The Negative Effect of Traffic Noise on House Prices: A Landscape Hedonic Analysis

Abstract

One consequence of the expanding road network and its associated traffic is increased levels of traffic noise. While the hedonic literature has consistently shown a negative effect of this phenomenon on the real estate market, research in the United States has often relied on crude measures of traffic noise. Here, we reduce the measurement error of traffic noise exposure through a detailed model of noise propagation over the landscape. Additionally, we estimate the impact on single family home transactions throughout the St. Paul, Minnesota, urban area using spatially explicit local regression techniques to allow for spatial non-stationarity in the hedonic function.

1 Motivation and Past Research

Prolonged exposure to traffic noise affects people in a number of ways, ranging from simple annoyance (Miedema and Oudshoorn, 2001; Ouis, 2001; Ohrstrom et al., 2007; de Kluizenaar et al., 2013; Weinhold, 2013), to sleep disturbance (Netherlands, 2004), to increasing risk for stroke (Sorensen et al., 2011), hypertension (Jarup et al., 2008; Bodin et al., 2009), myocardial infarction (Babisch et al., 2005), and overall quality of life (Shepherd et al., 2013). The noise level at which such effects are observed does not have to be high. It has been shown that people exposed to traffic noise with a 24-hour average of 55 decibels (dBA) are found to be at a higher risk for hypertension (Barregard et al., 2009; Bodin et al., 2009), and those exposed to 60 dBA or greater are found to be at a higher risk for stroke (Sorensen et al., 2011).

Automobiles are already perhaps the greatest source of noise in residential neighborhoods (Barber et al., 2010) and traffic noise was estimated to cause around three billion dollars worth of external damage to the United States economy alone in 1991 (Delucchi and Hsu, 1998). Traffic noise and its attendant problems are greater now than they were in the 1990s and predicted to worsen in the future (Goines and Hagler, 2007). Globally, Dargay et al. (2007) predict the number of personal vehicles to grow from roughly 800 million in 2002 to over two billion units by 2030. Additionally, the global trend of increased urbanization is expected to continue, with the proportion of the population living in cities growing from roughly half now to over two-thirds by the middle of the century (United Nations, 2012).

One way to estimate some of the costs of traffic noise (and therefore the economic benefits of policies that reduce noise) is through hedonic regression, examining the relationship between house prices and noise levels. Nelson (1982) was the first to review some of the earliest noise hedonic work in the United States and Canada. The review of nine studies found an average estimated impact of traffic noise to be a reduction in house prices of approximately 0.4 percent per additional decibel. More recently, Nelson (2008) notes that many of these early studies (such as Gamble et al., 1974; Langley, 1976) are plagued by issues of small sample sizes (in the hundreds) and limited spatial extents (one or a handful of cities).

Reviews of the research by Bateman et al. (2001) found that the negative effect of traffic noise on house prices ranged from 0.08 to 2.22 percent per decibel with a mean around 0.55 percent per decibel, while Navrud (2002) found an average decrease of 0.64 percent per decibel. Much of the recent research in this literature has tended to focus on the European experience with noise, often due to the availability of high quality traffic noise data made available from the government, especially the United Kingdom (Day et al., 2007; Blanco and Flindell, 2011), Sweden (Wilhelmsson, 2000; Andersson et al., 2010), Switzerland (Baranzini et al., 2010), Spain (Duarte and Tamez, 2009), Netherlands (Theebe, 2004), and Germany (Brandt and Maennig, 2011).

Instead of focusing on the impact of automobile traffic noise, hedonic work in the United States has tended to concentrate on noise from airports (Espey and Lopez, 2000; McMillen, 2004; Cohen and Coughlin, 2008) or indirect measures of traffic noise, such as proximity to highways (Matthews and Turnbull, 2007; Chernobai et al., 2009; Li and Saphores, 2012), or traffic counts (Hughes Jr. and Sirmans, 1992; Larsen, 2012). One exception is Cheng (2008), who estimated that year 2007 housing prices in Louisville, Tennessee were, on avarege, 0.34 percent lower for each additional decibel of traffic noise. However, the generalizability of this research is limited given that it is based on less than a thousand house transactions in just one city.

A major reason for the thinness of the scholarly literature on this topic in the US lies in the difficulty of modeling traffic noise propagation over the landscape. To properly analyze the spatial association between real estate prices and exposure to traffic noise, it is necessary to create a noise surface map at sufficiently detailed spatial resolution to account for the complex and heterogeneous interaction between the noise source and the resistance of the landscape to noise propagation. Implementing such a model can be very difficult. The data needed for the model is very extensive and may not even be readily available (e.g., building footprint and height data). Furthermore, it is computationally very intensive. Fortunately, recent developments in Geographic Information Systems and distributive computing have reduced these difficulties, making it much easier to create a noise surface map at landscape level with high spatial resolution.

In this paper, we estimate a hedonic relationship between traffic noise and single family house prices across 44 cities and several hundred square miles in the St. Paul, Minnesota, urban region by taking advantage of a uniquely detailed set of noise estimates. The data represent perhaps the most accurate regional noise estimates ever used in the hedonic literature in the United States to date due to the high spatial resolution of the estimates (10 m x 10 m) and the incorporation of the effects of nearby buildings and land cover on noise propagation. More detail about the noise model inputs and procedures are presented in the next section of the paper.

We construct a flexible hedonic model that allows for a non-stationary relationship between house prices and our explanatory variables over geography and time. Such Locally Weighted Regression (LWR) models have become increasingly common in published work (see Duarte and Tamez, 2009; Carruthers and Clark, 2010; Sunding and Swoboda, 2010; Nappi-Choulet and Maury, 2011) and have been shown to be more accurate than other common spatial econometric models under common real world circumstances (McMillen, 2012). Thus, in addition to using new data that reduces the measurement error of the impact of traffic noise, the spatial and temporal variation in our data and hedonic modeling allow us to present the first estimates in the United States of how the marginal effect of traffic noise varies over space and time.

We find strong evidence to support hedonic modeling choices that allow for spatial and temporal heterogeneity. Our paper also presents the first estimates of the implicit price of traffic noise after the housing market collapse and economic recession of 2008-09. Contrary to work such as Cho et al. (2011), which found that environmental amenities mattered *less* after the housing crash, our estimated noise coefficients roughly double in magnitude after the recession compared to before. We conclude with a discussion of how future research can expand and improve upon this work.

2 Study Area and Data

The 2010 US Census lists the population of the Twin Cities Metropolitan Region (Minneapolis and St. Paul and their surrounding areas) as almost 3 million residents spread over seven counties. This study examines single family residential home transactions in the Census-defined urban areas of three of the seven counties: Dakota, Ramsey and Washington County (see Figure 1). We obtained sales data from approximately forty thousand transactions between 2005 and 2010 (n=42,095) from the 2010 MetroGIS Regional Parcel Dataset published by the Metropolitan Council. Figure 1 shows an overview of the study area as well as the spatial distribution of the house sales prices.

In addition to the geographic location and date of the house sale, we collected or calculated structural and locational variables commonly used in the hedonic literature, such as the size of the house, lot size, and architectural style of the house, as well as distances to amenities like the Central Business District, parks, and lakes. Table 1 provides a brief description of the variables in our data and some basic summary statistics. Table 2 displays a simple correlation matrix of the quantitative variables.

[Figure 1 about here.]

[Table 1 about here.]

[Table 2 about here.]

2.1 Noise Data

To determine the relationship between real estate price and road traffic noise levels, we first created a traffic noise exposure surface by calculating the propagation of traffic noise over the landscape using the FHWA (Federal Highway Authority) 1978 standard (Barry and Regan, 1978). The entire methodology has already been explained in detail in (CITE two NEGA PAPERS) What follows is a summary. The modeling of traffic noise has three key components: choosing the mathematical function for noise propagation, assembling the input data, and assessing the accuracy of the model prediction. These will be briefly discussed. Because it is currently used by the state of Minnesota, we used the FHWA 1978 standard model for calculating the noise level. According to this standard, the noise level at any given location on the landscape can be expressed:

$$\begin{aligned} L_{EQ}(i) &= \bar{L}_0(i) + 0.115\sigma_i^2 + 10log\frac{N_i \pi D_0}{T * S_i} + 10log\left[\frac{D_0}{D}\right]^{1+\alpha} \\ &\quad +10log\left(\frac{\psi_{\alpha(\phi_1,\phi_2)}}{\pi}\right) + \Delta_{gradient} + \Delta_{shielding} \qquad (1) \end{aligned}$$

where $L_{EQ}(i)$ is A-weighted hourly energy equivalent noise level in decibels (dBA), which is calculated for each class i of vehicle (automobile and trucks); $\bar{L}_0(i)$ is the mean Sound Pressure Level (SPL) at the reference distance for class i; σ_i is the standard deviation of the SPL for each class of vehicle; N_i is the number of vehicles of the i^{th} class passing during the relevant hour; D_0 is the reference distance (usually 15 m); D is the perpendicular distance from the road center line to the receiver; α is a site parameter (soft and hard surface), $0 < \alpha < 1$; S_i is the mean speed of the i^{th} class; T is the duration, usually 1 hour; ϕ_1 and ϕ_2 are the angles from the perpendicular of the limits of the observer's view of a section of the road. They are used to account for only the energy coming from a portion of the roadway; $\Delta_{gradient}$ is an adjustment for road surface gradient; $\Delta_{shielding}$ is a shielding adjustment (land cover, buildings, noise barriers).

We used several data sources as inputs in the study. A 2007 road centerline and the associated traffic characteristics (volume and proportion of trucks and vehicles) for the region were obtained from the Minnesota Department of Transportation (MNDoT). We converted the posted speed of each road segment into a GIS layer. We used three data sources for the shielding adjustment: buildings, foliage, and noise barriers. We extracted the perimeter and height of 818,500 buildings using a combination of LiDAR data and aerial photography. Foliage shielding is accounted by extracting forest polygons that are at least 10 m high and have an area $\geq 10,000m^2$ from a 2005 land use and 2001 National Land Cover Data. We obtained a noise barrier layer from the MNDoT that contains the locations, width, and height of each noise barriers to account for its shielding effect.

We used the SoundPlanTM noise modeling software to implement equation1. Predicted noise output is calculated at a grid resolution of 10 m^2. We conducted a preliminary validation of the model output by comparing it with observed noise levels, which were sampled at 134 locations along a major highway. The relationship between the mean observed and predicted noise levels was linear and moderately strong with a correlation of 0.76.

In order to incorporate noise produced by aircraft into the overall noise exposure, we obtained the aircraft noise contour lines of the region from the Twin Cities Metropolitan Airports Commission. The contour line noise values coincident with the traffic noise surface then were added following rules of logarithmic addition. For example, adding 52 dBA from aircraft contour line into the noise surface where the noise value is 51 yields 53.5 dBA.

A map of our noise surface can be seen in Figure 1. Once the noise surface for the entire region was developed, we extracted the noise surface for the present study area by overlaying the housing parcel boundaries and calculating the mean noise level within the parcel.

2.2 Structural Attributes

According to a review by Wilhelmsson (2000), the five most common structural attributes included in real estate hedonic pricing studies are living area, number of bathrooms, age, garage and lot size. The 2010 MetroGIS Regional Parcel Dataset includes structural data on living area, age, garage presence, lot size, owner-occupancy and house architectural style for every transaction. However, standard additional control variables like the number of bathrooms, bedrooms, and size of garage were not available through this source. We were able to obtain additional data for the number of bedrooms, bathrooms, and size of the garage for a subset of our house sales from the Dakota County Assessor's Office. In section 5 we show that the inclusion of these variables in our model does not significantly change our results for these areas. Therefore, we feel confident in our results even without these independent variables for most of our study area.

2.3 Other Locational Attributes

A common real estate adage states that the three most important real estate attributes are: location, location and location. Knowing where each house is located allows us to also construct a vector of other attributes associated with the sales transaction. For instance, using GIS software we are able to calculate the Euclidean distance to numerous points of interest within the dataset, such as the nearest central business district, shopping centers, parks, and lakes. Additionally, three demographic variables denoting the median household income for the surrounding census tract, the percentage of the population that identified their race as "white", and the percentage of the population that is under the age of 18 were created through the use of TIGER shapefiles from the 2010 Census Bureau and data from the 2010 American Community Survey. Lastly, we associate each transaction with its elementary school and include the average 3rd grade Minnesota Comprehensive Assessment (MCA) score for the local elementary school during the year of purchase.[1]

2.4 Time

We have sales transactions across six years (from 2005 to 2010) for our study area. Such temporal variation will allow us to estimate the hedonic price function over time and test for differences before, during, and after the economic recession. Table 3 displays the mean variable values by year of sale to look for differences over time. The mean sale price in 2009 (roughly $230,000) is almost 20 percent lower than the mean sale price in 2006 (roughly $280,000) and the number of sales transactions in 2009-10 is less than the number of transactions in 2005. However, most other variables have small or no differences across time.

[Table 3 about here.]

We also created univariate density plots of the variables over time to check for differences in the shape or spread of the distributions over time (see Figure 2). It is not the case that after the recession the distribution "hollowed out" in the middle and became bimodal with many high- and low-priced houses selling with few mid-priced homes selling. Additionally, maps of the locations of sales transactions over time reveal no discernable patterns, either (see Figure 3). The geographic dispersion of sales in earlier years looks surprisingly similar to that of later years. Thus, we can find no evidence of only certain types or locations of houses transacting in later years compared to earlier years in our data.

