

Does Cleanup of Hazardous Waste Sites Raise Housing Values? Evidence of Spatially Localized Benefits

Shanti Gamper-Rabindran*
Christopher Timmins†

Working Paper EE 12-03
February 2012

*Graduate School of Public and International Affairs, University of Pittsburgh

†Department of Economics, Duke University

The Duke Environmental Economics Working Paper Series provides a forum for Duke faculty working in environmental and resource economics to disseminate their research. These working papers have not undergone peer review at the time of posting.

DOES CLEANUP OF HAZARDOUS WASTE SITES RAISE HOUSING VALUES? EVIDENCE OF SPATIALLY LOCALIZED BENEFITS

Shanti Gamper-Rabindran
Graduate School of Public and International Affairs
University of Pittsburgh

Christopher Timmins
Department of Economics
Duke University

Version: Jan 19, 2012

Abstract

Economists often rely on publicly available data provided at coarse geographical resolution to value spatially localized amenities. We propose a simple refinement to the hedonic method that accommodates this reality: specifically, we measure localized benefits from the cleanup of hazardous waste sites at the sub-census tract level by examining the *entire within-tract housing value distribution*, rather than simply focusing on the tract median. Doing so, we find significantly larger benefits from NPL listing at lower percentiles. We find large effects of construction complete and deletion across the housing value distribution, although these effects are also larger at lower percentiles. We confirm these results with restricted access census block data, and use proprietary housing transactions data to show that cheaper houses within a census tract are indeed more likely to be closer to a hazardous waste site, explaining the greater impacts they receive from the cleanup process.

JEL codes H41, C2, Q5

Keywords: Hedonic property method, localized public goods, heterogeneous treatment effects, hazardous waste remediation

Funding for this research is from the U.S. Environmental Protection Agency under Purchase Order Number EP09W001911, the University of Pittsburgh's Center on Race and Social Problems and University Center for Social and Urban Research. Support for this research at the Triangle RDC is from NSF (ITR-0427889). We use data from the EPA, the Census Research Data Center and additional data compiled and published by Michael Greenstone and Justin Gallagher. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

DOES CLEANUP OF HAZARDOUS WASTE SITES RAISE HOUSING VALUES?

EVIDENCE OF SPATIALLY LOCALIZED BENEFITS

Abstract (146 words)

Economists often rely on publicly available data provided at coarse geographical resolution to value spatially localized amenities. We propose a simple refinement to the hedonic method that accommodates this reality: specifically, we measure localized benefits from the cleanup of hazardous waste sites at the sub-census tract level by examining the *entire within-tract housing value distribution*, rather than simply focusing on the tract median. Doing so, we find significantly larger benefits from NPL listing at lower percentiles. We find large effects of construction complete and deletion across the housing value distribution, although these effects are also larger at lower percentiles. We confirm these results with restricted access census block data, and use proprietary housing transactions data to show that cheaper houses within a census tract are indeed more likely to be closer to a hazardous waste site, explaining the greater impacts they receive from the cleanup process.

JEL codes H41, C2, Q5

Keywords: Hedonic property method, localized public goods, heterogeneous treatment effects, hazardous waste remediation

1. Measuring Localized Benefits with Tract-Level Housing Data

A growing number of studies document that amenities or disamenities are highly localized (e.g., at the sub-tract level) with effects that decline rapidly with distance. For example, Davis (2011) detects strong adverse effects of power plants on prices of houses that are located within two miles, weaker effects between two to five miles, and no effect beyond five miles. Campbell, Giglio and Pathak (2011) find that each nearby foreclosure lowers the price of a house by about 2% if it takes place at zero distance and 1% if it takes place at a distance of 0.05 miles. Because fine-resolution data at the house-level or block-level are often inaccessible,¹ these benefits from spatially localized amenities are often estimated using publicly available tract, zip-code, or county-level median housing value data.² Using tract-level median housing values to capture benefits that are localized at the sub-tract level can, however, result in a failure to detect the true underlying benefits. Our study proposes a simple refinement to the hedonic method that overcomes this problem. Specifically, we recover localized effects by examining the *entire within-tract housing value distribution*, rather than simply focusing on the tract median.

We apply our method to estimating the benefits from the cleanup of hazardous waste under the Superfund program. Under that program, the most severely contaminated sites are placed on its National Priorities List (NPL) (Sigman 2008; Sigman and Stafford 2011) and

¹ Block-level Decennial Census data are restricted to users at the Census Data Centers (Davis, 2011) and (ii) individual house transactions data are typically proprietary, expensive (Oberholzer-Gee and Mitsunari, 2006), or limited in their coverage to specific counties (Davis 2004; Linden and Rockoff 2008; Pope 2008; Zabel and Guignet 2010).

² Recent examples include Bui and Meyer (2003), Chay and Greenstone (2005), Hanna (2007), Greenstone and Gallagher (2008), Grainger (2010), and Sanders (2011). Early papers from the hedonics literature that relied upon coarse geographical housing data include Ridker and Henning (1967), Deyak and Smith (1974), Nelson (1975), Harrison and Rubinfeld (1978), and Nelson (1978).

cleanup is undertaken for a subset of these sites. Restricting our analysis to sites that have similar risk scores, we compare owner-occupied housing values in neighborhoods located within three miles around sites that have been cleaned up with corresponding neighborhoods around sites that have not been cleaned. The appreciation in housing values in response to cleanup (measured with the deletion milestone in the Superfund process) is defined relative to the pre-proposal baseline. In response to cleanup, our tract-level analysis detects larger appreciation at the 10th percentile of the within-tract housing value distribution (24.5%) than at the median (20.3%) and the 90th percentile (18.6%). A cost-benefit analysis based on these results indicates that cleanup under the Superfund program yields net benefits for 35 out of 55 sites that have been deleted from the NPL by 2000.

The pattern of within-tract heterogeneity we find with deletion is even more evident considering listing on the NPL – i.e., the Superfund milestone indicates that the site should be cleaned at some point in the future. Listing on the NPL results appreciation of 9% to 5.6% at the 10th to the 40th percentiles of housing values, but estimates at the median and higher percentiles are no longer statistically significant.

We find that examining the full housing value distribution can have important policy implications for the results of a hedonic analysis. In our example of valuing Superfund cleanup, a focus on the median housing value would have understated the larger effects at the lower tails of the housing value distribution. This understatement would have been most problematic at an intermediate stage of the Superfund process (i.e., at the listing stage). One can imagine other situations in which the distribution of benefits is such that a focus on the mean or median could lead to a failure to detect any treatment impacts, should those impacts exist only in the tails of the distribution of housing values.

Our analysis of high geographical resolution data (i.e., restricted-access block data) supports this finding. We re-run our analysis for those blocks contained in the tracts lying within three miles of these sites and find that housing values appreciate by 14.7% with deletion. We do not expect to recover identical estimates of the effects of cleanup from the block and tract analyses, as the neighborhoods' exposure to sites at various stages of cleanup cannot be defined identically in these separate analyses. The block level analysis does, however, provide a valuable check on our tract-level results.

As further supporting evidence, we examine geocoded proprietary housing transactions data from ten different states. These data show explicitly that it is, in fact, the cheaper houses within each tract that are more likely to be exposed to waste sites within one kilometer. This pattern is less evident when we consider exposure at greater distances (e.g., two or three kilometers). The results of this transaction-level analysis are particularly useful in explaining the greater appreciation from site listing in the lower tail of the housing value distribution that we find in our tract-level analysis.

Our proposed hedonic refinement proves to be important when the analysis of coarse resolution data results in a failure to detect localized benefits. We note that many hedonic studies are forced to rely on coarse resolution data because of its nationwide coverage and public availability, but demonstrate that the benefits of spatially localized amenities could be both substantial (particularly in densely populated areas) and likely missed by analyses focused on mean or median values. While our refinement cannot detect all forms of heterogeneity across housing markets that could, for example, be evident with transaction-level data, it does avoid an important source of bias without saddling the researcher with difficult (often prohibitive) data expenses.

2 Potential Benefits from Superfund Cleanup

In the late 1970's, events at Love Canal and the Valley of Drums raised public concern over the health and environmental risks associated with contaminated waste sites.³ In response to these and other similar incidents, the US Congress enacted the 1980 Comprehensive Environmental Response, Compensation and Liability Act (CERCLA). Under that law, the most hazardous sites are placed on the National Priorities List (NPL). There are four major milestones in the NPL process – proposal, listing, construction complete, and deletion – at which the EPA publicizes information about the site, sometimes entering information into the Federal Register and soliciting public comment. These milestones, by providing information to the housing market, have the potential to influence housing values.

The NPL process begins with a preliminary assessment and site inspection; based on that assessment, the EPA may propose a site to the NPL in the Federal Register. Information collected during the preliminary assessment and site inspection is used to calculate a Hazard Ranking System (HRS) score.⁴ The EPA then lists the site on the final NPL if it meets at least one of three criteria – (i) the HRS is of sufficient magnitude, (ii) the state environmental authority designates the site to be a top priority, or (iii) the US Public Health Service recommends removing all people in close proximity to the site. The construction complete designation indicates the physical construction phase of the cleanup process has been completed and immediate public health threats have been addressed, though other remedial actions have yet

³ A recent study documents that mothers' residence close to a Superfund site before cleanup is associated with a 20 to 25% increase in the risk of congenital anomalies (Currie, Greenstone, Moretti, 2011).

