Real Estate Finance

VOLUME 37, NUMBER 1 SUMMER 2020

Highway Noise and Elevation Effects on Nearby Home Prices: Spatial Econometrics Using LIDAR-Derived Data

By Michael A. McElveen, Brian E. Brown, and Charles M. Gibbons

Michael A. McElveen, MAI, CCIM, CRE *is president and founder of Urban Economics in Tampa, FL. He has a Bachelor of Arts in Finance from the University of South Florida and a Bachelor of Science in Risk Management/Insurance, Real Estate, and Legal Studies from Florida State University. Professional practice is in real estate econometrics, machine learning, and litigation support. He has been a contributor to Assessment Digest, Property Tax Journal, and Journal of Housing Research. He may be contacted at* mcelveen@urbaneconomics.com.

Brian E. Brown, MA, *is a data scientist at Urban Economics. He has a Bachelor of Science in Economics from Florida State University and a Master of Arts in Economics from the University of South Florida. Professional practice is in real estate econometrics, machine learning, and litigation support. He may be contacted at* bbrown@urbaneconomics.com.

Charles Gibbons, MAURP, GISP *is a GIS Professional at Urban Economics, where he is lead geospatial analyst and data engineer. He has a Bachelor of Science in Geography and a Master of Arts in Urban and Regional Planning from the University of Florida. He may be contacted at* cgibbons@urbaneconomics.com.

The study of traffic noise and the sales price effect on nearby homes has been on-going since the early 1970s as a result of a mandate in the Federal Aid Highway Act of 1970,[1] which required that highway project decisions be made considering the costs to eliminate or minimize the adverse effects of air, noise, and water pollution and to measure those factors that relate to the social and environmental impacts (traffic noise) of the highway project.

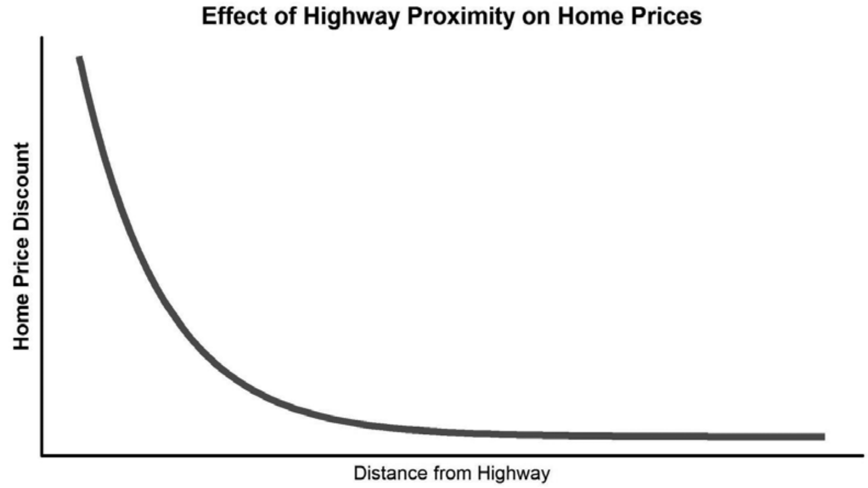
The subjective annoyance of traffic noise is the result of a complex mix of traffic speed, vehicle mix, time of day, wind direction, elevation of the highway and noise barriers. Traffic noise creates a sorting of home buyers, with homes that have a higher level of noise purchased by individuals with a low willingness to pay for a reduction of noise and quiet homes purchased by individuals with a high willingness to pay for a reduction of noise.

Aside from being an annoyance, traffic noise can have adverse effects on the health of nearby residents who are continuously exposed to it. Tetreault[2] notes that indirect traffic noise can increase cardiovascular disease. He posits that exposure to traffic noise increases stress responses from neural activation or cognitive interpretation, and traffic noise has been shown to disrupt sleep patterns another stress inducer. A person's avoidance of stress-inducing traffic-generated noise may be a factor in a homebuyer's capitalization of traffic noise into the sales price of a home.

Every home is characterized by a combination of structural, neighborhood, and environmental characteristics that can have a positive, negative, or no effect on sales price. It is well-recognized in empirical studies that traffic noise is a negative environmental externality. An externality is one that affects only those in close proximity to the externality. Positive externalities have a beneficial effect on price, and these include nearby parks, golf courses, retail centers, and recreation areas; negative externalities can include nearby landfills, blighted areas, and manufacturing plants.

The housing consumer is a wealth maximizer, always trying to choose a combination

Exhibit 1



of housing, neighborhood, and environmental characteristics that maximize utility, subject to the consumer's housing budget. If a home's proximity to traffic-generated noise and the ensuing noise level is perceived as a negative characteristic of a home's location (environmental characteristic), the magnitude of the market's capitalization of the negative externality into the sales price of the home will be revealed by the lower price paid for this bundle of environmental characteristics.

Costless freedom to move is crucial to fully measure traffic externality effects by a sales price-driven hedonic regression model; however, home transactions are anything but costless, as exemplified by the existence of real estate agents, home loan brokers, closing agents, and transaction taxes. There also is the social cost of moving personal belongings and uprooting from established local connections. Theoretical and empirical studies have shown that high transaction costs in the housing market create a lock-in effect; therefore, price differentials derived from a sales price-driven hedonic regression model will be understated.

