CHAPTER 3

Nonparametric policy analysis: An application to estimating hazardous waste cleanup benefits

James H. Stock

1 Introduction

This chapter investigates the viability of nonparametric techniques in a setting in which conventional parametric procedures often yield ambiguous results. The substantive problem with which I will be concerned is inferring the value of cleaning up a contaminated hazardous waste disposal site from housing value data. The data set contains observations on housing sales price, housing characteristics, and distances to the known hazardous waste sites in the general area. The data are for transactions completed during 1978–81 in eleven Massachusetts towns west and northwest of Boston. The estimation procedure, developed in Stock (1989), uses kernel regression to estimate the change in the price of a specific house in the hypothetical case that some of the attributes of that house (in particular, those related to hazardous waste facilities) are altered as a result of an exogenous policy experiment. The estimator of the mean benefit of the policy is the average of these estimated price changes over all the houses in the data set.

Although the main purpose of this chapter is to examine this nonparametric technique in an actual data set, the application is important in its own right. Cleaning up unsafe hazardous waste sites is potentially one of the most costly environmental problems currently faced in the United States. One natural approach to ascertaining the benefits of cleaning up a contaminated site is to estimate the effect the cleanup would have on

The author thanks David Harrison, Arthur Lewbel, and David Wise for helpful comments on an earlier draft and Rick White for research assistance. This work was supported in part by National Science Foundation grant no. SES-84-0879. This research uses the data set developed by Harrison and Stock (1984), which was supported in part by U.S. Environmental Protection Agency Cooperative Agreement no. 68-809-702-01. The opinions expressed in this paper do not necessarily reflect the views of the Environmental Protection Agency.

housing values, assuming that any predicted increase in price would represent the present value of the benefits of the cleanup to current and future occupants of the home. This type of analysis has, however, been criticized on several grounds. Key among these is the difficulty of specifying a functional form relating price and housing attributes, the "hedonic pricing equation." Because economic theory provides little guidance in specifying the pricing surface, the issue of selecting a regression specification constitutes a major hurdle to hedonic analysis.

A conventional econometric approach to estimating the mean benefits of a proposed policy is to specify a parametric regression model, to estimate its parameters, and to compute the implied average benefits by simulating the effect of the policy on each observation in the data set. If the regression function is linear in the parameters, then this final step reduces to computing a linear combination of coefficients. But, as White (1980, 1981) and others have noted, ordinary least squares can provide a local but not a global approximation to a nonlinear regression (i.e., conditional expectation) function. If the regression function is misspecified, the result is an inconsistent benefits estimator.

The nonparametric estimator considered here uses kernel regression to avoid this potential inconsistency. The dependent variable is assumed to be some variable of policy interest (housing prices, employment status, level of air pollution). By assumption, the proposed policy changes the vector of exogenous variables ($\mathbf{X}$) for the ith observation from $\mathbf{X}_i$ to $\mathbf{X}_i^*$. The effect of the policy on observation i is estimated by the kernel estimator of $E(Y_i|\mathbf{X}_i^*)-E(Y_i|\mathbf{X}_i)$. The nonparametric benefits estimator is the sample average of these estimated differences.

Section 2 of the chapter briefly discusses the use of hedonic analysis to estimate the value of local public goods. The nonparametric estimation procedure is summarized in Section 3. Section 4 describes the data set, and Section 5 presents the empirical results. Conclusions are summarized in Section 6.

2 Hedonic pricing analysis

The objective of hedonic analysis is to use prices of a bundle of goods (a house) to infer the willingness of individuals to pay for the individual items comprising the bundle (air quality, age of the house, living area), for which there are no separate markets. Since its original applications (e.g., in Ridker and Henning 1967, Harrison and Rubinfeld 1978, Polinsky and Rubinfeld 1978, and Polinsky and Shavell 1975, 1976), however, the hedonic pricing approach has been criticized on several points.

To examine these criticisms, it is useful first to set out a simple economic model of housing prices in the context of the hazardous waste problem. Let $\mathbf{z}$ and $\mathbf{w}$ respectively denote vectors of nonwaste and waste-related housing attributes and let x be a scalar composite of all other goods and services. Let x be the numeraire, and let $p(\mathbf{z},\mathbf{w})$ denote the price (in units of x) at which a house with attributes $(\mathbf{z},\mathbf{w})$ can be purchased. The utility $U(x,\mathbf{z},\mathbf{w},S)$ of the household depends on the goods x, $\mathbf{z}$, and $\mathbf{w}$, and on the health status S of the individuals in the household. The health status is assumed to be random but to depend in part on the observable waste-related attributes $\mathbf{w}$; let the distribution of S given $\mathbf{w}$ be $F(S\mid\mathbf{w})$, with density $f(S\mid\mathbf{w})$. In this atemporal model, the household spends all its income I on either x or housing. With these assumptions, the household maximizes its expected utility subject to a budget constraint:

$$\begin{aligned} &\text{Maximize:} \quad E(U\mid\mathbf{w}) = \int U(x,\mathbf{z},\mathbf{w},S)\,f(S\mid\mathbf{w})\,dS \\ &\text{subject to:} \quad I = x + p(\mathbf{z},\mathbf{w}). \end{aligned} \tag{1}$$

Let $U_x = \partial U(x,\mathbf{z},\mathbf{w},S)/\partial x$, $EU_x = \int U_x(x,\mathbf{z},\mathbf{w},S)\,f(S\mid\mathbf{w})\,dS$ (and similarly for $U_\mathbf{z}$, $U_\mathbf{w}$, $\mathbf{p}_\mathbf{z}$, and $\mathbf{p}_\mathbf{w}$), and let $f_\mathbf{w}(S\mid\mathbf{w}) = \partial f(S\mid\mathbf{w})/\partial\mathbf{w}$. Then the first-order conditions corresponding to (1) imply that in equilibrium each homeowner consumes $(x,\mathbf{w},\mathbf{z})$ to satisfy the marginal conditions,

$$\mathbf{p}_\mathbf{z} = \frac{EU_\mathbf{z}}{EU_x} \tag{2a}$$

$$\mathbf{p}_\mathbf{w} = \frac{EU_\mathbf{w} + \int U(x,\mathbf{z},\mathbf{w},S)\,f_\mathbf{w}(S\mid\mathbf{w})\,dS}{EU_x}. \tag{2b}$$

This expected utility framework generates a somewhat different formulation of the hedonic housing price function than the conventional treatment (e.g., Rosen 1974 or Quigley 1982) because of the terms depending upon the subjective probability density $f(S\mid\mathbf{w})$. If S and $\mathbf{w}$ are independent, then (2) reduces to the familiar condition that the marginal price of $\mathbf{z}$ (or $\mathbf{w}$) equals the marginal rate of substitution between $\mathbf{z}$ (or $\mathbf{w}$) and x. More generally, however, the value placed on $\mathbf{w}$ in the hedonic pricing surface will depend on the disutility of becoming ill and on the perceived link between $\mathbf{w}$ and illness.

A similar analysis for producers yields a "supply" side of the market; in equilibrium, both producers' and homeowners' first-order conditions are satisfied. In the housing pricing literature, however, it is conventional to assume that supply (the housing stock) is fixed in the short run, so that movements in prices correspond to movements that satisfy the homeowners' marginal conditions. This simplifies the analysis substantially and is realistic for the housing market. It implies, however, that the long-run benefits of a project can differ substantially from the short-run benefits computed using only (2).