⁴ The HRS score serves as a numerically based screening device that uses information from initial, limited investigations. Sites with an HRS score of 28.5 or greater are eligible for listing on the NPL and require the preparation of an HRS scoring package. The story behind the 28.5 cutoff is described in detail below.

to be completed. Finally, deletion of a site from the NPL requires that the necessary actions for remediation have been completed and the site no longer poses a threat to human health.

There are two channels – one direct and one indirect – through which deletion from the NPL can increase housing values. First, cleanup reduces health risks and other disamenities associated with a site. Second, cleanup may prompt further development in the area surrounding a site, including the potential for re-zoning from a lower-value commercial use to higher-value residential (even luxury) development.⁵ As long as this sort of development occurs *conditional* on cleanup being undertaken, our study correctly considers the benefits from it to be part of the benefits from Superfund cleanup. Such an outcome would represent one mechanism through which remediation can be translated into higher housing prices. Our analysis would be biased if causality went in the opposite direction – e.g., if developers decided to build a luxury resort (which was going to raise nearby housing prices regardless of EPA actions), and the EPA responded by moving the site through the remediation process more rapidly. Our review of the literature does not suggest that this is the case. Moreover, our analysis employs (i) sample restrictions to ensure that we are making comparisons among tracts that are similar to one another aside from their receipt of cleanup, and (ii) panel methods to control for time-invariant unobservable differences in tracts.

Any hedonic estimates of the benefits of Superfund cleanup come with five caveats. First, benefits are understated if homeowners ignore downstream benefits. Second, the appreciation of housing values reflects homebuyers' perceptions of risk reductions, and these

⁵ For example, the Empire Canyon Daly West Mine Superfund site in Utah underwent extensive remediation under the Superfund program. After remediation, the landowners leased that site for the development of a luxury resort, including a hotel, spa and condominium project (EPA 2008).

perceptions, though influenced by the information that EPA provides,⁶ may not fully reflect true reductions in risks. Third, we cannot account for changes in housing values that result from the cleanup of nearby sites undertaken outside the Superfund program because data describing these sites are unavailable.⁷ Fourth, like previous studies, we treat Superfund cleanup as a marginal change to the overall housing market.⁸ With the assumption that the hedonic price schedule does not shift, we can interpret our capitalization results in a marginal willingness-to-pay framework (Kuminoff and Pope 2010). Without this assumption, our results can be interpreted as capitalization effects, which are also important to policy makers. Finally, we follow the majority of the hedonics literature and simply analyse the value of marginal changes along the hedonic price function. We do not attempt to identify the marginal willingness to pay function, given the difficulties inherent in such a task; see Kuminoff, Smith and Timmins (2010) for a discussion. Bishop and Timmins (2011) describes these difficulties in more detail and suggest empirical approaches to deal with them.

While our study focuses on the effects of deletion from the NPL, we note that at least three other Superfund milestones can also influence nearby housing values *measured relative to the pre-proposal stage*. Proposal of a site to the NPL may reduce neighborhood housing prices when this action provides new information to the housing market that contamination is severe enough to warrant the potential listing of that site on the NPL (although, if the housing market

⁶ For example, Gayer, Hamilton and Viscusi (2000) find that residents around seven NPL and non-NPL sites in Grand Rapids Michigan updated their perceptions of risks when the EPA released information about those sites from remedial investigations.

⁷ The EPA does not maintain a list of verified coordinates of non-NPL sites. This data limitation has constrained other studies (Kiel and Williams 2007; Noonan, Krupka and Baden 2007; Greenstone and Gallagher 2008).

⁸ We assume that the cleanup of a single Superfund site may not significantly alter the hedonic price function, unlike a policy that, for example, affects the quality of a high percentage of the schools in a district.

expects that proposal signals that the site is likely to be remediated, this countervailing factor will dampen the extent of that depreciation). Housing prices have been found to decline due to perceived increases in health risks (Hamilton and Viscusi 1999)⁹ and stigma (Fischhoff 2001; Messer et al. 2006) associated with a contaminated site.¹⁰ Unlike proposal and deletion, listing of an NPL site is associated with two countervailing forces. (i) It may reduce housing prices by confirming the severe nature contamination of that site, but (ii) it may also increase housing prices by signaling that a site has been placed on the path towards remediation. The construction complete designation, which indicates the physical construction phase of the cleanup process has been completed and immediate threats have been addressed, is likely to raise housing values. In the case the market is forward looking and treats listing as indicative that the site will be cleaned, the additional appreciation experienced at the construction complete and deletion milestones would be smaller than it would be otherwise.

2.1 Previous Studies on Valuing Superfund Benefits

The large literature that seeks to measure the value of Superfund site remediation has been exhaustively reviewed in Schultze et al. (1995), Kiel and Williams (2007), Sigman (2008), EPA (2009) and Sigman and Stafford (2011). We briefly describe the hedonic approach that examines median housing values in locations that vary in the number or characterization of sites

⁹ Davis (2004) finds that information on health risks are capitalized into housing values – i.e. the emergence of a cancer cluster resulted in the depreciation in housing values in a Nevada county relative to those in a nearby county.

¹⁰ Messer et al. (2006) note that “when residents or potential buyers are extraordinarily fearful of a site, they may respond by shunning the site... If risks are perceived as being excessive, people replace calculations of risk versus benefit with a simple heuristic of shunning, the avoidance of the stigmatized object.”

contained within.¹¹ Greenberg and Hughes (1992) study seventy-seven communities in New Jersey and find that sale prices of houses in Superfund communities appreciate by less than those in non-Superfund communities. Noonan, Krupka and Baden (2007) study the effect of Superfund remediation activities on housing values measured at the block-group level using a national sample, and employ an instrumental variables approach to separate direct and indirect effects of cleanup. Their comparison of those block groups that are close to waste sites with other block groups across the contiguous US, however, could lead to bias because unobservables are likely to differ systematically across these two sets of block groups.¹²

We build most directly upon Greenstone and Gallagher (2008) (hereafter GG) who examine how tract *median* housing prices vary depending upon whether they contain a site that has been listed on the NPL or one that has narrowly missed being listed on the NPL. GG's important methodological contribution is to restrict their comparison to sites that are similar in their risk scores, but that differ in their Superfund status. As described in GG, in the early years of the Superfund program, budget constraints forced the EPA to choose only 400 sites to list on the NPL (out of 690 potential sites that the EPA had identified). The EPA employed the HRS ranking to choose those sites that posed the greatest risks. It turned out that an HRS score of 28.5, as recorded in 1982, served as the cutoff between the 400th listed and 401st non-listed sites. GG argues that the comparison should be made among (i) sites with 1982 HRS scores; and (ii) sites whose 1982 HRS scores are 12 points above or below the 28.5 regulatory cutoff. Their Regression Discontinuity (RD) analysis exploits the dichotomous treatment (listing versus non-

¹¹ A second approach takes a particular site and determines how distance from it impacts the selling price of nearby homes. That effect is measured with a distance gradient that typically varies with site status (Kiel and Zabel, 2001; Kiel and Williams, 2007).

¹² Table II in Greenstone and Gallagher (2008) shows that tracts that host and do not host waste sites differ significantly in their observables and by extension, are likely to differ in their unobservables.

listing) at the 28.5 regulatory cutoff, while assuming that the unobservables are continuous across that cutoff.

GG's analysis examines (i) 487 sites out of 690 sites with 1982 HRS scores and (ii) 227 sites out of 332 sites as their units of observations. They drop 203 out of the 690 sites and 95 out of 332 sites, respectively.¹³ GG's instrumental variable strategy, which compares sites that are listed versus sites that narrowly missed listing, concludes that *cleanup* of Superfund sites has little and no effect on median housing values. One important drawback in that study is its examination of "listing" as the milestone to capture the benefits of cleanup instead of deletion, which is the milestone that marks the completion of cleanup activities. Their comparison of "listed sites" (which combines listed and deleted sites) with sites that have not yet reached the listing designation leads a downward bias because listing has ambiguous overall effects on housing prices, while deletion is likely to raise housing prices (Smith 2006).¹⁴ In contrast to their approach, our study estimates the effect of deletion from the NPL to capture the benefits from remediation, and we measure the effect of deletion separately from other Superfund milestones.¹⁵

¹³ GG's specification relates 2000 prices to listing status in 2000, with 1980 covariates as explanatory variables. The 1980 covariates are unavailable for tracts associated with these 95 sites in the RD sample. Covariates also are unavailable for tracts associated with the 203 sites in the 1982 HRS sample (GG 2008).

¹⁴ This combination is necessitated by their instrument. GG use the 1982 HRS score to instrument for the variable indicating that a site has been listed on (or deleted from) the NPL by 2000; that one variable cannot separately instrument for the two milestones of listing and deletion.

