The Influence of Leaking Underground Storage Tanks on Nearby House Values

Hans R. Isakson and Mark D. Ecker*

ABSTRACT. Some households are willing to pay a premium to live farther away from a disamenity, such as a neighborhood gas station with a leaking underground storage tank (LUST). In this study, submarkets are constructed to allow for this premium to be a function of both the distance to the nearest LUST and the intensity of multiple LUSTs. We find evidence that households within ¼ mile of multiple LUSTs do not have a statistically significant aversion to living near them. However, households more than a ¼ mile from a LUST are willing to pay 9.29% more for a house located 10% farther away. (Q51, Q53, R21)

I. Introduction

Households are often willing to pay a premium to live farther away from a disamenity, such as a gas station, airport, landfill, sewage treatment plant, or high voltage lines. This premium varies not only by the type of disamenity, but also across multiple housing submarkets in an urban area. Factors such as the proximity to a disamenity, intensity of multiple, clustered disamenities, public awareness of the nature of the disamenity, and household aversion to living near a disamenity would reasonably affect this premium. This study segments housing sales data into submarkets to better understand the premium associated with living near a particular disamenity: a leaking underground storage tank (LUST) frequently found under gasoline stations.

Rosen's (1974) hedonic model has become the standard model used to estimate the premium associated with environmental features (see Nelson, 2004, Simons and Saginor, 2006, and Boyle and Kiel, 2001). The Rosen model is popular because the premium (to live farther away from a disamenity) is derived from the statistically estimated coefficients of the model.

---

*Hans Isakson, Department of Economics, University of Northern Iowa, Cedar Falls, IA 50614-0129. E-mail: hans.isakson@uni.edu
Mark Ecker, Department of Mathematics, University of Northern Iowa, Cedar Falls, IA 50614-0129.

A. SUBMARKET

The possibility of market segmentation can create problems for hedonic models. If market segmentation exists, Freeman (1979), Islam and Asami (2009), and Goodman and Thibodeau (2003) point out that a single hedonic model fitted using aggregated sales data will produce biased estimates of the implicit prices. Straszheim (1975) observes that the coefficients of some independent variables, such as proximity to an amenity/disamenity, can vary substantially between submarkets. If market segmentation exists, Freeman (1979, p. 164) recommends using a separate hedonic price function for each submarket.

Goodman and Thibodeau (1998, 2003) study housing submarkets stratified by census track, zip code, and elementary school districts. Jones, Leishman, and Watkins (2003) construct housing submarkets markets using stratified geographical areas. Tu (2007, p. 387) defines a submarket as a collection of one or more, not necessarily contiguous, neighborhoods "across which [Rosen type] hedonic prices have no significant differences." Tu (2007) creates housing submarkets by combining multiple neighborhoods defined in terms of similarities in structural features of a house (living area, lot size, number of bathrooms, etc.) as well as similarities in household preferences or tastes (age and number of children, ethnicity, religious preference, etc.). Tu (2007) combines the similar (in terms of hedonic prices) non-contiguous neighborhoods into a larger housing submarket. As long as the hedonic prices among the neighborhoods of a submarket have no significant differences, the submarket can be composed of multiple, non-contiguous neighborhoods. For example, Kuethe and Keeney (2012) define non-contiguous housing submarkets stratified by the selling prices of houses.

In general, there may be numerous, non-contiguous subdivisions (neighborhoods) with very similar houses and households (in terms of age, family size, income, etc.) spread around the urban area that naturally form a housing submarket because buyers and sellers would consider houses within the multiple neighborhoods as substitutes for each other. In all of these aforementioned studies, the creation of these housing submarkets significantly improved the hedonic model.

B. DISTANCE AND INTENSITY

The premium households are willing to pay to live farther away from a disamenity is generally modeled as a function of distance between the disamenity and the house. We propose that this premium is a function of not only distance, but also the intensity of the disamenity site. Specifically, we propose that $\pi = f(d, I)$, where π = premium, d = distance to the (nearest) disamenity, and I = intensity of the disamenity.

Modeling the premium as only a function of distance (i.e. ignoring intensity) can produce a washout effect, partially because many more sales will naturally be found farther away from a disamenity site. In addition, when distance is the only disamenity variable and sales are combined into a single model, then the intensity of multiple, very close disamenity sites biases the estimated premium. Therefore, it is appropriate to explore the market for the presence of multiple submarkets as a function of both distance and intensity.

The purpose of this study is to examine the implicit prices in a hedonic model when housing sales data are divided into submarkets based upon how many disamenity sites (LUSTs) are located near each house. Others (discussed in the next Section) have estimated this premium primarily as a function of proximity to a LUST. However, we allow this premium to be a function of both the proximity to and intensity of the LUST.

II. Previous Lust Literature

Previous LUST studies draw from only two geographical areas: (1) Cuyahoga County including Cleveland, Ohio and (2) Fredrick and Baltimore Counties including the city of Baltimore. Simons, Bowen and Sementelli (1997) use linear and Box-Cox forms of a hedonic model of Cuyahoga County housing sales and find that houses within 300 feet of a registered LUST experience a 17% reduction in their selling price. Simons and Sementelli (1997) find that LUST sites, themselves, in Cuyahoga County are half as likely to sell and less likely to obtain mortgage financing. Simons, Bowen, and Sementelli (1999) find a 14% to 15% difference between a hedonic estimate of price and the actual selling price of a house in Cuyahoga County due to the presence of a nearby LUST. They also report a 16% reduction in selling prices of high-priced houses near a LUST in a new subdivision.

Simons and Winson-Geideman (2005) survey 1,115 households in eight states using a contingent valuation questionnaire and report a 25%

to 33% reduction in bid prices for houses contaminated by a nearby LUST. Guignet (2012) surveys 303 households in Maryland using a stated preference questionnaire and reports a 18% to 25% reduction in bid prices for houses contaminated by a LUST.

