DISCUSSION PAPERS IN ECONOMICS

Working Paper No. 09-15

Valuing public goods with changing implicit prices

Megan Lawson

University of Colorado

December 2009

Department of Economics

University of Colorado at Boulder
Boulder, Colorado 80309

Valuing public goods with changing implicit prices

Megan Lawson
University of Colorado-Boulder

November 21, 2009

Abstract

Traditional panel data methods applied to estimate hedonic models assume that the implicit prices of housing attributes remain constant over time. In this paper I demonstrate that this assumption may not hold true when there are large changes in the supply of the public good over time, and failure to account for changing implicit prices can lead to biased estimates of the value of changes in public goods. I use air quality in southern California, measured as exposure to toxic air emissions, to demonstrate the effect of violating the assumption of constant implicit prices on estimates of the implicit price for air quality improvements.

1 Introduction

If toxic air pollution is a disamenity, hedonic theory tells us that the prices of houses with more exposure to pollution should be lower than houses with less pollution. However, empirical evidence has not always supported this hypothesis. One possible explanation for this inconsistent evidence is that implicit prices for air quality improvements may change over time, and models that do not account for these changes will produce biased estimates of the value of changes in pollution levels. In this paper I ask the following research question: in hedonic models, what portion of observed changes in housing prices is attributable to changes in the implicit price of that public good? When they are not identified separately from per-unit values, changes in the implicit prices of an amenity can bias estimates of the effect of changes in amenities on housing prices. In this paper I evaluate the relative magnitude of these two components of price change and demonstrate the effect of their relative sizes on inference about capitalization, using exposure to toxic air pollution.

Panel data and difference-in-difference methods are increasingly popular methods of estimating amenity capitalization in the hedonic price valuation literature (Gautier, Siegmann, & Van Vuuren, 2009; Pope, 2008b, 2008a; Anderson & West, 2006). These methods are popular because they allow researchers to control for unobserved differences between properties, avoiding unobserved variable bias and endogeneity problems associated with cross-section analysis. These methods, however, are only effective if the underlying attributes and implicit prices of these attributes are constant over time, an assumption that has received relatively little attention in the hedonic price literature (see McMillen (2008) and Redfearn (2009) for exceptions to this), and has yet to be applied to research on the capitalization of environmental amenities in housing prices. In this paper I demonstrate that failing to account for changes in both the supply and price of public goods over time can bias estimates of their implicit value.

Consider a house with high air quality at a time when many houses in the region have poor air quality. High air quality is a relatively scarce amenity, so this home would demand a higher price. If air quality throughout the region improves, the effect on this home's price will depend upon the relative changes in the implicit price of air quality and the levels of pollution. In general, when the supply of a good increases, its price decreases. If more houses have high air quality, air quality can be expected to command a smaller price premium per unit. If the price elasticity of supply for air quality is sufficiently large, then the house's price could actually decrease despite the air quality improvement.

Conversely, consider a home that begins with low air quality. When air quality improves over the study area, this home's air quality will also improve. However, because increases in its supply lead to a lower implicit price in the second time period, the overall price increase is smaller than it would have been had the house experienced the improvement in the first time period. Because traditional panel data methods only account for the total price change for a house, if implicit prices decrease over time, then these methods - used in most hedonic studies - will yield biased estimates of the value attributable to improved air quality. I avoid this bias by decomposing the change in house price into changes in the level and the implicit price of air quality. In the Results section, I predict price changes over a range of baseline exposure and changes in exposure, to illustrate the effect of changing implicit prices.

In southern California, the area studied in this analysis, between 1990 and 2000 the average exposure to toxic air pollution was cut in half, thereby increasing the number of homes with relatively high air quality. When more houses have high air quality, air quality may command a smaller price premium, thereby decreasing the implicit price of air quality over time. I find strong evidence that changes in the implicit price of air quality is the primary component of pollution-related changes in house price. This change in the implicit price of a public good could result from changes in its supply, particularly if there are non-marginal changes over the time period.

I use house sales data from five counties in southern California in 1990 and 2000, along with data on toxic air emissions from the Toxic Release Inventory (TRI), to demonstrate the components of changes in housing prices over time. Results from a basic first differences model suggests a per-household capitalization of $520 from changes in pollution in the study area. However, once I control for changes in implicit price, the per-household capitalization increases to $1,553 due to a large change in the implicit price air quality improvements.

In the remaining sections of this chapter I discuss the methods that have been used to address problems of unobserved variable bias in hedonic price models; my empirical strategy; the data used to address the research question; and model results and implications for inferences about total price changes.

2 Literature Review

While hedonic models are a widely-used tool for valuing goods and attributes not traded in markets, researchers have long been concerned about unobserved variables and their effect on estimated implicit prices. When these unobserved variables are correlated with housing prices, parameter estimates - and implicit prices - will be biased. Specifically related to questions of pollution, pollution levels (and subsequent risks from the pollution) are often positively correlated with business cycle variables such as employment and urbanization levels. These variables are also positively correlated with housing prices, which can cause biased parameter estimates. Researchers have used several strategies - including instrumental variables, difference-in-difference, and price decomposition approaches - to manage the problem of unobserved variable bias in hedonic models. Each of these is discussed in turn below.

2.1 Instrumental variables

Several authors have addressed the endogeneity problem using an instrumental variables approach. Gayer (2000) examines whether a community's marginal willingness to pay for risk reduction from Superfund site cleanup, and subsequent welfare gain from that reduction, varies according to the community's sociodemographic characteristics. He uses an instrumental variables approach to control for the fact that environmental risk both determines and is endogenous to housing values. The IV estimation equations are based on a structural equations model, using the probability of collective action as a proxy for risk, assuming that collective action is likely to reduce the likelihood that polluters locate in the neighborhood. He quantifies collective action as the voter turnout rate from the 1988 presidential election, the proportion of votes for the Democratic candidate, and the proportion of homeowners. He finds that the OLS specification is likely to be biased due to endogenous risk levels, and results from the IV approach suggest that household MWTP differs according to community sociodemographic characteristics.