¹⁵ Other studies have measured the distinct effects of these various milestones (Kiel and Zabel 2001; Cameron and McConaha 2006; Kiel and Williams 2007) or treated these milestones as distinct (Sigman 2001).

3 Estimation Method

Our approach to measuring *localized* benefits is three-fold. First, we demonstrate a refinement of the hedonic method that is aimed at providing more accurate estimates of localized benefits when analysts are restricted to using publicly available tract-level (or other similarly geographically coarse) data. We examine numerous points along the within-tract distribution of housing values (including, but not limited to, the median) in order to measure the heterogeneous appreciation of housing values in response to cleanup. Second, we recover these benefits directly using high geographic resolution data measured at the census *block level*. The block analysis reveals that benefits from cleanup are sizable but highly localized. Finally, we provide supplementary analysis using geo-coded house-level data to document the spatial pattern of housing values within tracts and their proximity to Superfund sites. Our tract-level finding that cleanup causes greater appreciation at the lower percentiles of the within-tract house value distribution is consistent with our finding that Superfund sites are in closer proximity to the lower-value houses within each tract.

To identify the effect of Superfund milestones on housing prices, we rely on two complementary strategies: (i) we restrict our comparison to sites whose 1982 HRS scores are within a narrow interval, as in GG; and (ii) we rely on a panel model to examine how the changes in the exposure of neighborhoods to various Superfund milestones affect changes in the housing prices. We restrict our analysis to 321 out of 322 sites whose 1982 HRS scores are within the narrow (16.5 to 40.5) interval (we drop one site for which geocoordinates are unavailable). The progression of these sites through the Superfund milestones is shown in the Online Appendix Figure A1. Our observations are tracts that fall at least partially in a 3 mile buffer around each of these sites (our block-level analysis uses all blocks contained in these

tracts). Our choice of the 3 mile buffer is based on panel data studies on the association between hazardous waste sites and housing prices that have detected effects at a maximum distance of 2 to 2.5 miles with a mean estimated price effect of 7.4% (reviewed in Jenkins et al. 2006).

Our tract (and block) analyses take snapshots of the NPL status of each site in 1990 and 2000.¹⁶ We compare changes in owner-occupied housing values in census units lying in 3 mile buffers surrounding sites between 1990 and 2000 to changes in exposure to (i) sites that are proposed for the NPL but not listed, (ii) sites that are listed on the NPL but where construction is not yet completed, (iii) sites where construction is completed, but which are not yet deleted, and (iv) sites that are deleted from the NPL. Our study measures the cleanup “treatment” by examining the effect of *deletion from the NPL*. To summarize, we rely on the sample restriction, i.e., comparing neighborhoods near sites that are listed with those near that missed listing, primarily to identify the effect of listing. To identify the effects of construction complete and deletion, we rely on the panel methods to further control for time-invariant unobservables.

4 Regression Models

4.1 Census Tract - Specification

We begin with a basic hedonic regression model relating owner-occupied housing prices to the characteristics of the house and the neighborhood, including exposure to the 331 sites.

$$(1) \quad \ln H_{k,t}^{\theta} = \beta_{1,t}^{\theta} P_{k,t} + \beta_{2,t}^{\theta} L_{k,t} + \beta_{3,t}^{\theta} C_{k,t} + \beta_{4,t}^{\theta} D_{k,t} + \beta_{5,t}^{\theta} X_{k,t} + \nu_k^{\theta} + \varepsilon_{k,t}^{\theta}$$

¹⁶ Ideally, we would examine changes over a long enough time period to detect changes in housing prices, but over a short enough time-period so that parameters of the hedonic price function are stable. Like other decennial census based studies, we are constrained by the decadal frequency in data collection. More frequently collected census data, such as the American Community Survey, is not collected at a sufficient level of geographic density for our analysis.

The subscript k indexes tracts that lie within a 3 mile buffer of a site. A tract is included as long as any part of it falls within the 3 mile buffer. $\ln H_{k,t}^\theta$ is the natural log of the θ^{th} percentile of owner-occupied housing values in tract k in year t ($t = 1990, 2000$). X is a vector containing characteristics of the housing stock along with the socioeconomic and demographic attributes of the tract.¹⁷ These variables and the housing value distributions are summarized in Table 1. v_k^θ are time-invariant tract-level unobservables specific to houses in the θ^{th} percentile, and $\varepsilon_{k,t}^\theta$ is a tract-percentile-year unobservable.

Our main variable of interest is the exposure of the tract in 1990 or 2000 to sites that are deleted by that time period. Other variables of interest are exposure of the tract to sites that are proposed, listed, or where construction has been completed. Exposure is defined as the share of the land area in a tract that falls into 3 mile buffers surrounding sites.¹⁸ Specifically, we first use GIS to draw 3 mile buffers around each site. A tract's exposure to sites at each stage of remediation is then defined as the ratio of its area of overlap with the 3 mile buffers drawn around sites at that stage to its total area.¹⁹ $D_{k,1990}$ represents the exposure of tract k to sites that are deleted by 1 Jan 1990, and $D_{k,2000}$ represents the corresponding measure for 1 Jan 2000.

¹⁷ Our estimates of the benefits from cleanup examine the 'direct effects' on housing values. Covariates control for the other changes that might occur in response to Superfund cleanup that in turn affect housing values. Gamper-Rabindran and Timmins (2011) documents changes in neighborhood socio-demographics associated with Superfund cleanup using block-level data.

¹⁸ Note that EPA defines site location by the geocoordinates of the site's centroid. Sites may vary greatly in size, however, and we would expect the geographic "reach" of larger sites to be greater. Without specific GIS information describing the boundaries of all sites, our best option is to use centroid geocoordinates to indicate location.

¹⁹ Further detail on the calculation of tract exposures, including illustrative maps, is included in Appendix A2 in Gamper-Rabindran, Mastromonaco and Timmins (2011). We also describe in more detail below how we handle situations in which a tract is simultaneously exposed to multiple sites at the same stage of the remediation process.

$P_{k,1990}$ and $P_{k,2000}$ correspond to proposed status; $L_{k,1990}$ and $L_{k,2000}$ correspond to listed status; and; $C_{k,1990}$ and $C_{k,2000}$ correspond to construction complete.

Next, we take the difference between the 1990 and 2000 regression models (restricting parameters to be constant over time), thereby removing the effect of time-invariant tract-percentile unobservables.²⁰

$$(2) \quad \ln H_{k,2000}^{\theta} - \ln H_{k,1990}^{\theta} = \beta_1^{\theta} (P_{k,2000} - P_{k,1990}) + \beta_2^{\theta} (L_{k,2000} - L_{k,1990}) + \\ \beta_3^{\theta} (C_{k,2000} - C_{k,1990}) + \beta_4^{\theta} (D_{k,2000} - D_{k,1990}) + \\ \beta_5^{\theta} (X_{k,2000} - X_{k,1990}) + (\varepsilon_{k,2000}^{\theta} - \varepsilon_{k,1990}^{\theta})$$

The coefficient β_4^{θ} measures the appreciation of house values at the θ^{th} percentile as a result of a one unit (i.e., 0 to 1) increase in exposure of the tract to deleted site(s). Recognizing the log dependent variable, a positive β_4^{θ} indicates that house values appreciate by $100 \left[\exp \left(\beta_4^{\theta} - \frac{1}{2} V(\beta_4^{\theta}) \right) - 1 \right]$ percent as a result of a one unit (i.e., 0 to 1) increase in the exposure to deleted sites (Kennedy 1981). In practice, this transformation has little impact on our block and tract results, so we ignore it in order to simplify the discussion of our estimates. The changes in exposure to (i) proposed sites, (ii) listed sites and (iii) construction completed

²⁰ Our conservative interpretation of the coefficients in the panel analysis is that they measure the capitalization into the housing values resulting from the cleanup (Kuminoff and Pope 2010). Capitalization into housing values is in itself valuable information for policymakers in judging the benefits from Superfund cleanup and affects the local economy including the property tax base. If the coefficients are, in fact, stable over time, the estimates can be further interpreted as measures of willingness-to-pay. Without access to some other form of quasi-experimental variation in the data, it is not possible to test this assumption of stability of coefficients.

Looking across deciles, we assume only that the tract-level unobservable affecting the θ^{th} percentile house in 1990 has to be the same tract-level unobservable affecting the θ^{th} percentile house (whatever house that may be) in 2000. We do not take the restrictive interpretation that the θ^{th} percentile house in 1990 has to be the same θ^{th} percentile house in 2000.

sites similarly capture the changes in house values associated with these steps in the remediation process. Table 2 summarizes the interpretation of all the coefficients.

We weight observations in our preferred specification by the number of owner-occupied housing units in each census tract. Section 6.4 shows that our main conclusions are not sensitive to this decision.