Zabel and Guignet (2012) examine 136,816 house sales using a difference-in-difference hedonic Simultaneous Auto-Regressive (SAR) model and report no clear overall price effect attributable to nearby LUSTs. However, they do find a 12.4% decline for houses within 1,000 meters of a particular, heavily contaminated gas station. Guignet (2013) studies 244,169 house sales from Maryland using a difference-in-difference hedonic Spatially Autoregressive Combined (SAC) model and finds no effects of a LUST as close as 500 meters. But, Guignet (2013) does report a 11% decline of house prices that received well testing results from the Maryland Department of the Environment reporting LUST related contaminants.

These studies reveal mixed results concerning the influence of LUSTs on the selling prices of nearby houses. Zabel and Guignet (2012) and Guignet (2013) find no clear evidence of a LUST effect in three Maryland market areas. Using the Cuyahoga housing sales data, Simons Bowen and Sementelli (1997 and 1999) report a 17% decline of house prices within 300 feet of a LUST as well as a 16% price reduction in a new, high-priced subdivision. Similarly, Guignet (2013) finds an 11% decline in house price when the household receives a notice from the Maryland Department of the Environment reporting LUST related contamination of their well water.

Only two types of information regarding LUSTs, other than proximity, are included in these studies. In the Cuyahoga County data, LUSTS are classified as leaking/not leaking and registered/not registered. In the Fredrick and Baltimore Counties data, LUSTs are either open cases being investigated by the state or closed cases. So, additional studies using LUST data from states, other than Ohio and Maryland, that provide additional LUST information (other than proximity) will expand our understanding of the impact of LUSTs on the prices of nearby houses.

III. The Study Framework

To check for a non-constant premium associated with LUSTs, we make use of (1) a sufficient number of house sales surrounding disamenity sites, (2) a means of disaggregating housing sales data into submarkets based upon intensity of the disamenity sites, and (3) models that allow for varying coefficients within each submarket.

Dividing housing sales data into submarkets based upon the intensity of the disamenity sites can be challenging, because there are no examples to draw from in the literature. In this study, we use a partitioning approach (Kauko 2002) to segment the data into three submarkets based on the intensity of multiple, clustered LUSTs. The three submarkets are defined as follows:

- GP2+s consist of houses located very near two or more LUST sites;
- GP1s consist of houses located very near exactly one LUST site; and
- GP0s consist of houses that are not very close to any LUST site.

The F-test suggested by Ott and Longnecker (2010) reveals that the coefficients of the hedonic model are statistically different among the three groups.[1] We define very close as ¼ mile and then check the sensitivity of our results by using definitions a bit less (0.2 mile) and a bit more (0.3 mile) than ¼ mile. This approach allows the premium (and all of the coefficients of the other independent variables) to vary between the three submarkets.

IV. The Empirical Model

Similar to Ihlanfeldt and Taylor (2004), the intensity of disamenities (LUSTs) relevant to any particular house is measured using a Count variable: the number of very close LUSTs. A distance of ¼ mile is chosen to define "very close." That is, the Count variable is equal to the number of LUSTs within ¼ mile (about three city blocks) of the house.

After classifying the data into submarkets (GP2+, GP1s and GP0s) based on the number of LUSTs very close to each house, we use a Rosen-type hedonic approach to model the relationship between the price of a house and a set of independent variables. Key aspects of the hedonic model are discussed below, namely, (1) the specific set of variables (dependent and independent) used in the model, (2) the functional relationship among these variables, and (3) the structure of the error term of the model.

A. THE VARIABLES

We use the selling price of a house as the dependent variable in our empirical model. We use independent variables that reflect the (1) structural characteristics of the property including date of sale, (2) proximity of the property to points of influence (other than the LUST sites), (3) characteristics of the neighborhood in which the property is located, and (4) LUST specific data.

The structural characteristic data include: living area, lot size, number of rooms, and year built. Our model developed in the next section also includes a spatial correlation term that helps to mitigate the impact of omitted variables as well as any incorrectly specified variables (Brasington and Hite, 2005; Kim, Phipps, and Anselin, 2003; Cohen and Coughlin, 2008). In addition, the date of sale is included as a continuous variable to control for the effects of temporal price trends. The local housing market experienced gradually rising prices during our study period from 2000 to 2004.

We also include variables that measure the distance to various potential points of influence (other than the 50 LUST sites), including the sewage treatment plant (STP), the Cedar River, the nearest public school, and the nearest highway. The STP is located just south of a small downtown retail area that occasionally emits offensive odors which dissipate with increasing distance. The Cedar River is a major feature of the community, and may influence the selling prices of houses located near it. Proximity to a highway can save commuters travel expenses and time, but living too close to a highway may be undesirable due traffic and noise. Proximity to a public school might also affect house prices. We include all of these distance variables to directly capture spatial effects that may influence house prices. Other spatial effects are captured by the error term of our model.

Characteristics of the neighborhood in which a house is located (census block groups) can also affect a house's selling price. We include median rent, median year that houses were built, percent of housing units that are owner occupied, the number of housing units (a proxy for housing density), and median household income (a proxy for neighborhood quality) of the census block group in which the house is located as additional independent variables. We expect that the selling prices of houses will be positively related to the median rent, median year that houses were built in the neighborhood, the percent of housing units that are owner occupied, and the median household income of the neighborhood. We expect selling prices of houses to be negatively related to housing density (number of housing units in the neighborhood).

B. LUST VARIABLES

We also include three LUST specific variables. The Euclidian distance (*d*) from each house to the nearest LUST, the Count of LUSTs within a ¼ mile of each house, and the Condition (*good* or *poor*) of the nearest LUST. The Count variable is included to capture the intensity of multiple LUSTs.