This formulation provides a framework in which to discuss two criticisms of conventional hedonic analysis. The first is that any parametric specification of the hedonic pricing surface is arbitrary. In principle one could solve (2) for the hedonic price surface, although assumptions are required about the distribution of endowments and preferences across the housing market participants and about the characteristics of the outstanding stock of houses. In practice, however, even simple examples result in complicated pricing surface specifications; see, for example, Scotchmer (1985). As a result, current hedonic methodology entails estimating ad hoc flexible functional forms such as the Box–Cox transformation or the quadratic Box–Cox proposed by Halvorsen and Pollakowski (1981).

A second criticism of hedonic analysis arises from its use to evaluate large projects that have nonmarginal effects on individual houses. A recipient of a nonmarginal benefit from a project in general would be unwilling to pay as much for this benefit as the full increase in the house price, in the sense that the compensating variation is less than the increase in the house price. Thus, if transactions costs are small, the individual will be induced to move. If the project has a localized effect, then areawide pricing surface will not change. Thus the individual would be able to move into a house like his prior to the project, thereby fully capitalizing the value of the improvement. If, however, the project has an areawide effect, then the entire equilibrium housing surface will change, and the individual will in general be unable fully to capitalize the improvement. Thus knowledge of the pricing surface alone is insufficient for calculating benefits from widespread nonmarginal improvements: One also needs to know the utility functions of the homeowners. Rosen (1974) suggested estimating utility functions by equating the (parameterized) marginal utility to the derivative of the (parameterized) hedonic pricing surface, as is suggested by (2a). However, as Brown and Rosen (1982) pointed out, this approach is particularly sensitive to the choice of functional form: Identification of the utility function obtains through the parametric assumptions made on the pricing surface.

This second criticism is important, because hedonic analysis is typically used to study widespread policies with nonmarginal effects, such as installing citywide sewer systems or reducing the level of air pollution. In this chapter, however, the effect of the project, while nonmarginal, is localized. Because the data set contains eleven hazardous waste sites, in principle the owners of a home near a newly decontaminated site could capitalize their gains by finding a new home with attributes similar to theirs prior to the cleanup. For further discussion of the validity of hedonic analysis in this context (referred to as the "small open city" model), see Yinger (1982), Polinsky and Shavell (1975, 1976, 1978), Harrison and Rubinfeld (1978), Scotchmer (1985), and Freeman (1979).

3 Estimation procedure

This section briefly summarizes the nonparametric estimator developed in Stock (1989). Let $\mathbf{X}$ denote the vector of independent variables; in the notation of the previous section, $\mathbf{X}=(\mathbf{z}'\ \mathbf{w}')'$. Assume that the hedonic price equation can be written as the sum of an unknown function of the continuous independent variables $\mathbf{X}$, fixed-effects terms corresponding to the community in which the house is located, and an additive error that is independent of $\mathbf{X}$. Let $\mathbf{d}_i$ be a vector of dummy variables indicating the town from which observation i was drawn, let $\mathbf{A}$ be the vector of town effects, and let Y_i denote the house price. The hedonic pricing equation is assumed to have the form,

$$Y_i = g(\mathbf{X}_i) + \mathbf{A}'\mathbf{d}_i + u_i, \quad i=1,\dots,n, \tag{3}$$

where $g(\mathbf{x})$ is assumed to be continuous in $\mathbf{x}$. In addition, assume that $(\mathbf{X}_i, \mathbf{d}_i, Y_i)$ are independently and identically distributed (i.i.d.), with $E(u_i|\mathbf{X}_i,\mathbf{d}_i)=0$ and $E(u_i^2|\mathbf{X}_i,\mathbf{d}_i)=\sigma^2(\mathbf{X}_i,\mathbf{d}_i)$, $0<\sigma_0^2\le\sigma^2(\mathbf{X}_i,\mathbf{d}_i)\le\sigma_1^2<\infty$ for all i.

The mean benefit of the cleanup policy (which shifts $\mathbf{X}$ to $\mathbf{X}^*$) is

$$B = E[g(\mathbf{X}_i^*)+\mathbf{A}'\mathbf{d}_i] - E[g(\mathbf{X}_i)+\mathbf{A}'\mathbf{d}_i] = Eg(\mathbf{X}_i^*) - Eg(\mathbf{X}_i), \tag{4}$$

where the expectation in (4) is taken over $(\mathbf{X}_i, \mathbf{d}_i)$. The second equality obtains by assuming that the policy does not alter the observational cell, an assumption that rules out cleanup policies that entail moving houses from one town to another.[1]

The fixed-effects coefficients $\mathbf{A}$ are estimated by a procedure analogous to one by which the OLS estimator of $\mathbf{A}$ could be obtained were $g(\mathbf{x})$ known to be linear in $\mathbf{x}$. In the special case of a linear regression function, the OLS estimator $\hat{\mathbf{A}}$ can be written as the coefficient obtained from regressing the residuals of the regression of Y_i on $\mathbf{X}_i$ against the residuals of the regression of $\mathbf{d}_i$ on $\mathbf{X}_i$. That is, $\hat{\mathbf{A}}=(\mathbf{D}'\mathbf{M_X}\mathbf{D})^{-1}\mathbf{D}'\mathbf{M_X}\mathbf{Y}$, where $\mathbf{D}$, $\mathbf{Y}$, and $\mathbf{M_X}\equiv\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ respectively refer to $\mathbf{d}_i$, Y_i, and the orthogonal projection matrix of $\mathbf{X}_i$ in the usual matrix notation. A

[1] Although houses have occasionally been moved in cases of severe contamination, relocation is the exception rather than the rule. Even when the cleanup does entail relocation, these moves are but a small part of a larger program.

similar idea is used in the nonparametric model (3). Let η_i be the residuals from the kernel regression of Y_i on $\mathbf{X}_i$ and let ξ_i be the residuals from the kernel regression of $\mathbf{d}_i$ on $\mathbf{X}_i$.[2] Then $\mathbf{A}$ is estimated by

$$\mathbf{A}_n=\left(\sum_{i=1}^{n}\xi_i\xi_i'\right)^{-1}\sum_{i=1}^{n}\xi_i\eta_i. \tag{5}$$

The nonparametric estimator B_n of B is the nonparametric analog of (4) – that is, the sample average of the n kernel regression estimates of $E[Y_i|\mathbf{X}_i^*,\mathbf{d}_i]-E[Y_i|\mathbf{X}_i,\mathbf{d}_i]$. Let $g_n(\mathbf{x})$ denote the nonparametric estimator of $g(\mathbf{x})$. Then

$$B_n=n^{-1}\sum_{j=1}^{n}[g_n(\mathbf{X}_j^*)-g_n(\mathbf{X}_j)]. \tag{6}$$

An alternative expression for B_n, derived in the appendix to this chapter, is as a weighted average of the Y's after subtracting off the town effects, where the weights are estimates of the ratio of the density of $\mathbf{X}$ prior to the policy experiment to the density of $\mathbf{X}$ after the experiment:

$$B_n=n^{-1}\sum_{j=1}^{n}\gamma_n(\mathbf{X}_j)(Y_j-\mathbf{d}_j'\mathbf{A}_n), \tag{7}$$

where

$$\gamma_n(\mathbf{x})=\sum_{i=1}^{n}\left[w\left(\frac{\mathbf{x}-\mathbf{X}_i^*}{b_n}\right)\Big/\sum_{j=1}^{n}w\left(\frac{\mathbf{X}_j-\mathbf{X}_i^*}{b_n}\right)\right]$$
$$-\sum_{i=1}^{n}\left[w\left(\frac{\mathbf{x}-\mathbf{X}_i}{b_n}\right)\Big/\sum_{j=1}^{n}w\left(\frac{\mathbf{X}_j-\mathbf{X}_i}{b_n}\right)\right].$$