Chay and Greenstone (2005) use an instrumental variables approach to estimate household WTP for changes in air quality, using features of the Clean Air Act to create an instrument for household exposure to pollution. Under the Clean Air Act, a county can be designated "non-attainment" if ambient air pollution levels exceed either an annual mean standard or the second-worst day exceeds a separate acute standard. The authors find that attainment status is strongly correlated with changes in pollution over time, but is not correlated with housing prices and is therefore a strong instrument for air quality. They estimate several different models, comparing OLS to IV estimates and demonstrating that OLS estimates of the effect of pollution on housing prices is biased towards zero.

Strong instruments for endogenous variables can eliminate endogeneity bias in cross-section and panel data models. However, they do not account for changes in implicit prices over time, which remains a concern when using panel data.

2.2 Panel Data methods

Panel data methods - including first differences, difference-in-difference (D-in-D), and fixed effects regressions - have become increasingly popular in hedonic models because of their usefulness in controlling for some types of unobserved variability within study areas. Figlio and Lucas (2004) investigate the effect of school report cards on housing prices. In their study, the omitted variable bias concerns the relationship between "better neighborhoods" and "better schools." To avoid this problem they used repeat sales data and both house-specific fixed effects and neighborhood fixed effects to control for larger neighborhood trends. Their results are identified by variation in an individual house's price as a function of changes in school quality, controlling for unobserved differences in inter-neighborhood quality. Davis (2004) uses similar methods to investigate the effect of pediatric leukemia risk on housing prices in a small Nevada town using repeated-sales data.

Greenstone and Gallagher (2006) use differencing in the context of a regression discontinuity design to estimate the effect of Superfund site cleanup on nearby housing prices. Utilizing a feature of Superfund site designation, which resulted in very similar neighborhoods either being cleaned up or not making the list, they were able to control for unobserved variation between the communities. Using this quasi-experimental approach, they found that site clean-up had no significant effect on housing prices.

2.3 Decomposing Price Changes

While panel data methods are useful tools for mitigating unobserved variable bias, their effectiveness hinges on the assumption that these unobserved variables are constant over time. If this assumption is not met, then observed price changes are a function of changes in both attribute levels and the implicit prices. Several authors have adapted methods developed in the labor economics literature to address this issue and decompose the changes in housing prices.

Along these lines, McMillen (2008) decomposes the changes in the distribution of housing prices. Analyzing Census tract-level price indices that are a function of the average house's physical attributes in the tract, he found that most of the change in housing prices was attributable to a change in the coefficient distribution, not a change in the variable distribution. He only considered a house's physical characteristics, not its level of public goods.

Redfearn (2009) used locally-weighted regression methods on two cross-sections of data to identify differences in implicit prices for light rail access between census tracts. By exploiting a quasi-experiment of the timing of new rail station openings, he was able to compare price changes between different neighborhoods with different light rail access. He strongly rejected the assumption of constant implicit prices for attributes across neighborhoods near light rail stations and found that forcing constant implicit prices rendered coefficient estimates of light rail access highly unstable.

Building on McMillen's and Redfearn's insights regarding the inappropriateness of the assumption of constant implicit prices, I extend their insights to the question of the valuation of intertemporal changes in public goods. The analysis considers explicitly the implications for non-market valuation when the supply of the public good changes substantially over time.

3 Empirical Strategy

Hedonic price theory assumes that a house's price in a given time period is a function of the level of its attributes:

$$y_{i,t} = \beta X_{i,t} + \psi_n + \varepsilon_i, \tag{1}$$

where y is house price, X is a vector of house attributes, β is a vector of coefficients describing the marginal effect of X on price, ψ is a vector of time-invariant unobserved neighborhood-specific characteristics (indexed by n), and ε represents time-invariant unobserved house-specific heterogeneity (indexed by i). Let βX be the sum of different categories of explanatory variables, such that $\beta X = \lambda A + \gamma P$, where A is a vector of house attributes and P is a house's pollution exposure. First differencing exploits multiple observations on a given parcel over time to eliminate time-invariant parcel characteristics such as unobserved heterogeneity between neighborhoods - ψ - and parcels - ε. Assume there are two time periods, 0 and 1.

$$y_{i,1} - y_{i,0} = \beta(X_{i,1} - X_{i,0}) \tag{2}$$

Therefore the effect on y of the change in X between time periods 0 and 1 is equal to β.

However, it is clear that the power of first differencing depends on a constant value of β and ψ between time periods. Suppose this is not the case. Oaxaca (1973) was one of the first to separate the elements of intertemporal change in what is known as the "Oaxaca decomposition", studying differences between male and female wages. Assume that the vector of coefficients, β, differs between the first and second time periods, and is designated by β_0 and β_1, respectively. By adding and subtracting the term $\beta_1 X_{i,0}$, which is the baseline value for X multiplied by the current value for β to Equation 2, it is apparent that the change in house price is comprised of two components (with the i subscript dropped for brevity):

$$y_1 - y_0 = (\beta_1 X_1 - \beta_1 X_0) + (\beta_1 X_0 - \beta_0 X_0), \tag{3}$$

This can be simplified to present changes in y as a function of changes in coefficients and attributes, in what I call the "corrected" first differences model:

$$y_1 - y_0 = \beta_1 (X_1 - X_0) + (\beta_1 - \beta_0) X_0 \tag{4}$$

Using this framework, I identify the change in capitalization attributable to a change in X (β_1) and the change in prices due to a change in implicit prices ($\triangle\beta$).

By separating X into its constituent components, A and P, I arrive at the equation I will estimate:

$$\triangle ln(price) = [\lambda_1(\triangle A) + (\triangle\lambda)A_0] + [\gamma_1(\triangle P) + (\triangle\gamma P_0)] \tag{5}$$

Thus the change in the log of price is a function of the change in attributes, $\triangle A$, and pollution, $\triangle P$, and the change in the implicit prices for A and P: $\triangle\lambda$ and $\triangle\gamma$. In models that use panel data, a house's physical characteristics would remain constant over time (assuming implicit prices for these remain constant as well), so the term $\lambda_1(\triangle A)$ would drop out of Equation 5. As I explain in the next section, in lieu of repeat sales data, I construct price indices at the Census block level, similar to the approaches used by McMillen and Redfearn. Because these price indices represent the average house in the block, they vary over time, and the attribute expression in Equation 5 does not drop out in estimation.