4.2 Census Block - Specification

Our tract analysis contains all tracts that have some overlap with the 3 mile buffer surrounding the sites. Correspondingly, our block analysis examines all blocks contained in these tracts.²¹ The cross-section and panel regression models for census blocks are defined analogously to equations (1) and (2), except that (i) block median values replace within-tract percentiles of the house value distribution, and (ii) exposure is defined by a count of Superfund sites at each stage of remediation lying within 2 mile of the centroid of each block. The exposure variables are counts of sites located less than 2 miles from the centroid of census block k at time t that are proposed ($P_{k,t}$), listed ($L_{k,t}$), construction completed ($C_{k,t}$), and deleted ($D_{k,t}$).

5 Data

Restricted-access census block data for 1990 and 2000 are from the US Census Bureau. Proprietary housing transactions data are from Dataquick Information Systems and are used under a licensing agreement with the Duke Department of Economics. Census tract data come from the Geolytics Neighborhood Change Database, which has reapportioned census data from 1980, 1990 and 2000 into census tract boundaries that are fixed in 2000. The Decennial Census

²¹ We also run a tract analysis which includes only those tracts that lie within a 2 mile buffer of the 321 sites. See section 6.4.

provides counts of houses with owners' stated values in various intervals, allowing us to calculate the discrete distribution of house values within each tract.^{22,23} We use straight lines to connect the midpoints of these intervals portrayed in a cumulative distribution function histogram; we then read the cumulative distribution function of house values in each census tract from those lines. Percentiles read off of these distribution functions are then used as dependent variables in our empirical analysis. Data on sites are from the EPA. The 1982 HRS scores come from the dataset compiled and published by GG. The Consumer Price Index used to deflate housing prices is compiled by the Bureau of Labor Statistics and is based upon a 1982 Base of 100.

6 Results

6.1 Summary Statistics

Table 1 reports summary statistics for census tracts.²⁴

6.2 Tract Results: Evidence of Localized Benefits from Deletion

Observations are weighted by tract counts of owner-occupied housing units, and robust standard errors account for heteroskedasticity. Overall, our tract analysis indicates that the appreciation of housing values varies within the tract, with greater percentage appreciation in the lower tail of the housing price distribution. Table 3 presents results from the tract analysis using 3 mile buffers around the sites. The results indicate that the deletion of a site from the NPL

²² For details on the intervals, see Online Appendix Table A1.

²³ We only consider the effect of NPL site remediation on residential property values. See Ihlanfeldt and Taylor (2004) for an analysis of the effects on commercial real estate values.

²⁴ Similar summary statistics for the block sample were not released by the Census Bureau due to confidentiality concerns.

raises nearby housing values, but the appreciation, in percentage terms, is more prominent at the lower deciles of the within-tract housing value distribution. As seen in Panel A, carrying a site through the remediation process to deletion raises house values by 24.5% at the 10th percentile, 20.3% at the median, and 18.6% at the 90th percentile. All estimates are statistically significant at the 1% level. The pattern of declining point estimates from the lower to the higher percentiles of housing values is evident from Panel A, although we note that the standard errors indicate that these estimates are not statistically different from one another at the 95% level.

Panel B presents the results using housing value levels as the dependent variable. Appreciation attributable to deletion increases as one moves from the 10th percentile (\$9,240) to the 90th percentile (\$17,553); as shown in Panel A, however, these increases are not rapid enough to prevent appreciation as a percentage of housing value from falling across the percentiles. Interestingly, even in levels, appreciation attributable to listing and construction complete falls as one moves from the 10th to the 90th percentile, and is not significant above the 10th percentile in the case of listing, or the 30th percentile in the case of construction complete.

6.3 Other Superfund Milestones

For all milestones, we find heterogeneity in the within-tract house value distribution. In particular, the magnitudes of changes in housing values are largest for the lower tails of the distributions. This pattern is consistent with our finding in section 6.8 that lower-value houses tend to be located closer to Superfund sites and are therefore more impacted as those sites progress through the Superfund milestones. We next consider each of the individual milestones preceding deletion – proposal, listing, and construction completion – in turn.

We find that proposal to the NPL leads to a statistically significant and sizable depreciation in housing values at the lower percentiles of the within-tract house value distribution. In particular, proposal to the NPL reduces housing values by 12.7% and 12.4% at the 10th and 20th percentiles, respectively.²⁵ Importantly, estimates at the 30th percentile and higher are not statistically significant. Depreciation in nearby housing values in response to the proposal of a site to the NPL can be explained by two channels. First, the proposal of the site provides new information to the market about the presence of a harmful site, or about its severity. Second, even if the market is already aware of the site and the extent of contamination, the proposal of the site to the NPL may further decrease housing values by stigmatizing the neighborhood (Messer et al. 2006).

Conversely, listing to the NPL leads to a sizable and statistically significant appreciation in housing values at the lower percentiles of the within-tract house value distribution. Specifically, listing on the NPL raises housing values by 9% to 5.6% between the 10th and 40th percentiles. Estimates at and above the median are not statistically significant. The smaller magnitude of appreciation from listing compared to deletion can be explained by the countervailing pressures on housing values that arise when a site is listed – i.e., listing reduces housing values by confirming the severe nature of site contamination, but it also increases housing values by signaling that the site will be remediated. Nevertheless, the promise of cleanup associated with final listing appears to outweigh the effect of confirming a site's

²⁵ To be clear, even though proposal leads to depreciation in housing values, the Superfund remediation process, taken in its entirety, leads to an overall appreciation in housing values even at the bottom of the within-tract house value distribution. As described in section 6.2, our estimated coefficients on deletion, which measure the effect of deletion on housing values relative to values at the pre-proposal stage, indicate that the Superfund remediation process, taken in its entirety, leads to an overall appreciation in housing values.

contamination level at the bottom of the within-tract housing value distribution. This latter result could be partly responsible for the finding in previous research that combined listing-deletion treatment did not have a significant effect on housing values at the median – we return to this point in section 7.

Achieving the construction complete milestone leads to additional appreciation in housing values. These effects are statistically significant at all parts of the within-tract house value distribution, but are larger at lower percentiles – 13.8% at the 10th percentile, 11.3% at the median and 6.9% at the 90th percentile. As might be expected, completion of construction leads to a smaller appreciation in housing values than does deletion. This can be explained by the additional value the market places on moving the site from the completion of the physical infrastructure required for the cleanup to the stage where all remedial actions have been completed.

6.4 Deletion of sites from the NPL – Sensitivity Analysis

First, we estimate an unweighted specification. Results from those regressions, presented in Table 4 columns 1-5 are comparable to those from the weighted regressions. We continue to find that carrying a site through the remediation process to deletion leads to larger percentage appreciation at the lower deciles of housing values than at the upper deciles. In particular, the unweighted regressions indicate that house values appreciate by 30.7% at the 10th percentile, 24.0% at the median and 19.5% at the 90th percentile.

Second, to address the possibility that errors are spatially correlated, we estimate clustered standard errors. Ideally, we would cluster the standard errors over groups of tracts that lie in close proximity, as these errors are likely to be spatially correlated. Practically, because

the next available level of geographical identifier is the county, we estimate standard errors clustered at the county level. The drawback of this approach is that clustering on too aggregate a geographical region will lead to overly large standard error estimates. As seen in Table 4 columns 6-10, we continue to find statistically significant appreciation in response to deletion with clustering, although the estimates are now statistically significant at the 5% level for the 10th percentile of housing values and at the 10% level for the rest of the deciles of housing values.

Third, to explore the spatial extent of the effects of deletion, we repeat our analysis using 2 mile buffers. Table 5 reveals that using the narrow definition of neighborhoods near sites yields larger point estimates of appreciation in response to deletion. House values appreciate by 26.2 % at the 10th percentile, 25.9 % at the median and 20.8% at the 90th percentile. Comparison of these results with our earlier results from Table 3, where neighborhoods are defined more broadly using 3 mile buffers, suggests that the larger neighborhood lumps nearby affected houses with more distant unaffected houses, thereby diluting the effects of deletion.

6.5 Sensitivity Analysis: Exposure to Overlapping Site Buffers

Consider a tract that can be divided into three sections. The first is exposed only to the buffer surrounding site A ($x\%$ of the area of the tract), the second is exposed only to the buffer surrounding site B ($y\%$), and the third is exposed to buffers surrounding both sites A and B ($z\%$). To fix ideas, suppose this tract has progressed from proposed status in 1990 to deleted status in 2000. Based on our definition of exposure, the tract's exposure to deletion will be $(x+y+z)\%$. This definition treats the exposure of $y\%$ of the tract similarly whether that section of the tract is exposed to one or more site.

In order to check the implications of this assumption, we limit our analysis to tracts that are exposed to the 2 or 3 mile buffer of only one site (i.e., we drop all tracts that are exposed to two or more sites). For the 3 mile sample, we go from having 321 sites to 299 sites, and our observations decline from 3,584 tracts to 2,993 tracts. We find larger impacts of deletion in the restricted sample (Table A2) than in the original sample (Table 3). Deletion leads to appreciation of 32.7%, 27.8% and 21.1% at the 10th, 50th and 90th percentile in the restricted sample. In contrast, deletion leads to appreciation of 24.5%, 20.3% and 18.6% at the 10th, 50th and 90th percentile in the original sample.