[1] Test scores were obtained from the Minnesota Department of Education website. The school district and elementary school attendance boundary spatial information were obtained from the Minnesota Geospatial Information Office Clearinghouse Data Catalog.

[Figure 2 about here.]

[Figure 3 about here.]

3 Basic Econometric Model

Our aim is to estimate the marginal willingness to pay for different attributes, in particular changes in the traffic noise associated with the house. Consistent with past research, this study implements a semi-logarithmic hedonic pricing model. Before implementing the Locally Weighted Regression model, we present the results of a simpler econometric model given by equation (2).

$$ln \text{ Sale Price}_i = \beta_0 + \beta_1 \text{Noise}_i + \beta_2 S_i + \beta_3 N_i + \beta_4 T_i + \text{error}_i \tag{2}$$

Noise_i is the noise level for house i, S_i is a vector of the house's structural attributes, N_i is a vector of the neighborhood attributes, and T_i is a vector of time fixed effects. Due to the semi-logarithmic functional form of the model, we can interpret the regression model coefficients as the price semi-elasticities of the underlying attributes. For instance, we can interpret the coefficient on noise as the percentage increase in price for a one decibel increase in the traffic noise associated with the transaction in our dataset.

[Table 4 about here.]

Model (A) in Table 4 displays the results of a basic Ordinary Least Squares regression model. In addition to the important structural and neighborhood variables, we also include fixed effects for each city and month of sale in the dataset to allow for month-to-month shocks to prices. The traffic noise coefficient suggests houses in our data that are similar in all respects but a one decibel difference in noise will, on average, have sales prices roughly 0.27 percent lower at the noisier location.

Further exploratory analysis suggests the assumption of temporal and geographic stationarity of the hedonic price function is overly restrictive and needs to be relaxed. We first tested for temporal changes in the coefficients for house size, lot size, and noise by creating interaction terms with these variables and the month*year year fixed effects. In all three cases the improved model fit was significantly more than expected due to random chance.[2] Columns (B) - (D) of Table 4 show the results of restimating Model (A) using only sales within three separate years of data. Subsequent analysis also suggests that many of the important regression coefficients vary across the different cities within our data. We are able to reject the null hypothesis that the interaction terms between cities and important variables (house size, lot size, noise, etc.) are zero, even after limiting the analysis to only houses sold within a given year (for instance, by adding the interaction terms to model specifications (B) - (D)).[3]

These results should be taken with appropriate skepticism, as the models are simple Ordinary Least Squares regressions and the spatial and temporal nature of the data likely violate assumptions of the Classical Linear Regression Model. However, the initial findings seem to suggest that the relationship between house prices and many other hedonic characteristics may vary over space and time. While methods exist for parameterizing this variation (such as spatial expansion as suggested by Casetti, 1972), we have no a priori knowledge of how to parameterize the variation. As such, we turn to a semi-parametric form of hedonic regression, a flexible modeling approach which lets the data reveal how relationships vary, rather than specifying them beforehand. Such a modeling choice will allow us to perform additional tests of the hypotheses of spatial and temporal stationarity in the regression model.

[2]The results of specific hypothesis tests are available upon request.
[3]Results of specific F-tests are available upon request.

4 Locally Weighted Regression Model

Locally Weighted Regression (LWR) techniques (also known as Geographically Weighted Regression) are described in detail by Cleveland and Devlin (1988), Brunsdon et al. (1998), Fotheringham et al. (2002), and others. It is a weighted least squares methodology in which regression coefficients are estimated over space as a function of the local data as described in Equation (3),

$$\hat{\beta}_i = (X'W_iX)^{-1}X'W_iY, \tag{3}$$

where X is a $n \times m$ matrix of independent variables, W_i is the $n \times n$ weights matrix, and Y is the $n \times 1$ vector of dependent variable values. The weights matrix, W_i is a diagonal matrix where element w_{jj} denotes the weight that the j^{th} data point will receive in the regression coefficients estimated at location i in the dataset. We employ a bi-square weights function and a k-nearest neighbor bandwidth approach as described in equation (4),

$$w_{jj} = \left[1 - \left(\frac{d_{ij}}{d_k}\right)^2\right]^2 \text{ if } d_{ij} < d_{ik}, \text{ otherwise} = 0, \tag{4}$$

where d_{ij} denotes the distance between observations i and j, and d_{ik} is the distance from observation i to the k^{th} nearest observation. This function assigns weights close to 1 for data points near observation i, weights positive but closer to zero for observations farther away, and zero for all $n-k$ observations farther away than the k^{th} nearest observation. We estimate LWR coefficients using bandwidths ranging from as small as $k = 50$ observations and as large as 10,000 observations. We choose the LWR bandwidth my minimizing the Generalized Cross Validation score as detailed in equation (5),

$$n * \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{(n - v_1)^2}, \tag{5}$$

where y_i is the dependent variable value, $\hat{y}_i$ is the predicted dependent variable value for observation i, and v_1 is the "effective number of model parameters." The value $v_1 = \text{tr}(\mathbf{S})$, where the matrix $\mathbf{S}$ is the "hat matrix" which maps y onto $\hat{y}$,

$$\hat{y} = \mathbf{S}y,$$

and each row of $\mathbf{S}$, r_i is given by:

$$r_i = X_i(X'W_iX)^{-1}X'W_i.$$

The GCV score is a convenient model selection metric that rewards models that provide a good fit to the data, while penalizing models with a greater number of model parameters (Loader, 1999; McMillen and Redfearn, 2010). Model selection via this strategy has been shown to discern whether spatially varying relationships exist, accurately estimate spatially varying coefficients, even outperform other spatial econometric techniques, and do so without wasting degrees of freedom (Paez et al., 2011; McMillen and Redfearn, 2010; McMillen, 2012).

Similar to the basic econometric model described in Section 3, the LWR model estimates the logged sales price of a house as a function of structural and neighborhood variables, location and time. In order to account for changing market conditions in our data as manifested in regression coefficients that vary over time, when estimating LWR coefficients we use only houses sold within the past 12 months and also include month fixed effects to account for seasonality. Thus, the coefficient estimates at a particular house are estimated using data from other sales nearby in both time and geography.

4.1 Improved Model Performance

The use of the Locally Weighted Regression model significantly increases model performance as measured by the generalized cross validation score. Figure 4 displays the Generalized Cross Validation score for three different model specifications estimated at the "global" level as well as with varying numbers of observations included in the regression bandwidth. We estimate three different hedonic function specifications with different levels of additional spatial control variables. Model (1) includes the noise variable, the basic structural characteristics (size of the house in square feet, size of the lot in acres, a categorical variable denoting the architectural style of the house, and whether or not the house is owner-occupied) plus month of sale*year fixed effects. Model (2) includes all of the str adds neighborhood and locational variables to the previously described model (test scores at the local elementary school, census tract median income, percent of the census block population that is white, percentage of the census block that is under age 18 and distances to the central business district, nearest park, nearest shopping center, and nearest lake). The third model adds city fixed effects to the second model. Thus Model (3) explicitly includes both common important characteristics of neighborhood quality while also allowing for spatial heterogeneity in the regression coefficients, and even city-level fixed effects to control for omitted variables that vary at the municipal level.

Two important conclusions can be drawn from these results. Model (3) consistently outperforms the other two models– slightly better than Model (2) and considerably better than Model (1). Thus, even conditional on having modeled the data with local regression, the inclusion of more neighborhood-level variables substantially improves the model fit. Second, local analyses perform significantly better than global models. The minimum GCV score is obtained using a bandwidth of 500 nearest houses and a model including structural, neighborhood, and locational variables in addition to time and city fixed-effects.

[Figure 4 about here.]

[Table 5 about here.]

Table 5 shows the results for our LWR Model (3) with four different bandwidths: the GCV minimizing bandwidth of 500 nearest sales as well as three larger bandwidths: 650, 1,000, and 2,000 nearest sales. For each bandwidth, we display the mean LWR coefficient as well as the 10^{th} and 90^{th} percentile values. In addition to having smaller GCV scores than the global model, the LWR models substantially reduce the amount of spatial autocorrelation in our regression residuals (the Moran's I statistic for the global model is seven times larger than for the 500 nearest sales model).

The mean LWR coefficient for the noise variable is similar across the different bandwidths (approximately a third of a percent difference in house sales prices for each additional decibel of noise, all other varibles held constant). Unsurprisingly, the spread of the coefficient distribution is greater for the smaller bandwidths. We now turn to investigating whether this increased dispersion appears from the stochastic nature of smaller sample sizes, or due to heterogeneity in the hedonic function. In the next three subsections we describe the results of hypothesis tests that lead us to believe that LWR modeling is an appropriate and beneficial research strategy.

4.2 Does the Hedonic Function Vary over Space?

We first conducted a Monte Carlo simulation recommended in Fotheringham et al. (2002) to see if the improved model fit with our LWR models was due to random chance. The intuition of the test is as follows - if the improved model fit isn't significant, then we should be able to see similar improvements in model fit by using the LWR model with the same houses but in a different spatial arrangement. Failing to recreate a spatial arrangement that yields a similar model fit as to that obtained with our actual data is then evidence that our spatial arrangement and the improved model fit obtained from a spatially heterogeneous model is significant.

We conducted 100 different experiments in which we randomly resampled (without replacement) the locations of our house transactions. In each case we then estimated the LWR model using 10 different bandwidths ranging from the nearest 100 to 4,000 sales just as before and calculated the GCV score. Figure 5 shows the distribution of the GCV scores we obtained for the simulated data for four selected bandwidths relative to the score we obtained with our actual data. Note that the GCV distributions for the simulated data a strongly related their bandwidth - smaller bandwidths have larger GCV scores. That is, with our randomized location data, the GCV score is inversely proportional to the bandwidth. Additionally, note that no GCV scores from the simulated data are near our GCV score with the actual data. In every single instance the smallest Generalized Cross-Validation score was obtained at the largest bandwidth. That is, in 100 consecutive opportunities to fit our data at a local level of a couple hundred nearest sales up to a couple thousand nearby sales, the GCV score was always smallest for the largest possible bandwidth. In other words, after running a thousand different local regression models on the shuffled data, we never obtained model performance anywhere near the LWR model with our true locational data. This seems to be strong evidence that the increased ability to predict house prices with a local model is not due to chance and that there is spatial non-stationarity in the hedonic function.

[Figure 5 about here.]

4.3 Which Coefficients Vary over Space?

The previous section presented evidence in support of the claim that there exists significant spatial variation in the hedonic function. However, we have not yet established which variables have spatially varying coefficients. Figures 6 – 8 show the spatial variation in the estimated LWR coefficients for three important variables, lot size, house size, and noise. For comparison, we present the results across four different levels of local analysis, ranging from LWR bandwidths of 500 nearest houses to the nearest 2,000 sales.

A few patterns are clear in Figures 6 – 8. For instance, across all four bandwidths we see that the hedonic coefficient[4] on lot size (measured in acres) is largest near the central business district of St. Paul. Thus, consistent with standard urban economic theory, to the extent that the hedonic coefficient is correlated with the price of land, our results suggest the price of land is highest near the central business district and declines with distance from the CBD. This tendency is consistent across the four different bandwidths presented (other specific results available upon request). The LWR coefficients on house size (measured in thousands of square feet of finished space) also seem to exhibit spatial variation across the different bandwidths. In the west-central portion of our study area (a collection of notably desirable neighborhoods) we see a cluster of large LWR coefficients, while those areas in the southernmost portion of our region (suburban areas) have the smallest coefficients.

In Figure 8 we map the estimated LWR noise coefficients from four different bandwidths. In this figure the red areas represent large negative coefficients, while yellow areas are characterized by values closer to zero. It is interesting to note that for the smaller bandwidths, the area with the most negative noise coefficients is ringed by some of the coefficients closest to zero (and even greater than zero in some cases). As the bandwidth increases to 1000 or 2000 nearest houses, the variation in noise coefficients decreases substantially. In contrast to the previous two figures, we have largely eliminated the "hot spots" that appear in the smaller bandwidths.

[Figure 6 about here.]

[4]Remember that the dependent variable is the log of sales price.

[Figure 7 about here.]

[Figure 8 about here.]

We now conduct a second series of simulation experiments aimed at testing whether the spatial variation exhibited in particular coefficients is significant or due to random chance. The mechanics of the second simulation are very similar to the those described in the previous section. We randomly change the spatial arrangment of our house sales and estimate the LWR model. However, in this simulation we compare the *spread* (rather than the degree of model fit) of the regression coefficient estimates under the random arrangment of the data to the spread found in our original LWR models with the actual spatial arrangement, conditional on bandwidth. If an explanatory variable in our hedonic regression function exhibits a stationary relationship with sales prices, then we should see similar levels of dispersion in the randomly arranged data as in our true data. However, if the dispersion we see in our randomized data is *smaller* than the spread we observed in our actual data, that is evidence consistent with the variable having a non-stationary relationship. That is, the variation in regression coefficients obtained in our data is higher than we observe from random chance.