Highway noise decreases rapidly, by three to six decibels for each doubling of distance from the traffic noise source, if all other aspects that facilitate the travel of noise over space are constant. The elevation of the noise transmitter, "in this instance vehicles on the highway", above the surrounding landscape increases the distance and volume of the noise because the noise is less obfuscated by natural or manmade barriers that otherwise would have a limiting effect. The effect of the elevated noise source is analogous to the effect of an elevated stage or podium on the clarity and volume of a public speaker or music. The relationship between distance from the highway and its effect on price should be similar to Exhibit 1.

LITERATURE REVIEW

Empirical literature regarding the effect of traffic noise on the prices of nearby homes has been prevalent in real estate economic literature for almost half a century. In the 1970s and 1980s, data limitations, computational power and constrained modeling techniques limited the scope of earlier studies relative to later studies; however, significant advances in the understanding of transportation-related externalities were nonetheless made during this period.

Ambient noise (traffic noise) was the focus of much of the early work on price effects of highway proximity, and typically was found to be a primary driver of sales price variability. Vaughan and Huckins,[3] Anderson and Wise,[4] and Allen[5] used the decibel level of traffic noise at the home sale and found an inverse relationship between decibel level of traffic noise and the revealed price of the home, while Hall *et al.*[6] obtained similar results but with mixed statistical significance. Langley[7] and Bailey[8] studied highway noise, using the home's distance from the highway as the variable of interest, finding that the price of homes approximately adjacent to the highway sold at a discount relative to home sales outside of the highway's area influence. Nelson[9] found negative home price effects of highway proximity in a comprehensive economic analysis of transportation noise abatement.

Hughes and Sirmans[10] found that the reduction in the price of a home was correlated with traffic count (the premise is that higher traffic counts generate more noise than a lower traffic count) at the nearby highway, with price effects on homes in urban areas exceeding price impacts on suburban homes by nearly double.

More recent studies on this topic have benefited from more data and a wider spectrum of data parameters, which, in turn, have permitted and have provided analysis using a broader assortment of parameters to be studied, as well as providing opportunities to utilize more sophisticated modeling techniques. The collective result of more data and a wider array of parameters has been a discovery of nuances about the effect of traffic noise that previously were not obvious or significant because of the sparsity of data and analytics available.

Carey[11] conducted an analysis of the price effects of a home's proximity to the Superstition Freeway in Phoenix, Arizona. The regression model was constructed using methods intended to mitigate impacts of traffic noise, air pollution, and visual externalities including depressed grade construction of the highway (constructing the freeway such that the freeway elevation is lower than the elevation of abutting properties), vegetated right-of-way barriers, and a noise-reducing barrier wall adjacent to the freeway. Home sales in the study were grouped into three zones: (1) a zone of highway adjacency, (2) a broader impact zone, and (3) a control group outside of the impact zone. Carey found that homes adjacent to the freeway incurred a significant reduction in price, and home sales in the broader impact zone incurred a reduction in price approximately half that of the adjacent home sales.

Kilpatrick *et al.*[12] examined the effect of a home's proximity to transit corridors, controlling for the presence or lack of access to and from the corridor. Their regression analytics indicate that the sales prices of homes within 300 feet of a transit corridor are negatively affected by the externalities of the nearby corridor.

Chernobai *et al.*[13] analyzed the effect of a newly completed highway extension in Los Angeles, California on the prices of nearby homes using a spline regression technique. Their analysis allowed spatial and temporal price effects of the highway construction to be examined concluding that homes closest to the highway extension appreciated in price over time at a slower rate than homes located further from the highway extension.

Using spatially-weighted modeling techniques, Allen[14] analyzed the price effects of highway proximity on homes along the Wekiva Parkway in Orange County, Florida. Between the original ordinary-least-squares (OLS) model, the spatial lag model, and the spatial error model, the price effects of the highway proximity variables were all significant and negative.

The aforementioned studies solidly support of the opinion that proximity to a highway has a negative effect on the price of a home. These studies have been conducted using a variety of statistical modeling methods across a variety of geographic areas, and all have reached the same conclusion: that the proximity of a home to a highway and its resulting negative externalities of noise, pollution, and vibration have a negative effect on the price of a home. Additionally, these studies examined various indicators of the price effects of highway proximity, such as visual disamenity, noise, and general proximity stigma, but none of this previous research considered the height of the highway above the home.

It is clear from the collective results of the many studies going back nearly 50 years that highway proximity is a negative externality on home prices, and this phenomenon holds across time and locale. The literature also provides us with a predicate on which to base our analysis—clearly, hedonic regression modeling is a common empirical real estate technique and is appropriate to apply in our examination of the price effect of a highway's proximity to a home and of the height of the highway above a home lot. Allen *et al.* is particularly insightful, as their use of the spatial lag and spatial error models offers a predicate to the spatially-weighted modeling techniques employed in our study.

CONTRIBUTIONS OF THIS STUDY

The use of a sales price-powered hedonic regression model to determine the effect of the traffic externality on the price of a home is well-covered territory. All of the empirical studies used either the decibel level of traffic noise or the distance of the home from the highway as a proxy for the level of traffic noise. Many factors affect the level of traffic noise at a home, and one of the more impactful factors is the relative height of the highway (noise transmitter) above the home (noise receiver). All things being equal, there are fewer objects, trees, walls, and buildings to impede noise from an elevated highway versus a highway near the grade of the home. A common method used to reduce the effect of traffic noise on a home is installation of a barrier wall between the noise source and the home, but the function of the barrier wall is defeated if the elevation of the highway is above the height of the barrier wall— the noise just goes over the wall unimpeded.