6.6 Block Results: Direct Evidence of Localized Effects

The analysis of median housing values at the block level provides direct evidence for the localized benefits from Superfund cleanup. These results are described in Table 6. Taking a site through the remediation process from proposal to deletion results in statistically significant appreciation of median house values by 14.7% in blocks lying less than 3 miles from the site. Moreover, this estimate continues to be statistically significant at conventional levels when the standard errors are clustered at the tract-level. These block results complement our tract results that deletion from the NPL raises housing values by 24.5% at the 10th percentile, 20.3% at the median, and 18.6% at the 90th percentile of tract-level housing values. We attribute the smaller magnitudes to the different definition of exposure used in the block-level analysis. In particular, in the tract-level analysis, variation in the exposure to deleted sites is captured by the ratio of the area of the tract that overlaps with the 3 mile buffer surrounding the deleted sites to the total area of the tract. At the block level, exposure of the 59,055 block observations to deleted sites is measured using counts of sites located within 2 miles from the block centroid.

Looking at the other milestones, the block results indicate a comparable level of depreciation (-8.6%) associated with proposal. They also indicate a sizable appreciation in housing values relative to the pre-proposal stage occurs when the site is listed, with only small additional appreciation occurring at construction complete and deletion.²⁶ In contrast, when we carried out the tract-level analysis, we found that a sizable appreciation in housing values occurred at listing for houses in the lower percentiles, but that there was appreciation that occurred at construction complete for all percentiles. Compared with the tract-level results, the block-level results therefore suggest that the market is more forward looking in treating the listing of a site on the NPL as a strong indication that the site will be cleaned. Again, these differences may be ascribed to differences in the definitions of exposure used in the two analyses.

6.7 Other Potential Estimation Issues

We discuss two potential estimation issues that, upon further examination, are not likely to negate our inference in the tract and block analyses. First, identifying the effect of deletion will be made more difficult if sites that are eventually deleted differ systematically from sites that do not reach this milestone in our study period. However, Sigman's (2001) study of the pace of progress at Superfund sites suggests that the extent of bias on our estimate of deletion is likely to be limited. In particular, our panel approach deals with time-invariant unobservable variables. The following variables are modeled as time-invariant in Sigman (2001): socioeconomic characteristics, voter turnout, the technical complexity of the cleanup, and the

²⁶ From our tests of the equality of coefficients, we are able to conclude that the estimates for listing and construction complete are statistically different, as are the estimates for construction complete and deletion.

presence of potentially responsible parties (PRPs). Our panel approach is not able to control for time-varying correlated unobservables, but factors modeled as time-varying in Sigman's study are found to have limited or no influence on sites' progress. Public funding, for example, does not influence the progress from listing to the Record of Decision (ROD), and legislative influence does not affect the sites' progress from listing through ROD to construction complete. While public funding does influence the pace of progress from ROD to listing, Sigman (2001) notes that most funding for cleanup at this stage comes from PRPs under their agreement with the EPA.

A second related issue is whether sites that receive the cleanup treatment are likely to have received systematically more intensive cleanup than comparison sites that have yet to receive cleanup. If this concern is valid, then our estimates are larger on average than those that would be realized by the cleanup of other sites in general. Our estimation strategy addresses this issue using fixed effects to control for time-invariant unobservables, and by comparing sites that are similar in terms of their 1982 HRS scores (which reduces the possibility for variation in the extent of cleanup to arise from time-varying unobservables). Moreover, previous studies suggest that the extent of cleanup does not vary systematically with observed neighborhood characteristics. For instance, the EPA did not choose less permanent cleanup options for sites with lower median household income or with greater shares of non-white residents at the zip-code level (Gupta et al. 1996). Similarly, the expenditure to avert an average cancer case in NPL sites was not influenced by mean income or minority population within a 1-mile ring of NPL sites; among the less hazardous sites, however, variation can arise from constituents' political activity (Hamilton and Viscusi 1999). Hamilton and Viscusi (1999) note that although EPA's directive set a baseline for cleanup standards, cleanup is set at more stringent levels in

states with stricter standards.²⁷ However, the state-level source of variation in the extent of cleanup does not bias our study because we do not systematically compare cleanup in sites located more stringent states relative to sites yet to be cleaned located less stringent states.

6.8 Supporting Evidence for Tract Analysis From House-Level Data

Our tract results are consistent with the observation that NPL sites are located closer to the lower-value houses within each tract. We provide direct evidence for this spatial distribution using geo-coded transactions data from Dataquick Information Systems drawn from ten states.²⁸ Table 7 summarizes the distribution of housing data and Superfund sites by states and site status. Superfund sites included in this analysis are sites that were scored in 1982 that had the same status in both 1990 and 2000. We have 68 sites that never reached the proposal stage, 94 sites that were listed in both years, and 2 sites that were deleted in both years. Considering only sites where status did not change over the course of that decade, it is a simple matter to assign each house in the tracts surrounding these sites to deciles of the within-tract price distribution without worrying that changing house prices (caused by changing site status) would alter that assignment.

We begin by taking all houses that transacted during the period 1990 - 2000 in the 3 mile buffer surrounding each site. We subtract the mean of the prices of all houses sold in a particular

²⁷ Viscusi and Hamilton (1999) provide details on this point. The 1991 EPA directive set a baseline of cleanup standards – the cumulative carcinogenic site risk to an individual based on reasonable exposure for both current and future land use is less than 10^{-4} and the non-carcinogenic hazard quotient is less than one. In practice, the cleanup goal is more stringent. Variation in environmental cleanup targets can arise from state-level variation in environmental standards. The 1986 Congress directed that remedial actions must meet federal standards and states' standards if stricter.

²⁸ Specifically, these include Arizona, California, Connecticut, Massachusetts, New Jersey, North Carolina, Oregon, Rhode Island, Tennessee and Washington.

tract-year from the price of each house sold in that tract-year. We then pool all of these mean-differenced transaction prices within each tract over the course of the decade. Next, we allocate each house to a decile of the within-tract distribution of mean-differenced transaction prices. Finally, we calculate the percentage of houses in each decile that are within X km of a Superfund site (X = 1, 2, and 3 km). We then normalize by the average probability of exposure in the entire sample (i.e., approximately 5%).

Figures 1 through 4 describe the results of this exercise. Figure 1 uses all sites regardless of status (as long as site status was the same in 1990 and 2000). The first panel shows that houses in the lowest decile are 38.5% more likely than those in the highest decile to be exposed to a site at 1 km. 95% confidence intervals show that this difference is statistically significant. The remaining two panels show that this difference disappears as we consider larger exposure buffers.

Figures 2, 3 and 4 repeat this exercise using only sites that are pre-proposal, listed, or deleted in both 1990 and 2000, respectively. We find patterns similar to those in Figure 1. Considering exposure defined at 1 km, Figure 2 shows that houses in the lowest decile are 26.3% more likely to be exposed than houses in the highest decile. Figure 3 puts this difference at 44.7% for listed sites, and Figure 4 reveals a difference of 46.3% for deleted sites (although this last difference is not statistically significant, owing to the smaller sample size for deleted sites). In all cases, these differences disappear as we consider exposure defined at larger distances.

Together, these figures provide direct evidence that exposure to Superfund sites is heterogeneous within tracts. This explains the patterns revealed by our tract-level results, and suggests that a focus on within-tract medians might therefore be misleading.

7. Comparison of Our Results with Greenstone and Gallagher (2008)

Our conclusions stand in contrast with GG's conclusion that cleanup leads to no or little effect in raising median housing values. Our study differs from GG's in at least two important ways. First, we examine deletion, which signals the end of cleanup, separately from listing; because of this, we rely on panel methods instead of instrumental variables to control for time-invariant unobservables.²⁹ In contrast, GG examine the effect of a variable that combines two distinct milestones in the Superfund remediation process – listing and deletion. This combination, which allows for their instrumental variables strategy, comes at the cost of biasing downward their estimate of cleanup; listing has ambiguous overall effects on housing prices, while deletion is likely to raise housing prices.³⁰ We argue that our approach of measuring cleanup using the deletion milestone (relying on panel variation instead of the GG IV strategy) will incur less bias than GG's approach of measuring cleanup by conflating the deletion and listing milestones.

Indeed, we do not find a statistically significant effect of listing at the within-tract median. Recognizing that there are a significant number of sites that only achieve the listing designation (69) compared with the number achieving construction complete (89) or deleted (57) status by 2000, this could explain the difference in the results of the two studies. In particular, our tract-level analysis reveals that listing on the NPL per se does not lead a statistically significant appreciation at the tract-level *median* housing values. The point estimate of 5.3% is not statistically significant at conventional levels. Instead, the appreciation at the median tract

²⁹ We describe evidence from previous studies as an indirect strategy to address concerns from time-varying unobservables in section 6.7.

