapped, marginal effects of TSP changes are overlaid on the Gi map. By doing so, only negative marginal effects among clusters of negative Gi statistics and only positive marginal effects among clusters of positive Gi statistics are mapped. The clusters of negative marginal effects represent the areas with significant increases in housing prices from reductions in TSP emission density. The clusters of positive marginal effects represent the areas with significant increases in housing prices from increases in TSP emission density. The major clusters focusing on positive amenity values of TSP improvement are characterized by local differences in economies, geographies, demographics, and institutions.

Study area and data

The study area includes the entire continental United States, which consists of 3,107 counties and county equivalents in the 48 States and District of Columbia. After excluding missing observations from five counties, 3,102 counties are used in the empirical model. The empirical model uses four county-level datasets in a geographical information system (GIS): (a) TSP (PM-10, which includes only those particles with aerodynamic diameter smaller than 10 μm) attainment status of 1995 and TSP emission density in 1990 and 2000 from the U.S. EPA (2008b), (b) socioeconomic and housing variables from the 1990 and 2000 County and City Data Books (2003) and the GeoLytics Census CD, (c) 1999 Natural Amenities Scales and Rural–Urban Continuum Code from U.S.D.A. Economic Research Service (U.S. Department of Agriculture (USDA) 2003; McGranahan 1999), and (d) regional dummy variables and their interactions with the TSP variable based on Census Bureau Regions and Divisions with State Federal Information Processing Standard (FIPS) codes (U.S. Census Bureau 2008). All of these variables are joined together by means of county FIPS codes.

Definitions and descriptive statistics for the variables used in the model are presented in Table 1. The change in TSP emission density is chosen to reflect air quality improvement because (1) TSP is the most visible form of air pollution, (2) TSP has the most harmful health effects of all the pollutants regulated by the CAA amendments (Graves et al. 1988; Palmquist and Isangkura 1999), and (3) TSP is highly correlated with other major pollutants.[1] Unlike monitoring data that come from EPA's Air Quality System (EPA 2008c), TSP emission density (ton/km^2) measures all possible sources, for instance, fuel combustion, industrial processes and highway vehicles, and is provided for every U.S. County by EPA's National Emission Inventory (NEI) database (EPA 2008d). They are expressed in ton/km^2 because each county has a different land area. The change in median housing value during the 1990s, instead of the median housing value for 1990 or 2000, is used as the dependent variable because first-differencing the data absorbs the county permanent effects under the framework of the hedonic model (Chay and Greeonstone 2003). Accordingly, all explanatory variables except the natural amenity scale, regional dummy variables, and their interactions with the TSP change are measured as changes between 1990 and 2000.

One concern with using aggregate housing value at the county level instead of at the individual level is that the aggregate values could mask considerable spatial heterogeneity within the county that may be critical to measuring the attributes of housing value. This spatial heterogeneity could introduce some bias. Nevertheless, Chay and Greeonstone (2005) made a case that the aggregation to the county level may not be an important source of bias for the following reasons. First, their estimates generated by aggregation to the county level were similar to the results based on more disaggregated data summarized in Smith and Huang (1995). Second, the aggregation does not lead to the loss of substantial variation in TSP emission density; thus, the bias generated by the spatial heterogeneity of TSP emission density within the county should not be significant. Using the availability of readings from multiple monitors in most counties, they found that only 25% of the total variation in 1970–80 TSP changes was attributable to within-county variation.

Changes in socioeconomic conditions that may affect the change in housing value are represented by changes in income, unemployment, employment in manufacturing, population density, white ratio, senior ratio, population with high school degree, population with college degree, urban population ratio, poverty ratio, and per capita tax. Changes in housing characteristics that may affect the change in housing value include changes in the percentage of houses built in the last 10 years, percentage of houses built 10–20 years ago, percentage of houses built more than 40 years ago, percentage of houses without plumbing, percentage of vacant houses, and percentage of owner-occupied houses. The natural environment and regional characteristics that may influence housing values consist of the natural amenity scale and the rural–urban continuum code. These variables are chosen on the basis of the general hedonic specification.