Previously, it was not cost-effective or possible to measure or obtain the lot elevation of a large number of home sales or the height of the nearby highway for a large dataset to drive a hedonic regression model. This limitation has been mitigated by the proliferation of spatially accurate Light Detection and Ranging (LIDAR) data. LIDAR is defined by The National Oceanic and Atmospheric Administration (NOAA) as "a remote sensing method that uses light in the form of a pulsed laser to measure ranges (variable distances) to the Earth."[15] Using LIDAR elevation data, we obtained highly accurate lot elevations and the elevation of the highway at the closest Euclidean distance cost-efficiently and expeditiously.

Elevations and distances based on LIDAR Digital Elevation Model data and raw LIDAR data are highly accurate, and LIDAR is affected by vegetation but it can be accounted for by means of classification, thus LIDAR has many applications in the study of real estate, such as 3D modeling of the built environment, viewshed analysis, distance measuring, and site planning. We believe this is the first implementation of LIDAR data combined with distance to measure the effect of an externality on sales prices.

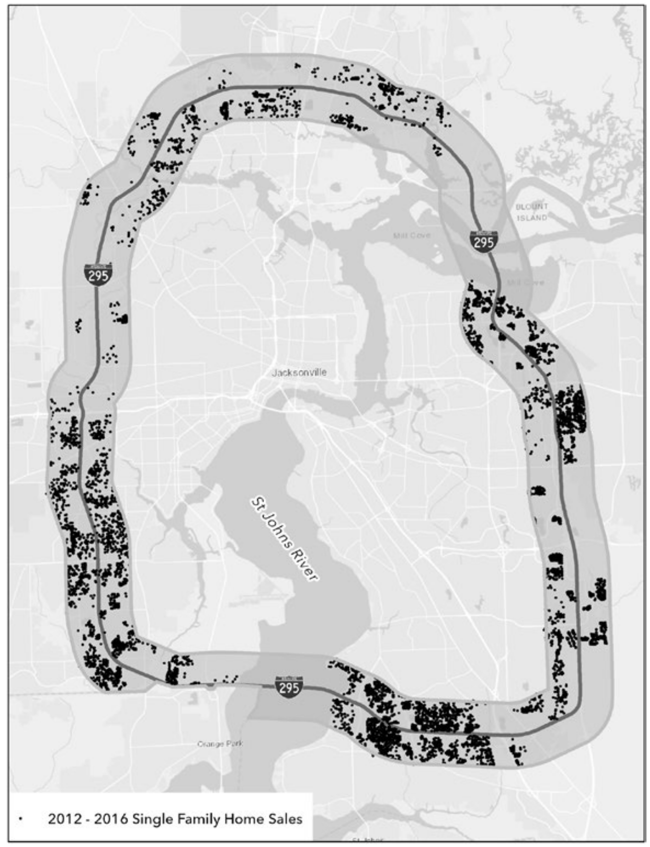
DATA ACQUISITION

Our study area is the one-mile buffer along Interstate 295 in Jacksonville, Florida. Jacksonville, Florida is a medium-sized port city in northeast Florida. It is bisected by the St. Johns River, and it has an estuary with the Atlantic Ocean. Jacksonville has a reasonably well-defined central business district located along the banks of the St. Johns River and in the approximate center of the Interstate 295 ring road.

We obtained from the Duval County Florida Property Appraiser their database of qualified home sales that occurred between 2012 and 2016. "Qualified" in this case means the Office of the Property Appraiser has determined that the sale conforms to the tenets of an arm's length transaction. This database contained structural and transactional characteristics of the home sales such as lot size, living area, home type (attached vs. detached), garage size, date of sale, sale price, presence of a pool or spa, number of bedrooms and bathrooms, and year built.

Using ArcGis version 10.4.1, we calculated the Euclidean distance between each home sale and the centerline of the closest travel lane of Interstate 295. The study area of home sales extends out to one mile from this datum line of Interstate 295. Within the one-mile buffer there are 8,010 qualified single-family home sales that occurred during the study period. Exhibit 2 is a locational reference of the study area with the location and density of home sales data.

Exhibit 2



The elevation difference between the highway and the home sale lot was captured by the difference in elevation between each lot and the nearest datum line on Interstate 295. We calculated the elevation of each lot relative to the elevation of the closest-by-linear-distance travel lane centerline on Interstate 295. The elevation of the lot was obtained from LIDAR-derived Digital Elevation Model data, and the elevation of the highway centerline was derived from raw LIDAR data.

Lot elevations were based on the FLIDAR MOSAIC dataset from the University of Florida—GeoPlan Center. This dataset represents a five-meter cell size Digital Elevation Model (DEM) covering the State of Florida. The elevation units are expressed in feet. The vertical datum is NAVD88, meters, and the projection is Albers Equal Area Conic HARN, meters. The DEM was created by mosaicking data from four different sources, with the following order of priority: (1) NWFWMD District DEM, (2) FLIDAR Coastal DEM, (3) Statewide FWC DEM, and (4) Contour Derived DEM.

The highway centerline elevations were derived from the 2007 Florida Division of Emergency Management (FDEM) Lidar Project: Duval County dataset. The LIDAR data was collected in March 2007 and published in 2009 via NOAA. This dataset represents actual land cover elevations for the entirety of Duval County. The elevations returned are in feet and accurate within 0.3 feet RMSE as compared to ground control points. The LIDAR data was transformed into a traditional elevation raster. This raster was used to calculate the average height of the highway along centerline for a given 350-foot-long segment.