³⁰ GG use the 1982 HRS score to instrument for the variable indicating that a site has been listed on (or deleted from) the NPL by 2000; that one variable cannot separately instrument for the two milestones of listing and deletion.

housing values occurs mainly after some cleanup measures are undertaken, both after the completion of physical construction and after the completion of all remedial actions. Relative to pre-proposal baseline, we find tract median housing values appreciate by 11.3% and by 20.3% in response to the construction complete and deletion milestones, respectively (Table 3).

The analysis of listing on the NPL yields two important lessons. First, an analysis of median housing values ignores within-tract heterogeneity and misses the effects at the lower tails of the house price distribution. While a focus on median housing values would have concluded there is no appreciation in response to listing, further examination of the within-tract heterogeneity would have detected statistically significant appreciation in response to listing. The within-tract appreciation amounts to 9%, 7.3%, 6% and 5.6% at the 10th to the 40th percentile of tract-level housing values (Table 3). This result is confirmed by our block-level analysis. Second, the analysis of median tract housing values can fail to detect effects evident in data at finer geographical resolution. As seen in Table 6, median block housing values appreciate by 11.8% in response to listing.

Nevertheless, our study and GG's study reinforce each other in one important way – taken together, they rule out the case that the benefits from a cleanup, measured as capitalization into housing values, are manifest across a large area. Our finding that benefits are highly localized within the tract – relative to the case if benefits were to appear over a larger area – may make it more difficult for the aggregate benefits of a cleanup to exceed the costs. A more detailed comparison of our study and GG's study can be found in the Online Appendix.

8. Cost-Benefit Analysis

We estimate benefits associated with cleanup for the 57 sites that have been deleted from the NPL as of 2000. For each site, we consider tracts within the 3 mile buffer drawn around the site. The benefit per tract for houses at the θ^{th} percentile is obtained by multiplying the share of the tract exposed to the deleted sites, the deletion coefficient for θ^{th} percentile (from Table 3), and one-tenth of the number of owner occupied housing units in the tract. We then sum the benefits within each tract for houses from the 10th to the 90th percentiles. Finally, we sum the benefits for all tracts within the 3 mile buffer of a given site.

For comparison, we use actual cleanup costs by site published by GG (2005), based on their “best effort to calculate the actual amount spent on remedial action at each site by the EPA, state governments, and responsible parties.”³¹ GG note that their cost figures may not account for all state costs, and hence may be underestimates. (GG 2005, Data Appendix) Still, they are the best cost estimates that we were able to obtain. Next, we compare the estimated benefits and actual cleanup costs by site, for the 55 out of 57 sites for which costs data are available.

The estimated benefits and the actual cleanup costs across these sites show considerable heterogeneity. The mean and standard deviation of the estimated benefits are \$72 million and \$121 million respectively; while the mean and standard deviation for costs are \$9 million and \$13 million respectively. The histogram of the net estimated net benefits plotted in Figure 5. For 35 out of 55 sites, we find that cleanup yields positive net benefits. These results suggest that for the majority of the deleted sites, the Superfund program does pass a simple cost-benefit

³¹ GG reports that these costs are “the sum of the EPA actual costs and the PRP estimated costs. Direct costs include remedial action and operations and management costs. Indirect costs are the EPA’s estimate of the portion of the Superfund program costs (personnel wages, travel costs to inspect the sites, etc.) that are attributed to each site. These are EPA estimates for additional Potential Responsible Party costs.” (GG 2005, Data Appendix)

test, where benefits are estimated from the appreciation in housing values. We note several outliers with very large net benefits, which may be implausibly large. These outliers may have arisen from our approach of estimating the average treatment effects of cleanup across sites, suggesting the need for future work to examine the heterogeneity of cleanup effects across sites.

9. Conclusion

Our study shows that benefits from Superfund remediation activities can be highly localized. Our proposed refinement to the hedonic method – i.e., consideration of the entire distribution of the housing values within the tract – enables the possible detection of localized benefits at the sub-tract level using tract-level data. When we apply this refined method to the evaluation of the benefits from Superfund remediation, we detect significant within-tract heterogeneity. Cleanup causes greater appreciation (in percentage terms) of housing prices at the lower deciles – i.e., by 24.5% at the 10th percentile, 20.3% at the median, and 18.6% at the 90th percentile of the within tract distribution. Restricted access block data, observed at a fine level of geographical resolution, confirms this point; in particular, we find that owner-occupied housing values appreciate by 14.7% for blocks lying less than 3 miles from remediated sites.

The pattern of within-tract heterogeneity is most evident in the response to the listing on the NPL, the intermediate Superfund milestone which indicates the site will be cleaned at some point in the future. Listing on the NPL results in appreciation of 9% to 5.6% at the 10th to the 40th percentiles of housing values, but estimates at the median and higher percentiles are not statistically significant. The analysis of listing using median tract housing values can therefore fail to detect effects evident in data at finer geographical resolution. As seen in Table 6, median block housing values appreciate by 11.8% in response to listing. One can imagine other

situations in which the distribution of benefits is such that a focus on the mean or median could lead to a failure to detect any treatment impacts, should those impacts exist only in the tails of the distribution of housing values.

Further investigation reveals that within-tract heterogeneity is partly explained by the spatial distribution of Superfund sites. Geo-coded housing transactions data from ten states provide direct evidence of the proximity of these sites to the cheaper houses within each tract. Finally, our cost-benefit analysis indicates that cleanup under the Superfund program yields positive net benefits for 35 out of 55 sites that have been deleted from the NPL by 2000.

The localized nature of these benefits (e.g., at the sub-tract level) has important methodological implications for analysts who are forced to rely on coarse-resolution, publicly available (e.g. tract-level) data. In particular, the analyst must consider heterogeneity within those units, paying particular attention to the tails of the housing value distribution. Otherwise, the standard hedonic approach of focusing on *median* housing values may understate or fail to detect these benefits altogether.

Our proposed method has its limits. Not all localized benefits detectable with finer resolution block- and house-level data may be found by an analysis of the tract-level housing price distribution. Nor can analyses like that conducted here detect heterogeneity in the valuation of cleanup across housing markets that may be evident in house-level data (Kiel and Williams 2007). However, given the reality that many hedonic studies are forced rely on tract, zip-code, or county-level data (the advantages of which include accessibility and nationwide coverage), our extension avoids an important source of bias that results from a narrow focus on mean or median housing values.

Bibliography

Baum-Snow, Nathaniel, and Matthew Kahn, "The Effects of New Public Projects to Expand Urban Rail Transit," *Journal of Public Economics*, 77 (2000), 241-263.

Bishop, Kelly and Christopher Timmins, "Hedonic Prices and Implicit Markets: Consistent Estimation of Marginal Willingness to Pay for Differentiated Products Without Exclusion Restrictions," Nicholas Institute Working Paper, Duke University, 2011.

Black, Sandra E., "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*, 114 (1999), 577-599.

Bui, Linda T. M and Christopher J. Mayer, "Regulation and Capitalization of Environmental Amenities: Evidence from the Toxic Release Inventory in Massachusetts," *The Review of Economics and Statistics*, 85(3) (2003), 693-708.

Cameron, Trudy Ann, and Ian T. McConnaha, "Evidence of Environmental Migration," *Land Economics*, 82 (2006), 273-290.

Campbell, John Y., Stefano Giglio, and Parag Pathak, "Forced Sales and House Prices." *American Economic Review*, 101(5) (2011), 2108–31.

Chay, Kenneth Y., and Michael Greenstone, "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 113 (2005), 376-424.

Currie, Janet, Michael Greenstone and Enrico Moretti, Superfund Cleanups and Infant Health *American Economic Review*, 101(3) (2011), 435-441.

Davis, Lucas W., "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster," *American Economic Review*, 94 (2004), 1693–1704.

Davis, Lucas W., "The Effect of Power Plants on Local Housing Values and Rents," *Review of Economics and Statistics*, 93(4) (2011), 1391–1402.

Deyak, T.A. and V.K. Smith, "Residential Property Values and Air Pollution: Some New Evidence," *Quarterly Review of Economics and Business*, 14(4) (1974), 93-100.

DiPasquale, Denise and C. Tsurriel, "Do House Price Indices Based on Transacting Units Represent the Entire Stock? Evidence from the American Housing Survey," *Journal of Housing Economics* 4 (1995): 195-229.

Environmental Protection Agency, "Promoting Responsible Redevelopment at the Empire Canyon Daly West Mine Superfund site," Publication 330-B-08-001 (2008).

Environmental Protection Agency, "Challenges in Applying Property Value Studies to Assess the Benefits of the Superfund Program," Office of Superfund Remediation and Technology Innovation (2009).

Environmental Protection Agency, “Advisory on Superfund Benefits Analysis” Scientific Advisory Body Superfund Benefits Analysis Advisory Panel, Publication EPA-SAB-ADV-06-002, 2006.

Fischhoff, Baruch, “Defining Stigma,” in *Risk, Media, and Stigma*, James Flynn, Paul Slovic and Howard Kunreuther, eds. (Sterling, VA: Earthscan, 2001).