The average distance to the nearest LUST (for all 2078 sales) is 0.48 miles with 27.6% of the sales (574/2078) having at least one LUST within ¼ mile. The hedonic regression coefficient of this variable is a primary emphases of this study. A positive and statistically significant coefficient for this distance variable is consistent with the LUST being a disamenity. We are also interested in variations in this coefficient across the three intensity-submarkets: GP2+s, GP1s and GP0s. We expect that this distance coefficient will be small and perhaps statistically insignificant in the GP2+ submarket, because these households would probably not mind living near an additional LUST (since they are already living near at least two). For the GP0 submarket, the distance coefficient should be positive and statistically significant, because we expect these households would prefer to live farther away from a LUST site. We expect that the distance coefficient for the GP1 submarket to be between the coefficients of the other two submarkets.

The number of LUSTs very close to a house is relevant only in the GP2+ submarket, since the Count of very close LUSTs is defined to be one for the GP1s submarket and zero in the GP0 submarket. However, in the GP2+ submarket, we anticipate that house prices will not be affected by the number of very close LUSTs, because these households are already living near (at least) two.

We also include the Iowa Department of Natural Resources (IDNR) condition of each LUST in our analysis to capture a potential publicity/information effect.[2] Six of the LUST sites are classified by IDNR as *high risk* (receptors exist within the transport plum of the contaminants, but the IDNR has not reported contamination of the city's water supply), four are classified as *low risk* (no receptors within the transport plum of the contaminants), and the remaining 40 are classified as *no further action* required (contaminants do not exceed delimited levels). We create a binary variable (called Condition) to capture the classification of the LUSTs as follows: Condition equals one if the LUST is classified as *low risk* or *high risk*; zero if it is classified as *no further action* required. We combine the *low* and *high risk* LUSTs because there are so few (four) *low risk* LUST sites. The condition of LUSTs is publically available at the IDNR's web site, but finding it takes considerable time and effort. Therefore, households may not know the condition of a LUST under any particular gas station. In addition, house sellers are not legally required to disclose the condition of the nearest LUST. If households are sensitive to the condition of a LUST (accessible at the IDNR website), then we expect the coefficient of the Condition variable to be negative and statistically significant. In contrast, a statistically insignificant coefficient for the Condition variable is consistent with (1) households not undertaking the search to find the condition information or (2) households know the condition, but it does not impact their decision to live near the LUST.

C. THE EMPIRICAL MODEL

The empirical model contains explanatory variables typically thought to explain house prices. Most broadly, these variables include (1) site level non-spatial characteristics of the home (house size characteristics including living area, number of rooms, parcel size, year built, and time of sale), (2) spatial characteristics (distances to the STP, a river, and the nearest highway), (3) neighborhood characteristics (median rent, median year that houses were built, percent of housing units that are owner occupied, the number of housing units, and median household income), and (4) LUST variables (distance to the nearest LUST; classification of the nearest LUST, and number of LUST sites within ¼ mile). Specifically, let

P = the selling price of the house,
S = lot size in acres,
H = the living area of the house,
d = distance to the nearest LUST, and
C = condition of the nearest LUST,
CT = number of LUSTs within ¼ mile of the house.
t = the date of the sale (t = 0 on January 1, 2000),

R = number of rooms in the house,
D = a vector of variables measuring distance to potential points of influence, and
N = a vector of neighborhood level variables.

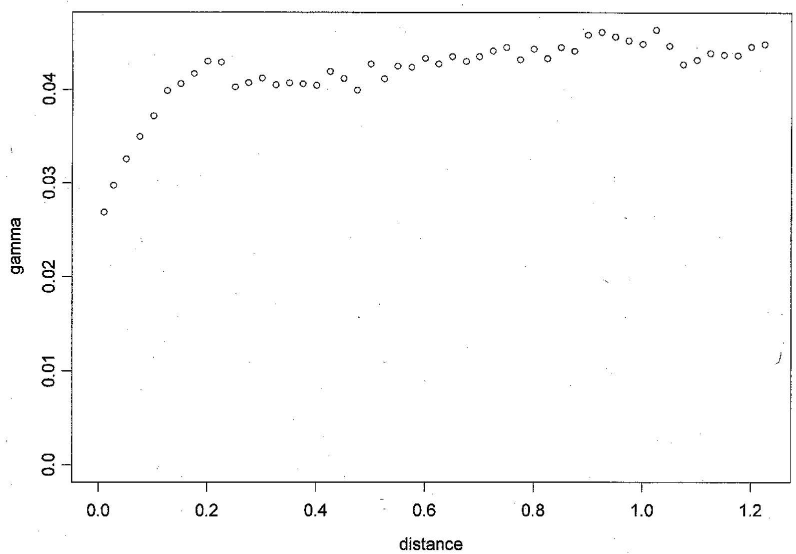
On the natural log scale, the hedonic model is:[3]

$$\ln P = \ln a + \beta_1 \ln S + \beta_2 \ln H + \beta_3 \ln d + \delta C + \lambda CT + \kappa t + \omega R \quad \theta D. \quad (1)$$

The lot size (S) and house size (H) parameters, β_1 and β_2, represents the (constant) house price elasticity of lot size and house size, respectively (see Isakson and Ecker, 2001 and Ecker and Isakson, 2005 for discussion). The key parameter, β_3 , represents the house price elasticity of distance to the nearest LUST. The parameters associated with the remaining explanatory variables represent the percentage change in price given a one unit change in the independent variable.

To estimate the parameters of (1), we include an error or disturbance term, ϵ. Statistical inference involving Ordinary Least Squares (OLS) regressions requires that $\varepsilon \sim N(0,\tau^2)$, where τ^2 is a variability parameter after accounting for site-level covariates. When dealing with spatial data, such as housing sales, the potential exists for the OLS parameter estimates to be biased, especially any spatial distance parameter estimates in the mean structure (such as distance to the STP and distance to the nearest LUST site). Visual examination of an empirical variogram (Cressie, 1993) of the OLS residuals in Figure 1 reveals the presence of spatial correlation. Therefore, efforts to model this source of bias will be beneficial.