School zone information was digitized from the location map of Duval County Public Schools. Traffic count information was obtained from the Florida Department of Transportation. Additional locational information was obtained via the ESRI online catalog.

THEORETICAL FRAMEWORK

The theoretical framework of this study is built on the principle that the sales price of a home reveals the sum of the values of the physical attributes of the home, the characteristics of the neighborhood, and the environmental characteristics, which include noise, near the home. If proximity to a highway is perceived by the housing market as a negative externality, the home buyer/seller will deduct from the value of the home a proportional amount to compensate for this negative aspect. Theoretically, the closer a home is to the negative externality, the higher the compensation offset and the lower the sale price, while the further the home sale is from the externality the less of an offset and the higher the sale price, all other factors held constant. The negative externality effect on sales price can also vary in magnitude. In this case, a greater difference in height between the expressway and the lot elevation implies a stronger negative externality; hence a greater negative sales price effect is expected. This difference in sales price can be combined with linear proximity to the expressway to achieve a comprehensive picture of the effect that the expressway has on the sale price of nearby homes.

To determine whether this hypothesis is reflected in the marketplace of home buyers and sellers, we compared the sales prices of homes that are adjacent to or very near Interstate 295, as well as home sales in an adjacent but farther removed area, with a control area that is outside the affected areas. An area of immediate impact was used in place of an adjacency zone, which has commonly been the parameter of interest in previous studies on this topic. This is due to geospatial characteristics of the dataset that render an unimpeded straight-line indicator an ineffective means of determining which homes are and are not primarily impacted by proximity to the highway. A better indicator of the area of immediate effect was found to be location within a band of fixed distance from the centerline of Interstate 295. The prevalence of immediately affected homes was most efficiently maximized at a distance of 500 linear feet from the Interstate 295 centerline of nearest travel lane. The wider affected zone included homes within 500 to 1,500 linear feet from the highway centerline of nearest travel lane, while the control group is all home sales in the dataset that are more than 1,500 linear feet from the Interstate 295 centerline of the nearest travel lane.

These locational attributes are interacted with the difference between the elevation of each home lot and the height of the highway at 20-foot distance intervals corresponding to each home in order to capture the effect of the elevation difference specific to each location zone.

VARIABLES

Exhibit 3 is a list and description of the independent variables used in this hedonic regression analysis.

The dependent variable, home sales price, is transformed to a natural logarithm form to better approximate a normal distribution and to improve the interpretability of the model coefficients.

Lot acreage and home square feet also were transformed to a natural logarithm to account for diminishing marginal utility of property size. Age of the home, number of bedrooms and bathrooms, and square feet of garage area were kept in their linear form as generally is appropriate in hedonic models of home price. The presence of a pool or spa was treated as a binary variable, as was the variable indicating whether the home was attached versus detached and the variable indicating that the home was a new build.

Neighborhood characteristics that influence home prices are primarily economic in nature and can include median household income, percentage of renters, and other variables. A sufficient proxy for these neighborhood characteristic variables is school district, which has a significant effect on home prices due to the prestige of the school itself, as well as the associated economic characteristics that are

Exhibit 3

Variables and Descriptions

Variable	Description
LN Price	The natural logarithm of the sale price in dollars (dependent variable).
Age	Actual age of the home.
Bath	Number of bathrooms.
Bed	Number of bedrooms.
Garage	Square feet of garage space.
LN Acres	The natural logarithm of the number of acres of land.
LN SF	The natural logarithm of the number of square feet of building area.
PoolSpa	Binary variable indicating the presence of a pool or spa.
Attached	Binary variable indicating whether the home was attached to a separate dwelling unit.
NewBuild	Binary variable indicating whether the sale was likely to be that of a newly-built home (determined to be between zero and one year of actual age at the time of sale).
Atlantic_Coast	Binary variable indicating that the home is located in the Atlantic Coast High School district, relative to the Westside Charter School district.
Ed_White	Binary variable indicating that the home is located in the Edward White High School district, relative to the Westside Charter School district.
Englewood	Binary variable indicating that the home is located in the Englewood High School district, relative to the Westside Charter School district.
First_Coast	Binary variable indicating that the home is located in the First Coast High School district, relative to the Westside Charter School district.
Mandarin	Binary variable indicating that the home is located in the Mandarin Charter School district, relative to the Westside Charter School district.
Ribault	Binary variable indicating that the home is located in the Jean Ribault High School district, relative to the Westside Charter School district.
Sandalwood	Binary variable indicating that the home is located in the Sandalwood High School district, relative to the Westside Charter School district.
Lee	Binary variable indicating that the home is located in the Robert E. Lee High School district, relative to the Westside Charter School district.
East_of_River	Binary variable indicating whether the home was located east of the St. Johns River.
Waterfront	Binary variable indicating whether the home had a waterfront location.
2013	Binary variable indicating whether the sale ocurred in 2013, relative to 2012.
2014	Binary variable indicating whether the sale ocurred in 2014, relative to 2012.
2015	Binary variable indicating whether the sale ocurred in 2015, relative to 2012.
2016	Binary variable indicating whether the sale ocurred in 2016, relative to 2012.
Wall	Binary variable indicating that a sound barrier wall exists between the home and the highway.
Height_Difference	Feet of difference in height between the home and the highway.
AADT1000	Average daily traffic count on Interstate 295 at the home's location, divided by 1,000.
D_0_500	Binary variable indicating that the home is located within 500 feet of the Interstate 295 centerline, relative to the control group (greater than 1,500 feet from the Interstate).
D_500_1500	Binary variable indicating that the home is located between 500 and 1,500 feet of the Interstate 295 centerline, relative to the control group (greater than 1,500 feet from the Interstate).
D_0_500_Diff	Interaction term indicating the difference in height between the home and the highway for homes located within the 0-500-foot distance band.
D_500_1500_Diff	Interaction term indicating the difference in height between the home and the highway for homes located within the 500-1,500-foot distance band.