According to the U.S. Census Bureau (2008), the continental United States is delineated into four regions that consist of nine divisions: New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, and Pacific. Eight regional dummy variables are used to

[1] Other major pollutants such as carbon monoxide, nitrogen dioxide, and volatile organic compound are highly correlated with TSP in the 2000 EPA data (correlation coefficients greater than 0.8).

Table 1 Definitions and descriptive statistics

Variable	Definition	Mean (std. dev.)
Dependent variable		
HVAL	Change in log median housing value during 1990s ($)	0.45 (0.18)
Variable of interest		
TSP	Change in TSP emission density during 1990s (tons/km^2)	2.17 (13.24)
Economic condition variables		
INCOME	Change in household income during 1990s ($)	8,683.37 (2,599.22)
UNEMP	Change in unemployment rate during 1990s	−2.45 (2.34)
MANF	Change in percentage of employment in manufacturing during 1990s	−0.03 (0.04)
Demographic and socioeconomic variables		
POPDEN	Change in population density during 1990s (population/km^2)	14.38 (76.83)
WHITE	Change in percentage of white during 1990s	−0.03 (0.03)
SENIOR	Change in percentage of age above 65 during 1990s	−0.13 (1.44)
HIGHSCH	Change in percentage of persons with high school graduate during 1990s	0.07 (0.03)
COLLEGE	Change in percentage of persons with college graduate during 1990s	0.09 (0.03)
URBAN	Change in percentage of urban population during 1990s	0.04 (0.11)
POVERTY	Change in percentage of persons in poverty during 1990s	0.01 (0.03)
Housing variables:		
BLTTEN	Change in percentage of houses built in last 10 years during 1990s	−0.02 (0.06)
BLTTWTY	Change in percentage of houses built 10–20 years ago during 1990s	−0.09 (0.05)
BLTOLD	Change in percentage of houses built before 1939 during 1990s	−0.03 (0.03)
PLUMB	Change in percentage of houses without plumbing during 1990s	0.00 (0.02)
VACANT	Change in percentage of vacant house during 1990s	−0.01 (0.04)
OWNER	Change in percentage of owner-occupied house during 1990s	0.01 (0.02)
Tax and neighborhood variables		
TAX	Change in per capita taxes ($) during 1990s	341.47 (1,423.48)
Natural environment		
AMENITY	Natural amenity scale	0.05 (2.28)
RURAL	Rural urban continuum code	5.59 (2.72)
Regional dummy variables		
New England	New England = 1, otherwise = 0	0.22 (0.15)
Middle Atlantic	Middle Atlantic = 1, otherwise = 0	0.05 (0.21)
East North Central	East North Central = 1, otherwise = 0	0.14 (0.35)
West North Central	West North Central = 1, otherwise = 0	0.20 (0.40)
South Atlantic	South Atlantic = 1, otherwise = 0	0.19 (0.39)
East South Central	East South Central = 1, otherwise = 0	0.12 (0.32)
Mountain	Mountain = 1, otherwise = 0	0.09 (0.29)
Pacific	Pacific = 1, otherwise = 0	0.04 (0.20)

account for spatial heterogeneity by capturing the regional variation in housing value. The West South Central division is chosen as the reference region (Kennedy 1992).[2]

Changes in the level of TSP emission densities between 1990 and 2000 are mapped in Fig. 1. During this period, TSP emission density declined by 0.07 (tons/km^2) over the continental United States. The TSP emission density decreased most in Kings County, New York by 87.30 (tons/km^2) whereas it increased most in Covington, Virginia by 87.72 (tons/km^2). Figure 2 shows the change in median housing price (2,000 dollars) during the 1990s. The median housing price doubled in 188 of 3,102 counties (6%). It increased between 50% and 100% in 1,842 counties (59%), establishing two major clusters in the eastern and the western regions. It increased by less than 50% in 1,019 counties (33%). It dropped in 53 counties during the 1990s.

[2] The choice of reference region of West South Central Division is arbitrary; however, the marginal effect of TSP calculation is invariant to the choice.