Under technical conditions given in Stock (1989), $\mathbf{A}_n$ and B_n are consistent and $n^{1/2}(B_n-E[B_n|\{\mathbf{X}_i,\mathbf{d}_i\}])$ has an asymptotic $N(0,\mathbf{V})$ distribution. When $\sigma^2(\mathbf{X}_i,\mathbf{d}_i)=\sigma^2$ for all i,

$$V=\left[\int\gamma(\mathbf{x})^2h(\mathbf{x})\,d\mathbf{x}+\mathbf{R}'\mathbf{M}^{-1}\mathbf{R}\right]\sigma^2, \tag{8}$$

where

$$\gamma(\mathbf{x})=h^*(\mathbf{x})/h(\mathbf{x})-1,$$
$$\mathbf{R}=E[\gamma(\mathbf{X}_i)\mathbf{d}_i],$$
$$\mathbf{M}=E[(\mathbf{d}_i-E(\mathbf{d}_i|\mathbf{X}_i))(\mathbf{d}_i-E(\mathbf{d}_i|\mathbf{X}_i))'],$$

where $h(\mathbf{x})$ is the density of $\mathbf{X}_i$ and $h^*(\mathbf{x})$ is the density of $\mathbf{X}_i^*$. The variance $\mathbf{V}$ can be consistently estimated by $\mathbf{V}_n$:

$$\mathbf{V}_n=n^{-1}\sum_j(\gamma_n(\mathbf{X}_j)-\mathbf{R}_n'\mathbf{M}_n^{-1}\pi_{nj})^2\hat{u}_j^2, \tag{9}$$

[2] For an introduction to kernel regression, see Prakasa Rao (1983).

where

$$\mathbf{R}_n = n^{-1} \sum_{i=1}^{n} \gamma_n(\mathbf{X}_i)\mathbf{d}_i,$$

$$\mathbf{M}_n = n^{-1} \sum_{j=1}^{n} \xi_j \xi_j',$$

$$\pi_{nj} = \xi_j - \sum_{i=1}^{n} \left[w\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{b_n}\right)\xi_i \Big/ \sum_{j=1}^{n} w\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{b_n}\right)\right],$$

and $\{\hat{u}_j\}$ are the residuals from the kernel regression,

$$\hat{u}_j = Y_j - g_n(\mathbf{X}_j) - \mathbf{d}_j'\mathbf{A}_n.$$

Note that the estimator $\mathbf{V}_n$ is robust to heteroskedasticity in $\{u_i\}$.[3]

The nonparametric estimator requires that the support of $\mathbf{X}_i^*$ be contained in the support of $\mathbf{X}_i$, and so is useful only if the policy does not entail extrapolation outside the range of the data. Parametric techniques can, of course, be used to extrapolate, but in general there is no compelling reason to believe that the selected functional form for the conditional expectation is valid outside the observed range of the data.

4 The hazardous waste data

The data set consists of observations on sales price, house characteristics, and waste-related variables for 324 single-family homes in eleven western and northwestern Boston suburbs.[4] Based on the 1980 U.S. Census, these observations constitute 0.67 percent of the 48,303 detached single-family, owner-occupied houses in these towns. The sales occurred between April 1978 and March 1981. The price was deflated to 1980 dollars using the annual National Association of Realtors Index of Housing Prices (existing one-family homes) for the northeastern United States, interpolated to a monthly basis using the Consumer Price Index. The housing characteristics

[3] Robinson (1988) independently obtained a $n^{1/2}$-consistent results for the parametric part of this model, $\mathbf{A}_n$, for the more general case that $\{\mathbf{d}_i\}$ consists of some discrete and some continuous variables. In computing B_n, $\mathbf{A}$ is here treated as a nuisance parameter.

[4] The suburbs are Acton, Ashland, Bedford, Bellingham, Concord, Framingham, Hollis-ton, Maynard, Natick, Sudbury, and Wayland. These data are a subset of the 2,182 observations from 80 towns used by Harrison and Stock (1984), who describe its collection and the transformations in more detail. The selling price and housing characteristics data were obtained from the Massachusetts Society of Real Estate Appraisers. The reduced number of observations was chosen to control the computational cost of the NPA estimator, which grows as the square of the number of observations. Computing B_n once with 324 observations requires 21 minutes on a 8-MHz IBM PC-AT; with 2,182 observations, it would require 16 hours.

Table 1. *Hazardous waste sites in the Boston SMA identified before 1982*

Site name	Town	Size (acres)	Date of discovery
W. R. Grace Co.	Acton	400	Dec. 1978
Nyanza, Inc.	Ashland	30	1967
BSAF Industries	Bedford	5	May 1978
Benzenoid Organics	Bellingham	4	Oct. 1980
W. R. Grace Co.	Cambridge	10	Mar. 1979
Indian Line Farm	Canton	25	Dec. 1980
Marty's GMC	Kingston	1	Apr. 1980
Salem Acres, Inc.	Salem	180	Sep. 1980
Agrico	Weymouth	10	May 1980
Industriplex 128	Woburn	300	June 1979
Wells G and H	Woburn	200 (plume)	Sep. 1979

Source: Massachusetts Department of Environmental Quality Engineering (1981a, 1981b) as reported in Harrison and Stock (1984).

in the data set are the size of the lot in square feet (LOTSZ), the living area in the house (LIV), a measure of the neighborhood status (the census-tract average of the proportion of the population with blue-collar jobs and the proportion of adults with at most a high school education, based on the 1980 census, STAT), the weighted average distance from the house to the center of local towns, weighted by the town population (ACCESS), and the age of the house (AGE).

Information on disposal sites was obtained primarily from two lists published in 1981 by the Massachusetts Department of Environmental Quality Engineering (1981a, 1981b) describing 367 waste sites in the state. These sites were classified into three categories: (1) sites containing waste classified as hazardous under the regulations promulgated by the U.S. Environmental Protection Agency pursuant to the Resource Conservation and Recovery Act; (2) industrial disposal sites believed to contain no hazardous wastes; and (3) sanitary landfills. Most of the hazardous waste sites are on-site lagoons used to store process wastes. Eleven hazardous waste sites had been discovered by the end of the sample period in the greater Boston area; these are listed in Table 1. Most of these sites received considerable media coverage during and after their discovery (particularly the Acton, Ashland, and Woburn sites), so there is reason to believe that nearby sellers and local real estate agents were aware of the sites. Whether buyers were aware of the sites presumably depended on their local knowledge and on their dealings with their realtors.