We conducted this simulation 100 times for the same four different bandwidths in Table 5- 500, 650, 1,000 and 2,000 nearest observations. Across all bandwidths we obtain similar results, so for brevity's sake we present the results for the 500 and 2,000 bandwidth sizes[5] in Figure 9. The solid red (blue) lines show the distribution of the coefficient standard deviations we obtained from the 100 simulations for the 2,000 (500) nearest sale bandwidths. Across all variables it is the case the the distribution of standard deviations is significantly larger for the smaller bandwidth. That is, when we use smaller bandwidths we tend to see greater dispersion in regression coefficients. We also plot the standard deviation of our coefficients from our actual

[5] Additional results available on request.

data as vertical dashed lines (again, color denotes the bandwidth size (red=2,000 and blue=500). For most variables, the standard deviation of the LWR coefficients obtained with our actual data is significantly greater than any standard deviation obtained with the randomized simulated data. That is, across our simulated data we never saw anywhere near the level of dispersion in regression coefficients for house size, lot size, etc. that we observe with our actual data. This seems to be strong evidence that many of our regression coefficients in our hedonic function exhibit spatial non-stationarity. However, for some variables (note the distributions for noise, owner occupancy and percent under age 18) for which our observed spread in the LWR coefficients is similar when we used the simulated data. This simulation does not present evidence that these variables exhibit a non-stationary relationship in our hedonic function.

[Figure 9 about here.]

4.4 Does the Noise Coefficient Change over Time?

We know that house prices in the United States fell dramatically during and after the Great Recession of 2008-09. We know less about the patterns of change in the hedonic price functions during and after the fall in house prices. Cho et al. (2011) found evidence from hedonic analysis in Nashville, Tennessee, that consumers' willingness to pay for environmental amenities (such as proximity to open space and water views) decreased during the recession as compared to the previous economic boom. Previous research has suggested that the impact of noise can change with time and economic conditions. For instance, Wilhelmsson (2000) found that the traffic noise penalty to be stronger in the 1990s near Stockholm, Sweden compared to the 1980s. Cohen and Coughlin (2009) also found the (airport) noise penalty to be larger in the early 2000s compared to the late 1990s in Atlanta, Georgia.

Figures 10–12 show the spatial distribution of some our LWR hedonic coefficients for bandwidths of 500 and 2,000 nearest sales within the last year for the sale years 2006–2010. For both bandwidths presented[6], the figures show that the estimated LWR coefficients are different over time. In particular, the lot size and house size coefficients tend to decrease over time. These results are consistent with standard economic theory, as wealth and income decrease, consumer willingness to pay for normal goods decreases. What is surprising, however, is that the negative noise coefficient estimates also tend to be more negative in later years of our study. Figure 12 shows that the spatial patterns are similar across years (the most negative coefficients tend to be in the same locations across years), but the average noise coefficient tends to be more negative for the later years.

[Figure 10 about here.]

[Figure 11 about here.]

[Figure 12 about here.]

We also find statistical evidence consistent with the estimated noise coefficients being more negative in the later time periods. This finding is robust to multiple different formulations and specifications. Tables 6 and 7 present the results of regressions using our estimated LWR noise coefficient as the dependent variable and time and spatial characteristics as explanatory variables for LWR bandwidths of 500 and 2,000 nearest sales. Both tables are structured similarly. Column (1) finds evidence of a negative linear trend in the LWR noise coefficients over time - predicting a LWR noise coefficient of approximately twice the magnitude for the end of our study period as for the beginning. Column (2) includes a quadratic term and again predicts noise coefficients much more negative towards the end of our study period compared to the beginning. We repeat the regressions from (1) and (2) in (3) and (4) but this time also include city fixed effects in order to control

[6] Additional results available upon request.

for any potential locational differences of house transactions over time. Again, we see very similar results - the noise coefficient is more negative in later time periods. We also limit our analysis to just the city of St. Paul and repeat the analysis from (1) and (2) in columns (5) and (6). Again, we estimate the noise coefficient to be substantially more negative in later time periods.

[Table 6 about here.]

[Table 7 about here.]

5 Discussion

The hedonic analysis results presented suggest that noise exhibits a significant negative effect on house prices and that the hedonic function varies over space and time. These results and analysis rest on important assumptions, which if violated, may change the results. This section describes three important considerations worthy of discussion and needed to contextualize our results: causation vs. correlation, potential temporal problems with our data, and omitted variable bias.

In a stronger research design, we would observe the same houses over time and randomly "treat" some houses with more or less noise to observe the change in price relative to the control houses. Such a study is not feasible and we are therefore limited to a hedonic analysis with observational data. Our conclusions, then, should be interpretted like most other hedonic analysis results and, while we find a strong negative correlation between noise and house prices, we cannot definitely state that noise *causes* lower sales prices.

Future work may want to further investigate this correlation for more evidence consistent with causation. For instance, by collecting data on the quality of house windows one could test for differential effects of more soundproof windows in noisier areas vs. quieter areas. Unfortunately, such house characteristics are unavailable for this study.

Our work faces other data limitations. Notably, while our house sales data are collected over the course of six years, the traffic noise estimates are for the year 2007. To the extent that traffic flows and composition significantly changed over time, our traffic noise variables may be inaccurate for those other time periods. However, Table VM2 from the US. Department of Transportation Office of Highway Policy Information website reports that in the year 2007 there were approximately 30 million vehicle miles traveled in the urban areas of the state of MN, while for years 2008, 2009, and 2010 the respective values were 32.5, 32.3, and 32.0 million vehicle miles. Overall, this represents a change of less than 10 percent. [7] Such small changes in total vehicle miles traveled are unlikely to yield significant differences in noise levels. Additionally, we do not believe the changes are likely to manifest significant spatial variation in traffic noise, which is ultimately responsible for our identification of the noise regression coefficient. However, given enough time for changes in driving habits and technological advancements to gain significant market penetration (for instance, quieter hybrid cars), future work should seek to create time-series noise data in order to obtain even better estimates of the impact of traffic noise over time. Given the computational complexity of re-estimating the landscape noise surface, such work is beyond the scope of this paper, but may be a worthwhile investment for future work seeking to test the robustness of our results.

Our previously reported results also lack some important control variables in the hedonic function. As mentioned previously, we do not have data on the actual level of noise purchasers were exposed to inside the house, which could be altered depending on the quality of insulation, windows, and even time of day, or day of week. Locations exposed to more noise due to their proximity to high traffic areas may also be less safe due to the potential for accidents. Without including a proxy variable for safety, our noise coefficient could be biased - making it appear that home prices are reflecting noise exposure when in fact it could be safety issues.

[7] http://www.fhwa.dot.gov/policyinformation/quickfinddata/qftravel.cfm

Future research aimed at measuring the impact of noise should continue to seek out and control for other confounding variables.

Our data also omits other structural characteristics and to the extent that variables like the number of bedrooms, bathrooms, garage size, or construction quality co-vary with other variables in our dataset, our regression coefficients will suffer from omitted variable bias. We are comforted, however, because we do have some of these important additional variables for a subset of our data. We were able to obtain the number of bedrooms, bathrooms, and garage area for our houses located within Dakota County from the Dakota County Assessor's Office. Our LWR analysis was then repeated on this geographic subset of the data in order to compare our noise coefficients across the different models. We found strikingly similar traffic noise coefficient estimates when these additional structural variables were included and omitted. As shown in Figure 13, a simple linear regression of the noise coefficient estimates obtained when these additional structural variables were excluded on the estimates obtained when these variables were included yields a line of best fit that is almost indiscernable from the 45 degree line and has $R^2 = 0.96$.

[Figure 13 about here.]

6 Conclusion

We estimated a hedonic price function for single family houses using Locally Weighted Regression techniques in the St. Paul, Minnesota, urban area. Specifically, we estimated semi-logarithmic regressions at each house within our dataset using only information contained in "local" (both geographically and temporally) house sales. We find strong evidence that the hedonic function in our study area varies over space and time and such flexible models represent significant improvements over conventional parametric models.

Monte Carlo simulations suggest that the better goodness-of-fit provided by the local models are not due to chance and that many hedonic implicit prices vary over space within our study area. When the location of our data was randomly assigned and our LWR model was re-estimated across more than a dozen different bandwidths, in 100 consecutive simulations the smallest GCV score was obtained when the data were analyzed at a pooled/global level rather than local. That is, after trying thousands of different combinations of varying levels of local analysis with the spatially redistributed data, we never came close to estimating our observed housing sales prices as well as we can with the local analysis on the actual data. Additionally, re-estimating the LWR model using local bandwidths but randomly assigned locations yielded substantially smaller standard deviations for the majority of our regression coefficients. Such differences suggest that the variation exhibited by most of our regression coefficients was not due to simple random chance, but instead is consistent with spatial non-stationarity.

While our results show very strong evidence that the coefficients on lot size and living space vary over space, the noise coefficient was one of a few variables to exhibit similar standard deviations to the simulated distributions obtained from Monte Carlo experiments. We do, however, find significant temporal variation in the impact of traffic noise in our data.

Precise estimates of the impact of traffic noise are important for efficient implementation of noise mitigation projects. The US Federal Highway Administration reports that 47 states have constructed noise barriers to reduce noise propagation, spending over a half a billion dollars from 2008-2010 (USFHWA, 2012). Delucchi and Hsu (1998) used an estimate of 0.85 percent reduction in house prices per additional decibel in their cost-benefit analysis of traffic noise. While their preferred estimate of the total damage cost to the United States in 1991 is \$3 billion, they report a possible range of less than \$100 million to over \$40 billion due in part to uncertainty surrounding the degree of noise propagation over space in the urban built environment and the effect of noise on house prices. Indeed, one of the main conclusions of their work is that more research should be done to better understand the impact of noise on house prices and that better noise data are needed for such work to occur.

Future research should be conducted in more housing markets, as the potential to apply the results of analysis in one set of geographical and economic circumstances may be limited. Second, "mixed" regression techniques (in which some regression coefficients are constrained to remain constant across the study area while others are allowed to vary) may allow researchers to obtain more precise estimates of the impact of hedonic characteristics by increasing the degrees of freedom in the regression. Future work may also seek to re-estimate traffic noise models over time to better account for changes in traffic flows associated with macroeconomic conditions. Lastly, researchers should take advantage of the recent suggestions in Carruthers and Clark (2010) and use the spatial variation in regression coefficients obtained from LWR models to estimate the second-stage hedonic regressions to identify consumer demand curves for these characteristics.

References

The Influence of Night-time Noise on Sleep and Health. Technical report, Health Council of the Netherlands, 2004.

World Urbanization Prospects: The 2011 Revision. Technical report, United Nations Department of Economic and Social Affairs, 2012.

Summary of Noise Barriers Constructed by December 31, 2010. Technical report, US Department of Transportation Federal Highway Administration, 2012.

Henrik Andersson, Lina Jonsson, and Mikael Ögren. Property Prices and Exposure to Multiple Noise Sources: Hedonic Regression with Road and Railway Noise. *Environmental and Resource Economics*, 45(1):73–89, 2010.

Wolfgang Babisch, Bernd Beule, Marianne Schust, Norbert Kersten, and Hartmut Ising. Traffic Noise and Risk of Myocardial Infarction. *Epidemiology*, 16:33–40, 2005.

Andrea Baranzini, Caroline Schaerer, and Philippe Thalmann. Using measured instead of perceived noise in hedonic models. *Transportation Research Part D: Transport and Environment*, 15(8):473–482, December 2010.

Jesse R Barber, Kevin R Crooks, and Kurt M Fristrup. The costs of chronic noise exposure for terrestrial organisms. *Trends in ecology & evolution*, 25(3):180–9, March 2010.

Lars Barregard, E Bonde, and E Ohrström. Risk of hypertension from exposure to road traffic noise in a population-based sample. *Occupational and Environmental Medicine*, 66:410–415, 2009.

T. M. Barry and J. A. Regan. FHWA Highway Traffic Noise Prediction Model: Report No. FHWA-RD-77-108. Technical report, Federal Highway Administration, Washington, DC, 1978.

Ian Bateman, Brett Day, Iain Lake, and Andrew Lovett. The Effect of Road Traffic on Residential Property Values : A Literature Review and Hedonic Pricing Study. Technical Report January, 2001.

J C Blanco and I H Flindell. Property prices in urban areas affected by road traffic noise. *Applied Acoustics*, 72(4):133–141, 2011.

Theo Bodin, Maria Albin, Jonas Ardö, Emilie Stroh, Per-Olof Ostergren, and Jonas Björk. Road traffic noise and hypertension: results from a cross-sectional public health survey in southern Sweden. *Environmental Health*, 8:38, January 2009.

Sebastian Brandt and Wolfgang Maennig. Road noise exposure and residential property prices: Evidence from Hamburg. *Transportation Research Part D: Transport and Environment*, 16(1):23–30, January 2011.

Chris Brunsdon, Stewart Fotheringham, and Martin Charlton. Geographically weighted regression-modelling spatial non-stationarity. *Journal of the Royal Statistical Society Series D The Statistician*, 47(3):431–443, 1998.

John I. Carruthers and David E. Clark. Valuing Environmental Quality: a Space-Based Strategy. *Journal of Regional Science*, 50(4):801–832, April 2010.

Emilio Casetti. Generating Models by the Expansion Method: Applications to Geographical Research. *Geographical Analysis*, 4(1):81–91, 1972.

Hao Cheng. *Using Micro Data to Assess Traffic Noise Impact on Residential Property Value: an Application of the Hedonic Model*. Phd, University of Louisville, 2008.

Ekaterina Chernobai, Michael Reibel, and Michael Carney. Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices. *The Journal of Real Estate Finance and Economics*, 42(3):348–370, October 2009.