delineated by the boundaries of these districts. The school district variables are in relation to the Westside Charter School district, which is located in the southwestern portion of the study area.

Time of sale is captured with binary variables indicating the year of sale. These binary variables are relative to 2012.

Average daily traffic count (AADT1000) is the average daily traffic count on Interstate 295 nearest the home sale location. The traffic count is divided by 1,000 to facilitate simpler interpretation of the variable's coefficient.

A wall binary variable is included in the model to control for the effect of a sound barrier wall between the highway and the home on the sales prices of nearby homes. The sound barrier wall is a means of reducing the effect of highway traffic noise on nearby homes; the actual effect of the variable as a noise suppressor or as a visual disamenity remains to be seen.

The Height_Difference variable indicates the difference in feet between the elevation of the home lot and the height of the expressway at the nearest linear distance from the home.

Distance from the highway functions as an indicator of the zone of highway influence, and has been defined by distance bands, the closest of which is a distance band of zero to 500 feet from the Interstate 295 centerline, while the middle-distance band is 500 feet to 1,500 feet from the highway centerline. The distance bands included in the model are relative to the area located between 1,500 feet from the highway centerline and the edge of the study area (one mile from the highway centerline of the nearest travel lane)—this area contains the control group of observations.

Because Height_Difference is expected to have a varying effect on price according to distance from the highway, interaction terms between the distance band variables and the Height_Difference variable are used. This interaction will enable the model to more precisely capture the price effect of the height difference of and proximity to the highway.

METHODOLOGY

Hedonic regression is a well-established methodology appropriate for economic analyses of real estate data. Housing as a consumer product consists of a bundle of structural, neighborhood, environmental, temporal, and external characteristics. The sales price of a home indicates the buyer's revealed sum of the values of these characteristics. Hedonic regression estimates the price of each individual price-influencing characteristic, holding all else constant.

The first model used to represent the data in this analysis is ordinary least squares (OLS) estimation. The OLS model is specified as follows:

$$ln(P) = \alpha + H_{SLT}\beta + E\theta + Z\delta + \gamma\phi + \varepsilon \tag{1}$$

where $ln(P)$ is the natural logarithm of the sale price of a home; α is the regression constant, H_{SLT} is a matrix of structural, locational, and transactional characteristics; E is the elevation difference between the highway and the lot of the home at the shortest linear distance; Z is the zone of highway influence; γ is the interaction of elevation difference and zone of highway influence; β, θ, δ, and ϕ are coefficient vectors, and ε is the error term.

The OLS model serves as the starting point for the spatial analysis. Econometric analyses of real estate data tend to be subject to a high degree of spatial autocorrelation, which can render parameter estimates unreliable. Thus, to properly specify econometric models with real estate components, the spatial distribution of the data must be considered.

To determine whether spatial autocorrelation exists in the model, we employ Moran's I and Lagrange multiplier tests and examine their significance. Moran's I is a misspecification test that indicates whether a null hypothesis of no spatial autocorrelation must be rejected on the basis of spatially clustered residuals from the initial OLS model. If this null hypothesis is rejected, it can be inferred that the model is mis-specified, and that a spatial component must be included in the model to properly account for the spatial distribution of the data. The Lagrange multiplier tests for spatial autocorrelation can then indicate the function in which the spatial autocorrelation most likely will be captured best in a hedonic model. (*See* Addendum at end of article for details on the Lagrange multiplier tests and Moran's I.) This function will then be expressed and controlled for using the appropriate models—the spatial lag model and/or the spatial error model.

The spatial lag model and the spatial error model address spatial autocorrelation via fundamentally distinct assumptions regarding the spatial distribution of the data and the spatial relationships within each model. The spatial lag model assumes that the spatial autocorrelation exhibited in the model is a function of spatial clustering with respect to values of the dependent variable; thus, under such conditions, the proper estimation of the dependent variable of any given observation considers the values of the independent variables as well as the spatially lagged values of the dependent variable of the observation's spatial neighbors. The spatial lag model is specified as follows:

$$\begin{gathered} ln(P) = \alpha + \rho W\left[ln(P)\right] + H_{SLT}\beta + E\theta + Z\delta + \gamma\phi + \varepsilon \\ \varepsilon \sim N(0, \sigma^2 I_n) \end{gathered} \tag{2}$$

where ρ is the autoregressive parameter and W is a row-standardized connectivity matrix defined as follows:

$$W = w_{ij}^{*} = w_{ij}/\Sigma_j w_{ij} \tag{3}$$

The spatial weights matrix is standardized to allow the weights to be applied proportionally across all observations, which in turn allows spatially isolated observations to be included in the spatial process according to the magnitude of their spatial influence.