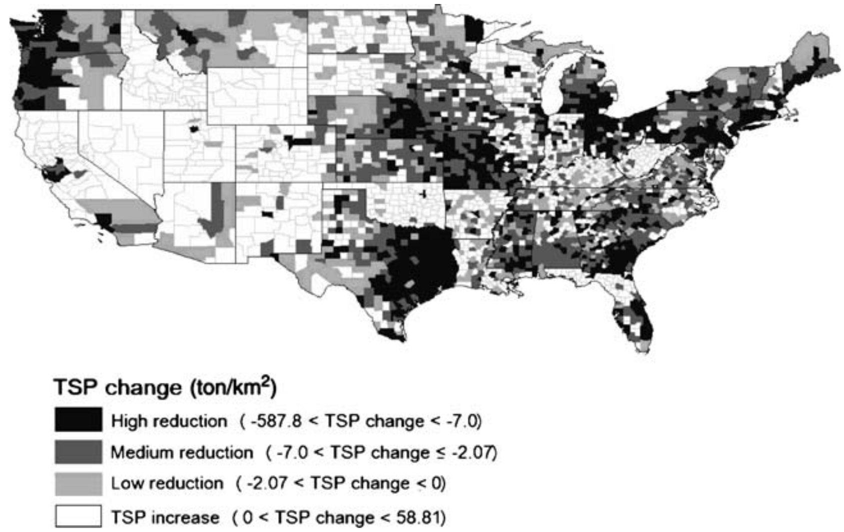

Fig. 1 Changes in TSP emission density between 1990 and 2000 (ton/km^2)

Empirical results

Overall findings

The overall performances of the three models are compared in Table 2. The LM test for spatial error shows that residuals of the OLS and GWR models are spatially autocorrelated. The spatial error LM value (175.0) based on the residuals of the GWR model is reduced by 80% compared to the LM value (896.0) based on the residuals of the OLS model. However, spatial autocorrelation still remains in the residuals of GWR model. The null hypothesis of no spatial error autocorrelation could not be rejected in the GWR-SEM model. The adjusted R^2 for the GWR-SEM model is 0.85, higher than for the OLS model (0.53), and slightly lower than for the GWR model (0.87). The error sum of squares for the GWR-SEM model is 14.4, lower than for the OLS model (47.0) and slightly greater than for the GWR model (13.5). The global F test comparing the global and local models confirms that the local models of GWR and GWR-SEM outperform the global OLS model. The overall fit of the GWR model is slightly better than the GWR-SEM model. However, the GWR-SEM model effectively controls for spatial error autocorrelation, which is still present in the residuals of GWR model.

Parameter estimates for the three models can be obtained by request to the authors. The effect of the TSP change is not trivial due to interactions with the regional fixed effects, and more insight can be gained by calculating the marginal effects of the TSP change by region. The marginal effects of the TSP change on housing prices across all regions for the three models are presented in Table 3.

The marginal effects for the TSP change in the OLS model are insignificant. The median value for the marginal effects from the GWR model is fairly close to zero for all counties; however they vary between −0.837 and 0.811, showing significant spatial variation. The median value for