Seong-Hoon Cho, Seung Gyu Kim, and Roland K. Roberts. Values of Environmental Landscape Amenities during the 2000 - 2006 Real Estate Boom and Subsequent 2008 Recession. *Journal of Environmental Planning and Management*, 54(1): 71–91, January 2011.

William S Cleveland and Susan J Devlin. Locally Weighted Regression : An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, 83(403):596–610, 1988.

Jeffrey P. Cohen and Cletus C. Coughlin. Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices. *Journal of Regional Science*, 48(5):859–878, December 2008.

Jeffrey P. Cohen and Cletus C. Coughlin. Changing Noise Levels and Housing Prices Near the Atlanta Airport. *Growth and Change*, 40(2):287–313, 2009.

Joyce Dargay, Dermot Gately, and Martin Sommer. Vehicle Ownership and Income Growth, Worldwide : 1960-2030. *The Energy Journal*, 28(4):143–170, 2007.

Brett Day, Ian Bateman, and Iain Lake. Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic property price model. *Environmental and Resource Economics*, 37(1):211–232, 2007.

Yvonne de Kluizenaar, Sabine a Janssen, Henk Vos, Erik M Salomons, Han Zhou, and Frits van den Berg. Road traffic noise and annoyance: a quantification of the effect of quiet side exposure at dwellings. *International journal of environmental research and public health*, 10(6):2258–70, January 2013.

Mark Delucchi and Shi-Ling Hsu. The External Damage Cost of Noise Emitted from Motor Vehicles. *Journal of Transportation and Statistics*, 1(3):1–24, 1998.

Carlos Marmolejo Duarte and Carlos Gonzalez Tamez. Does Noise Have a Stationary Impact on Residential Values? *Journal of European Real Estate Research*, 2 (3):259–279, 2009.

Molly Espey and Hilary Lopez. The Impact of Airport Noise and Proximity on Residential Property Values. *Growth and Change*, 31(3):408–419, 2000.

A. Stewart Fotheringham, Chris Brunsdon, and Martin Charlton. *Geographically Weighted Regression: the analysis of spatially varying relationships*. John Wiley & Sons, West Sussex, England, 2002.

Hays B. Gamble, Owen H. Sauerlender, and C. John Langley. Adverse and Beneficial Effects of Highways on Property Values. *Transportation Research Record*, (508): 37–48, 1974.

Lisa Goines and Louis Hagler. Noise Pollution : A Modern Plague. *Southern Medical Journal*, 100:287–294, 2007.

William T. Hughes Jr. and C.F. Sirmans. Traffic Externalities and Single-Family Housing Prices. *Journal of Regional Science*, 32(4):487–500, 1992.

Lars Jarup, Wolfgang Babisch, Danny Houthuijs, Göran Pershagen, Klea Katsouyanni, Ennio Cadum, Marie-Louise Dudley, Pauline Savigny, Ingeburg Seiffert, Wim Swart, Oscar Breugelmans, Gösta Bluhm, Jenny Selander, Alexandros Haralabidis, Konstantina Dimakopoulou, Panayota Sourtzi, Manolis Velonakis, and Federica Vigna-Taglianti. Hypertension and Exposure to Noise Near Airports: the HYENA Study. *Environmental Health Perspectives*, 116(3):329–333, December 2008.

C. John Jr. Langley. Adverse Impacts of the Washington Beltway on Residential Property Values. *Land Economics*, 52(1):54–65, 1976.

James E. Larsen. Surface street traffic volume and single-family house price. *Transportation Research Part D: Transport and Environment*, 17(4):317–320, June 2012.

Wei Li and Jean-Daniel Saphores. Assessing Impacts of Freeway Truck Traffic on Residential Property Values Southern California Case Study. *Transportation Research Record*, (2288):48–56, 2012.

Clive Loader. *Local Regression and Likelihood*. Springer-Verlag, New York, NY, 1999.

John W. Matthews and Geoffrey K. Turnbull. *Neighborhood Street Layout and Property Value: The Interaction of Accessibility and Land Use Mix*, volume 35. July 2007.

Daniel P. McMillen. Airport expansions and property values: the case of Chicago O'Hare Airport. *Journal of Urban Economics*, 55(3):627–640, May 2004.

Daniel P. McMillen. Perspectives on Spatial Econometrics: Linear Smoothing With Structured Models. *Journal of Regional Science*, 52(2):192–209, May 2012.

Daniel P. McMillen and Christian L. Redfearn. Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions. *Journal of Regional Science*, 50(3):712–733, April 2010.

Henk M E Miedema and Catharina G M Oudshoorn. Annoyance from Transportation Noise : Relationships with Exposure Metrics DNL and DENL and Their Confidence Intervals. *Environmental Health Perspectives*, 109(4):409–416, 2001.

Ingrid Nappi-Choulet and Tristan-Pierre Maury. A Spatial and Temporal Autoregressive Local Estimation for the Paris Housing Market. *Journal of Regional Science*, 51(4):732–750, October 2011.

Stale Navrud. The State-Of-The-Art on Economic Valuation of Noise: Final Report to European Commission DG Environment. Technical report, 2002.

Jon P. Nelson. Highway Noise and Property Values: A Survey of Recent Evidence. *Journal of Transport Economics and Policy*, 16(2):117–138, 1982.

Jon P Nelson. Hedonic Property Value Studies of Transportation Noise : Aircraft and Road Traffic. In Andrea Baranzini, Jose Ramirez, Caroline Schaerer, and Philippe Thalmann, editors, *Hedonic Methods in Housing Market Economics*, number October 2007, chapter 3, pages 57–82. Springer, New York, NY, 2008.

Evy Ohrstrom, Lars Barregard, Eva Andersson, Annbritt Skarberg, Helena Svensson, and Par Angerheim. Annoyance due to single and combined sound exposure from railway and road traffic. *The Journal of the Acoustical Society of America*, 122(5):2642–52, 2007.

D Ouis. Annoyance From Road Traffic Noise: a Review. *Journal of Environmental Psychology*, 21:101–120, 2001.

Antonio Paez, Steven Farber, and David Wheeler. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. *Environment and Planning A*, 43(12):2992–3010, 2011.

Daniel Shepherd, David Welch, Kim N Dirks, and David McBride. Do quiet areas afford greater health-related quality of life than noisy areas? *International journal of environmental research and public health*, 10(4):1284–303, April 2013.

M. Sorensen, M. Hvidberg, Z. J. Andersen, R. B. Nordsborg, K. G. Lillelund, J. Jakobsen, a. Tjonneland, K. Overvad, and O. Raaschou-Nielsen. Road traffic noise and stroke: a prospective cohort study. *European Heart Journal*, 32(6): 737–744, January 2011.

David L. Sunding and Aaron M. Swoboda. Hedonic analysis with locally weighted regression: An application to the shadow cost of housing regulation in Southern California. *Regional Science and Urban Economics*, 40(6):550–573, November 2010.

Marcel Theebe. Planes, Trains, and Automobiles: The Impact of Traffic Noise on House Prices. *Journal of Real Estate Finance and Economics*, 28(2/3):209–234, 2004.

Diana Weinhold. The Happiness-Reducing Costs of Noise Pollution. *Journal of Regional Science*, 53(2):292–303, May 2013.

Mats Wilhelmsson. The Impact of Traffic Noise on the Values of Single-Family Houses. *Journal of Environmental Planning and Management*, 43(6):799–815, 2000.

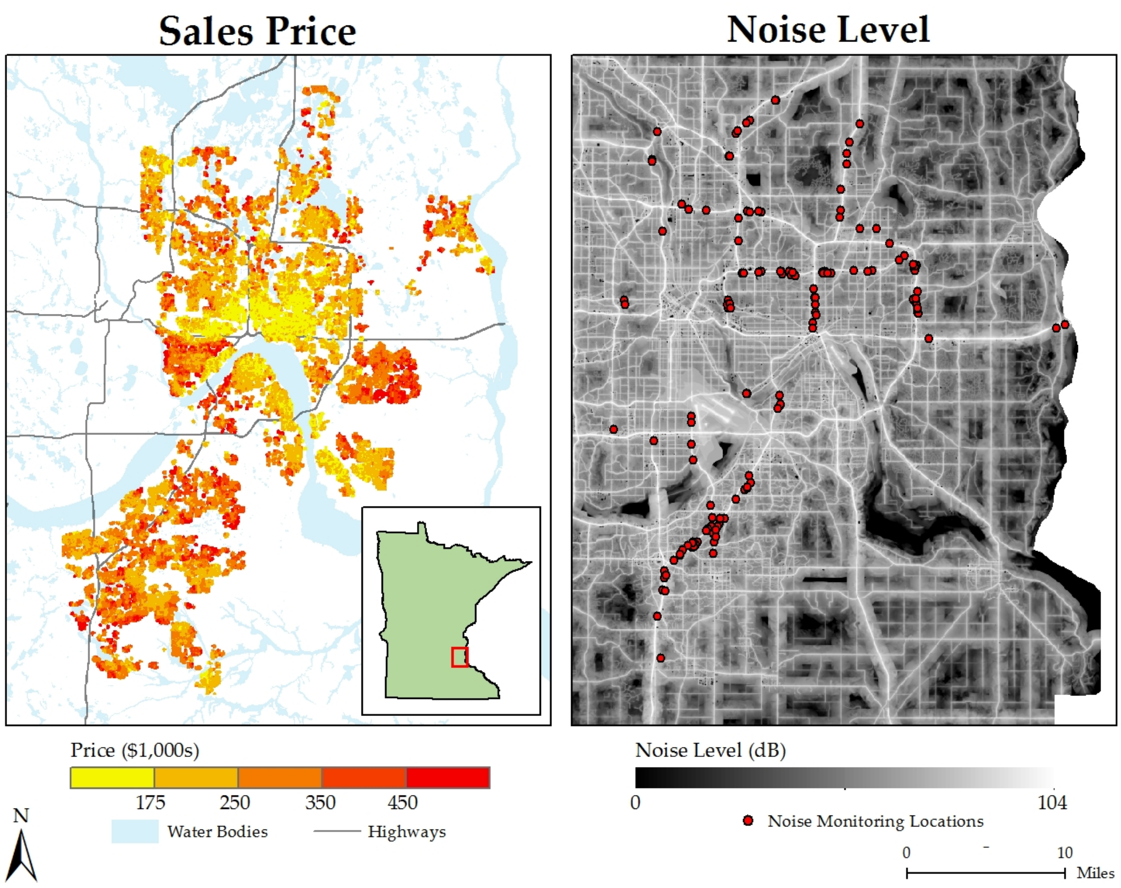

Figure 1: This figure shows the spatial extent of our study area, the significant spatial variation in single family house sales prices, as well as our noise data.

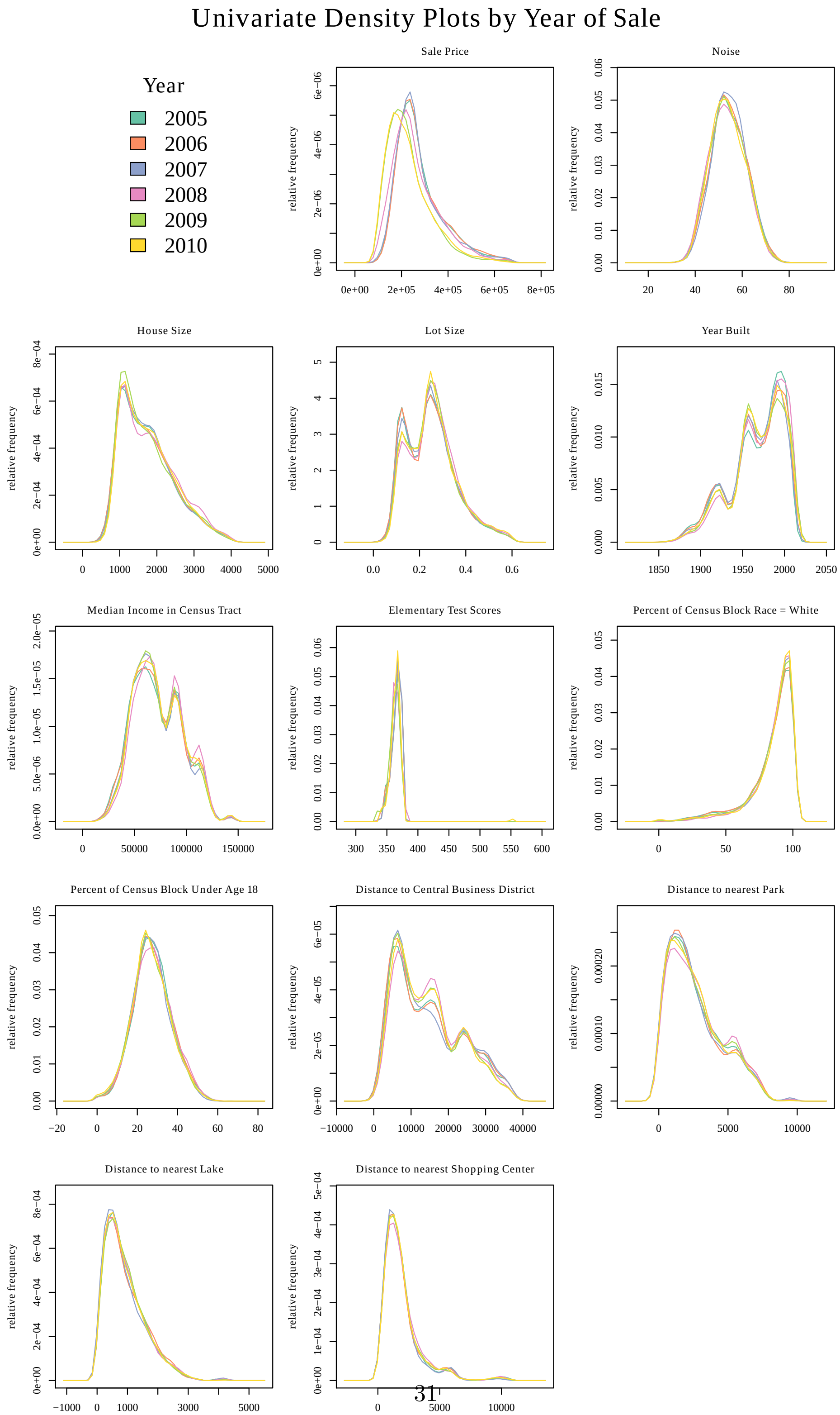

Figure 2: This figure shows the distribution of our variables over time. Note that there are no discernable differences in mean, standard deviation, or shape of the distribution for most variables other than sales price.