Implicit in Equation 5 is the assumption that the neighborhood's unobserved heterogeneity remains constant over time. As McMillen (2008) demonstrates, this does not always hold and if ψ_n varies over time, first differencing will not control for unobserved heterogeneity. To control for the changes in spatially correlated unobserved heterogeneity, I include tract fixed effects in Equation 5.[1]

Identification of changes in sale prices is based on within-Census tract differences in changes in average Census block covariate values. I expect that this small spatial scale ensures relative homogeneity in unobserved neighborhood characteristics. I also assume that the distribution of types of houses sold is constant between time periods.

4 Data

To approximate repeat sales data I create Census block level price and attribute indices by averaging house attributes and pollution exposure over all the houses in each Census block. There are, on average, 10 individual house observations per block per year, with a median of seven. Creating price and attribute indices facilitates a comparison of the same location - the individual census block - between time periods, holding time-invariant unobserved differences between the blocks constant. In lieu of data on repeated sales of the same house, this approach of creating small-scale price indices is the next best means of controlling for spatially-correlated unobserved variable bias.

[1] It would also be possible to estimate changes in these tract fixed effects between time periods to further control for changes over time. However, due to the high cost in terms of degrees of freedom, I was unable to estimate a model using this approach.

4.1 TRI Emissions

Data on TRI facilities comes from U.S. EPA's database on TRI emissions and includes only air emissions, either directly from stacks (point) or fugitive (nonpoint). When the TRI program first began in 1987, firms were required to report emissions if they fulfilled one or more of the following criteria:

1. Firm falls under SIC codes 20 - 39 or is a federal facility,
2. Firm manufactures or emits at least 75,000 pounds of chemicals on the TRI list, or
3. Firm manufactures or emits at least 10,000 pounds of any one chemical on the TRI list.

The reporting threshold decreased over time, from 75,000 in 1987 to 50,000 in 1988, and has remained at 25,000 pounds from 1989 on. The 1987 mandatory chemical list included 320 chemicals and chemical categories, referred to as the "1988 Core Chemicals" list. This list was expanded in 1991 and 1995. For consistency, this analysis only uses emissions of the mandatory chemicals required in 1988 and counts any emissions exceeding 25,000 pounds, the 1989 reporting standard.

This data includes the latitude and longitude of the facilities,the type and level of emissions, and the 4-digit SIC code. Figure 1 shows where the TRI facilities are located in Southern California. I include proximate facilities located outside the five counties where the housing transactions occurred, assuming that facilities near the study areas would impact housing prices.

Figure 2 shows the distribution of the housing sample, with red indicating those locations with positive exposure to toxic air emissions in 1990 and blue indicating those with no exposure in 1990. While more non-exposed properties are located outside of the densest urban areas, the inset (Figure 2a) shows that in the densest locations (both in terms of popula-

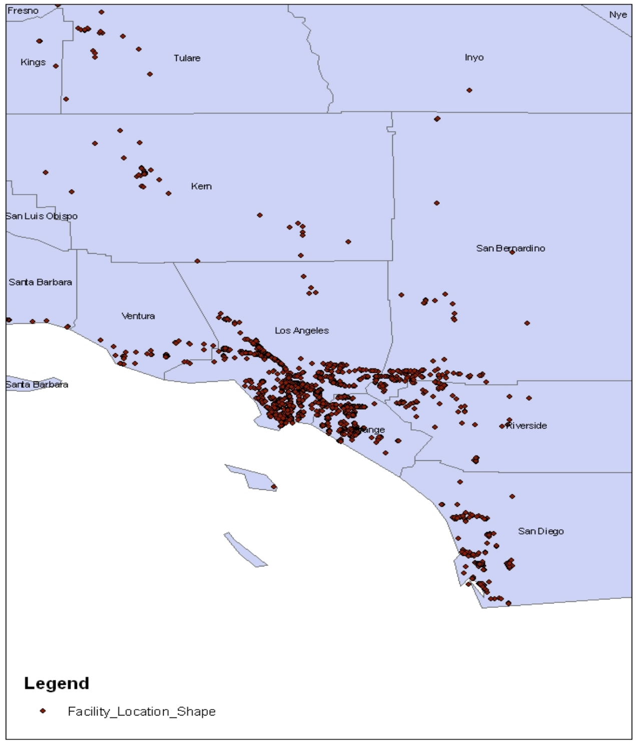


Figure 1: Facilities emitting air pollutants requiring reporting in the Toxic Release Inventory in Southern California, 1990-2000

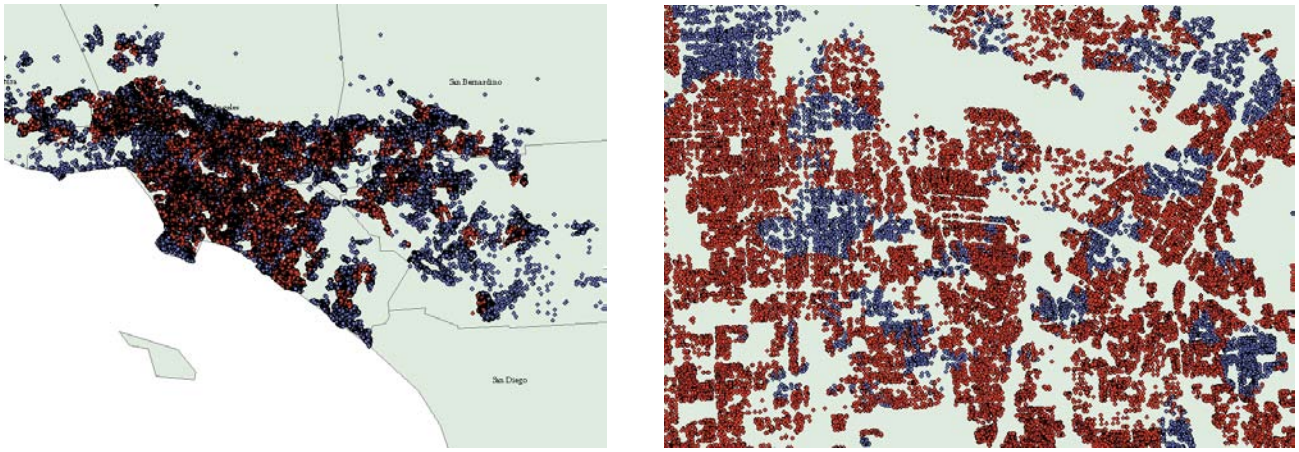
tion and observations in the dataset), exposed and non-exposed properties are distributed relatively evenly.