Gamper-Rabindran, Shanti and Christopher Timmins, “Hazardous Waste Cleanup, Neighborhood Gentrification, and Environmental Justice: Evidence from Restricted Access Census Block Data,” *American Economic Review Papers and Proceedings*, May 2011.

Gamper-Rabindran, Shanti, Ralph Mastromonaco, and Christopher Timmins, “Valuing the Benefits of Superfund Site Remediation: Three Approaches to Measuring Localized Externalities,” National Bureau of Economic Research Working Paper 16655, 2011.

Gayer, Ted, James T. Hamilton, and W. Kip Viscusi, “Private Values of Risk Tradeoffs at Superfund Sites: Housing Market Evidence on Learning about Risk,” *Review of Economics and Statistics*, 82 (2000), 439-451.

General Accounting Office, “Superfund: Funding and Reported Costs of Enforcement and Administration Activities,” Publication GAO-08-841R (2008).

General Accounting Office, “Superfund: EPA's Estimated Costs to Remediate Existing Sites Exceed Current Funding Levels, and More Sites Are Expected to Be Added to the National Priorities List,” Publication GAO-10-380 (2010).

Grainger, Corbett A., “Do Renters Fully Pay for Cleaner Air?” Department of Agriculture and Resource Economics, University of Wisconsin-Madison, Working Paper (2010).

Greenberg, Michael, and John Hughes, “The Impact of Hazardous Waste Superfund Sites on the Value of Houses Sold in New Jersey,” *The Annals of Regional Science*, 26 (1992), 147-153.

Greenstone, Michael, and Justin Gallagher, “Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program,” NBER Working Paper No. 11790 (2005).

Greenstone, Michael, and Justin Gallagher, “Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program,” *The Quarterly Journal of Economics*, 123 (2008), 951-1003.

Gupta, Shreekant, George Van Houtven, and Maureen Cropper, “Paying for Permanence: An Economic analysis of EPA's Cleanup Decisions at Superfund Sites,” *The Rand Journal of Economics*, 27 (1996), 563-582.

Hamilton, James T. and Kip W. Viscusi (1999). “Are Risk Regulators Rational? Evidence from Hazardous Waste Cleanup Decisions.” *American Economic Review*. 89(4), 1010-1027.

Hanna, Bríd Gleeson, “House Values, Incomes, and Industrial Pollution,” *Journal of Environmental Economics and Management*, 54 (2007), 100-112.

- Harrison, D. and D.L. Rubinfeld, "Hedonic Housing Prices and the Demand for Clean Air," *Journal of Environmental Economics and Management*, 5 (1978), 81-102.
- Ihlanfeldt, Keith R., and Laura O. Taylor, "Estimating the Economic Impacts of Environmentally Contaminated Properties in an Urban Area," *Journal of Environmental Economics and Management*, 47 (2004), 117-139.
- Jenkins, Robin, Elizabeth Kopits, and David Simpson, "Measuring the Social Benefits of EPA Land Cleanup and Reuse Programs," National Center for Environmental Economics, Working Paper No. 06-03, 2006.
- Kiel, Katherine A., and Michael Williams, "The Impact of Superfund Sites on Local Property Values: Are all Sites the Same?" *Journal of Urban Economics*, 61 (2007), 170-192.
- Kiel, Katherine A., and Jeffrey E. Zabel, "The Accuracy of Owner Provided House Values: The 1978-1991 American Housing Survey," *Real Estate Economics*, 27 (1999), 263-298.
- Kiel, Katherine A., and Jeffrey E. Zabel, "The Economic Benefits of Cleaning Superfund Sites," *Journal of Real Estate Finance and Economics*, 22 (2001), 163-184.
- Kennedy, Peter E., "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 71 (1981), 801.
- Kuminoff, Nicolai V., and Jaren C. Pope, "Hedonic Equilibria, Land Value Capitalization, and the Willingness to Pay for Public Goods," Arizona State University, Working Paper, 2010.
- Kuminoff, Nicolai V., V. Kerry Smith and Christopher Timmins (2010). "The New Economics of Equilibrium Sorting and Its Transformational Role for Policy Evaluation," NBER Working Paper #16349, 2010.
- Linden, Leigh L., and Jonah E. Rockoff, "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws," *American Economic Review*, 98 (2008), 1103-1127.
- Mastromonaco, Ralph, "Hazardous Waste Hits Hollywood: Superfund and Housing Prices in Los Angeles," Department of Economics, Duke University working paper (2010).
- Messer, Kent D., William D. Schulze, Katherine F. Hackett, Turdy A. Cameron, and Gary H. McClelland, "Can Stigma Explain Large Property Value Losses? The Psychology and Economics of Superfund," *Environmental and Resource Economics*, 33 (2006), 299-324.
- Nelson, J. P., "The Effects of Mobile-Source Air and Noise Pollution on Residential Property Values," Report to the U. S. Department of Transportation (1975).
- Nelson, J. P., "Residential Choice, Hedonic Prices, and the Demand for Urban Air Quality," *Journal of Urban Economics*. 5 (1978), 357-369.
- Noonan, Douglas S., Douglas J. Krupka, and Brett M. Baden, "Neighborhood Dynamics and Price Effects of Superfund Site Clean-up," *Journal of Regional Science*, 47 (2007), 665-692.

Pope, Jaren C., "Fear of Crime and Housing Prices: Household Reactions to Sex Offender Registries," *Journal of Urban Economics*, 64 (2008), 601-614.

Ridker, R.G. and J.A. Henning, "The Determinants of Residential Property Values With Special Reference to Air Pollution," *Review of Economics and Statistics*, 49 (1967), 246-257.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens III, "Housing Externalities," *Journal of Political Economy*, 118 (2010), 485-535.

Sanders, N., "Toxic Assets: How the Housing Market Responds to Environmental Information Shocks." Stanford Institute for Theoretical Economics Summer Workshop (2011), Stanford University

Schultze, William, Gary McClelland, Ed Balistreri, Rebecca Boyce, Michael Doane, Brian Hurd, and Ronald Simenauw, "An Evaluation of Public Preferences for Superfund Site Cleanup. Volume 1: Preliminary Assessment" Environmental Protection Agency: Office of Policy, Planning, and Evaluation, Publication EE-0255A-01, (1995).

Sigman, Hilary (ed.), *Economics of Hazardous Waste and Contaminated Land*, (Cheltenham, UK: Edward Elgar, 2008).

Sigman, Hilary, "The Pace of Progress at Superfund Sites: Policy Goals and Interest Group Influence," *Journal of Law and Economics*, 44 (2001), 315-344.

Sigman, Hilary and Sarah Stafford, "Management of Hazardous Waste and Contaminated Land," *Annual Review of Resource Economics*, 3(1) (2011), 255-275.

Smith, Kerry V., "Methods for Estimating the Social Benefits of EPA Land Cleanup and Reuse Programs," Summary Report of the Workshop on Methods for Estimating the Social Benefits of EPA Land Clean Up and Reuse Programs, Environmental Protection Agency, Publication EE-0506-01 (2006).

Staiger, Douglas and James H. Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65(1997):557-586.

Zabel, Jeffrey and Dennis Guignet, "A Hedonic Analysis of the Impact of LUST Sites on House Prices in Frederick, Baltimore, and Baltimore City Counties," National Center for Environmental Economics Working Paper #10-01, Environmental Protection Agency (2010).

Table 1: Summary statistics for 3 mile tracts around 321 sites (No. obs.=3,584)

Variable	1990	1990	2000	2000
	Mean	Std Dev	Mean	Std Dev
<u>Housing value distribution</u>				
10 th percentile	\$46,918	\$37,940	\$48,222	\$38,483
20 th percentile	\$56,291	\$43,197	\$57,212	\$45,645
30 th percentile	\$63,344	\$47,136	\$64,178	\$50,251
40 th percentile	\$69,790	\$50,731	\$70,664	\$54,495
50 th percentile	\$76,225	\$54,448	\$77,358	\$59,180
60 th percentile	\$83,310	\$58,729	\$84,651	\$64,370
70 th percentile	\$91,520	\$63,249	\$93,386	\$70,419
80 th percentile	\$102,410	\$69,174	\$105,424	\$78,586
90 th percentile	\$120,717	\$78,996	\$127,079	\$94,576
<u>Share of tract exposed to a Superfund milestone</u>				
Proposal	0.007	0.080	0.006	0.068
Listing	0.316	0.414	0.168	0.343
Construction Complete	0.024	0.145	0.111	0.283
Deletion	0.018	0.122	0.097	0.275
<u>Other variables</u>				
% units occupied	92.2	7.0	92.8	6.7
% owner occupied	64.8	21.9	65.2	22.7
Housing unit density	0.001	0.002	0.001	0.002
Population Density	0.002	0.004	0.002	0.004
% Black	12.3	22.7	14.3	23.1
% Hispanic	6.8	13.3	10.0	16.6
% under 18 years old	25.0	6.2	25.2	6.3
% high school dropout	25.5	13.6	20.4	12.8
% college educated	19.3	14.4	23.1	16.7
% below poverty line	12.6	12.0	12.6	10.9
% public assistance	8.0	7.9	8.7	7.6
% female head of HH	23.866	16.040	25.865	15.729
Mean HH income	\$38,733	\$16,996	\$55,744	\$24,863
% attached homes	7.7	16.7	8.3	16.7
% mobile homes	5.7	11.1	5.4	10.8
% 0-2 bedrooms	28.8	16.7	28.7	16.8
% 3-4 bedrooms	66.4	16.2	66.5	16.3
% units built within 5 years	7.9	10.4	6.8	9.3
% units built within 10 years	13.8	16.2	12.3	14.3
% living in the same house in the last 5 years	55.7	12.7	56.1	12.3

Notes: Housing unit density and population density are in counts per m².