Figure 1. Empirical Variogram of the OLS Residuals



Spatial linear models assume that the errors are not independent, that is, two comparable homes that are closer in space sell for a more similar price than two comparable homes farther apart. For example, houses located near each other are also near the same neighborhood amenities/disamenities, and often are similar in terms of their physical features (and price). In this case, the selling prices of nearby, comparable houses tend to be more highly correlated than comparable houses farther apart. We build spatial correlation into the model by assuming that

$$\varepsilon \sim N(0, \tau^2 + \sigma^2) \qquad (2)$$

where τ^2 is called the "nugget", i.e. a micro-scale or measurement error variability, in the geostatistical literature (Cressie, 1993). The sum $\tau^2 + \sigma^2$ in equation (2) is termed the spatial variability of the spatial process or "sill" (the variability of the home prices after adjusting for individual house characteristics). Finally, for two house sales with errors ε_i and ε_j, their spatial correlation is modeled as a function of their Euclidean distance apart, d_{ij}. Specifically, we adopt the spherical correlation structure, i.e.,

$$Corr(\ln(\in_i), \ln(\in_j)) = \begin{cases} \frac{1}{2}(\phi^5 d_{ij}^5 - 3\phi d_{ij} + 2) & \text{if } d_{ij} \le \frac{1}{\phi} \\ 0 & \text{if } d_{ij} > \frac{1}{\phi} \end{cases} \quad (3)$$

The parameter ϕ directly controls the spatial correlation in the dataset and is termed the "range" (technically, $\frac{1}{\phi}$ is the exact value of the range in equation (3) - see Ecker (2003) and Tu (2007) for the spherical correlation structure). Thus, any two houses that are separated by a distance of more than the range have selling prices that are essentially uncorrelated.

We model the housing submarkets by combining the traditional hedonic model, (1), together with a spatial correlation term, given by (2) and (3). The spatial correlation terms (2) and (3) are random effects models designed to capture extra spatial variability not explained in the mean structure of the model.[4] We estimate the parameters of the model separately for each submarket (GP2+, GP1 and GP0).

V. The Study Area and Data

This study uses data from Cedar Falls, Iowa, a medium sized, Midwestern college town. The college (University of Northern Iowa) has about 12,000 students, over 95% of which live on or very near campus. The city has two junior high schools and several elementary schools of nearly equivalent quality that feed students into the city's only public high school. All houses in the city are in the same school district.[5]

The city is highly decentralized with an average commute time of only 14 minutes. There is no dominant destination point, such as a central business district. Instead, jobs and retail shops are disbursed throughout the city. Preliminary analysis of the data reveals that households are not willing to pay a premium to live closer to destination points (including the college) in the city, largely due, we suspect, to the short commute times. Several small retail strip malls exist as well as a downtown area containing several restaurants, boutique retail shops, banks, a small hotel, and a small theater. The downtown area is also within a mile of the city's sewage treatment plant (STP), a source of occasional noxious odors depending on weather conditions. The city has a newer (1995) industrial park located on its far south side containing retail businesses and warehouses.

The city of Cedar Falls provides utilities to its residents, including electricity, gas, refuse collection, cable TV, internet, and water. All houses are connected to the city's water supply; no houses have private water wells. The city's potable water comes from eight ground-water wells drawing from the Silurian-Devonian limestone aquifer. The wells range in depth from 147 to 275 feet. The water from this source is of high enough quality that little treatment is needed; only chlorine and fluoride are added at the well sites. Contamination of potable water supplies by the LUSTs in this study has not been reported by the IDNR, most likely due to the extreme depth of the aquifer source of city water.

A. HOUSING SALES AND CENSUS DATA

Initially, the housing sales dataset contained every sale of a single-family house in the city of Cedar Falls from January 2000 to November 2004.[6] This time period was selected to avoid the effects of the recent housing spike and subsequent crash of 2006-07. These sales were parsed by using only those identified as "arm's length transactions" by the county tax assessor's office. Additional refinement consists of choosing only those sales with a selling price greater than $32,000 or less than $400,000, houses with at least three but less than 12 rooms, at least 500 square feet of living area, and a lot size less than 3 acres.[7] Several houses were repeatedly sold (but not enough to perform a repeat sales analysis) in this time frame; we use only the most recent sale. Table 1 contains the summary statistics of the housing sales data.

TABLE 1–Summary Statistics of Sales Data

	ALL SALES N = 2,078		GP2+ N=365		GP1 N=209		GP0 N=1504	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Sale Price	$144,995	71,415	$108,350	39,838	$97,717	27,606	$160,415	75,398
Structural Variables								
Living Area	1332.8	489.5	1172.1	412.5	1073.5	309.7	1407.8	505.4
Year Built	1960.6	28.5	1941.2	26.1	1938.9	23.6	1968.3	25.7
Number of Rooms	5.9	1.4	5.7	1.3	5.5	1.2	5.9	1.5
Lot Size (acres)	0.27	0.16	0.23	0.14	0.21	0.10	0.28	0.16
Date of Sale (0=1/1/2000)	2.56	1.39	2.55	1.40	2.54	1.46	2.57	1.37
Distances (miles)								
Distance to STP	3.43	1.17	3.07	0.78	3.12	0.78	3.56	1.27
Distance to River	1.20	0.72	0.82	0.37	0.85	0.36	1.35	0.77
Distance to School	0.44	0.27	0.39	0.16	0.33	0.16	0.47	0.30
Distance to Highway	0.73	0.46	0.63	0.53	0.68	0.44	0.76	0.43
LUST Variables								
Distance to LUST (miles)	0.48	0.31	0.15	0.05	0.20	0.04	0.60	0.28
Count	0.64	1.29	3.05	1.40	1	NA	0	NA
Condition	0.11	0.31	0.15	0.36	0.19	0.39	0.09	0.28
Census Block Variables								
Median House-hold Income	40,817	12,015	37,878	10,197	38,927	8,661	41,793	12,665
Number Occupied Units	695.3	423.3	482.0	319.2	492.2	267.2	775.3	435.7
Percent Owner Occupied	59.3	22.4	55.8	22.0	62.5	26.0	59.7	21.8
Median year Build	1962.3	12.8	1953.8	11.5	1954.0	8.4	1965.6	12.1
Median Rent	474.3	189.9	472.7	173.3	455.6	154.8	477.2	197.9
Difference Median Year	-1.7	20.9	-12.6	21.8	-15.1	21.5	2.7	18.7