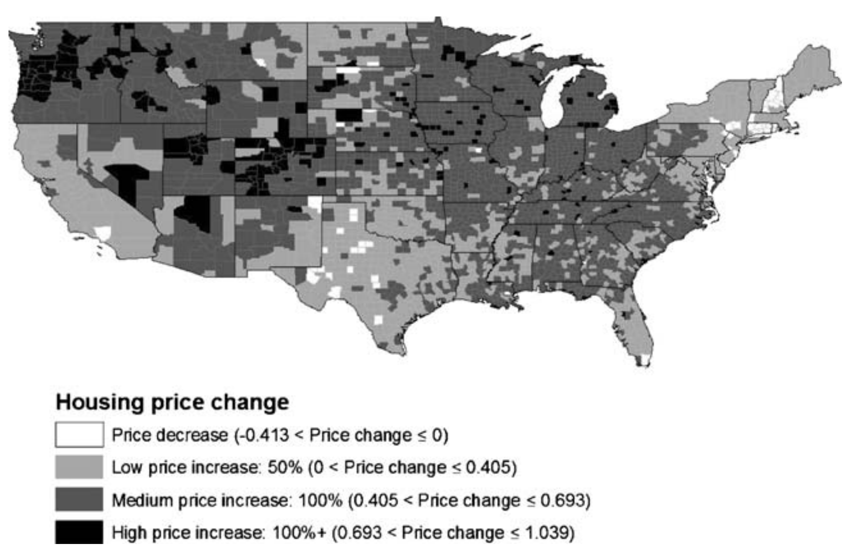

Fig. 2 Changes in log housing price between 1990 and 2000

Table 2 Comparison of overall performance of three models

Statistic	OLS	GWR	GWR-SEM
Adjusted R square	0.53	0.87	0.85
Error sum of squares	47.0	13.5	14.4
Effective parameters [tr(H)]	37	1,196	1,232
Improvement over OLS		33.5	32.6
Degrees of freedom (improvement)		1,158.8	1,195.0
Global F test for global vs. local models		4.1[a]	3.5[a]
Spatial error LM test	896.0[a]	175.0[a]	1.9

OLS, GWR, GWR-SEM represent the estimation results using ordinary least square, geographically weighted regression, and geographically weighted regression corrected for spatial autocorrelation, respectively

[a] Indicates significance at the 0.01% level

the marginal effects of the GWR-SEM model are also close to zero but the variation is smaller than that of the GWR model, between −0.705 and 0.569. The GWR and GWR-SEM models consistently show that the median marginal effect for the lower 48 states is close to zero, whereas local marginal effects vary significantly with different signs in different regions. These different relationships are obscured in the global model. Without the results at the local level from the GWR and GWR-SEM models, the variation in effects associated with the TSP change on housing prices would not be captured. The marginal effects of the TSP change are higher than the average median value in the Middle Atlantic, East North Central, South Atlantic, and East South Central regions.

Cluster analysis

The estimates from the GWR-SEM model are used to identify clusters of areas where the marginal effects of the TSP change on housing price are significantly different from others by overlaying the Gi statistics on the map of marginal effects. The spatial clusters of significant marginal effects at the 5% level are mapped in Fig. 3. The threshold distance of 1.46 decimal degrees (approximately 160 km^2) is used. Four major clusters of areas are found with positive amenity values from TSP reductions. Brief descriptions of each cluster with regards to median housing value, employment in manufacturing, and income are summarized in Table 4.

The largest cluster is in East Kentucky and most of Georgia in the Southern Appalachian area. In this cluster, a decrease in TSP emission density by 1 ton/km^2 increases the average median housing price (hereafter, average housing price) by 5.23%. This cluster can be characterized by successful TSP reductions coupled with a fast-growing economy. The cluster is one of the fastest-growing areas in the United States, with a 61% increase in median housing value, 35% increase in household income, and 21% increase in population density over the period between 1990 and 2000. In spite of a vibrant and growing economy, employment in the manufacturing sector decreased by almost 5% over the same period. The significant reduction in manufacturing likely contributed to the 0.39 ton/km^2 reduction in TSP emission density between 1990 and 2000. For example, no county in Georgia had TSP nonattainment status in 1995. From 1992 through 2002, Georgia participated in the Southern Appalachian Mountains Initiative (SAMI; 2002), a decade-long federal-state collaboration aimed at improving air quality in the Appalachian premier natural areas. Positive impacts on the housing market from TSP reductions in the cluster of counties around the Southern Appalachian area are in accordance with the expected benefits (SAMI 2002) from developing and evaluating potential incentive-based approaches to reducing emissions in the SAMI region.