A census block's exposure to TRI emissions is measured by counting the total number of facilities within 1/2 mile of its centroid.[2]

4.2 Housing Data

Housing transaction data is taken from 5 counties in Southern California - Los Angeles, San Bernardino, Orange, Ventura, and Riverside - for the years 1990 and 2000. A price and attribute index is constructed for each Census block, averaging sale price, lot size, house size,

[2] I also estimated these models using the sum of emissions and the toxicity-weighted emissions, with similar results.



(a) Distribution of all houses

(b) Close-up

Figure 2: Houses with and without TRI facilities within 1/2 mile, 1990

year built, month sold, and the number of facilities releasing TRI-listed pollutants within a half-mile of the block's centroid in 1990 and 2000.

The dataset includes sale price and characteristics about the property, including home and lot size and number of bedrooms, bathrooms, and the year the house was built. Approximately 6% of these observations have zero entered for the number of bedrooms, number of bathrooms, and the year built; these observations were dropped to control for outliers. Observations associated with the upper and lower 1% of sale prices were also removed. Table 1 compares the distribution of pollution, price, and attribute values between 1990 and 2000, as well as the quarter when the house was sold.

Across the distribution of values, sale price is consistently higher in 2000 than in 1990, with 13% higher mean price in 2000. Mean TRI exposure decreased by 58%, but the distribution is more difficult to evaluate because in both years, relatively few blocks have any exposure. The distribution shifts further left in 2000 as only 15% of the blocks were exposed to any TRI emissions while 28% were exposed in 1990. The dramatic change in exposure to

Table 1: Summary statistics for 1990 and 2000

Variable	Mean	Std Dev	p5	p25	p50	p75	p95
Block average values for 1990 sales, n=1,206							
Sale price	$229,095	$105,550	$104,091	$154,055	$216,167	$273,639	$431,114
# of TRI facilities	0.65	1.57	0	0	0	1	3
Lot size	8,128	5,639	3,449	5,722	7,043	8,887	15,865
House size	1,768	454	1,185	1,459	1,695	1,999	2,584
Bedrooms	2.26	0.44	1.58	2	2.25	2.5	2.92
Bathrooms	3.38	0.48	2.5	3.09	3.4	3.71	4
Year built	1975	11	1955	1967	1976	1984	1989
q1	0.19	0.18	0	0	0.18	0.27	0.5
q2	0.26	0.19	0	0.13	0.25	0.4	0.6
q3	0.28	0.19	0	0.17	0.25	0.4	0.6
q4	0.25	0.19	0	0.12	0.25	0.33	0.6
Block average values for 2000 sales, n=1,206							
Sale price	$259,026	$118,457	$113,071	$175,402	$245,216	$312,623	$461,190
# of TRI facilities	0.27	0.91	0	0	0	0	1
Lot size	8,030	4,606	3,535	5,672	7,050	9,057	15,799
House size	1,727	418	1,152	1,438	1,661	1,951	2,484
Bedrooms	2.15	0.49	1.2	1.9	2.2	2.46	2.88
Bathrooms	3.31	0.47	2.5	3	3.33	3.62	4
Year built	1975	11	1956	1968	1976	1984	1992
q1	0.21	0.17	0	0.09	0.2	0.29	0.5
q2	0.27	0.16	0	0.17	0.25	0.38	0.56
q3	0.27	0.17	0	0.17	0.25	0.39	0.6
q4	0.25	0.16	0	0.14	0.25	0.33	0.5

Table 2: Summary statistics for exposed versus non-exposed census blocks

Variable	Mean	Std Dev	p5	p25	p50	p75	p95
Census blocks with no TRI facilities within 1/2 mile, n=1,898							
Sale price	243371	118370	106914	158373	221560	297689	460177
Sale year	1995	5	1990	1990	2000	2000	2000
Lot size	8431	5542	3535	5807	7320	9253	16871
House size	1758	445	1161	1449	1686	1994	2565
Bedrooms	2.21	0.48	1.33	2	2.23	2.5	2.92
Bathrooms	3.33	0.49	2.5	3	3.38	3.67	4
Year built	1976	11	1955	1970	1977	1985	1990
q1	0.2	0.18	0	0.04	0.19	0.29	0.5
q2	0.27	0.18	0	0.17	0.25	0.4	0.6
q3	0.28	0.18	0	0.17	0.25	0.4	0.6
q4	0.25	0.18	0	0.13	0.25	0.33	0.57
Census blocks with at least 1 TRI facility within 1/2 mile, n=514							
Sale price	246608	91439	113985	189771	247446	284856	405959
Sale year	1993	5	1990	1990	1990	2000	2000
Lot size	6782	2972	3230	5362	6270	7519	11929
House size	1708	403	1152	1438	1643	1909	2484
Bedrooms	2.18	0.4	1.5	2	2.19	2.4	2.83
Bathrooms	3.4	0.44	2.67	3.17	3.41	3.67	4.11
Year built	1971	9	1957	1966	1970	1977	1986
q1	0.19	0.16	0	0.04	0.2	0.29	0.5
q2	0.26	0.18	0	0.13	0.25	0.36	0.58
q3	0.29	0.18	0	0.17	0.27	0.4	0.6
q4	0.25	0.17	0	0.14	0.25	0.33	0.56

toxic air emissions supports the hypothesis that changes in the implicit price are attributable to an increase in the supply of clean air across the region.