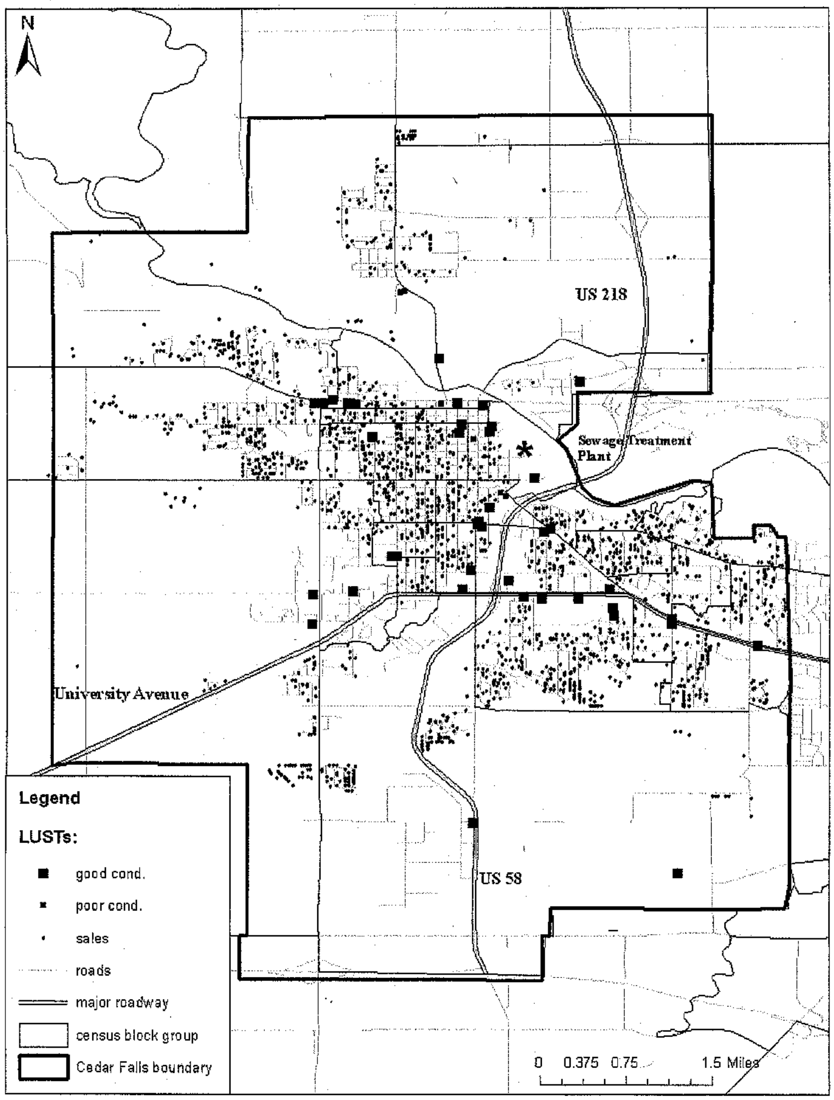
B. LUST DATA

This study focuses on 50 LUST sites (mostly neighborhood gasoline stations) that have been identified by the IDNR as containing a LUST. We note that none of them are located near heavy industrial sites. In fact, there are no heavy industrial sites anywhere in the city (see Deaton and Hoehn (2004)). Information for the 50 LUST sites was obtained from the Iowa Department of Natural Resources website (http://www.iowadnr.gov/land/LUST/LUSTdbindex.html). Figure 2 shows the location of all 2,078 sales together with the STP, the 50 LUST sites, the Cedar River, major highways, and Census block group boundaries.

Registration of LUSTs is required by federal law (40 CFR parts 280 and 281). In Iowa, all LUSTs, except certain small, residential LUSTs found in rural areas, must be registered with the Iowa DNR (Iowa Code 455.B.474). The condition of these tanks is monitored by the IDNR. All underground storage tanks must be inspected for their likelihood of contamination of the surrounding land.[8] If the site containing the LUST has evidence of contaminants that do not exceed certain delimited levels, then it is classified as requiring *no further action*. If the onsite contaminants exceed the delimited levels, but no receptors (aquifer, plastic water lines, basement, sewers, or surface water) are within the transport plum of the contaminants, then it is rated as a *low risk* and must be tested annually. If receptors exist within the transport plum of the contaminants, then it is classified as *high risk* and faces more intensive monitoring, testing, and possible remediation. The IDNR data includes the state-plane coordinates (centroid) of each LUST site as well as the LUST's condition. We define the Condition of each LUST as either good (*no risk*) or poor (*low or high risk*), as indicated in Figure 2 using different sized boxes.

VI. Model Fitting and Results

To fit the model given by equations (2) and (3) simultaneously with the mean structure coefficients in equation (1), we use SAS Proc Mixed (a maximum likelihood technique), and the results are reported in Table 2. Some of the coefficients for the independent variables in Table 2 are not consistent (in sign, magnitude, or significance) across all three submarkets. This lack of consistency statistically supports, along with the formal F-test (see footnote 1), the segmentation of the housing sales data into the three distinct submarkets.