The data set contains the distance from each house to each of the eleven sites.[5] To reduce the dimension of the waste-related variables, two indices of disamenities and health risks were constructed from the distance and area data. These indices reflect the several ways that hazardous waste can affect well-being. First, contaminated soil can come into contact with the skin and hazardous chemicals can be ingested. Second, volatile hazardous chemicals can be inhaled. Third, the presence of a nearby hazardous waste site can impose additional psychological costs associated with risk and uncertainty. It was assumed that these disamenities and risks depend on the size of the site and, inversely, on the square of the distance to the site. Thus two proxy variables, RISK1 and RISK2, were constructed, which for the ith house are

$$\text{RISK1}_i = \sum_{j=1}^{11} I_{ij}(d_{ij})^{-2}, \tag{10a}$$

$$\text{RISK2}_i = \sum_{j=1}^{11} I_{ij}A_j(d_{ij})^{-2}, \tag{10b}$$

where A_j is the area of site j, d_{ij} is the distance from house i to site j, and I_{ij} is an indicator that equals 1 if site j had been discovered as being hazardous by the time of sale. A fourth source of risk – contaminated drinking water – can be another important factor in determining house prices. However, in this sample, residents of the same town (and, occasionally, of several towns) typically obtain their drinking water from a common source. Thus this source of risk cannot be distinguished from the town effects.

The simulated policy experiment is the elimination of the risk associated with a given site, say site k. Thus the waste-related variables $\mathbf{w}_i^*$ after cleaning up the site are

$$\text{RISK1}_i^* = \sum_{\substack{j=1 \\ j\neq k}}^{11} I_{ij}(d_{ij})^{-2}, \tag{11a}$$

$$\text{RISK2}_i^* = \sum_{\substack{j=1 \\ j\neq k}}^{11} I_{ij}A_j(d_{ij})^{-2}. \tag{11b}$$

For discussion of this and alternative policy experiments, see Harrison and Stock (1984).

[5] The distance from each house to each of the sites was computed using the latitude and longitude of each of the houses in the data base. For houses near Boston, the latitude and longitude was determined using a computer-readable map of the Boston SMSA developed by the Bureau of the Census (the GBF/DIME file). For houses not on this map, latitude and longitude were manually ascertained from street maps.

Table 2. *The data: Summary statistics*

Variable	Units	Mean	Std. dev.	Minimum	Maximum
Price	1980 dollars	56,475	19,721	21,771	1,554,800
LOTSZ	sq. feet	25,454	35,186	1,277	3,550,200
LIV	sq. feet	1,446	520	617	4,501
STAT	index	1.040	.019	1.000	1.116
ACCESS	index	10.9	1.95	7.89	17.91
AGE	years	23.7	18.9	0.0	80.0
RISK1	index	.234	.472	.006	3.673
RISK2	index	41.6	139.4	.553	1,462.2
Simulated cleanup of Acton site:					
RISK1*	index	.140	.349	.006	3.395
RISK2*	index	3.83	8.57	.120	102.2
Simulated cleanup of Ashland site:					
RISK1*	index	.156	.408	.003	3.668
RISK2*	index	39.2	139.7	.447	1,462.1

Note: See the text for definitions of the variables.

In summary, the hedonic pricing surface relates price to the seven continuous variables (five nonwaste attributes and two waste-related attributes) and the fixed town effects. Summary statistics on these variables are presented in Table 2. Two separate policy experiments are analyzed: cleaning up the Acton site and cleaning up the Ashland site. Summary statistics for RISK1* and RISK2* under these experiments are presented in the final four rows of Table 2.

Because an objective of this chapter is to investigate the sensitivity of kernel and bandwidth decisions in an actual data set, numerical considerations made it necessary to restrict the number of observations and therefore regressors. These seven **X** variables represent a subset of the variables available in the original data set, and it is possible to imagine at least three channels through which omitted variable bias might affect the results: omitted measures of industrialization, omitted variables that capture shifts in demand over time, and omitted additional measures of housing characteristics. The selection of these particular seven variables was guided by earlier parametric results based on the full data set, reported in Harrison and Stock (1984). Specifically, measures of distance to industrial centers are correlated with the RISK variables, so that excluding industrial variables matters when town effects are also excluded. Once town effects are included, however, dropping industrial measures did not affect the coefficients on the RISK variables. Similarly, variables that reflect changing demand over time (local employment, quarter-specific dummy variables) often entered significantly but did not change substantially the waste-related coefficients or cleanup benefits. Finally, additional detail on each house (type of heating, fraction of the basement that is furnished, number of fireplaces, a subjective index of construction quality, pupil–teacher ratio by school district, etc.) helped to predict price, but there was no indication that these measures altered the coefficients on the RISK variables after controlling for town effects and the five nonwaste attributes in Table 2. Although these findings might be an artifact of using parametric techniques, I take them as indicating that omitting these three classes of variables from the nonparametric analysis is unlikely to introduce important biases.

5 Empirical results

For purposes of comparison, a conventional parametric analysis is presented prior to the nonparametric analysis.

OLS results

Table 3 presents OLS estimates of the parameters in a specification that includes linear and squared terms of the two hazardous-waste variables, linear terms of the five nonwaste housing attributes, and fixed town effects. The three columns report the results for log–log, level–log, and level–level specifications. The coefficients on the nonwaste characteristics have large *t*-statistics and generally have the expected sign (recall that a large value of STATUS corresponds to a blue-collar, low-education neighborhood). In contrast, the effects of the waste-related variables are imprecisely measured. The final two rows of Table 3 present estimates of the average benefits of cleaning up two of the sites in the sample. The imprecise measurement of the cleanup benefits reflects the imprecise measurement of the effects of the waste-related variables.

Summary statistics for these regressions, augmented by quadratic terms, are presented in Table 4. The *F*-statistics in Panel A indicate that, while the terms interacting the waste-related variables with the nonwaste variables are not statistically significant, in each of the three columns all the other blocks of coefficients are significant at the 1 percent level using the usual *F*-statistics. For each of the columns, then, this suggests adopting the specification (II) in the table, which has a total of 36 estimated coefficients (including the coefficients on the town dummy variables). As indicated in Panel B, the estimated cleanup benefits vary across specification by a factor of 2 for regression (II). These benefits are estimated quite

Table 3. *Preliminary OLS regression results*

Regressor	(1) Dep. vble.: ln(price) Regressors: logs	(2) Dep. vble.: price Regressors: logs	(3) Dep. vble.: price Regressors: levels
RISK1	−.199	−8,637	−6,669
	(−1.96)	(−1.23)	(−.97)
RISK2	$.508 \times 10^{-3}$	19.2	15.7
	(1.46)	(.80)	(.66)
RISK1×RISK2	$-.269 \times 10^{-2}$	−81.5	−116
	(−1.46)	(−.642)	(−.92)
$(\text{RISK1})^2$	.120	4,278	4,465
	(1.64)	(.84)	(.89)
$(\text{RISK2})^2$	$.601 \times 10^{-5}$	.179	.265
	(1.44)	(.62)	(.93)
Lot size	.065	4,623	.043
	(4.41)	(4.58)	(2.46)
Liv. area	.464	31,623	20.5
	(15.48)	(15.32)	(17.29)
Status	−2.575	−96,094	−116,570
	(−3.98)	(−2.15)	(−2.80)
Access	.359	23,772	2,268
	(2.24)	(2.15)	(2.35)
Age	−.053	−306	−339
	(−10.45)	(−8.86)	(−10.13)
SEE	.1476	10,155	10,016
Estimated cleanup benefits per house (1980 dollars):			
Acton	$63.4	$107.93	$83.0
	(264.7)	(322.4)	(318.6)
Ashland	$652.5	$487.11	$421.3
	(303.0)	(369.0)	(363.6)

Notes: t-statistics (computed using the usual OLS formula) are in parentheses. Based on OLS regressions with fixed town effects using 324 observations. Dollar values of benefits for column (1) were computed by exponentiating the predicted log benefit and adjusting for its variance. "Regressors: logs" means that logarithms of LOTSZ, LIV, STATUS, and ACCESS were used.

imprecisely; only for the log/log specifications can nonpositivity of the benefits be rejected at the 5 percent level.[6]

[6] Using all 2,182 observations, Harrison and Stock (1984) obtained more precise estimates of these cleanup benefits.