The spatial error model differs from the spatial lag model in that the underlying assumption about the nature of the spatial autocorrelation in the model is that the autocorrelation is due to unobserved explanatory factors embedded within the error term. The spatial error model is specified as follows:

$$ln(P) = \alpha + H_{SLT}\beta + E\theta + Z\delta + \gamma\phi + \varepsilon$$
$$\varepsilon = \lambda W\varepsilon + u \quad (4)$$

where λ is the autoregressive parameter, W is the row-standardized spatial weights matrix defined in equation (3), and u is spatially independent.

ρ and λ, the autoregressive parameters in the spatial lag and error models, respectively, are interpreted as indicators of spatial dependence of their model's respective form. In other words, a statistically significant ρ indicates that the spatial lag model is indeed controlling for some amount of spatial dependence within the context of a spatially lagged dependent variable. Similarly, a statistically significant λ indicates that the spatial error model is controlling for some amount of spatial dependence in the model error.

While the spatial lag and spatial error models address the spatial dependence inherent in the sales data, they each operate on distinct underlying assumptions about the nature of the spatial autocorrelation exhibited in the original model. The relative advantage or disadvantage of each model's specification is determined by the Akaike information criterion (AIC). (*See* Addendum at end of article for details on the AIC.)

RESULTS

Exhibit 4 contains the summary statistics for the subject dataset, and Exhibit 5 contains the results of the OLS and spatial models. Heteroscedasticity-robust standard errors were used to calculate the t-statistics for each coefficient in the OLS model, following detection of unequal error variance via White's general test, which is common in hedonic regression models of real estate sale data. The OLS model had an R-squared value of 0.8567 and an adjusted R-squared value of 0.8562.

In the OLS model, the signs of each coefficient of the structural, transactional, and locational variables were as expected, with the exception of Bed, which was neither statistically nor economically significant. Wall was significant and positive, indicating that a barrier wall is associated with a price premium for affected homes. AADT1000 was significant and negative, which, like Wall, had effects throughout the study area, although at a low degree of economic significance. The proximity binary variables were significant and negative, as expected, with the closest locational variable having a greater negative effect than the middle location variable, relative to the control group. Height_Difference by itself was not statistically significant, as expected, as homes within the control group comprise 79.3 percent of the dataset and are thus unlikely to be affected by a visual disamenity beyond the range of sight. However, the interaction term combining elevation difference with proximity to the highway within 500 feet was significant and negative, while the interaction of elevation difference and the middle- distance band was not statistically significant. This indicates that there is an inverse relationship between home price and the elevation difference between the home and the nearby highway, but this relationship exists only within 500 feet of the highway.

The residuals of the OLS model were used to calculate the Moran's I statistic, which had a value of 0.271 with

Exhibit 4

Summary Statistics

Variable	Mean	Median	Std. Dev
Sale Price	165,767	144,500	111,031
Distance from Interstate (ft)	2,861	2,907	1,389
Difference in Elevation (ft)	7.13	5.28	14.67
Annualized Average Daily Traffic Count	83,840	77,500	28,695
Lot Acres	0.28	0.20	0.43
Living Area (SF)	1,794	1,656	645
Bedrooms	3.2	3.0	0.7
Bathrooms	2.2	2.0	0.6
Age	24	22	18
Garage Area (SF)	352	400	209

Exhibit 5

Results

OLS		Spatial Lag		Spatial Error	
Variable	Coeff.	Variable	Coeff.	Variable	Coeff.
Constant	6.470 ***	Constant	0.596 **	Constant	6.600 ***
Age	-0.007 ***	Age	-0.006 ***	Age	-0.006 ***
Bath	0.054 ***	Bath	0.045 ***	Bath	0.041 ***
Bed	-0.002	Bed	0.001	Bed	0.001
Garage	0.000 ***	Garage	0.000 ***	Garage	0.000 ***
LN_Acres	0.077 ***	LN_Acres	0.073 ***	LN_Acres	0.084 ***
LN_SF	0.688 ***	LN_SF	0.660 ***	LN_SF	0.671 ***
PoolSpa	0.124 ***	PoolSpa	0.125 ***	PoolSpa	0.117 ***
Attached	-0.235 ***	Attached	-0.251 ***	Attached	-0.274 ***
NewBuild	0.144 ***	NewBuild	0.150 ***	NewBuild	0.157 ***
Atlantic_Coast	0.314 ***	Atlantic_Coast	-0.062 ***	Atlantic_Coast	-0.244 ***
Ed_White	0.040 **	Ed_White	0.003	Ed_White	0.046 **
Englewood	0.135 ***	Englewood	0.004	Englewood	-0.187 ***
First_Coast	0.061 ***	First_Coast	-0.085 ***	First_Coast	0.257 ***
Mandarin	0.160 ***	Mandarin	-0.128 ***	Mandarin	-0.261 ***
Ribault	-0.030 **	Ribault	-0.116 ***	Ribault	0.288 ***
Sandalwood	0.046 ***	Sandalwood	0.035 ***	Sandalwood	-0.035
Lee	-0.009	Lee	-0.086 ***	Lee	-0.069 ***
East_of_River	0.273 ***	East_of_River	0.132 ***	East_of_River	0.967 ***
Waterfront	0.384 ***	Waterfront	0.402 ***	Waterfront	0.399 ***
2013	0.067 ***	2013	0.062 ***	2013	0.064 ***
2014	0.154 ***	2014	0.151 ***	2014	0.149 ***
2015	0.253 ***	2015	0.247 ***	2015	0.247 ***
2016	0.304 ***	2016	0.298 ***	2016	0.298 ***
Wall	0.037 ***	Wall	0.042 ***	Wall	0.018 **
Height_Difference	0.000	Height_Difference	0.000	Height_Difference	0.000
AADT1000	-0.001 ***	AADT1000	0.001 ***	AADT1000	0.000
D_0_500	-0.041 ***	D_0_500	-0.047 ***	D_0_500	-0.034 **
D_500_1500	-0.028 ***	D_500_1500	-0.041 ***	D_500_1500	-0.032 ***
D_0_500_Diff	-0.005 ***	D_0_500_Diff	-0.004 ***	D_0_500_Diff	-0.005 ***
D_500_1500_Diff	0.001	D_500_1500_Diff	0.001	D_500_1500_Diff	0.001
		Rho	0.515 ***	Lambda	0.986 ***