Note: Plot is in State Plane Coordinates

Figure 2. Cedar Falls sales (o) and LUST (■) locations

A. STRUCTURE AND NEIGHBORHOOD VARIABLES

The coefficients of the size variables of Living Area, Number of Rooms, and Parcel Size are all positive and strongly significant for all three submarkets. As expected, bigger houses sell at higher prices. Year built (minus the Median year built for the census block group in which the house presides) has a positive and significant coefficient in all three submarkets; newer houses sell for higher prices, compared to the average age of the houses in the census block group. Date of Sale has positive and significant coefficients in the three submarkets, indicating an approximate 5% per annum increase in house prices.

The distance to the STP coefficient is positive and strongly significant in the GP0 submarket. Due to the smaller sample sizes associated with GP2+ and GP1 sales, the distance to the STP coefficient is actually larger in both the GP2+ and GP1 submarkets, but only moderately significant (p-value in GP2+ = 0.0438 and p-value in GP1 = 0.0885). In all three submarkets, no one wants to live near the sewage treatment plant. Distance to the nearest school coefficient is significant and negative in the GP1 submarket only, indicating that these households are willing to pay a premium to live near a school. The distance to the nearest highway coefficient is positive and marginally significant in the GP2+ submarket (p-value=0.0848) and GP0 submarket (p-value=0.0631). Neither group desires to live near a highway. The distance to the river coefficient is negative and statistically significant only in the GP0 submarket; these households are willing to pay a premium to live close to the Cedar River.

For the census group block variables, the median household income coefficient is positive and statistically significant in the GP1 and GP0 submarkets, but small in magnitude. House prices and neighborhood quality (median household income) are, not surprisingly, highly correlated. The housing density (number of occupied units) coefficient is only statistically significant in the GP2+ submarket, while the median rent coefficient is not significant for any submarket. The percent owner occupied coefficient is negative but only statistically significant in the GP0 submarket, suggesting that these households prefer low housing density neighborhoods. Median Year Build is significant only in the GP0 submarket, indicating that these households are willing to pay more to live in a neighborhood with newer houses.

TABLE 2–Results

	GP2+ n=365	GP1 n=209	GP0 n=1504
Distance to Nearest LUST	0.0194 (0.5477)	0.0966* (0.1178)	0.0929**** (0.0068)
Count	-0.0020 (0.8794)	N/A	N/A
Condition	0.0415 (0.3049)	0.0397 (0.4795)	-0.0462 (0.3362)
Distance to STP	0.1149*** (0.0438)	0.1102** (0.0885)	0.0814**** (0.0001)
Distance to School	0.0013 (0.9888)	-0.4095**** (0.0011)	-0.0320 (0.5171)
Distance to Highway	0.1126** (0.0848)	0.1009 (0.3022)	0.0738** (0.0631)
Distance to River	0.0209 (0.7768)	0.0243 (0.7894)	-0.0796**** (0.0082)
Ln Living Area	0.4575**** (0.0001)	0.3274**** (0.0001)	0.3030**** (0.0001)
Date of Sale	0.0526**** (0.0001)	0.0544**** (0.0001)	0.0490**** (0.0001)
Number of Rooms	0.0576**** (0.0001)	0.0782**** (0.0001)	0.0648**** (0.0001)
Ln Parcel Size	0.0933**** (0.0001)	0.2266**** (0.0001)	0.1200**** (0.0001)
Yr Blt - Median Yr Blt	0.0045**** (0.0001)	0.0037**** (0.0001)	0.0059**** (0.0001)
Median HH Income	0.000002 0.2809	0.000007*** 0.0224	0.000008**** (0.0001)
Num Occ Units	0.0001*** (0.0366)	0.00002 (0.8752)	0.00003 (0.5209)
Pct owner Occ	-0.0609 (0.5994)	-0.0415 (0.7733)	-0.2284*** (0.0113)
Median Yr Blt	0.0022 (0.3266)	-0.0024 (0.4788)	0.0069**** (0.0002)
Median Rent	-0.00007 (0.4883)	-0.00007 (0.6065)	-0.00002 (0.8088)
Nugget	0.0119	0.0126	0.0095
Sill	0.0355	0.0213	0.0454
Range	0.0166	0.0333	0.0677

Note: Numbers in parentheses are p-values. Significance is denoted by * for $p< 0.15$, ** for $p < 0.1$, *** for $p < 0.05$ and **** for $p < 0.01$

B. LUST VARIABLES

Overall, the parameter estimates for the LUST variables are consistent with our expectations. The non-significant distance to the nearest LUST coefficient in the GP2+ submarkets suggests that these households do not have an aversion to living close to multiple LUST sites. The positive and weakly significant coefficient for proximity to the nearest LUST in the GP1 submarket (p-value = 0.1178, with n=209) indicates that these households may be willing to pay a higher premium (to be farther from the LUST) than households in the GP2+ submarket. The positive and strongly significant distance to the nearest LUST coefficient in the GP0 submarket indicates that these households are willing to pay a statistically significant premium to live farther away from the nearest LUST. In particular, the coefficient for distance to the nearest LUST in the GP0 submarket (.0929) is smaller than the same coefficient in the GP1 submarket (.0966), but it is about five times as large as that in the GP2+ submarket (.0194). The elasticity of house price with respect to proximity to a LUST in the GP0 submarket is statistically significant and indicates that these households are willing to pay a premium to live farther away from a LUST. In particular, households in the GP0 submarket are willing to pay 9.29% more for a house that is 10% farther away from the nearest LUST.