Table 4. *OLS specification tests and estimated benefits*

Regressions:

$$\text{(I)}\quad y=\sum_{i=1}^{k_1} z_i\beta_{1i}+\sum_{i=1}^{k_2} w_i\beta_{2i}+\sum_{i=1}^{k_2}\sum_{j=1}^{k_2} w_i w_j\beta_{3ij}+\sum_{i=1}^{k_1} z_i^2\beta_{4i}+\sum_{i=1}^{k_1}\sum_{j=i+1}^{k_1} z_i z_j\beta_{5ij}+\sum_{i=1}^{k_1}\sum_{j=1}^{k_2} z_i w_j\beta_{6ij}$$

$$\text{(II)}\quad y=\sum_{i=1}^{k_1} z_i\beta_{1i}+\sum_{i=1}^{k_2} w_i\beta_{2i}+\sum_{i=1}^{k_2}\sum_{j=1}^{k_2} w_i w_j\beta_{3ij}+\sum_{i=1}^{k_1} z_i^2\beta_{4i}+\sum_{i=1}^{k_1}\sum_{j=i+1}^{k_1} z_i z_j\beta_{5ij}$$

	(1) Dep. vble.: ln(price) Regressors: logs	(2) Dep. vble.: price Regressors: logs	(3) Dep. vble.: price Regressors: levels
A. Wald tests of exclusion restrictions and p-values:			
1. $\{\beta_{6ij}=0\}$ in (I)	1.11	1.16	.86
	(.356)	(.318)	(.573)
2. $\{\beta_{5ij}=0\}$ in (II)	10.24	15.74	13.67
	(.000)	(.000)	(.000)
3. $\{\beta_{4i}=0\}$ in (II)	3.48	9.19	8.97
	(.001)	(.000)	(.000)
4. $\{\beta_{4i}=0,\ \beta_{5ij}=0\}$ in (II)	6.82	14.24	13.89
	(.000)	(.000)	(.000)

B. Estimated cleanup benefits per house (1980 dollars):

	(1) Acton	(1) Ashland	(2) Acton	(2) Ashland	(3) Acton	(3) Ashland
1. (II), unrestricted	85.9	515.3	160.6	296.0	110.2	283.5
	(240.7)	(276.4)	(258.6)	(296.9)	(256.6)	(295.2)
SEE:	.1301		7,892		7,825	
2. (II), imposing $\{\beta_{4i}=0\}$	48.4	626.6	133.8	330.5	12.4	422.9
	(251.8)	(295.8)	(281.3)	(330.5)	(275.9)	(323.6)
SEE:	.1400		8,830		8,630	
3. (II), imposing $\{\beta_{5ij}=0\}$	144.2	401.3	164.1	187.5	176.2	204.5
	(245.0)	(280.0)	(285.6)	(326.4)	(281.3)	(323.5)
SEE:	.1354		8,912		8,809	

Notes: In regressions (I) and (II), z denotes the vector of five nonwaste housing attributes, and w denotes the variables RISK1 and RISK2, so $k_1=5$ and $k_2=2$. The parameters β_3, β_4, β_5, and β_6 represent the coefficients on the various quadratic terms in these regressions. In Panel A, the usual F-statistic is given, with its marginal significance level in parentheses. All regressions included fixed town effects. See the notes to Table 3.

Nonparametric estimation results

The nonparametric estimator was computed using price as the dependent variable and the seven independent variables in Table 2 as the independent variables. The estimator was computed using several different kernels. The first two are conventional nonnegative kernels, based respectively on the so-called bisquare function and on the Gaussian density:

$$\text{Bisquare:} \quad w(v) = (1-4v^2)^2 1(|v| < .25). \tag{12a}$$

$$\text{Gaussian:} \quad w(v) = \exp(-.5v^2). \tag{12b}$$

The kernels were evaluated using the inverse of the $\mathbf{X}$ covariance matrix to scale the data, so that $v = [(\mathbf{X}_i - \mathbf{x})' \hat{\Sigma}_{\mathbf{X}}^{-1} (\mathbf{X}_i - \mathbf{x})]^{1/2} / b_n$.

It remains to select the bandwidth b_n. Recent research has examined the properties of various automatic bandwidth selection procedures; see, for example, Devroye and Penrod (1984), Gasser et al. (1984), Li (1984), Stone (1984), Marron (1985), Muller (1984), and especially Rice (1984). One candidate is to choose the bandwidth that minimizes the sum of squared residuals of the regression equation,

$$\text{MSE} = n^{-1} \sum (Y_i - g_n(\mathbf{X}_i) - \mathbf{d}_i' \mathbf{A}_n)^2. \tag{13}$$

Because Y_i is used in computing $g_n(\mathbf{X}_i)$, the mean square error (MSE) (13) can be reduced arbitrarily by decreasing the bandwidth until effectively all weight in $g_n(\mathbf{X}_i)$ is placed on Y_i. An alternative procedure that circumvents this problem is to choose the bandwidth to minimize the sum of squared residuals from the cross-validated regression. In this approach, the bandwidth minimizes

$$\text{MSE}_{cv} = n^{-1} \sum (Y_i - g_{ni}(\mathbf{X}_i) - \mathbf{d}_i' \mathbf{A}_n)^2, \tag{14}$$

where

$$g_{nj}(\mathbf{X}_j) = \sum_{i \neq j} w\left(\frac{\mathbf{X}_i - \mathbf{X}_j}{b_n}\right)(Y_i - \mathbf{d}_i' \mathbf{A}_n) \Big/ \sum_{i \neq j} w\left(\frac{\mathbf{X}_i - \mathbf{X}_j}{b_n}\right).$$

The MSE_{cv} avoids the difficulty of the MSE in (13) by not using Y_j in estimating $g(\mathbf{X}_j)$. The asymptotic distribution theory for B_n in Stock (1989) is based on a fixed sequence of bandwidths. Minimizing the MSE_{cv} is an automatic procedure that provides a pointwise consistent regression estimator. Developing an asymptotic theory for B_n with data-dependent bandwidths remains a task for future research.

The standard error of the estimate (SEE) of the cross-validated regression ($\text{SEE}_{cv} = (\text{MSE}_{cv})^{1/2}$) is presented in Figure 1 for the bisquare kernel and in Figure 2 for the Gaussian kernel. Also presented in these figures