Another cluster of positive marginal effects on housing prices is in a few counties in Illinois. In this cluster, a

Table 3 Marginal effects of 1990–2000 changes in TSP pollution on changes in log housing prices

Region	OLS	GWR			GWR-SEM		
	Marginal Effect	Min	Median	Max	Min	Median	Max
Overall	0.027	−0.837	−0.001	0.811	−0.705	−0.001	0.569
New England	0.045	−0.062	0.000	0.045	−0.064	0.000	0.019
Middle Atlantic	0.051	−0.061	−0.001	0.135	−0.105	−0.001	0.138
East North Central	0.049	−0.161	−0.001	0.811	−0.141	−0.001	0.473
West North Central	0.045	−0.788	0.000	0.274	−0.279	0.000	0.298
South Atlantic	0.051	−0.142	−0.001	0.141	−0.215	−0.001	0.149
East South Central	0.038	−0.837	−0.001	0.102	−0.705	−0.001	0.076
West South Central	−0.142	−0.377	0.000	0.364	−0.369	0.000	0.569
Mountain	0.051	−0.144	0.000	0.225	−0.185	0.000	0.235
Pacific	0.049	−0.079	0.000	0.216	−0.079	0.000	0.181

OLS, GWR, GWR-SEM represent the estimation results using ordinary least square, geographically weighted regression, and geographically weighted regression corrected for spatial autocorrelation, respectively

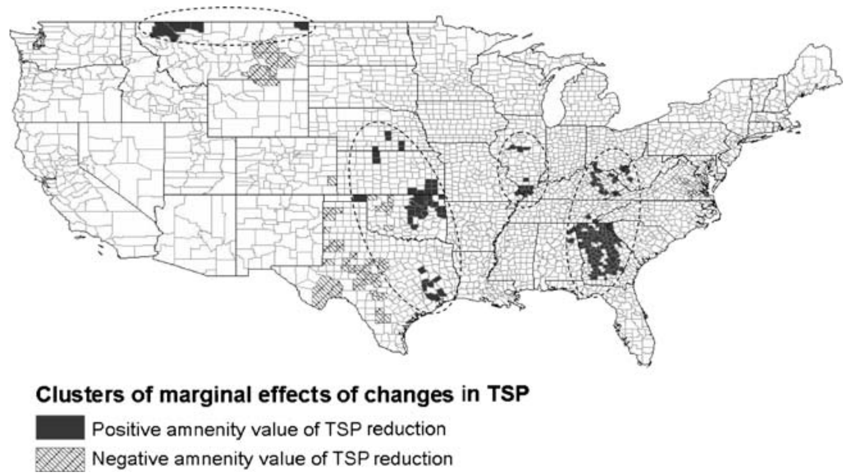

Fig. 3 Clusters of significant marginal effects from changes in TSP emission density between 1990 and 2000

decrease in TSP by 1 ton/km^2 increases the average housing price by 5.62%. This cluster can be characterized by an economy growing at a moderate pace with a 61% increase in median housing value, 44% increase in household income, and 1% increase in population density over the 1990–2000 period. Again, a decline in manufacturing of almost 2% was a likely contributor to the 8.5 ton/km^2 reduction in TSP emission density. Illinois has been regulating motor vehicle inspection and maintenance as a method to control air quality. Regulation of motor vehicles for air quality improvement may be an indication of higher public preference for improved air quality.

Another cluster with positive marginal effects on housing prices includes areas on the border of Oklahoma and Kansas, on the border of Kansas and Nebraska, and in east Texas. In this cluster, a decrease in TSP density by 1 ton/km^2 increases the average housing price by 11.32%. This cluster is characterized by moderate economic growth with a 47% increase in average housing price, 40% increase in household income, and 11% increase in population density. A possible explanation for this cluster is the preponderance of oil industry firms drawn by the natural gas basin in and around the area. Because of intense air pollution from oil processing, residents of the area may be more sensitive to TSP reductions and thus the marginal value of TSP reductions is significantly higher than most other areas. The cluster of counties along the border of Oklahoma and Kansas may be associated with beef-processing firms. This industry's emissions produced severe air pollution, damaging enough to affect the area's residents, thus the marginal value of TSP reductions is significantly higher than in most other areas. Although the TSP reduction in this cluster during the 1990s (−6.0 ton/km^2) is similar to the U.S. average (−5.7 ton/km^2), the TSP emission density