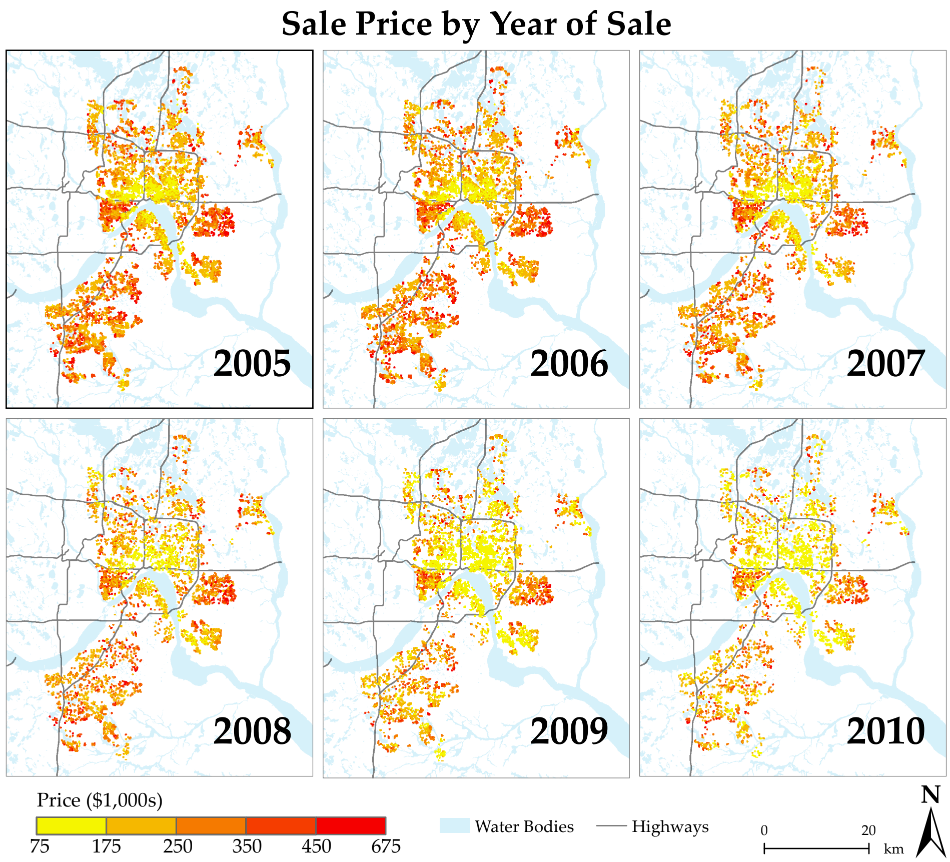

Figure 3: This figure shows the spatial distribution of observed sales prices over time. The change in color from 2005 to 2010 denotes the dramatic drop in sales prices associated with the recession of 2008-09. There are clear spatial patterns in house prices that persist over time (locations with relatively high vs. low prices in one year tend to be the same in other years). However, we see no evidence that sales shifted from one area of our study region to another (neighborhoods with many sales in 2005 also tend to have many sales in 2010).

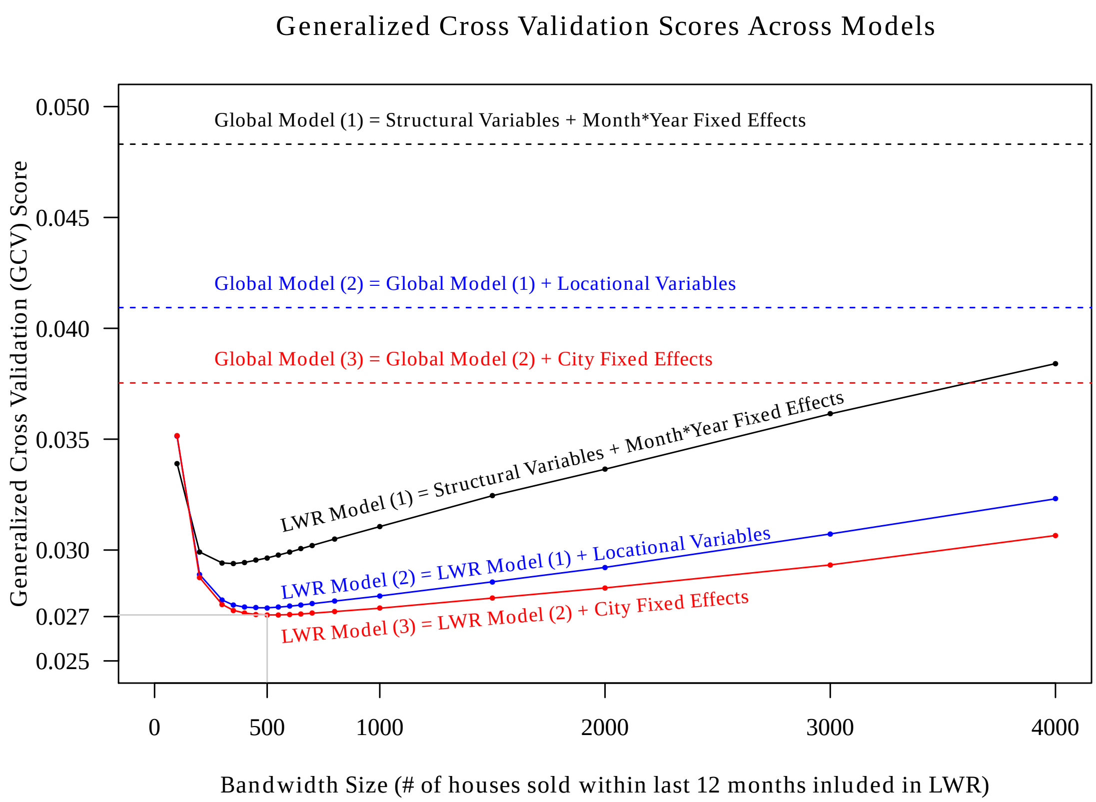

Figure 4: This figure shows the relationship between bandwidth size and the GCV score for three different Locally Weighted Regression (LWR) models. For comparison, the GCV scores for each model when estimated at a global scale are also shown. Note that the LWR models all have significantly smaller GCV scores than the global models. The minimum GCV score is obtained by LWR Model (3) at a bandwidth of 500 nearest houses.

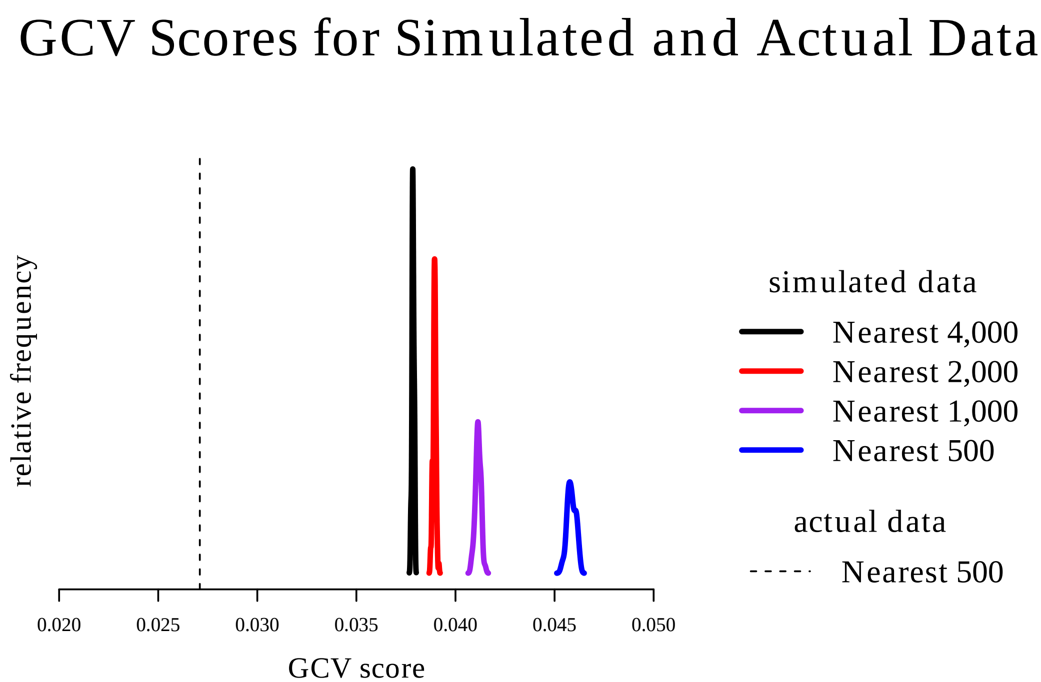

Figure 5: This figure shows the distribution of GCV scores obtained from simulations with spatially randomized data and the GCV score from our data. Each solid curve represents the smoothed density plot of 100 different instances of the simulation. The dashed line suggests that our actual minimum GCV score was not obtained by random chance as no other GCV scores come close.

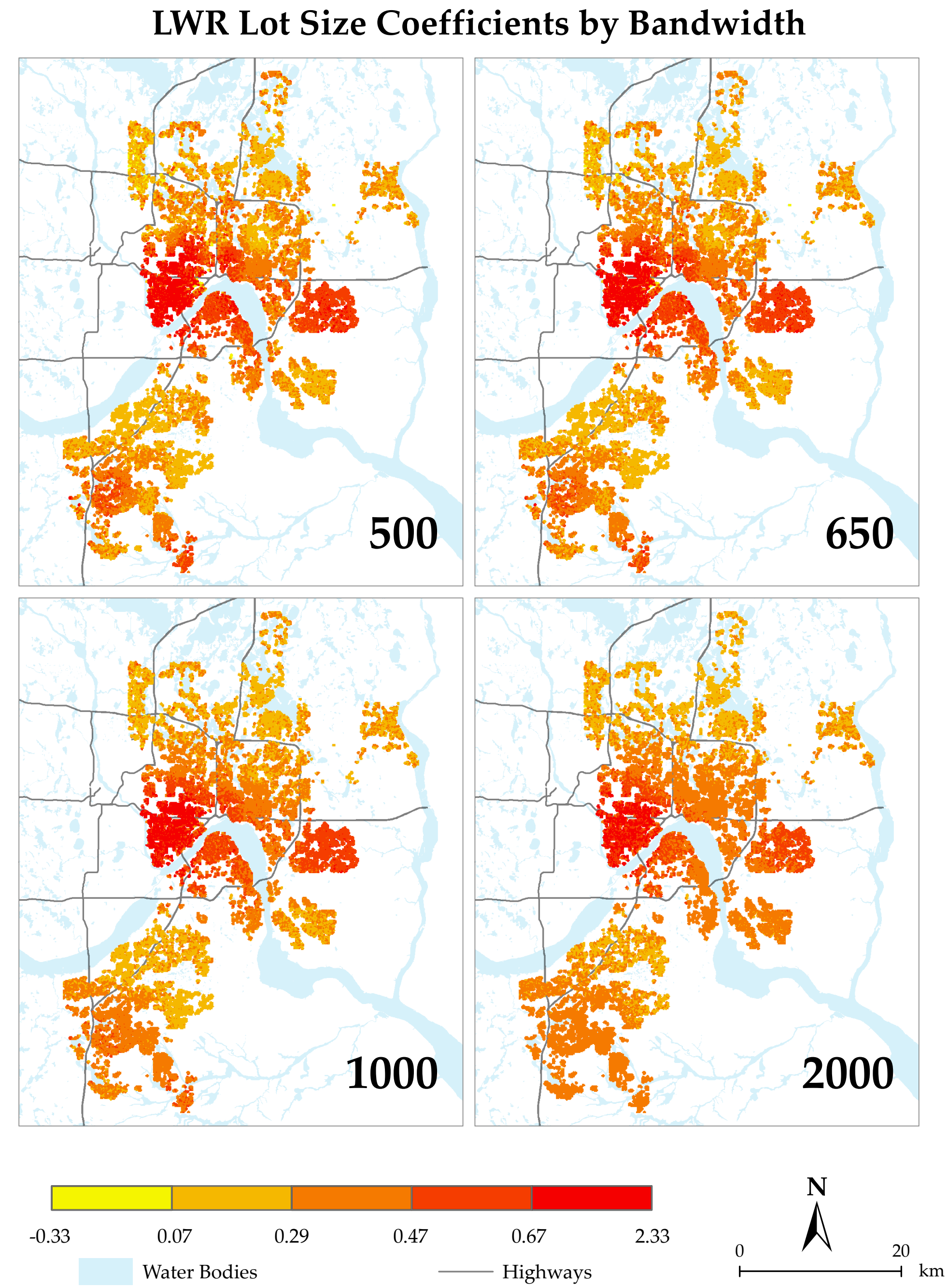

Figure 6: This figure shows an interpolated surface (using Inverse Distance Weighting) for the estimated LWR coefficients (dependent variable = ln(sales price)) for the lot size variable (in acres) obtained across four different bandwidths: 500, 650, 1,000, and 2,000 nearest observations. In all cases there is a clear inverse relationship between the coefficient magnitudes and distance to the central business district.