The distribution of all other house characteristics between the two years are very similar. The consistency in mean housing attribute values suggests that the types of houses selling in each year are relatively consistent and the block price indices are sufficiently small to allow for homogenous housing stock.

Summary statistics for prices, exposure, and house attributes for exposed and non-exposed census blocks are presented in Table 2. Houses in exposed blocks cost, on average,

$3,200 more than houses in non-exposed blocks, although this difference is not statistically significant. This is driven by higher prices in the lower half of the distribution for exposed blocks: the 5th, 25th, and 50th percentile sale prices are greater for exposed blocks, while sale prices are lower for the 75th and 95th percentiles. The range of house prices is more compressed in exposed blocks than non-exposed blocks.

Across much of their distribution, blocks with nearby TRI facilities generally are older, have smaller houses on smaller lots, and have fewer bathrooms. This pattern holds across much of the distribution for each variable, with the exception of the number of bathrooms through the 75th percentile.

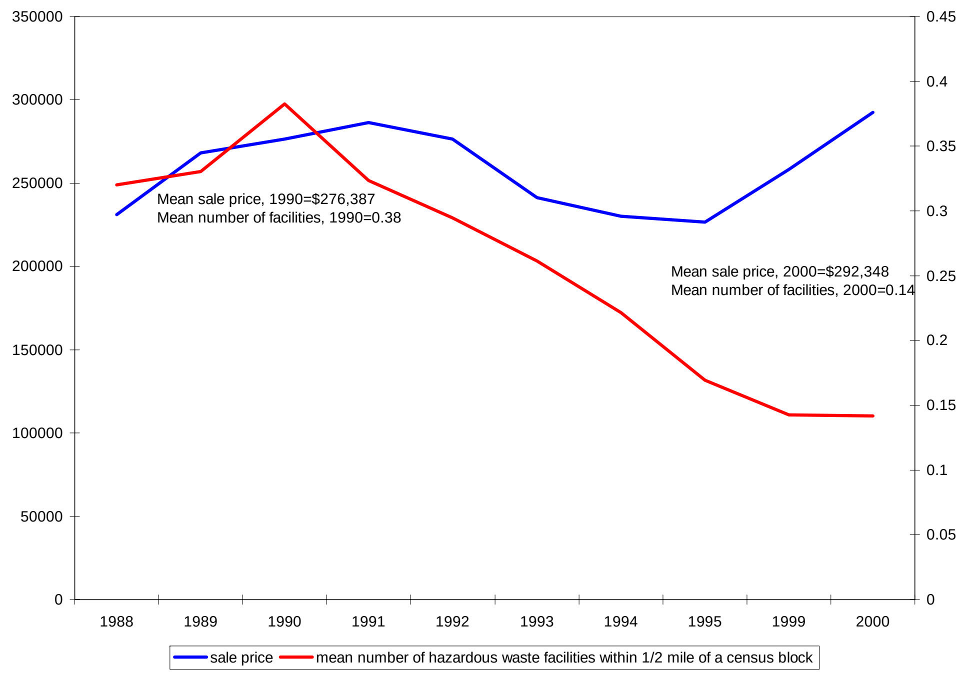


Figure 3: Mean house prices and toxic air pollution exposure in Southern California, 1990-2000

As Figure 3 shows, both mean exposure and sale price changed significantly over the study period. If community composition, household preferences for air quality, or supply of air quality changes over time, I have to be concerned about changes in the implicit price for air quality. As Figure 3 demonstrates, air quality changed dramatically between 1990 and 2000, making a strong case that any observed changes in implicit prices are attributable to changes in supply.

5 Results

Two sets of empirical models are implemented to evaluate the effect of allowing parameter estimates to vary over time in a first differences model. The first set consists of two cross-section models, one for 1990 and the other for 2000, to see the implicit prices for pollution exposure in both years. The second set consists of two first differences models, in which the first assumes constant implicit prices and the second allows them to vary between time periods. In each set I include one specification with Census tract fixed effects (and one without) to control for any spatially correlated, unobserved variables that remain.

5.1 Cross-section models

Table 3 reports results from cross-section models for 1990 and 2000, with a specification with and without Census tract fixed effects in each year. A comparison of the exposure coefficients in models with and without fixed effects highlights the importance of controlling for the unobserved heterogeneity between neighborhoods. However, how well these fixed effects control for unobserved variables remains an open question.

Despite their limitations, cross-section models are useful here to see if the two years yield different parameter estimates. If implicit prices change over time, then coefficients for the 1990 and 2000 cross-section models should differ. My results are consistent with the notion that the per-unit value for air quality has decreased as its supply increased: each additional TRI facility decreases average house prices by 1.4%, or $3,207 for the average-priced home in 1990, but the marginal effect in 2000 is only equal to 0.1% and is statistically insignificant.

Table 3: Regression results: cross-section models for 1990 and 2000

	1990				2000			
	With FEs		Without FEs		With FEs		Without FEs	
Variable	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
# TRI facilities	-0.014**	0.007	-0.001	0.005	0.001	0.006	0.021**	0.008
ln(Lot size)	0.140***	0.024	-0.084***	0.02	0.162***	0.011	-0.109***	0.018
ln(House size)	0.797***	0.07	1.367***	0.069	0.741***	0.036	1.342***	0.059
Bathrooms	-0.001	0.036	0.283***	0.036	0.050***	0.018	0.317***	0.021
Bedrooms	-0.033	0.024	-0.156***	0.024	0.012	0.012	-0.131***	0.022
Year built	0.006***	0.002	-0.014***	0.001	0.001*	0.001	-0.016***	0.001
Q2	0.048	0.047	0.01	0.055	0.035	0.025	0.08	0.055
Q3	0.115**	0.047	0.121**	0.055	0.059**	0.025	0.079	0.053
Q4	0.028	0.047	0.005	0.055	0.071***	0.026	-0.007	0.056
Intercept	-7.639**	3.067	31.075***	1.748	2.884**	1.398	34.030***	1.384
	R^2=0.30		R^2=0.59		R^2=0.33		R^2=0.66	
	n=1,206				n=1,206			

Significance levels : $*$: 10% $**$: 5% $***$: 1%
* Both models included fixed effects at the census tract level.