Table 4 Description of four clusters with positive air quality improvement

	Clusters							
	East Kentucky and most of Georgia around Southern Appalachian area		State of Illinois		Borders of Oklahoma and Kansas, and of Kansas and Nebraska, and east Texas		State of Montana	
Marginal effect of 1ton/km^2 decrease in TSP on median housing value	+5.23%		+5.62%		+11.32%		+10.23%	
Year	1990	2000	1990	2000	1990	2000	1990	2000
Median housing value	51,974	84,154	37,458	60,467	38,036	56,049	45,620	73,250
% of employment in manufacturing	25.14	20.29	16.40	14.42	15.60	15.30	5.88	4.40
Household income	23,361	31,597	21,162	30,387	21,351	29,858	22,145	28,828
Population density (population/km^2)	376.06	455.16	164.07	165.98	258.37	286.97	14.25	15.67
TSP emission density (tons/ km^2)	46.22	43.63	40.03	31.40	47.98	41.99	11.87	9.48
Number of counties	121		14		41		4	

(42.0 ton/km^2) is higher than the 2000 average (31.3 ton/km^2). The higher level of air pollution in this cluster may cause households to have higher willingness to pay for TSP reduction.

Another cluster with a positive amenity value for TSP reduction is found in a few counties in Montana. In this cluster, a decrease in TSP density by 1 ton/km^2 increases the average housing price by 10.23%. This cluster is characterized by relatively rural areas with population densities of 13 to 16 persons per km^2 and moderate economic growth with a 61% increase in median housing value, 30% increase in household income, and 25% increase in population density. This cluster can be explained by the 1.5% decline in manufacturing and because the marginal benefit of TSP reductions is higher in communities with relatively low pollution levels. Although the TSP reduction during the 1990s in this cluster (−2.3 ton/km^2) is smaller than the U.S. average, its TSP emission density (11.9 ton/km^2) is only about one third of the 2000 average. The better air quality in this cluster may cause households to have higher willingness to pay to maintain air quality.

Three distinctive clusters with negative amenity values of TSP reductions are found in east Virginia, west and central Texas, and southeast Montana. These clusters are more difficult to explain because they contradict the expectation of a positive marginal value of TSP reductions. Our general justification is that real estate markets are improving in these clusters, while air quality is diminishing. For example, the cluster in southeast Montana experienced significant deterioration in air quality with a TSP emission density increase of 2.8 ton/km^2 during the 1990s, in concert with a booming real estate market in Montana. This unexpected result may be due to the hedonic model not being able to capture all of the amenity values of economic growth, and the air pollution accompanying that economic growth.

Conclusions

Air quality has been evaluated with the hedonic housing price model by numerous researchers. In the midst of mixed results as to whether or not marginal willingness to pay for air quality improvement is significant, this study uses the hedonic model to estimate the amenity value of TSP reduction on changes in median housing values at the county level. An instrumental variables approach that combines geographically weighted regression and spatial autoregression methods is adopted to simultaneously account for spatial heterogeneity and spatial autocorrelation.

The median value for the marginal effects of TSP changes are close to zero in the GWR-SEM model but significant spatial variation exists in the marginal effects. Using a Gi Statistics of the marginal effects, the clusters of significant positive and negative amenity values from TSP reductions are identified. Positive amenity values from TSP reductions are found in four major clusters: (1) East Kentucky and most of Georgia around the Southern Appalachian area; (2) a few counties in Illinois; (3) on the border of Oklahoma and Kansas, on the border of Kansas and Nebraska, and in east Texas; and (4) a few counties in Montana.

The reasons for the clusters of significant positive amenity values may be different for different clusters. The first cluster is explained by successful TSP reductions coupled with a fast-growing economy and a significant decline in manufacturing; the second cluster is explained by awareness of diminishing air quality; the third cluster is explained by higher willingness to pay for TSP reductions in an area with poor air quality; and the fourth cluster is explained by higher willingness to pay for maintaining air quality in an area with good air quality. Surprisingly, negative amenity values of TSP reductions are found in the three distinctive clusters of eastern Virginia, western and central Texas, and southeastern Montana. This unexpected result may be explained by worsening air quality with a booming real estate market and the inability of the model to capture all of the amenity values of economic growth and the resulting air pollution.