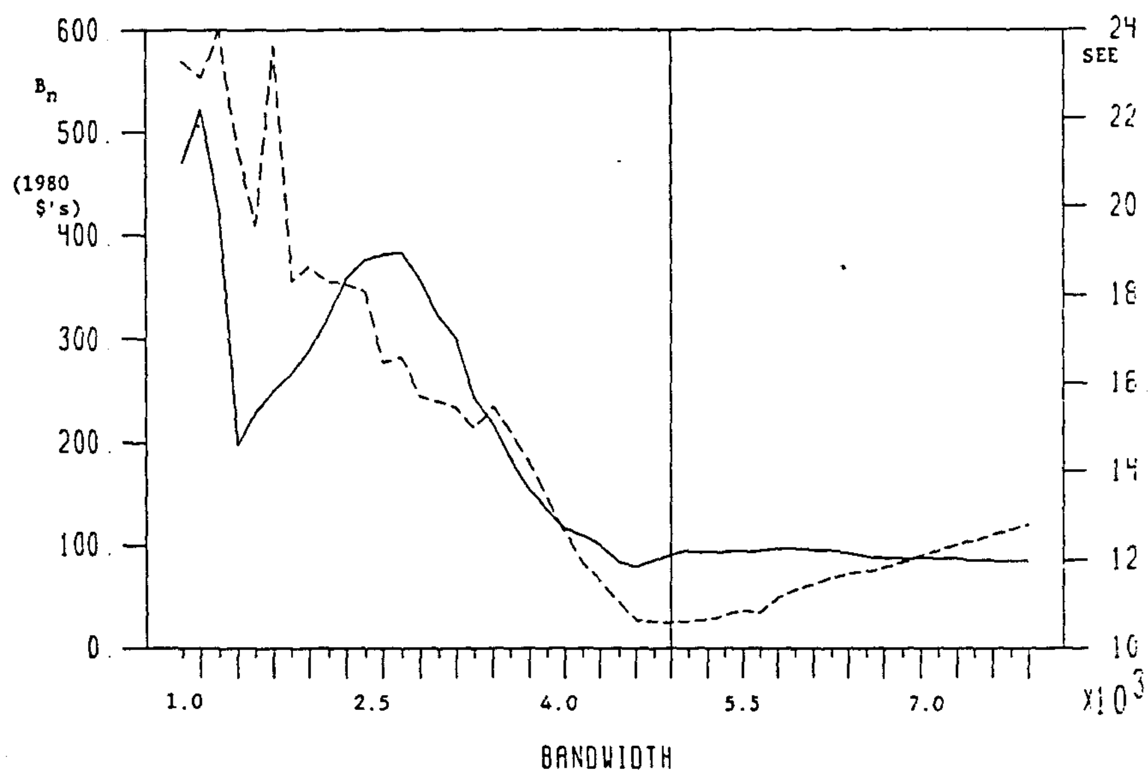

Figure 1. Cross-validated SEE (SEE_{cv}, dashed line) and estimated clean-up benefits (B_n, solid line) for the Ashland site, based on the bisquare kernel (12a).

is the benefits estimator B_n, evaluated for the Ashland site. The vertical line in the graphs denotes the value of the bandwidth that minimizes the MSE_{cv}.[7]

Figure 1 shows that the benefits estimator based on the bisquare kernel is relatively stable for bandwidths exceeding 4, ranging from $116 per house for $b_n = 4.15$ to $85 per house for $b_n = 7.9$. However, the estimator becomes unstable as b_n decreases below 4. The SEE_{cv} for the bisquare kernel exhibits a clear minimum and a fairly smooth shape near this minimum, although considerable instability for smaller bandwidths.

As in Figure 1, the benefits estimator based on the Gaussian kernel (Figure 2) is relatively insensitive to the bandwidth for values near the bandwidth that minimizes the SEE_{cv}. The benefits estimator and the SEE_{cv} exhibit substantial instability both as b_n decreases and, unlike the estimator based on the bisquare kernel, as b_n increases.

The instability in the nonparametric estimator for small bandwidths arises because, as the bandwidth decreases, a smaller fraction of the data

[7] The minimizing bandwidth was determined by a grid search, with a grid of 0.15 for the bisquare kernel and a grid of 0.05 for the Gaussian kernel.

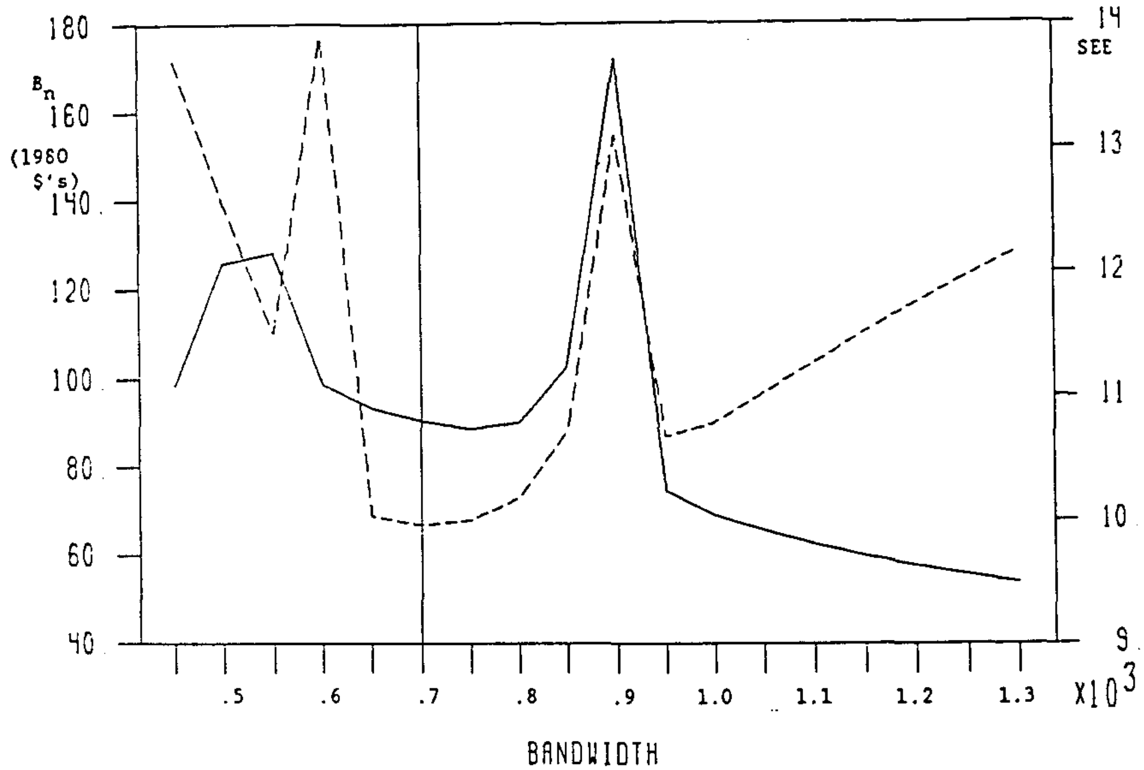

Figure 2. Cross-validated SEE (SEE_{cv}, dashed line) and estimated clean-up benefits (B_n, solid line) for the Ashland site, based on the Gaussian kernel (12b).

is used to estimate the pointwise conditional expectations. A rough measure of the amount of averaging involved in computing the nonparametric estimator is the ratio of the area in the space of $\tilde{\mathbf{X}} = \hat{\Sigma}_{\mathbf{X}}^{-1/2}\mathbf{X}$ over which the pointwise averaging occurs to the support of $\tilde{\mathbf{X}}$ variables. Approximating this support by a sphere of radius 2 and the neighborhood over which the averaging occurs by a sphere with radius b_n (appropriate for the Gaussian kernel), this ratio is $(\frac{1}{2}b_n)^k$, where k is the dimension of X. For $k=7$ and $b_n=0.6$, this ratio is 0.02 percent; for $b_n=1$ and 1.5, this ratio is 0.8 percent and 13.3 percent, respectively. Although the fraction of observations contained in the neighborhood of an observation would typically be larger than this because of the clustering of the X data, for small bandwidths the averages will still effectively be taken over relatively few observations.