The coefficients on the number of TRI facilities in 1990 and 2000 are significantly different from one another in both specifications. F-tests of the equality of coefficients equalled 4.74 (p=0.0295) for the model with Census tract fixed effects and 4.50 (p=0.0341) for the model without Census tract fixed effects. Including fixed effects has the largest effect on the coefficients for the number of TRI facilities, lot size, year built, and lot size. This indicates that these variables are correlated with unobservable characteristics at the census block level, which may include attributes such as access to employment centers, school quality, and proximity to industrial areas. By including tract fixed effects, I reduce the area over which the variation in price is estimated, thereby reducing the heterogeneity in unobserved attributes.

Considering only the specifications with fixed effects, the parameter estimates are consistent between the two years for all variables but the number of TRI facilities, the number of bedrooms, and the intercept. Even without controlling for unobserved differences between neighborhoods, which I accomplish in the next section by using first differences at the census block level, I see that the implicit price has changed between the two years, corresponding with a large decrease in pollution levels (see Figure 3)

5.2 First difference models

The cross-section results support my hypothesis of changing implicit prices, but the specification does not allow me to disentangle the effects on house prices attributable to changes in pollution versus changes in implicit prices. To separately identify changes in house price associated with changes in both pollution levels and implicit prices, I need to use the specification presented in Equation 5.

Tables 4 and 5 show results from two sets of first difference models - a standard first difference estimation (assuming constant implicit prices) and a corrected first difference specification that controls for the possibility of varying implicit prices. Table 4 presents models with Census tract fixed effects; Table 5 presents the models without fixed effects. In all models I regress the difference in log of mean sale price on the difference in covariates. House and lot size are included in log form to allow for nonlinear effects on house price as lot or house size increase; all other variables are included in level form. In the first row for each variable, the coefficients represent β_1 from Equation 4, so they capture the implicit price of the covariate in 2000. The second row for each variable shows the change in the implicit price for the covariate from 1990 to 2000, or $\triangle\beta$ from Equation 4. This is not recoverable in the standard first differences approach. Using the Oaxaca decomposition, I estimate the implicit exposure price in 1990 and 2000, with the price for 2000 equal to the coefficient on the change in the number of TRI facilities (γ_1 in Equation 5) and the price for 1990 equal to $(\gamma_1 - (\gamma_1 - \gamma_0))$.

In the first differences model, each additional TRI facility is predicted to reduce house prices by 0.6%, although the coefficient is not statistically different from zero.

The corrected first difference model tells a different story. In this specification, the change in house price is estimated as the sum of two separate components. The first component

Table 4: Regression results: first differences and "corrected" first differences with Census tract fixed effects

	First Differences		Corrected First Differences	
Variable	Coefficient	Std. Err.	Coefficient	Std. Err.
$\triangle$ # of TRI facilities	-0.006	0.01	0.025	0.018
$\triangle$ price for # of TRI facilities			0.025*	0.013
$\triangle$ ln(Lot size)	0.150***	0.025	0.165***	0.028
$\triangle$ price for ln(Lot size)			0.032	0.03
$\triangle$ ln(House size)	0.734***	0.065	0.704***	0.083
$\triangle$ price for ln(House size)			-0.062	0.096
$\triangle$ Bathrooms	0.01	0.029	0.026	0.039
$\triangle$ price for Bathrooms			0.036	0.051
$\triangle$ Bedrooms	0.004	0.022	0.029	0.028
$\triangle$ price for Bedrooms			0.038	0.033
$\triangle$ Year built	0.002	0.001	0	0.002
$\triangle$ price for Year built			-0.005**	0.002
$\triangle$ Q2	-0.002	0.038	-0.016	0.056
$\triangle$ price for Q2			-0.03	0.074
$\triangle$ Q3	0.033	0.037	-0.028	0.054
$\triangle$ price for Q3			-0.12	0.074
$\triangle$ Q4	0.022	0.038	0.023	0.058
$\triangle$ price for Q4			0.004	0.076
Implicit exposure price, 1990			-0.0008	0.010
Implicit exposure price, 2000			0.025	0.018
R^2	0.28		0.30	

Significance levels : $*$: 10% $**$: 5% $***$: 1%
n=1,206

of house price change is the value of the change in pollution exposure, measured using the implicit value of pollution in 2000 (coefficient reported in the first row of Table 4). The second component of the house price change is driven by the change in the implicit price for pollution (coefficient reported in the second row of Table 4). From Equation 4, this is the coefficient associated with baseline attribute levels, $\triangle\beta$. Because the change in the implicit price of pollution is estimated as the marginal effect of baseline pollution on the change in house price, the model predicts that as baseline pollution increases, the change in house price increases. Changes in house prices attributable to air pollution in Census blocks with

Table 5: Regression results: first differences and "corrected" first differences without Census tract fixed effects

	First Differences		Corrected First Differences	
Variable	Coefficient	Std. Err.	Coefficient	Std. Err.
△ # of TRI facilities	-0.020***	0.007	0.017	0.013
△ price for # of TRI facilities			0.021**	0.009
△ ln(Lot size)	0.117***	0.024	0.072***	0.026
△ price for ln(Lot size)			-0.084***	0.02
△ ln(House size)	0.760***	0.061	0.900***	0.07
△ price for ln(House size)			0.262***	0.073
△ Bathrooms	0.057**	0.022	-0.01	0.024
△ price for Bathrooms			-0.093***	0.035
△ Bedrooms	-0.007	0.022	0.027	0.026
△ price for Bedrooms			0.072***	0.025
△ Year built	-0.003**	0.001	-0.003**	0.001
△ price for Year built			-0.004***	0.001
△ Q2	0.024	0.038	-0.045	0.056
△ price for Q2			-0.132*	0.074
△ Q3	0	0.037	-0.045	0.053
△ price for Q3			-0.105	0.073
△ Q4	-0.032	0.038	-0.071	0.056
△ price for Q4			-0.104	0.074
Implicit exposure price, 1990			-0.0034	0.007
Implicit exposure price, 2000			0.017	0.013
R^2	0.30		0.34	

Significance levels : $*$: 10% $**$: 5% $***$: 1%

n=1,206

no TRI facilities in 1990 could only occur if new facilities started emitting in 2000 near that Census block. In other words, the change in implicit price for air pollution only affects the price of houses that were exposed in 1990.