The Count variable is only relevant in the GP2+ submarket, because Count is defined as one in the GP1 submarket and zero in the GP0 submarket. Households in the GP2+ submarket do not seem to mind if they live near two or more LUSTs, as indicated by the lack of statistical significance for the proximity-to-LUST coefficient. The Condition of a LUST does not seem to matter in any of the three submarkets. It would appear that the publicly available (at the IDNR website) information does not influence household decisions regarding living near a LUST. Nothing about a LUST influences households in the GP2+ submarket, while all that matters for households in the GP1 and GP0 submarkets is their proximity to the nearest LUST.

The spatial association parameters estimates of range, sill and nugget for the three submarkets are also reported at the bottom of Table 2. For all three submarkets, the nugget and sill estimates are fairly close, while range parameter is 166 feet (0.0166 times ten thousand feet) in the GP2+ submarket, 333 feet in the GP1 submarket and 677 feet in the GP0 submarket. Therefore, the prices of two houses in the GP0 submarket will be spatially uncorrelated if separated by more than 677 feet.

C. SENSITIVITY ANALYSIS

Lastly, we use sensitivity analysis to explore the choice of ¼ mile as the definition of being "very close" to a LUST. The Count variable is the number of LUSTs that are very close (1/4 mile) to a house, which, subsequently divides the sales data into three submarkets: GP2+, GP1 and GP0. We redefined "very close" (the Count variable) using two other distances: 0.2 and 0.3 miles. These two other distances, subsequently, create different sets of houses in each of the three submarkets. Then, we fit the model for each new submarket. Table 3 reports the results for the LUST variables.

Inspection of the results in the two redefined and original GP0 submarkets in Table 3 shows that the coefficient for distance to the nearest LUST increases as the distance that defines being "very close" increases. This result is consistent with a stronger LUST aversion in the GP0 submarket. Table 3 also shows that no LUST variable, including Count itself, is significant in the GP2+ submarkets, regardless of the choice of definition (0.2, 0.25, or 0.3 miles) of "very close". The distance coefficient in the GP1 submarkets is positive and weakly significant at ¼ mile; it is positive but not significant at the other two distances.

TABLE 3–Sensitivity analysis to the choice of the distance that defines the Count variable

Distance = 0.2 mile

	GP2+	GP1	GP0
Sample Size	202	189	1687
Distance to Nearest LUST	-0.0167 (0.7531)	0.0846 (0.1632)	0.0878**** (0.0005)
Count	-0.0332 (0.2267)	NA	NA
Condition	-0.0354 (0.7903)	0.1521**** (0.0030)	-0.0564* (0.1403)

Distance = 0.25 mile [from Table 2]

	GP2+	GP1	GP0
Sample Size	365	209	1504
Distance to Nearest LUST	0.0194 (0.5477)	0.0966* (0.1178)	0.0929**** (0.0068)
Count	-0.0020 (0.8794)	N/A	N/A
Condition	0.0415 (0.3049)	0.0397 (0.4795)	-0.0462 (0.3362)

Distance = 0.3 mile

	GP2+	GP1	GP0
Sample Size	550	187	1341
Distance to Nearest LUST	0.0357 (0.2762)	0.0350 (0.5329)	0.1088**** (0.0095)
Count	0.0009 (0.9336)	N/A	N/A
Condition	0.0400 (0.2790)	-0.1120 (0.2226)	-0.1404*** (0.0348)

Note: Numbers in parentheses are p-values. Significance is denoted by * for $p < 0.15$, ** for $p < 0.1$, *** for $p < 0.05$ and **** for $p < 0.01$

Lastly, Condition is only significant in the GP1 submarket at a distance of 0.2 and in the GP0 submarket at a distance of 0.3, indicating that households in these submarkets might potentially be aware of the condition of the nearest LUST. However, this awareness effect is not consistent in terms of statistical significance and changing signs across the submarkets and distance definitions.

VII. Summary and Conclusions

This study proposes and fits an empirical model that allows for a variable premium associated with LUSTs, primarily found under gas stations. This variable premium is modeled as a function of both the intensity of multiple LUSTs and the distance to the nearest LUST. Similar to Zabel and Guignet (2012) and Guignet (2013), we find evidence that households within ¼ of a mile of multiple LUSTs do not have a statistically significant aversion to living near these LUSTs. Neither the number of LUSTs nearby (Count), nor the Condition of the nearest LUST has an influence on selling prices. Households living near (within ¼ mile of) exactly one LUST have a weakly significant (willingness-to-pay) premium to live even farther away from the LUST. Lastly, households living more than ¼ mile from a LUST are willing to pay a statistically significant premium to live even farther away from it. The magnitude of this willingness to pay premium (9.29% more for a house 10% farther away) is consistent with the findings of Simons, Bowen, and Sementelli (1997 and 1999).

It is not surprising that some households (GP0s) have a strong aversion to living near a LUST. It should also not be surprising that some households (GP2+s) have little to no aversion to living near multiple LUSTS. In addition, a third group of households (GP1s) would naturally exist between these two extremes. Creating submarkets that to distinguish between these three types of households allows us to understand better the influence that LUSTs have upon the prices of nearby houses and assist policy makers in estimating the external benefits of LUST remediation.

References

Boyle, M.A. and Kiel, K.A. (2001). "A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities." *Journal of Real Estate Research*, 9(2), 117-134.