Note:
**Significant at the 95% confidence level.
***Significant at the 99% confidence level.

a z-score of 64.045, leading us to reject the null hypothesis that spatial autocorrelation is not present in the OLS model. Lagrange multiplier tests for spatial lag and spatial error were both highly indicative of spatial autocorrelation, leading us to consider both model specifications in our effort to capture and control for spatial autocorrelation. Accordingly, we proceeded with the spatial lag and spatial error models.

Some divergence between the OLS model and the spatial models in terms of coefficient signs and statistical significance is immediately apparent. This divergence is most pronounced among the school zone variables in the spatial error model, indicating that accounting for the spatial processes within the unobserved error term precipitates much different indications of neighborhood effects within each model. The same is true to a lesser degree for the spatial lag model, which accounts for spatial processes via a spatially lagged dependent variable. Bed also changed sign but remained statistically insignificant. AADT1000 changed sign in both spatial models and lost significance in the spatial error model, although the economic significance of this variable remained low.

The magnitude of the middle-distance band relative to the closest distance band narrowed significantly between the OLS model and each spatial model, while the overall magnitude of each proximity indicator was diminished in the spatial error model. Interestingly, the statistical significance of the closest distance band diminished somewhat from 99 percent to 95 percent confidence, while the middle-distance band remained significant at 99 percent confidence. D_0_500_Diff remained negative and significant at 99 percent. D_500_1500_Diff remained insignificant in each spatial model.

The autocorrelation parameters are statistically significant in each spatial model, indicating that spatial autocorrelation is prevalent in the dependent variable as well as the errors of the original OLS model. Thus, overall model fit has been improved relative to the OLS model, by virtue of having controlled for the autocorrelation indicated by the Moran's *I* statistic. Furthermore, the log-likelihood statistic shows improvement from 1060.508 in the OLS model to 1287.829 and 1451.818 in the spatial lag and spatial error models, respectively. Model fit can be further assessed using the AIC, which is -2511.658 for the spatial lag model and -2841.636 for the spatial error model, which indicates that the spatial lag model provides a more accurate representation of the data relative to the spatial error model. To extrapolate, it is implied that the spatial autocorrelation is more strongly associated as a direct function of the spatially-lagged dependent variable, relative to that of the error term.

CONCLUSIONS

The models, irrespective of their spatial components, suggest that homes that are approximately adjacent to an elevated highway can be expected to sell for less than homes that are further from the highway, all else held constant. Within this zone of adjacency the elevation between the lot of the home and the highway is inversely associated with expected sale price. This interaction effect dissipates quickly outside of the zone of adjacency, which is expected, as the prominence effect of the highway likewise diminishes quickly with distance.

Spatial autocorrelation was at issue in our data, and required consideration and mitigation. Two spatially autoregressive models were employed, which improved model fit and mitigated spatial dependence relative to the original OLS model. Spatial dependence was found to be a function of both the spatially-lagged dependent variable and the unobserved error term; thus, spatial dependence was not entirely mitigated in either of the two autoregressive models, an issue that merits further research. However, our analysis elucidates several important issues – that LIDAR data can be used to effectively analyze real estate data using GIS software and techniques, that the magnitude of visual and auditory disamenities can be estimated using this technique, and that these quantified magnitudes interact with distance from said disamenities.

Further research using geospatial viewshed analysis combined with a LIDAR built environment model to measure the visual prominence of the highway from a home could be helpful in ascertaining how much of the highway negative effect on price is a result of noise, dust and vibration and how much, if any, is a result of the visual impairment of the highway.

ADDENDUM

Anselin[16] developed Lagrange Multiplier (LM) statistics specific to spatial lag and spatial error models. Each associated LM tests the null hypothesis that the autoregressive coefficient (ρ in the spatial lag model and λ in the spatial error model) equals zero, against the alternative hypothesis that the coefficient does not equal zero. In effect, the LM tests whether the standard OLS model should be rejected as an estimator against the alternative of each respective spatial model.