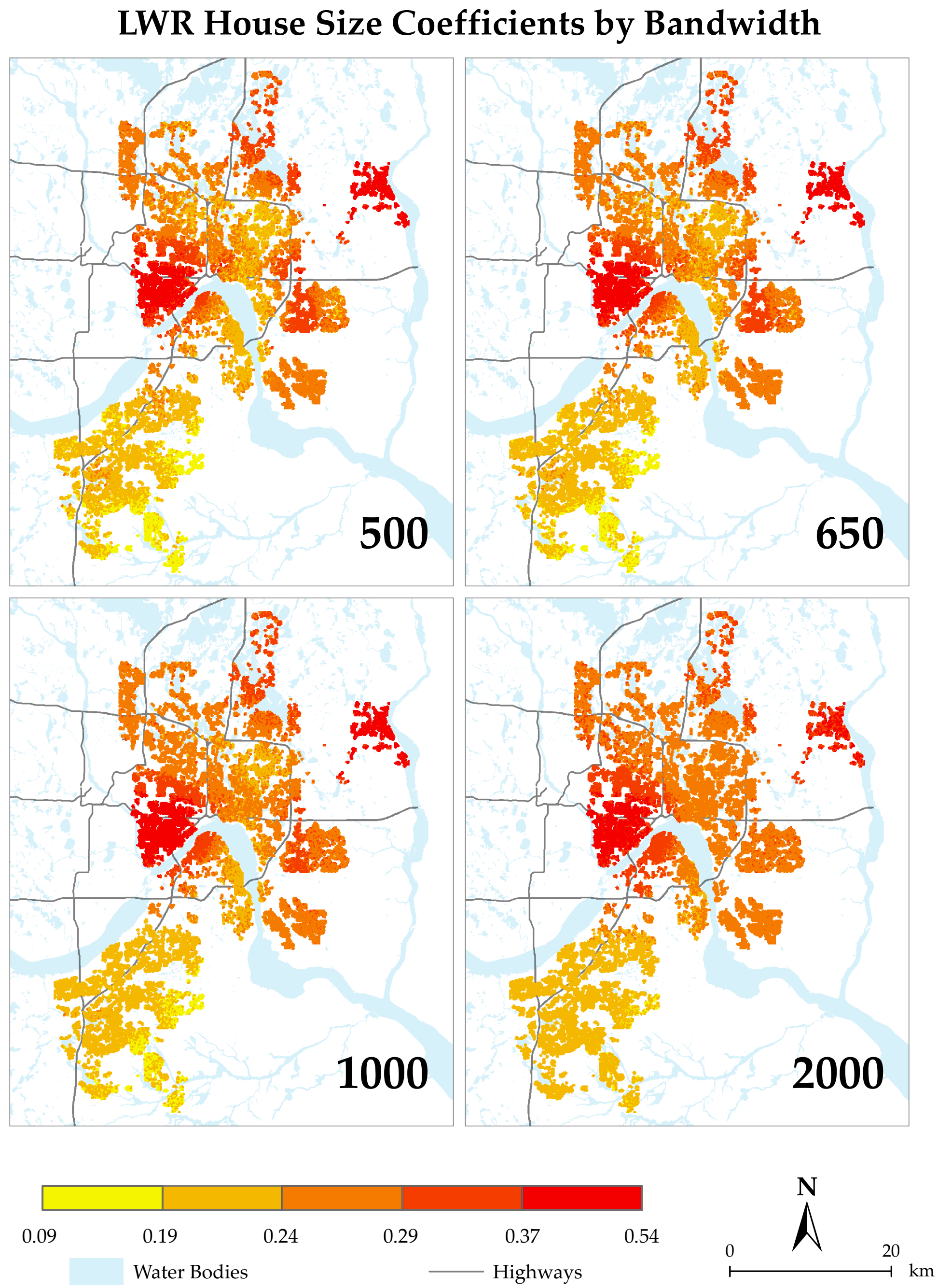

Figure 7: This figure shows an interpolated surface (using Inverse Distance Weighting) for the estimated LWR coefficients (dependent variable = ln(sales price)) for the finished house size variable (in 1,000s of square feet) obtained across four different bandwidths: 500, 650, 1,000, and 2,000 nearest observations. In all cases some of the highest coefficient values are in the west-central portion of our study area and the lowest coefficients tend to be in the far south.

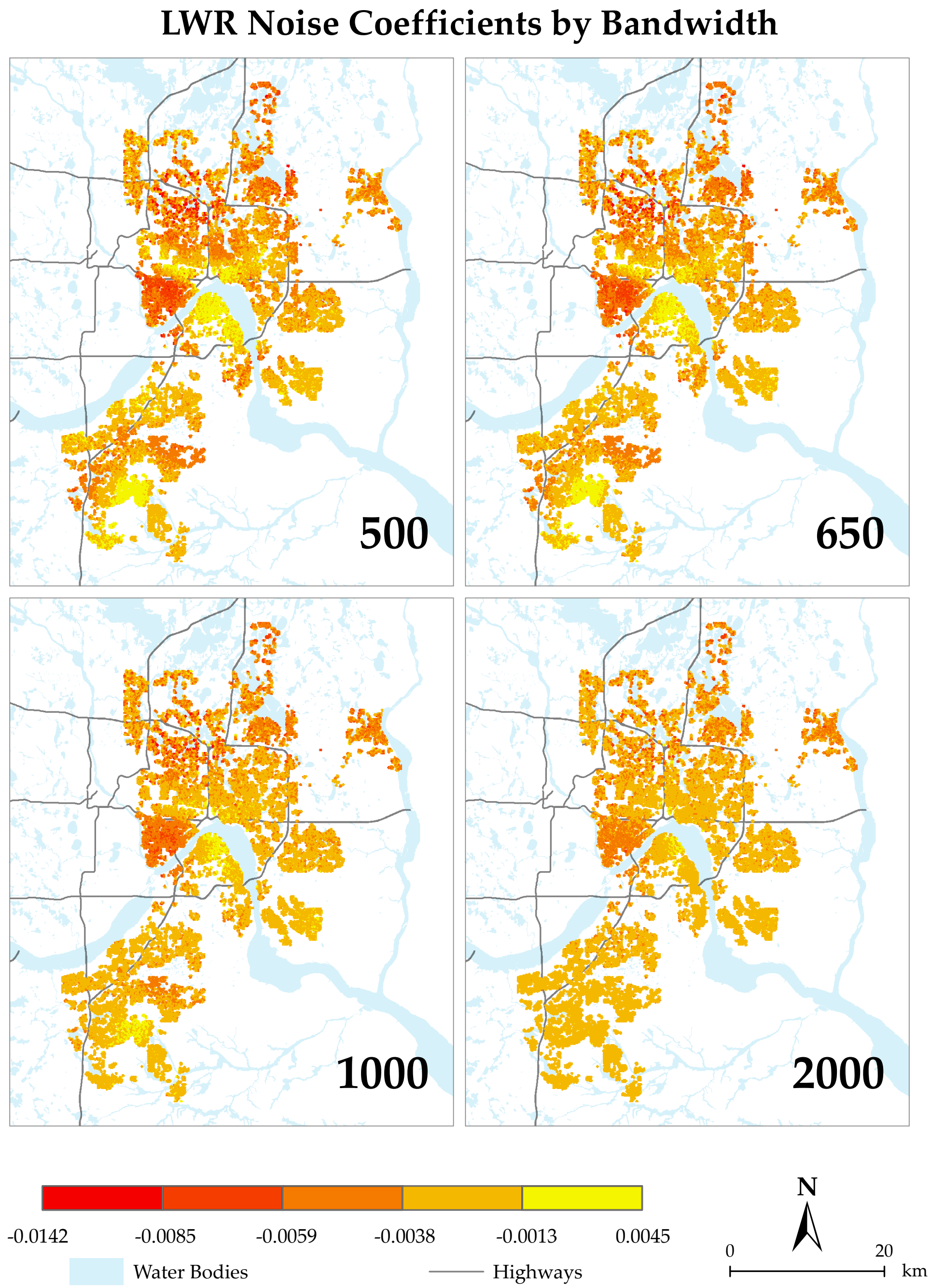

Figure 8: This figure shows an interpolated surface (using Inverse Distance Weighting) for the estimated LWR coefficients (dependent variable = ln(sales price)) for the noise variable (in dB) obtained across four different bandwidths: 500, 650, 1,000, and 2,000 nearest observations. In all cases some of the most negative coefficient values are in the west-central portion of our study area with the smallest in magnitude coefficients tending to be near the central business district.

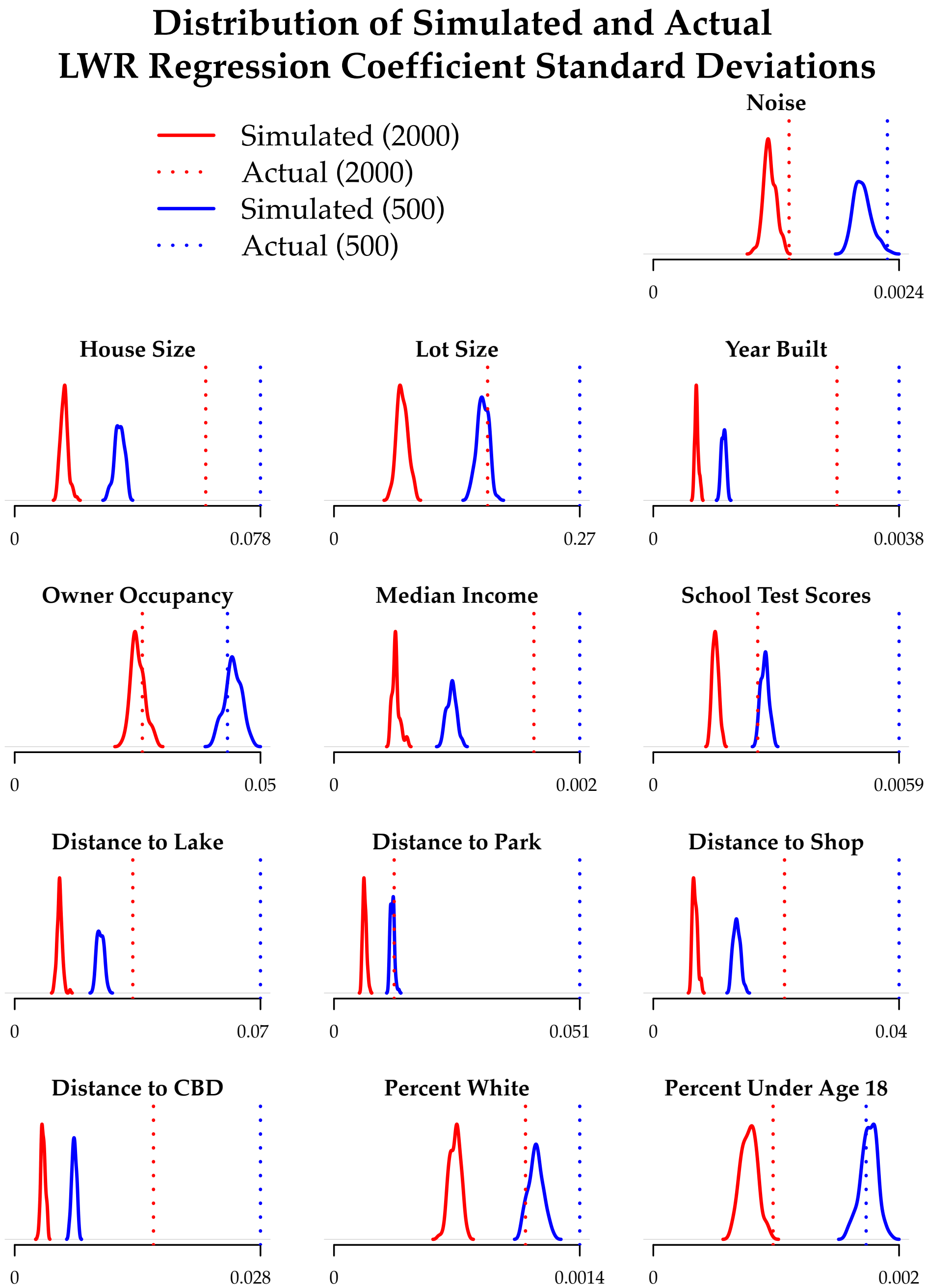

Figure 9: This figure shows the actual standard deviation in LWR coefficients as compared to those obtained from 100 different instances of spatial randomization. Results are presented for two different bandwidths (nearest 500 and 2,000) for our regression variables. In all cases the distribution of standard deviations for the smaller bandwidth is larger. In most cases the standard deviation of LWR coefficients for the actual data (the dashed line) is substantially higher than any standard deviation obtained with simulated data.