The finding of spatial heterogeneity in the marginal amenity values from TSP reductions in this study could be used to amend the CAA at the local level, i.e., state or regional level. Historically, the CAA amendments are federally imposed mainly because of equity issue. However, a need exists to update the CAA at the local level. According to 'Update on Clean Air Act Issues', the Midwest and Southern states claim that air quality problems are local, and more stringent nitrogen oxide emission standards announced by the EPA in September, 1998 are an unfair cost burden (American Geological Institute 2000). Local estimates of the marginal willingness to pay for TSP reductions, such as those generated in this study, should prove useful as input into future debates about amendment of the CAA at the local level. For example, higher marginal effects of TSP reductions mean higher marginal willingness to pay for TSP reductions. This finding implies that households in areas of higher marginal willingness to pay may be more amenable to supporting more stringent air quality requirements.

Two caveats should be mentioned when applying the hedonic model using the GWR-SEM framework with aggregate housing value at the county level. First, the clusters are found mostly in the eastern regions of the United States. A contributing factor to this phenomenon might be that the relatively small counties in the eastern regions fit the spatial hedonic model better than the larger western counties. Because the locations of specific counties are proxied by county centroids in establishing the weight matrix in the GWR-SEM model, the centroids of larger counties represent larger areas. The larger the area represented by the centroid, the wider and the larger the area represented by the optimal bandwidth, and the smaller the spatial heterogeneity inherent in the variables. Uneven county sizes may be a disadvantage of spatial analysis with county-level data. Further analysis may compare differences between the states with similar county sizes; one for the western and one for the eastern United States. Second, the hedonic model is only able to capture those benefits from air quality improvements that are jointly consumed with housing. Air quality improvements captured at other places, such as the workplace and when commuting to work, may not be captured in the price of the house. Also, commuters from other air basins will enjoy some of the benefits if they work in the area without capitalizing those benefits into housing prices. In other words, the hedonic method only captures a portion of the air quality benefits.

References

Alberini A, Cropper M, Fu T-T, Krupnick A, Liu J-T, Shaw D, Harrington W (1997) Valuing health effects of air pollution in developing countries: the case of Taiwan. J Environ Econ Manage 34:107–126

American Geological Institute (2000) Update on clean air act issues (http://www.agiweb.org/gap/legis106/cleanair.html)

Anselin L (1998) Spatial econometrics: methods and models. Kluwer, Dordrecht

Anselin L (2005) Exploring Spatial Data with GeoDa^{TM}: A Workbook (revised version). Spatial analysis laboratory, Department of Geography, University of Urbana-Champaign, Urbana, IL

Anselin L, Le Gallo J (2006) Interpolation of air quality measures in hedonic house price models: spatial aspects. Spatial Econ Anal 1:31–52

Anselin L, Lozano N (2007) Errors in variables and spatial effects in hedonic house price models of ambient air quality Working Paper 2007-01, University of Illinois

Basu S, Thibodeau TG (1998) Analysis of spatial autocorrelation in house prices. J Real Estate Finance Econ 17:61–85

Bayer P, Keohane N, Timmins C (2006) Migration and hedonic valuation: the case of air quality. National Bureau of Economic Research, Inc in its series NBER Working Papers with number 12106

Berry BJL (1976) Ghetto expansion and single-family housing prices: Chicago, 1968–1972. J Urban Econ 3:397–423

Bruecker JK (1990) Growth controls and land values in an open city. Land Econ 66:237–238

Capozza DR, Helsley RW (1989) The fundamentals of land prices and urban growth. J Urban Econ 26:295–306

Carson RT, Mitchell RC (1993) The issue of scope in contingent valuation studies. Am J Agric Econ 75:1263–1267

Chattopadhyay S (1999) Estimating the demand for air quality: new evidence based on the Chicago housing market. Land Econ 75:22–28

Chay KY, Greeonstone M (2003) Air quality, infant mortality, and the clean air act of 1970 NBER Working paper 10053

Chay KY, Greeonstone M (2005) Does air quality matter? Evidence from the housing market. J Polit Econ 113:376–424