In this specification I find that the effect on house price from the change in number of TRI facilities is not statistically significant (p=0.17), but it is positive. However, it appears that the more important effect related to pollution exposure comes from the change in implicit prices. The results presented here imply that the implicit price for pollution exposure has increased by 2.5% since 1990; because the estimated implicit price in 1990 was negative, this result suggests that pollution exposure decreased prices in 1990, but had little effect in 2000. These findings are consistent with the results from the simple cross-section models presented in Table 3.

All other covariates I included had the expected sign, with blocks with larger average lot size and larger average house size having higher average sale prices. The number of bedrooms nor bathrooms has no effect on sale price. I find no significant difference in the timing of sales throughout the year. These findings are robust regardless of whether fixed effects are included, with a relatively small change in magnitude and the same statistical significance. I find no significant changes in implicit price between years, supporting my conclusion from the summary statistics that the mean attribute values between years has not changed, indicating that the distribution of attributes for houses sold in each year has remained relatively consistent. This supports my construction of price and attribute indices at the Census block level as a means of approximating repeat sales data: the attributes that do not vary over time in repeat sales data are also consistent here.

In Table 5, the pattern in coefficient values between the first differences and corrected first differences is consistent with the results in Table 4, despite its omission of tract fixed effects. The biggest difference between the fixed effect and non-fixed effect model is the statistical significance of the covariates. From these robust estimates I conclude that differencing over time, within Census blocks, appears to eliminate much of the unobserved time-invariant, spatially-correlated heterogeneity across the study area. Relatively little unobserved heterogeneity remains within each Census tract after differencing at the Census block level.

The change in the implicit price for exposure to a facility can be explained by the increased supply of homes with relatively high air quality, and associated smaller premium for these high air quality homes. However, this change in price could also be explained by a change in the amount of actual pollution exposure at each facility. In 1990, reporting facilities emitted an average of 43,828 pounds of chemicals. In 2000, there were far fewer facilities, and the facilities remaining emitted an average of 14,378 pounds. Both possible reasons for the change in implicit price have interesting policy implications, but I leave the disentanglement of the exact reason for the price change for future research.

5.3 Inferring overall price changes

These results suggest that the implicit price for pollution exposure increased between 1990 and 2000 in southern California, thereby decreasing the price premium for air quality. This shift was likely in response to lower TRI emissions throughout the study area which increased the supply of properties with high air quality. In this section I evaluate the conditions under which the change in implicit prices exceeds the change in pollution, causing an overall decrease in prices.

Figure 4 presents the predicted percentage point change in house price over the observed range of baseline pollution levels, relative to a house with no pollution exposure in 1990. In other words, the vertical axis shows how much house prices change for exposed houses versus unexposed houses in 1990. These were calculated using the expression in Equation 5, where $\triangle ln(saleprice) = \gamma_1(\triangle P) + (\triangle \gamma P_0)$. In this section I only consider changes in house price attributable to changes in pollution, keeping house attributes constant.

The red line shows the predicted change in price when 1 additional facility is added to the 1/2 mile buffer surrounding a block; the green line shows the predicted price change when there is one less facility in the block's buffer. The blue line shows the predicted change in price at the mean change in exposure for the sample: 0.38 facilities per buffer. The star is located at the mean baseline TRI exposure (0.65), and shows the average predicted change in house price. Finally, for reference, the orange line shows the predicted price change when a Census block's pollution level does not change over time; the change in house price along this line is attributable entirely to its baseline pollution level.

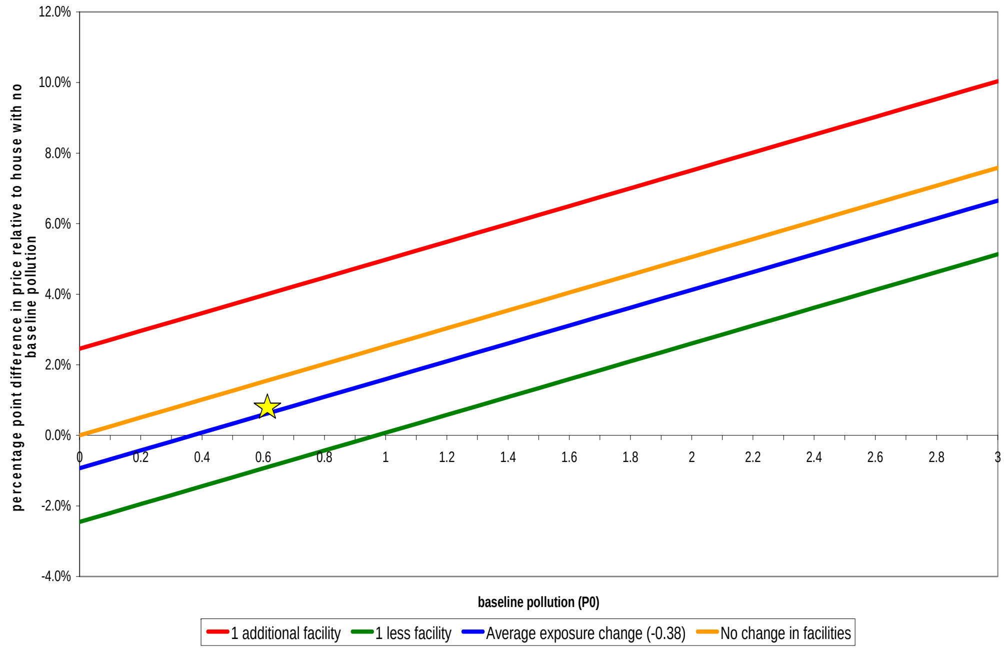


Figure 4: Changes in price for different baseline and changes in exposure, relative to houses with no baseline pollution.