In addition to these nonnegative kernels, the benefits estimator was computed for an infinite-order kernel, the "L" kernel discussed by Hall and Marron (1988):

$$L: \quad w(v) = v^{-2}\left\{\left[\sin\frac{v}{2}\right]^2 - \left[\sin\frac{v}{4}\right]^2\right\}. \tag{15}$$

Table 5. *Nonparametric cleanup benefit estimates*

Kernel	Bisquare	Gaussian	L
b_n	5.5	.7	.45
SEE_{cv}	10,755	9,956	14,754
Acton	−308.4	15.1	82.1
	(107.0)	(63.2)	(2,730)
Ashland	93.9	90.3	144.1
	(71.5)	(92.3)	(1,293)

Notes: Based on 324 observations from 11 towns in the Boston area. Standard errors, computed using (9), are reported in parentheses. The kernels and cleanup experiments are discussed in the text.

The advantage of the L kernel is that, by having higher moments equal to zero, in theory the bias in the pointwise kernel estimators and therefore in B_n can be reduced. The multivariate kernel was constructed as the k-fold product of univariate L kernels, each taking as an argument $v_j = (\hat{\Sigma}_{\mathbf{X}})_{jj}^{-1/2}|\mathbf{X}_{ji} - \mathbf{x}_j|/b_n$, where $\mathbf{X}_{ji} - \mathbf{x}_j$ denotes the jth element of the vector $\mathbf{X}_i - \mathbf{x}$ and where $(\hat{\Sigma}_{\mathbf{X}})_{jj}$ denotes the (j,j) element of $\hat{\Sigma}_{\mathbf{X}}$.

A small grid search was performed to find the bandwidth that minimized the SEE_{cv} for the L kernel; this yielded $b_n = 0.45$, with a SEE_{cv} of \$14,754. This SEE_{cv} is substantially larger than that for the other kernels or for the OLS regressions. The benefits estimators also exhibited much greater instability as a function of the bandwidth for the L kernel than for the nonnegative kernels, with values of B_n ranging from −\$1,436 for $b_n = 0.35$ to \$534 for $b_n = 0.65$.

The minimized SEE_{cv}, along with the corresponding bandwidth, are presented for the two kernels in the first two rows of Table 5.[8] The Gaussian kernel makes slightly better predictions than the bisquare, based on this measure. Both SEE_{cv}'s are larger than the SEE's reported for the OLS regressions in Table 4, columns (2) and (3). However, these statistics

[8] Although the bandwidths that minimized the SEE_{cv} for the various kernels appear to differ, in fact the shapes of the kernels with these automatically selected bandwidths are similar. Write each kernel as $w(u/b_n)$, where b_n is given in the appropriate column of Table 4. Then the greatest difference between the Gaussian and bisquare kernels occurs near $u = 1$; at $u = 1$, the bisquare kernel has a value of 0.71 and the Gaussian kernel has a value of 0.36. The values of the L kernel fall between the two nonnegative kernels for $|u| < 2$, at which point it takes on negative values.

are not directly comparable, since the SEE_{cv} could understate the predictive quality of the nonparametric regression.

The nonparametric estimates of the cleanup benefits, computed using the automatically selected bandwidth, are given in the final rows of Table 5. The estimates for Acton differ substantially from one kernel to another. In contrast, the point estimates for Ashland are relatively constant across the choice of the kernel, even though the estimated standard errors are not. The Ashland benefit estimates are lower than those computed using the OLS regression, although they are the same order of magnitude. Although the confidence regions for the estimates based on the nonnegative kernels are tighter than those based on the OLS results, one cannot reject the hypothesis that the cleanup benefits are zero at the 5 percent level for any of the nonparametric estimates.

An estimate of the total value of the cleanup is obtained by multiplying the per-house benefits by 48,303, the number of homes in the towns in the sample. The estimated total benefits for cleaning up the two sites are substantial. Based on the bisquare kernel, for example, the total benefits of the Ashland cleanup are estimated to be \$4.5 million. Based on the unrestricted version of model II in Table 4, the OLS estimates of these benefits range from \$13.7 million to \$24.9 million. Unfortunately, the range of these estimates is rather wide, as are the associated confidence intervals.

6 Conclusions

This application of kernel techniques to the housing price data set suggests several directions for future methodological work. First, for either of the nonnegative kernels considered, the benefits estimators were relatively insensitive to the bandwidth when in the neighborhood of the bandwidth that minimized the cross-validated MSE. However, outside of that range (and particularly for smaller bandwidths) the estimators became quite unstable. This emphasizes the need to extend the work on automatic bandwidth selection procedures from existing results on pointwise estimation to the new set of nonparametric procedures that use the entire sample, such as the nonparametric benefits estimator considered here. Second, the benefit estimates for the hypothetical Acton cleanup (but not the Ashland cleanup) differ substantially across the two nonnegative kernels, even at the automatically chosen bandwidth. This suggests the need for additional theoretical guidance in kernel choice. Third, a limited investigation with a higher-order kernel produced estimates that were highly sensitive to bandwidth choice and very imprecise. As a general observation, the range of estimates obtained from the nonparametric procedures based on the nonnegative kernels was comparable to the range of estimates obtained using different parametric models.

Both the parametric and nonparametric analysis resulted in imprecise estimates of the value of cleaning up the hazardous waste sites. There are several possible explanations for this imprecision. First, the number of observations (chosen to reduce the computations in the nonparametric analysis) is small. Second, including fixed town effects controls for a number of important town-specific attributes (such as the level of industrialization), but it eliminates the possibility of exploiting price variations across towns that arise from hazardous waste sites. Third and perhaps most important, the effect of interest here – the reduction of housing values as a result of a local hazardous waste site – is likely to be subtle for all but the closest homes, and in any case this analysis requires the home buyer to be aware of local sites.

These estimates ignore several potentially important sources of benefits. For example, they are based solely on the population of single-family residences in the eleven Boston suburbs considered, and do not include any benefits estimates for Boston itself. The estimates do not include benefits accruing to owners of multiple family residences, commercial, or industrial property, except to the extent that these benefits are reflected in the prices of single-family residences. Nor do these estimates include any assessment of the benefits of reducing health risks encountered through drinking water contamination, since these effects were included in the town-specific fixed effects estimated in both the nonparametric and the OLS applications. These arguments suggest that the approach used here will understate the total cleanup benefits. Despite these caveats, however, it is of interest that the range of \$4.5 million to \$25 million is consistent with the range of estimates of cleanup benefits obtained by other methods (see Desvousges and Smith 1983 and Harrison and Stock 1984 for discussion).