County and City Data Books (2003) The University of Virginia, Geospatial and Statistical Data Center. (http://fisher.lib.virginia.edu/collections/stats/ccdb)

Dubin R, Pace RK, Thibodeau TG (1999) Spatial autoregression techniques for real estate data. J Real Estate Liter 7:79–95

Environmental Systems Research Institute (2007) How hot spot analysis: Getis-Ord Gi* (Spatial Statistics) works. (http://edndoc.esri.com/arcobjects/9.2/NET/shared/geoprocessing/spatial_statistics_tools/how_hot_spot_analysis_colon_getis_ord_gi_star_spatial_statistics_works.htm)

Fotheringham A, Brunsdon C, Charlton M (2002) Geographically weighted regression: the analysis of spatially varying relationship. Wiley, West Sussex

Gillen K, Thibodeau TG, Wachter S (2001) Anistrophic autocorrelation in house prices. J Real Estate Finance Econ 23:5–30

Goklany IM (1999) Clearing the air: the real story of the war on air pollution. Cato Institute, Washington, DC

Graves P, Murdoch JC, Thayer MA, Waldman D (1988) The robustness of hedonic price estimation: urban air quality. Land Econ 64:220–233

Greenstone M (2004) Did the clean air act amendments cause the remarkable decline in sulfur dioxide concentrations? J Environ Econ Manage 47:585–611

Henderson JV (1996) Effects of air quality regulation. Am Econ Rev 86:789–813

Kelejian HH, Prucha IT (1999) A generalized moments estimator for the autoregressive parameter in a spatial model. Int Econ Rev 40:509–533

Kennedy PA (1992) Guide to econometrics, 3rd edn. MIT Press, Cambridge, MA, p 369

Krupnick A, Alberini A, Cropper M, Simon N, O'Brien B, Heintzelman M (2002) Age, health and the willingness to pay for mortality risk reductions: a contingent valuation survey of Ontario residents. J Risk Uncertain 24:161–186

McGranahan D (1999) Natural amenities drive rural population change FRED Report No. 781, US Department of Agriculture, Economic Research Service, Washington, DC

Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geograph Anal 27:286–306

Pace RK, LeSage JP (2004) Spatial statistics and real estate. J Real Estate Finance Econ 29:147–148

Pace RK, Barry R, Clapp JM, Rodriguez M (1998) Spatial autocorrelation and neighborhood quality. J Real Estate Finance Econ 17:15–33

Palmquist RB (1982) Estimating the demand for air quality from property value studies. Manuscript. North Carolina State Univ, Raleigh

Palmquist RB (1983) Estimating the demand for air quality from property value studies: further results. Manuscript. North Carolina State Univ, Raleigh

Palmquist R, Isangkura A (1999) Valuing air quality with hedonic and discrete choice models. Am J Agric Econ 81:1128–1133

Smith VK, Huang JC (1995) Can markets value air quality? A meta-analysis of hedonic property value models. J Political Econ 103:209–227

Southern Appalachian Mountains Initiative (SAMI) (2002) Final Report. (http://www.gefa.org/Modules/ShowDocument.aspx?documentid=265)

US Department of Agriculture (2003) Urban-rural continuum code (http://www.ers.usda.gov/Data/RuralUrbanContinuumCodes)

US Environmental Protection Agency (1997) The benefit and costs of the clean air act 1970-1990. Report to Congress Mimeograph

US Environmental Protection Agency (2008a) EPA's report on the Environment (ROE). (http://www.epa.gov/indicators/roe/html/tsd/tsdAir.htm)

US Environmental Protection Agency (2008b) Airdata (http://www.epa.gov/air/data/geosel.html)

US Environmental Protection Agency (2008c) EPA's air quality system (http://www.epa.gov/ttn/airs/airsaqs)

US Environmental Protection Agency (2008d) EPA's National Emission Inventory (NEI) database (http://www.epa.gov/ttn/chief/eiinformation.html)

US Census Bureau (2008) Census bureau region and division codes (http://www.census.gov/popest/geographic/codes02.html)

Whittle P (1954) On stationary processes in the plane. Biometrika 41:434–449