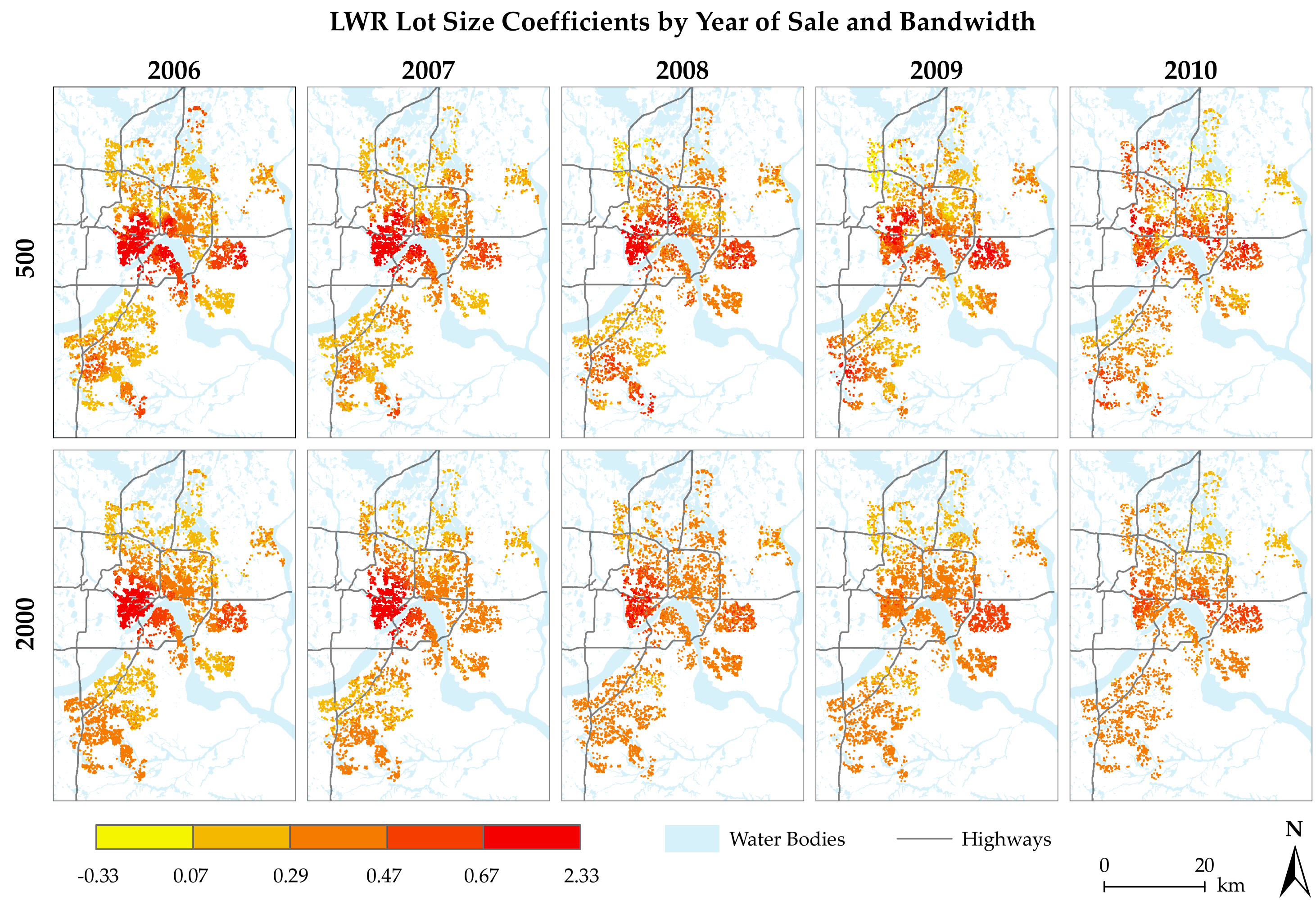

Figure 10: This figure shows the interpolated LWR lot size coefficients across space and time for two bandwidths (nearest 500 and 2,000). Note that the spatial pattern obtained is similar across time periods, but the average value tends to decline over time.

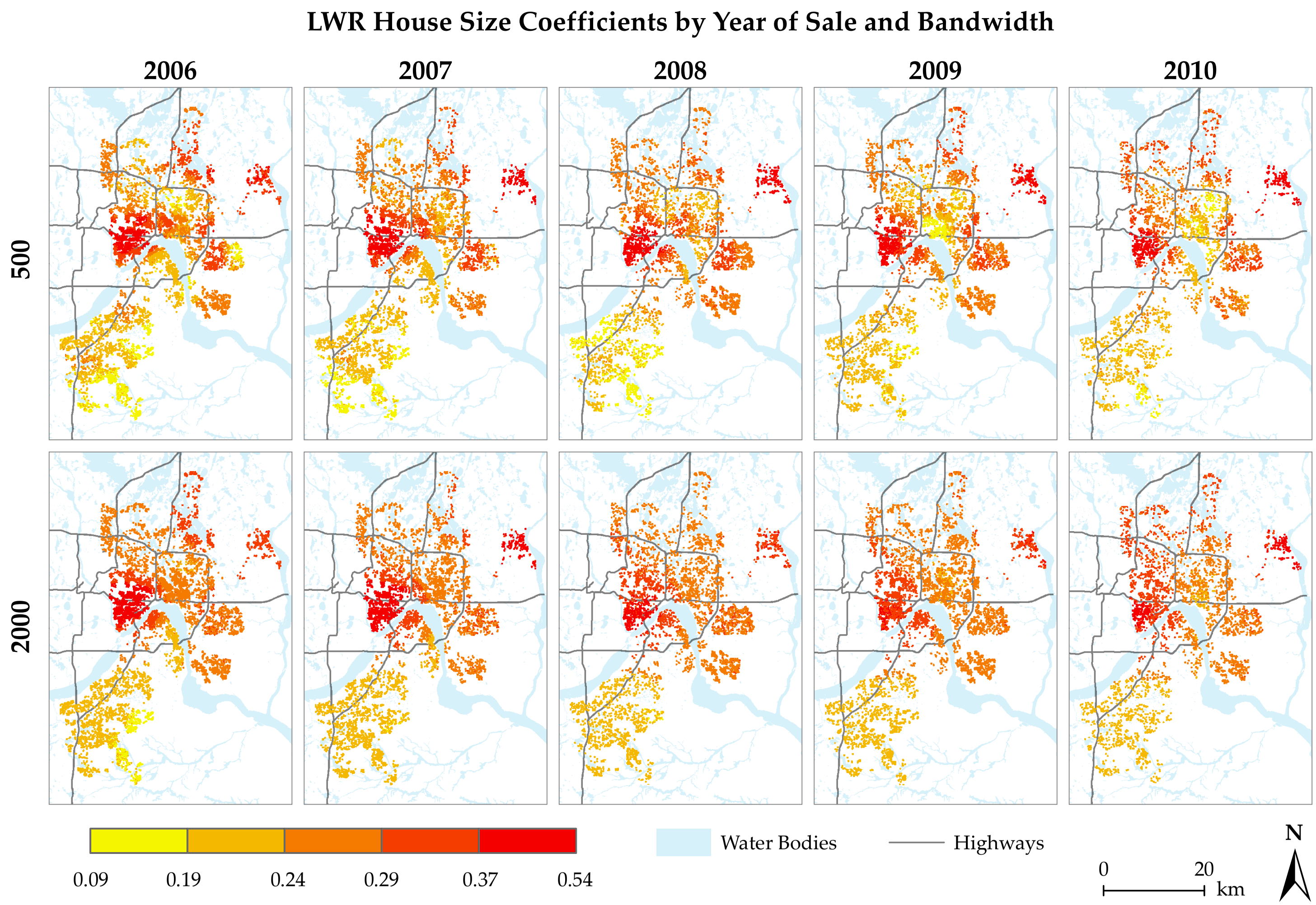

Figure 11: This figure shows the interpolated LWR finished house size coefficients across space and time for two bandwidths (nearest 500 and 2,000). Note that the spatial pattern obtained is similar across time periods, but the average value tends to decline over time.

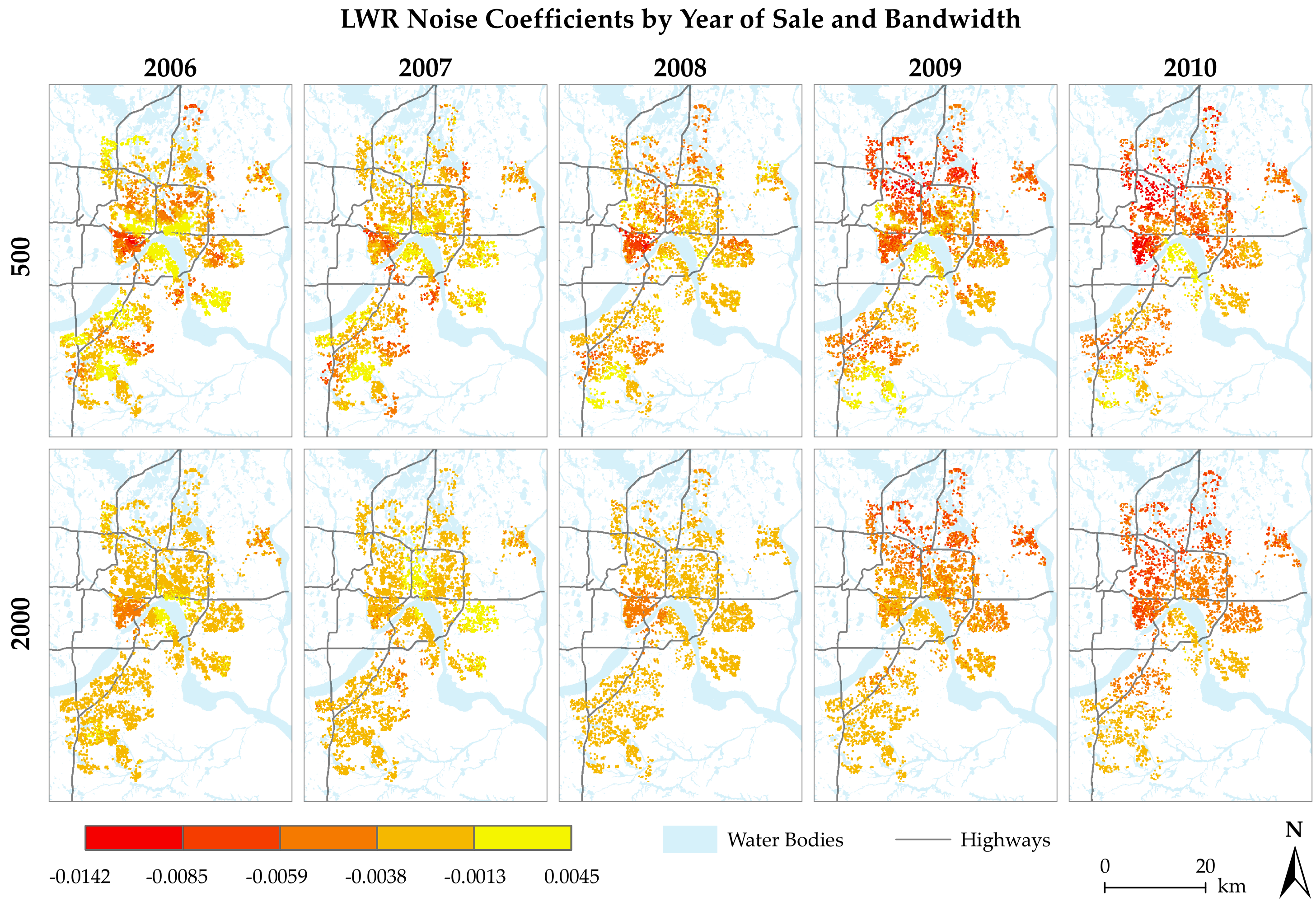

Figure 12: This figure shows the interpolated LWR noise coefficients across space and time for two bandwidths (nearest 500 and 2,000). Note that the spatial pattern obtained is similar across time periods, but the average value tends to decline over time.