The slope of the lines in Figure 4 is equal to $\triangle\gamma$. Because the implicit price of air pollution increased, the change in house price increases as baseline pollution levels increase. Relative to the houses with the least baseline pollution, houses with the most baseline pollution are predicted to experience greater gains from decreased pollution. The intercept, or the change in house price for houses with no pollution in 1990, is equal to γ_1. Because γ_1 is positive, more facilities shifts the line up, reflecting a larger price increase. Therefore, given two houses with identical baseline exposure, the house that added a nearby facility between 1990 and 2000 would have a greater price increase. A Census block with the average change in TRI exposure (0.38 fewer facilities within its buffer) and baseline exposure of 0.37 facilities or fewer would experience reduced house price.

An improvement in air quality would reduce the change in house prices, possibly resulting in negative changes in house prices for some houses. If a Census block had low initial exposure (or cleaner air quality), then improvements in air quality can reduce the home's price because the value per unit of air quality decreased. As air quality over the whole study area improves, the line is shifted down, and the range of baseline pollution levels over which improved air quality begets a lower house price expands.

In Table 6 I estimate the mean change in the value of houses in the five-county region between 1990 and 2000 that was attributable to the change in TRI exposure for the four model specifications: corrected and uncorrected first differences and the 1990 and 2000 cross-section estimates. I use parameter estimates from the models that included Census tract fixed effects. For each model I calculate the average of the predicted change in house price for all Census blocks and blocks with and without exposure in 1990.

Table 6: Estimated capitalization of the change in TRI exposure from 1990 to 2000.

	Corrected F.D.	Uncorrected F.D.	Cross-section 1990[a]	Cross-section 2000[a]
Mean $\triangle$ value, all blocks	$1,553	$520	$1,304	$83[b]
Mean $\triangle$ value, exposed blocks	$5,015	$1,988	$4,987	-$319[b]
Mean $\triangle$ value, non-exposed blocks	$210	-$49	-$124	$8[b]

[a] Cross-section estimates predicted using mean change (0.38) from 1990 to 2000
[b] Not significantly different from zero.

Because the change in the predicted implicit price of TRI exposure exceeds the predicted implicit price in 1990, the corrected first differences specification predicts a larger per-household capitalization from pollution than the uncorrected first differences specification. In the 1990 cross-section model, reducing pollution is predicted to increase house prices substantially. However, in 2000, reducing pollution has nearly zero effect on house price. The predicted per-household capitalization in the 1990 cross-section specification is lower than the corrected first differences specification likely because it does not take into account the effect of exposure in 1990. As Figure 4 shows, houses with higher baseline pollution tend to see larger increases in sale price.

The predicted capitalization for Census blocks that had at least one facility within 1/2 mile in 1990 is dramatically higher than blocks with no facilities within 1/2 mile. The change in price for non-exposed blocks reflects the fact that these blocks experienced a small (0.04) average increase in nearby facilities over the study period. These results illustrate the potential for large bias when the supply of a public good changes drastically, and this change in supply affects many households in the study area.

6 Conclusions

These results demonstrate that price responses to toxic air pollution in these southern California counties are largely attributable to changes in implicit prices for reduced air pollution, not from changes in the air pollution itself. As this time period coincided with a dramatic decrease in mean exposure to toxic air emissions, this change in implicit price for TRI exposure is potentially due to the large increase in the supply of homes with high air quality, which reduced the premium households were willing to pay for high air quality. The net effect on house price depends on both the baseline TRI exposure and the change from baseline. For the houses with the highest initial air quality, I find that home prices can actually decrease when air quality improves because value per unit of air quality decreases. The observed reduction in TRI exposure is worth $1,553 per household across all Census blocks in the study area.

These findings highlight the particular importance of allowing for varying implicit prices in the case of public goods, as changes in their levels are more likely to affect a much larger portion of the population than changes in private goods, which tend to occur on a smaller scale.

While this evidence supports the hypothesis that the observed change in implicit price is attributable to changes in the supply of the public good, in other applications the price change could arise for other reasons. For example, household preferences may change in response to a public education campaign, or the demographic composition of households may change within a neighborhood.

These findings illustrate the importance of using first differences methods with caution, as the assumption of constant implicit prices between time periods is not trivial. I find that parameter estimates for TRI facilities change dramatically when implicit prices are allowed to change between years, demonstrating that the traditional first differences result is biased in the presence of underlying change in coefficient values. First differences and panel data methods are commonly held as an ideal approach to control for unobserved variable bias, but the results presented here demonstrate that these methods are susceptible to misspecification as well. Researchers should be particularly concerned when there have been non-marginal changes in the supply of the public good or sale prices over time.

References

Anderson, S., & West, S. (2006). Open space, residential property values, and spatial context. *Regional Science and Urban Economic*, *36*, 773-789.

Chay, K., & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, *113*(2), 376-424.

Davis, L. (2004, December). The effect of health risk on housing values: Evidence from a cancer cluster. *American Economic Review*, 1693-1704.

Figlio, D., & Lucas, M. (2004, June). What's in a grade? School report cards and the housing market. *American Economic Review*, 391-604.

Gautier, P., Siegmann, A., & Van Vuuren, A. (2009). Terrorism and attitudes towards minorities: The effects of the Theo van Gogh murder on house prices in Amsterdam. *Journal of Urban Economics*, *65*, 113-126.

Gayer, T. (2000). Neighborhood demographics and the distribution of hazardous waste risks: An instrumental variables estimation. *Journal of Regulatory Economics*, *17*(2), 131-155.

Greenstone, M., & Gallagher, J. (2006). Does hazardous waste matter? Evidence from the housing market and the superfund program. *MIT Department of Economics Working Paper No. 05-27*.

McMillen, D. (2008). Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients? *Journal of Urban Economics*, *64*, 573-589.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693-709.

Pope, J. (2008a). Buyer information and the hedonic: The impact of a seller disclosure on the implicit price for airport noise. *Journal of Urban Economics*, *63*, 498-516.

Pope, J. (2008b). Fear of crime and house prices: Household reactions to sex offender registries. *Journal of Urban Economics*, *64*, 601-614.

Redfearn, C. (2009). How informative are average effects? Hedonic regression and amenity capitalization in complex urban housing markets. *Regional Science and Urban Economics*, *39*, 297-306.