The LM test against a spatial lag alternative is defined as:

$$LM_\rho = \frac{\left[\frac{e'Wy}{\hat{\sigma}^2_{ML}}\right]^2}{D} \sim \chi^2 \tag{5}$$

where e is a vector of OLS residuals, Wy is the spatial lag term, and

$$\hat{\sigma}^2_{ML} = \frac{e'e}{n} \tag{6}$$

The denominator of equation (5) has two components:

$$D = \frac{(WX\hat{\beta})'[I - X(X'X) - 1\ X'](WX\hat{\beta})}{\hat{\sigma}^2_{ML}} + T \tag{7}$$

The first term is the sum of the squared residuals of the spatially lagged predicted values. The second term is a trace expression defined as,

$$T = tr\left(WW + W'W\right) \tag{8}$$

The LM test against a spatial error alternative is defined as:

$$LM_\lambda = \frac{\left[\frac{e'We}{\hat{\sigma}^2_{ML}}\right]^2}{T} \sim \chi^2 \tag{9}$$

where e is a vector of OLS residuals, We is the spatial lag term, and σ^2_{ML} and T are as defined in equations X and X.

Moran's I is a test for spatial autocorrelation applied to the residuals from an initial non-spatial regression model. Rejection of the null hypothesis of no spatial autocorrelation indicates a diffuse alternative, in which the appropriate specification controls for spatial heterogeneity but is not specific in terms of the form the controls should take. Moran's I is defined as:

$$I = \frac{e'We/S_0}{e'e/n} \tag{10}$$

$$S_0 = \Sigma_i \Sigma_j w_{ij} \tag{11}$$

Measures of fit are difficult to define for stand-alone models, but comparisons of explanatory power between models can be easily obtained. Maximum likelihood (ML) estimators, which is the form of estimation of the spatial lag and spatial error models, can be compared to an equivalently-specified OLS estimator via a maximized log-likelihood function, defined as:

$$L = -\left(\frac{n}{2}\right) ln2\pi - \left(\frac{n}{2}\right) ln\hat{\sigma}^2_{ML} - \frac{n}{2} \tag{12}$$

Similar to an adjusted R-squared, the log-likelihood metric can be made to accommodate varying numbers of regressors. The Akaike Information Criterion (AIC) is one such metric, defined as:

$$AIC = -2L + 2k \tag{13}$$

L is the likelihood function defined in equation 12 and k is the number of independent parameters.

While the log-likelihood statistic indicates that the higher score indicates comparative superiority in terms of model fit, the AIC indicates superiority via the lower score.

NOTES

1. Federal-Aid Highway Act of 1970, Pub. L. No. 91-605, § 136, 1734 (United States Printing Office 1973).
2. Tetreault, L. F., Perron, S., and Smargiassi, A., "Cardiovascular Health, Traffic-Related Air Pollution and Noise: Are Associations Mutually Confounded? A Systematic Review," *International Journal of Public Health,* 58, 649–666 (2013).
3. Vaughan, R. J., and Huckins, L., *The Economics of Expressway Noise Pollution Abatement* (P-5475), Santa Monica, CA: The Rand Corporation (1975).
4. Anderson, R. J., and Wise, D. E., *The Effects of Highway Noise and Accessibility on Residential Property Values* (DOT-FH-11-8841), Springfield, VA: National Technical Information Service (1977).
5. Allen, G. R., *Relationships between Highway Noise, Noise Mitigation, and Residential Property Values* (FHWA/VA-81/1), Charlottesville, VA.: Virginia Highways and Transportation Research Council (1980).
6. Hall, F. L.; Breston, B. E.; and Taylor, S. M., *Effects of Highway Noise on Residential Property Values* (Transportation Research Record 686), Washington, D.C.: National Academy of Sciences, 38-43 (1978).
7. Langley, C. J., "Adverse Impacts of the Washington Beltway on Residential Property Values," *Land Economics,* 52(1), 54–65 (1976); Langley, C. J., "Highways and Property Values: The Washington Beltway Revisited," (Unpublished paper), University of Tennessee (1980).
8. Bailey, M. J., "Report on Pilot Study: Highway Noise and Property Values" (Unpublished Paper), University of Maryland (1977).
9. Nelson, J. P., Economic Analysis of Transportation Noise Abatemen, Cambridge, MA: Ballinger (1978).
10. Hughes, W. T., & Sirmans, C. F., "Traffic Externalities and Single-Family Home Prices," *Journal of Regional Science*, 32(4), 487-500 (1992).
11. Carey, J., and Semmens, J., "Impact of Highways on Property Values: Case Study of Superstition Freeway Corridor," Transportation Research Record: *Journal of the Transportation Research Board,* 1839, 128–135 (2003).
12. Kilpatrick, J. A., Throupe, R. L., Carruthers, J. I., and Krause, A., "The Impact of Transit Corridors on Residential Property Values," *Journal of Real Estate Research,* 29(3), 303–320 (2007).
13. Chernobai, E., Reibel, M., and Carney, M., "Nonlinear Spatial and Temporal Effects of Highway Construction on House Prices," *Journal of Real Estate Finance and Economics*, 42(3), 348-370 (2011).
14. Allen, M. T., Austin, G. W., and Swaleheen, M., "Measuring Highway Impacts on House Prices Using Spatial Regression," *Journal of Sustainable Real Estate,* 7(1), 83–98 (2015).
15. National Ocean Service, National Oceanic and Atmospheric Administration (2018, June 25). *What is LIDAR?* Retrieved from *https://oceanservice.noaa.gov/facts/lidar.html*.
16. Anselin, L., "Lagrange Multiplier test diagnostics for spatial dependence and spatial heterogeneity," *Geographical Analysis*, 20:1-17 (1988a); Anselin, L., *Spatial Econometrics: Methods and Models,* Kluwer Academic Publishes, Dordrecht (1988b).