Appendix

This appendix provides a derivation of equation (7). From (6),

$$B_n = n^{-1} \sum_{j=1}^{n} [g_n(\mathbf{X}_j^*) - g_n(\mathbf{X}_j)].$$

From (3), $E(Y|\mathbf{x}) = g(\mathbf{x}) + E(\mathbf{d}_i|\mathbf{x})'\mathbf{A}$. As in Stock (1989), define $f_1(\mathbf{x}) = E(Y|\mathbf{x})$ and $f_2(\mathbf{x}) = E(d|\mathbf{x})$, so $g(\mathbf{x}) = f_1(\mathbf{x}) - f_2(\mathbf{x})'A$. This suggests estimating $g(\mathbf{x})$ by $g_n(\mathbf{x}) = f_{1n}(\mathbf{x}) - f_{2n}(\mathbf{x})'\mathbf{A}_n$, where A_n is given in (5) and the kernel regression estimators of $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$, respectively, are $f_{1n}(\mathbf{x}) = \sum_{i=1}^{n} W_i(\mathbf{x}) Y_i$ and $f_{2n}(\mathbf{x}) = \sum_{i=1}^{n} W_i(\mathbf{x})\mathbf{d}_i$, where

$$W_i(\mathbf{x}) = w\left(\frac{\mathbf{X}_i - \mathbf{x}}{b_n}\right) \Big/ \sum_{i=1}^{n} w\left(\frac{\mathbf{X}_i - \mathbf{x}}{b_n}\right).$$

These definitions imply:

$$\begin{aligned} B_n &= n^{-1} \sum_{j=1}^{n} [(f_{1n}(\mathbf{X}_j^*) - f_{2n}(\mathbf{X}_j^*)'\mathbf{A}_n) - (f_{1n}(\mathbf{X}_j) - f_{2n}(\mathbf{X}_j)'\mathbf{A}_n)] \\ &= n^{-1} \sum_{j=1}^{n} \left[\sum_{i=1}^{n} W_i(\mathbf{X}_j^*) Y_i - \sum_{i=1}^{n} W_i(\mathbf{X}_j^*)\mathbf{d}_i' \mathbf{A}_n \right] \\ &\quad - n^{-1} \sum_{j=1}^{n} \left[\sum_{i=1}^{n} W_i(\mathbf{X}_j) Y_i - \sum_{i=1}^{n} W_i(\mathbf{X}_j)\mathbf{d}_i' \mathbf{A}_n \right] \\ &= n^{-1} \sum_{j=1}^{n} \left[\sum_{i=1}^{n} (W_i(\mathbf{X}_j^*) - W_i(\mathbf{X}_j))(Y_i - \mathbf{d}_i' \mathbf{A}_n) \right] \\ &= n^{-1} \sum_{i=1}^{n} \left[\sum_{j=1}^{n} (W_i(\mathbf{X}_j^*) - W_i(\mathbf{X}_j)) \right] (Y_i - \mathbf{d}_i' \mathbf{A}_n). \end{aligned}$$

Thus

$$B_n = n^{-1} \sum_{j=1}^{n} \gamma_n(\mathbf{X}_j)(Y_j - \mathbf{d}_j' \mathbf{A}_n),$$

where

$$\begin{aligned} \gamma_n(\mathbf{X}_j) &= \sum_{i=1}^{n} (W_j(\mathbf{X}_i^*) - W_j(\mathbf{X}_i)) \\ &= \sum_{i=1}^{n} \left[w\left(\frac{\mathbf{X}_j - \mathbf{X}_i^*}{b_n}\right) \Big/ \sum_{j=1}^{n} w\left(\frac{\mathbf{X}_j - \mathbf{X}_i^*}{b_n}\right) \right] \\ &\quad - \sum_{i=1}^{n} \left[w\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{b_n}\right) \Big/ \sum_{j=1}^{n} w\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{b_n}\right) \right], \end{aligned}$$

which, upon replacing $\gamma_n(\mathbf{X}_j)$ by $\gamma_n(\mathbf{x})$, is equation (7).

REFERENCES

Brown, J. N., and H. Rosen (1982), "On the Estimation of Structural Hedonic Price Models," *Econometrica,* 50, 765–8.

Desvousges, W. H., and V. K. Smith (1983), "The Benefits of Hazardous Waste Management Regulations," Technical Report, Research Triangle Institute, February.

Devroye, L., and C. S. Penrod (1984), "The Consistency of Automatic Kernel Density Estimates," *Annals of Statistics,* 12, 1231–49.

Freeman, A. M. (1979), *The Benefits of Environmental Improvement: Theory and Practice* (Baltimore: Johns Hopkins Press).

Gasser, T., H. G. Muller, W. Kohler, L. Molinari, and A. Prader (1984), "Nonparametric Regression Analysis of Growth Curves," *Annals of Statistics,* 12, 210–29.

Hall, P., and J. S. Marron (1988), "Choice of Kernel Order in Density Estimation," *Annals of Statistics,* 16, 161–73.
Halvorsen, R., and H. Pollakowski (1981), "Choice of Functional Form for Hedonic Price Equations," *Journal of Urban Economics,* 10, 37–49.
Harrison, D., Jr., and D. Rubinfeld (1978), "Hedonic Housing Prices and the Demand for Clean Air," *Journal of Environmental Economics and Management,* 5, 81–102.
Harrison, D., Jr., and J. H. Stock (1984), "Hedonic Housing Values, Local Public Goods, and the Benefits of Hazardous Waste Cleanup," Kennedy School of Government Discussion Paper, no. E-84-09, November.
Li, K.-C. (1984), "Consistency for Cross-Validated Nearest Neighbor Estimates in Nonparametric Regression," *Annals of Statistics,* 12, 230–40.
Massachusetts Department of Environmental Quality Engineering (DEQE) (1981a), *Management for Site Investigation 1981 Update.*
(1981b), *Inventory of Active and Inactive Solid Waste Landfills.*
Marron, J. S. (1985), "An Asymptotically Efficient Solution to the Bandwidth Problem of Kernel Density Estimation," *Annals of Statistics,* 13, 1019–23.
Muller, H. G. (1984), "Smooth Optimum Kernel Estimators of Densities, Regression Curves, and Modes," *Annals of Statistics,* 12, 766–74.
Polinsky, A. M., and D. Rubinfeld (1978), "The Long-Run Effects of a Residential Property Tax and Local Public Services," *Journal of Urban Economics,* 5, 241–62.
Polinsky, A. M., and S. Shavell (1975), "The Air Pollution and Property Value Debate," *Review of Economics and Statistics,* 57, 100–4.
(1976), "Amenities and Property Values in a Model of an Urban Area," *Journal of Public Economics,* 5, 119–29.
(1978), "Amenities and Property Values in a Model of an Urban Area: A Reply," *Journal of Public Economics,* 9, 111–12.
Prakasa Rao, B. L. S. (1983), *Nonparametric Functional Estimation* (New York: Academic Press).
Quigley, J. M. (1982), "Nonlinear Budget Constraints and Consumer Demand: An Application to Public Programs for Residential Housing," *Journal of Urban Economics,* 12, 177–201.
Rice, J. (1984), "Bandwidth Choice for Nonparametric Regression," *Annals of Statistics,* 12, 1215–30.
Ridker, R. G., and J. A. Henning (1967), "The Departments of Residential Property Values with Special Reference to Air Pollution," *Review of Economics and Statistics,* 49, 246–57.
Robinson, P. M. (1988), "Root-*N* Consistent Semiparametric Regression," *Econometrica,* 56, 931–54.
Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy,* 82, 34–55.
Scotchmer, S. (1985), "The Short-Run and Long-Run Benefits of Environmental Improvement," Harvard Institute of Economic Research Discussion Paper No. 1935, January.
Stock, J. H. (1989), "Nonparametric Policy Analysis," *Journal of the American Statistical Association,* 84, no. 406, 567–75.
Stone, C. J. (1984), "An Asymptotically Optimal Window Selection Rule for Kernel Density Estimates," *Annals of Statistics,* 12, 1285–98.
White, H. (1980), "Using Least Squares to Approximate Unknown Regression Functions," *International Economic Review,* 21, 149–70.

(1981), "Consequences and Detection of Misspecified Nonlinear Regression Models," *Journal of the American Statistical Association,* 76, 419–33.
Yinger, J. (1982), "Capitalization and the Theory of Local Public Finance," *Journal of Political Economy,* 90, 917–43.