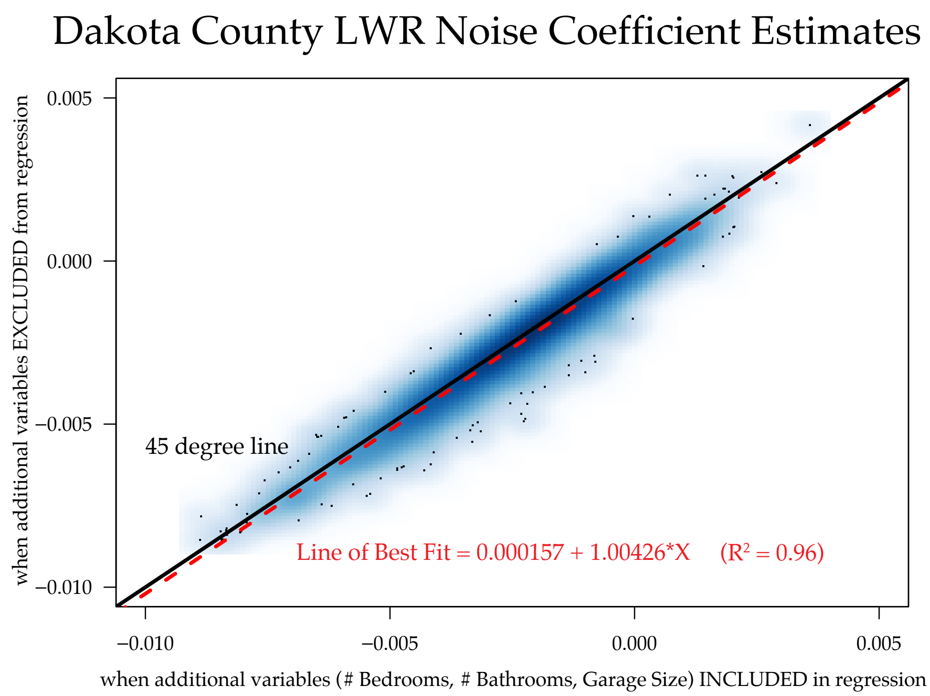

Figure 13: This figure shows the similarity between the LWR coefficients we obtained from our model in Dakota County when additional structural variables were excluded vs. included in the hedonic regression function. The black line shows the 45 degree line, while the red dashed line shows the estimated line of best fit, which yields a simple R^2 of 0.96

Table 1: Variable Description and Summary Statistics

	min	25%	50%	*mean*	75%	*max*	*stdev*
Sale Price ($1,000s)	99	195	241	266	315	675	103
Year House was Built	1850	1950	1973	1967	1993	2010	32
House Size (square feet)	390	1158	1628	1750	2189	4000	704
Lot Size (acres)	0.02	0.17	0.25	0.25	0.31	0.60	0.11
Owner Occupancy (Yes = 1, No = 0)	0.0	1.0	1.0	0.8	1.0	1.0	0.4
Traffic Noise (dB)	24.5	49.0	54.0	54.3	59.5	81.6	7.5
Distance to Central Business District (km)	1.1	6.8	13.2	14.7	21.8	37.1	8.9
Distance to nearest Park (km)	0.0	1.1	2.2	2.6	3.8	9.7	1.9
Distance to nearest Lake (km)	0.0	0.4	0.8	0.9	1.3	4.4	0.7
Distance to nearest Shopping Center (km)	0.0	0.9	1.5	1.9	2.3	10.8	1.6
Percent of Census Block Population Race = White	0	80	90	85	96	100	17
Percent of Census Block Population Under Age 18	0	21	27	27	33	67	9
Median Income in Census Tract ($1,000s)	14	55	69	73	90	143	24
Elementary Test Scores	336	360	365	364	370	552	11

Table 2: Matrix of Pearson Correlation Coefficients for Quantitative Variables

	Price	House	Lot	Built	Noise	CBD	Park	Lake	Shop	White	U18	Inc.	Test
Sale Price	1												
House Size	0.74	1											
Lot Size	0.38	0.45	1										
Year Built	0.45	0.52	0.47	1									
Noise	-0.24	-0.12	-0.04	-0.25	1								
Distance to CBD	0.31	0.45	0.43	0.58	-0.09	1							
Distance to Park	0.25	0.28	0.22	0.42	-0.17	0.57	1						
Distance to Lake	-0.04	-0.10	-0.25	-0.17	-0.01	-0.09	0.20	1					
Distance to Shop	0.27	0.27	0.15	0.37	-0.24	0.46	0.38	0.15	1				
Percent White	0.24	0.15	0.25	0.15	-0.09	0.34	0.16	-0.14	0.14	1			
Percent Under 18	0.26	0.30	0.11	0.36	-0.19	0.29	0.32	0.09	0.32	-0.22	1		
Income	0.52	0.50	0.41	0.57	-0.24	0.53	0.38	-0.12	0.36	0.35	0.32	1	
Elem. Test Scores	0.29	0.28	0.29	0.34	-0.09	0.35	0.22	-0.10	0.18	0.31	0.10	0.44	1

Table 3: Mean Variable Values by Year

Variable	Year of House Sale 2005	2006	2007	2008	2009	2010	All Years
Sale Price ($1,000s)	279	282	275	262	231	233	266
Year House was Built	1967	1966	1966	1971	1968	1968	1967
House Size (feet^2)	1737	1736	1743	1813	1727	1775	1750
Lot Size (acres)	0.25	0.25	0.25	0.26	0.25	0.26	0.25
Owner Occ (Y = 1, N = 0)	0.87	0.87	0.92	0.86	0.60	0.77	0.83
Traffic Noise (dB)	54.4	54.5	54.6	53.7	54.0	54.0	54.3
Dist CBD (km)	14.7	14.6	14.8	15.0	14.3	14.5	14.7
Dist to Park (km)	2.6	2.6	2.6	2.8	2.6	2.6	2.6
Dist to Lake (km)	0.9	0.9	0.9	0.9	0.9	0.9	0.9
Dist to Shop (km)	1.9	1.9	1.8	2.0	1.9	1.9	1.9
Percent Race = White	84	84	85	86	85	86	85
Percent Under Age 18	28	28	27	28	27	27	27
Income ($1,000s)	72	72	72	75	73	73	73
Elementary Test Scores	365	365	365	364	362	366	364
Number of Observations	10,990	8,885	6,549	5,498	5,961	4,200	42,083

Table 4: Basic Global Regression Results

	Dependent variable: ln Sale Price			
	Model (A) All Years	Model (B) Year=2006	Model (C) Year=2008	Model (D) Year=2010
Noise (dB)	-0.0027^{***}	-0.0020^{***}	-0.0033^{***}	-0.0040^{***}
	(0.0001)	(0.0003)	(0.0005)	(0.0006)
House Size (1,000s ft^2)	0.2914^{***}	0.2847^{***}	0.2810^{***}	0.3066^{***}
	(0.0022)	(0.0043)	(0.0067)	(0.0084)
Lot Size (acres)	0.2493^{***}	0.2719^{***}	0.2799^{***}	0.2135^{***}
	(0.0118)	(0.0221)	(0.0359)	(0.0447)
Owner Occupancy Dummy	0.0222^{***}	-0.0027	0.0478^{***}	0.0298^{***}
	(0.0026)	(0.0050)	(0.0085)	(0.0097)
Year House Built	0.0020^{***}	0.0025^{***}	0.0018^{***}	0.0011^{***}
	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Percent Under 18	-0.0001	0.00002	-0.0002	-0.0006
	(0.0001)	(0.0002)	(0.0004)	(0.0005)
Percent White	0.0037^{***}	0.0031^{***}	0.0040^{***}	0.0040^{***}
	(0.0001)	(0.0001)	(0.0002)	(0.0003)
Median Income ($1,000s)	0.0023^{***}	0.0021^{***}	0.0023^{***}	0.0029^{***}
	(0.0001)	(0.0001)	(0.0002)	(0.0003)
Elementary Test Scores	0.0009^{***}	0.0023^{***}	0.0040^{***}	0.0001
	(0.0001)	(0.0003)	(0.0005)	(0.0002)
Dist to CBD (km)	0.0078^{***}	0.0083^{***}	0.0053^{***}	0.0101^{***}
	(0.0005)	(0.0010)	(0.0016)	(0.0020)
Dist to Lake (km)	0.0230^{***}	0.0246^{***}	0.0021	0.0343^{***}
	(0.0017)	(0.0032)	(0.0054)	(0.0067)
Dist to Park (km)	0.0014	-0.0027	0.0099^{***}	0.0064^{*}
	(0.0009)	(0.0017)	(0.0027)	(0.0034)
Dist to Shop (km)	-0.0019^{*}	-0.0020	-0.0063^{*}	-0.0020
	(0.0010)	(0.0019)	(0.0033)	(0.0041)
Home Style Fixed Effects	Yes	Yes	Yes	Yes
Month*Year Fixed Effects	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes
Observations	42,083	8,885	5,498	4,200
R^2	0.742	0.785	0.691	0.668
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5: LWR Coefficient Mean and (10^{th} to 90^{th} Percentile) Values for Selected Bandwidths

	Dependent Variable = ln(Sales Price)			
	Nearest 500* sales	Nearest 650 sales	Nearest 1,000 sales	Nearest 2,000 sales
Noise (dB)	-0.0034 (-0.0064 to -0.0007)	-0.0034 (-0.0061 to -0.0009)	-0.0033 (-0.0057 to -0.0013)	-0.0032 (-0.0049 to -0.0017)
House Size (1,000s ft^2)	0.27 (0.18 to 0.39)	0.27 (0.19 to 0.39)	0.27 (0.20 to 0.40)	0.28 (0.20 to 0.37)
Lot Size (acres)	0.42 (0.17 to 0.73)	0.42 (0.19 to 0.71)	0.42 (0.21 to 0.67)	0.40 (0.25 to 0.58)
Year House Built	0.0046 (0.0004 to 0.010)	0.0045 (0.0005 to 0.0096)	0.0043 (0.0005 to 0.0090)	0.0038 (0.0004 to 0.0080)
Owner Occupancy	0.023 (-0.029 to 0.080)	0.024 (-0.024 to 0.076)	0.023 (-0.018 to 0.068)	0.021 (-0.012 to 0.056)
Percent White	0.0012 (-0.0003 to 0.0029)	0.0013 (-0.0001 to 0.0028)	0.0014 (0.0001 to 0.0028)	0.0016 (0.0003 to 0.0032)
Percent Under Age 18	0.00047 (-0.0016 to 0.0025)	0.00045 (-0.0014 to 0.0022)	0.00040 (-0.0011 to 0.0019)	0.00034 (-0.0008 to 0.0015)
Median Income ($1,000s)	0.0006 (-0.0014 to 0.0033)	0.0008 (-0.0010 to 0.0033)	0.0010 (-0.0006 to 0.0037)	0.0014 (-0.0002 to 0.0042)
Elementary Test Scores	-0.00002 (-0.0056 to 0.0055)	0.00008 (-0.0050 to 0.0049)	0.00029 (-0.0040 to 0.0042)	0.00044 (-0.0025 to 0.0037)
Distance to Lake (km)	-0.0150 (-0.091 to 0.058)	-0.0110 (-0.077 to 0.055)	-0.0076 (-0.063 to 0.052)	-0.0051 (-0.048 to 0.043)
Distance to Park (km)	-0.011 (-0.060 to 0.033)	-0.011 (-0.049 to 0.025)	-0.009 (-0.037 to 0.016)	-0.002 (-0.018 to 0.011)
Distance to Shop (km)	0.0092 (-0.033 to 0.052)	0.0096 (-0.027 to 0.047)	0.0100 (-0.021 to 0.044)	0.0078 (-0.014 to 0.036)
Distance to CBD (km)	0.0084 (-0.019 to 0.044)	0.0094 (-0.014 to 0.040)	0.0100 (-0.010 to 0.038)	0.0095 (-0.005 to 0.037)
Home Style Fixed Effects	Yes	Yes	Yes	Yes
Month*Year Fixed Effects	Yes	Yes	Yes	Yes
City Fixed Effects	Yes	Yes	Yes	Yes
Observations	31,737	31,737	31,737	31,737
GCV score	0.0271	0.0271	0.0274	0.0283
Moran's I statistic**	0.027	0.031	0.037	0.049

* A bandwidth of 500 implies that the nearest 500 houses sold within the past 12 months are included.
** For comparison, the Moran's I statistic for Model A in Table 4 is 0.188.

Table 6: LWR Noise Coefficient Over Time (bandwidth = 500 nearest sales)

	Dependent variable: LWR Noise Coefficient (bandwidth = 500 nearest sales)					
	(1)	(2)	(3)	(4)	(5)	(6)
Months since Jan 2005	-0.00005^{***}	0.00003^{***}	-0.00005^{***}	0.00003^{***}	-0.0001^{***}	0.00004^{***}
	(0.000001)	(0.000004)	(0.000001)	(0.000004)	(0.000002)	(0.00001)
(Months since Jan 2005)^2		-0.000001^{***}		-0.000001^{***}		-0.000001^{***}
		(0.00000005)		(0.00000004)		(0.0000001)
Constant	-0.0016^{***}	-0.0028^{***}	-0.0021^{***}	-0.0033^{***}	-0.0015^{***}	-0.0032^{***}
	(0.00003)	(0.0001)	(0.00003)	(0.0001)	(0.0001)	(0.0002)
City Fixed Effects	No	No	Yes	Yes	No	No
Observations	31,737	31,737	31,737	31,737	8,003	8,003
R^2	0.1273	0.1377	0.2931	0.3028	0.1566	0.1704

Table 7: LWR Noise Coefficient Over Time (bandwidth = 2,000 nearest sales)

	Dependent variable: LWR Noise Coefficient (bandwidth = 2,000 nearest sales)					
	(1)	(2)	(3)	(4)	(5)	(6)
Months since Jan 2005	-0.00005^{***}	0.00003^{***}	-0.00005^{***}	0.00003^{***}	-0.00005^{***}	0.0001^{***}
	(0.0000003)	(0.000002)	(0.0000003)	(0.000002)	(0.000001)	(0.000004)
(Months since Jan 2005)^2		-0.000001^{***}		-0.000001^{***}		-0.000001^{***}
		(0.00000002)		(0.00000002)		(0.0000001)
Constant	-0.0014^{***}	-0.0026^{***}	-0.0016^{***}	-0.0028^{***}	-0.0016^{***}	-0.0033^{***}
	(0.00001)	(0.00003)	(0.00002)	(0.00003)	(0.00003)	(0.0001)
City Fixed Effects	No	No	Yes	Yes	No	No
Observations	31,737	31,737	31,737	31,737	8,003	8,003
R^2	0.3828	0.4142	0.5128	0.5435	0.3454	0.3988