Brasington, D.M. and D. Hite (2005). "Demand for Environmental Quality: a Spatial

Hedonic Analysis." *Regional Science and Urban Economics*, 35(1), 57-82.
Cohen, J.P. and C.C. Coughlin (2008). "Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices." *Journal of Regional Science*, 48 (5), 859-878.
Cressie, N (1993). *Statistics for Spatial Data*. New York: John Wiley and Sons.
Deaton, B.J. and J.P. Hoen (2004). "Hedonic Analysis of Hazardous Waste Sites in the Presence of Other Urban Disameties." *Environmental Science and Policy*, 7, 499-408.
Ecker, M.D. (2003). "Geostatistics: Past, Present and Future." In *Encyclopedia of Life Support Systems* (EOLSS) . Developed under the Auspices of the UNESCO, EOLSS Publishers, Oxford, UK. [available at: www.eolss.net].
Ecker, M.D. and Isakson, H. (2005). "A Unified Convex-Concave Model of Urban Land Values." *Regional Science and Urban Economics*, 35, 265-277.
Freeman, A. M., III (1979). "Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues." *Scandinavian Journal of Economics*. 81, 154-173.
Goodman, A.C., Thibodeau, T.G. (1998). "Housing Market Segmentation." *Journal of Housing Economics*, 7, 121-143.
Goodman, A.C., Thibodeau, T.G. (2003). Housing Market Segmentation and Hedonic Prediction Accuracy." *Journal of Housing Economics*, 12, 181-201.
Guignet, D. (2012). "The Impacts of Pollution and Exposure Pathways on Home Values: A Stated Preference Analysis." *Ecological Economics*, 82, 53-63.
Guignet, D. (2013). "What do Property Values Really Tell Us? a Hedonic Study of Underground Storage Tanks." *Land Economics* 89:2, 211-226.
Ihlandfeldt, K.R. and L.O. Taylor (2004). "Externality Effects of Small-Scale Hazardous Waste Sites: Evidence from Urban Commercial Property Markets." *Journal of Environmental Economics and Management*, 47, 117-139.
Isakson, H. and Ecker, M.D. (2001). "An Analysis of the Influence of Location in the Market for Undeveloped Urban Fringe Land." *Land Economics*, 77(1): 30-41.
Islam, K. S. and Asami, Y. (2009). "Housing Market Segnation: A Review." *Review of Urban & Regional Development Studies*, 21: 93–109.
Jones, C., Leishman, C., and Watkins, C. (2003). "Structural Change in a Local Urban Housing Market." *Environmental and Planing*, 35, 1315-1326.
Kauko, T. (2002). "Modeling the Locational Determinates of House Prices: Neural Network and Value Tree Approaches." Utecht: Faculeti Ruimtelijke Wetenschappen, Universiteit Utrecht.
Kim, C.W., T.T. Phipps, and L. Anselin (2003). "Measuring the Benefits of Air Quality Improvement: a Spatial Hedonic Approach." *Journal of Environmental Economics and Management,* 45, 24-39.
Kuethe, T.H. and Keeney, R. (2012). "Environmental Externalities and Residential Property Values: Externalized Costs along the House Price Distribution." *Land Economics*, 88(2), 241-250.
Nelson, J.P. (2004). " Meta-analysis of Airport Noise and Hedonic Property Values: Problems and Prospects." *Journal of Transport Economics and Policy*, v (1), 1-27.
Ott, R.L. and M. Longnecker, (2010). *An Introduction to Statistical Methods and Data Analysis*, 6th edition. Brooks/Cole.
Rosburg, A., H. Isakson, M. Ecker, and T. Strauss (2017). "Analysis of the Impact of a Public School Closure on House Prices." *Journal of Housing Research*, 23:2,

Forthcoming.
Rosen, S. (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*. 82, 34-55.
Simons, R.A., Bowen, W.M, and Sementelli, A.J. (1997). "The Effect of Underground Storage Tanks on Residential Property Value in Cuyahoga County, Ohio." *Journal of Real Estate Research*, 14 (1/2): 29-42.
Simons, R.A., Bowen, W.M. and Sementelli, A.J. (1999). "The Price and Liquidity Effects of LUST Leaks from Gas Stations on Adjacent Contaminated Property." *Appraisal Journal*, April, 186-194.
Simons, R.A. and Saginor, J.D. (2006). "A Meta – Analysis of the Environmental Contamination and Positive Amenities on Residential Real Estate Values. *Journal of Real Estate Research*, 28 (1/2): 71-104
Simons, R.A. and Sementelli, A.J. (1997). "Liquidity Loss and Delayed Transactions with Leaking Underground Storage Tanks. *Appraisal Journal*, July, 255-260.
Simons, R. and K. Winson-Geideman (2005). "Contamination of Residential Property Buyers Using Contingent Valuation Surveys." *Journal of Real Estate Research*, 27:2, 193-220.
Straszheim, M.R. (1975). "An Econometric Analysis of the Urban Housing Market." National Bureau of Economic Research, New York.
Tu, Y. (2007). "Segmentation, Adjustment and Disequilibrium." Chapter 3 in O'Sullivan, T and Gibb, K (eds.) 2003. *Housing Economics and Public Policy*, Blackwell Science, Oxford, UK.
Zabel, J. and D. Guignet (2012). "A Hedonic Analysis of the Impact of Disamenity Sites on House Prices." *Resource and Energy Economics* (34), 549-564.

Endnotes

1. An Analysis of Covariance (ANCOVA), see Chapter 16 of Ott and Longnecker (2010), indicates that a separate analysis based upon intensity and distance is statistically warranted (F=2.54, p=.00000281).
2. Previous studies have not included the condition of a LUST in their models.
3. Equations 1 and 2 do not yet contain an error term. The structure of the error term is discussed later in this section.
4. Ihlanfeldt and Taylor (2004) use a SAR for the error term in their model.
5. See Rosburg, et. al., 2017, for more detailed discussion regarding the city's public school system.
6. The authors thank the Black Hawk County Board of Supervisors for providing the housing sales data used in this study. Of course, any opinions expressed in this study are strictly those of the authors.
7. Because the original data set contained all transactions of houses, it included a few very small structures (as small as 100 square feet) that sold for very low prices (as low as $7,500) as well as a few very large houses with large amounts of land that sold for as much as $2.5 million. These outliers were excluded because they were not representative of the typical house in Cedar Falls, Iowa.
8. The IDNR uses the Risk-Based Corrective Action process for classifying the riskiness of a LUST.