Table 2. Interpretation of Coefficients	
<u>Change in Superfund milestone</u>	<u>Estimated effect</u>
Not Proposed to Proposed	β_1
Not Proposed to Listed	β_2
Not Proposed to Construction Completed	β_3
Not Proposed to Deleted	β_4
Proposed to Listed	$\beta_2 - \beta_1$
Proposed to Construction Completed	$\beta_3 - \beta_1$
Proposed to Deleted	$\beta_4 - \beta_1$
Listed to Construction Completed	$\beta_3 - \beta_2$
Listed to Deleted	$\beta_4 - \beta_2$
No change in status	Omitted case

Notes: The coefficients refer to Equation 2 in Section 4.

Table 3: Panel analysis of tracts within 3 mile buffers of 321 sites whose 1982 HRS scores are in (16.5, 40.5)									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Percentiles	10	20	30	40	50	60	70	80	90
Panel A: Dependent variable: Δ Log price of owner occupied housing units at the θ th percentile									
Δ Proposal	-0.127*	-0.124*	-0.118	-0.110	-0.104	-0.095	-0.076	-0.063	-0.034
	(0.077)	(0.074)	(0.072)	(0.071)	(0.071)	(0.073)	(0.074)	(0.071)	(0.066)
Δ Listing	0.090**	0.073**	0.060*	0.056*	0.053	0.048	0.053	0.043	0.031
	(0.036)	(0.035)	(0.033)	(0.033)	(0.032)	(0.032)	(0.032)	(0.031)	(0.032)
Δ Construction Complete	0.138***	0.136***	0.127***	0.120***	0.113***	0.110***	0.113***	0.098***	0.069**
	(0.036)	(0.035)	(0.035)	(0.035)	(0.034)	(0.033)	(0.033)	(0.031)	(0.031)
Δ Deletion	0.245***	0.217***	0.213***	0.206***	0.203***	0.200***	0.195***	0.185***	0.186***
	(0.041)	(0.040)	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)
R-sqr	0.238	0.278	0.276	0.261	0.258	0.276	0.285	0.285	0.267
Panel B: Dependent variable: Δ Price of owner occupied housing units at the θ th percentile									
Δ Proposal	-2,523	-2,259	-2,541	-3,545	-3,991	-3,910	-3,348	-2,705	-1,169
	(3,190)	(3,528)	(3,749)	(4,021)	(4,423)	(4,798)	(5,375)	(5,581)	(5,790)
Δ Listing	3,517**	2,695	1,922	1,277	1,260	810	1,486	909	-2,468
	(1,572)	(1,792)	(1,973)	(2,204)	(2,351)	(2,504.162)	(2,795)	(2,900)	(3,386)
Δ Construction Complete	5,101***	4,828**	4,623**	3,972	3,936	3,611	4,554	3,820	-985
	(1,607)	(1,874)	(2,095)	(2,436)	(2,602)	(2,729)	(2,879)	(2,992)	(3,471)
Δ Deletion	9,240***	10,226***	11,417***	11,581***	12,292***	12,581***	12,129***	12,824***	17,553***
	(1,764)	(2,049)	(2,235)	(2,581)	(2,759)	(2,916)	(3,728)	(3,764)	(4,447)
R-sqr	0.129	0.170	0.179	0.170	0.171	0.188	0.209	0.226	0.234
Notes: Panel A Columns (1) to (9) represent 9 different regressions of changes in log of housing prices at the θ th percentiles on the change in exposure to Superfund milestones. Panel B Columns (1) to (9) represent 9 other regressions with changes in housing prices as the dependent variable. Exposure to deletion is measured using the ratio of the area of the tract that falls in the 3 mile buffer of deleted sites to the total area of the tract. The change in exposure to deletion is measured between 1990 and 2000. The change in exposure to other Superfund milestones are defined analogously. The control variables (in changes between 1990 and 2000) are listed in Table 1. The regression is weighted using the number of owner-occupied housing units. No obs=3,584. Robust standard errors in parentheses. Estimates are statistically significant at ***1%, **5% and *10%.									

Table 4: Panel analysis of tracts within 3 mile buffers of 321 sites whose 1982 HRS scores are in (16.5, 40.5)										
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Dependent variable: Log price of owner occupied housing units at the 0th percentile										
	Not weighted					Weighted with the no of owner occupied housing units				
	Robust standard errors					Standard errors clustered at the county-level				
Percentiles	10	30	50	70	90	10	30	50	70	90
Proposal	-0.118	-0.119*	-0.097	-0.074	-0.032	-0.127	-0.118	-0.104	-0.076	-0.034
	(0.081)	(0.072)	(0.069)	(0.071)	(0.067)	(0.287)	(0.263)	(0.254)	(0.272)	(0.243)
Listing	0.162***	0.104***	0.093***	0.084**	0.069**	0.090	0.060	0.053	0.053	0.031
	(0.040)	(0.036)	(0.034)	(0.035)	(0.034)	(0.091)	(0.092)	(0.090)	(0.092)	(0.086)
Construction	0.200***	0.171***	0.167***	0.161***	0.122***	0.138	0.127	0.113	0.113	0.069
Complete	(0.040)	(0.036)	(0.035)	(0.035)	(0.034)	(0.103)	(0.108)	(0.107)	(0.106)	(0.095)
Deletion	0.307***	0.246***	0.240***	0.221***	0.195***	0.245**	0.213*	0.203*	0.195*	0.186*
	(0.044)	(0.040)	(0.038)	(0.039)	(0.039)	(0.112)	(0.112)	(0.110)	(0.114)	(0.108)
R-squared	0.143	0.207	0.221	0.234	0.209	0.238	0.276	0.258	0.285	0.267

Notes: Each column represents a separate regression. These regressions are analogous to the main regressions in Table 3. Regression models in Columns 1-5 are not weighted. Regression models in column 6-10 have standard errors clustered at the county-level. No. obs.=3,584. Statistically significant at ***1%; ** 5% and * 10%.

Table 5: Panel analysis of tracts within 2 mile buffers of 321 sites whose 1982 HRS scores are in (16.5, 40.5)									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Percentiles	10	20	30	40	50	60	70	80	90
Dependent variable: Δ Log price of owner occupied housing units at the θ th percentile									
Δ Proposal	-0.248** (0.123)	-0.206* (0.118)	-0.193 (0.118)	-0.177 (0.118)	-0.180 (0.116)	-0.188 (0.118)	-0.169 (0.117)	-0.144 (0.114)	-0.118 (0.113)
Δ Listing	0.079 (0.055)	0.076 (0.053)	0.079 (0.051)	0.082 (0.052)	0.079 (0.051)	0.069 (0.050)	0.074 (0.049)	0.070 (0.048)	0.048 (0.052)
Δ Construction Complete	0.144** (0.059)	0.152*** (0.057)	0.154*** (0.056)	0.154*** (0.057)	0.148*** (0.055)	0.135** (0.055)	0.136** (0.053)	0.119** (0.052)	0.073 (0.055)
Δ Deletion	0.262*** (0.067)	0.262*** (0.064)	0.264*** (0.062)	0.264*** (0.063)	0.259*** (0.061)	0.244*** (0.060)	0.249*** (0.059)	0.240*** (0.058)	0.208*** (0.060)
R-squared	0.250	0.300	0.298	0.306	0.303	0.297	0.293	0.296	0.265

Notes: The regressions in this tables are analagous to the main regressions in Table 3. However, regressions in this table examine the tracts that overlap with 2 mile buffers around sites whose HRS scores are in (16.5, 40.5). Each column represents a separate regression of the change in the log housing price at the θ th percentiles on the change in exposure to Superfund milestones. The exposure to deletion is measured using the ratio of the area of the tract that falls in the 2 mile buffer of deleted sites to the total area of the tract. The change in exposure to deletion is measured between 1990 and 2000. The change in exposure to other Superfund milestones are measured analogously. The regressions are weighted using the number of owner-occupied housing units. No. obs.= 2,246. Robust standard errors are in parenthesis. Estimates are statistically significant at ***1%, **5% and *10%.

Table 6: Panel analysis of census block median housing values on counts of Superfund sites at various milestones.		
Dependent variable: Change in the log census block median housing price		
Variables of interest:		
Change in the counts of sites that are Proposed	-0.188	***
	(0.021)	
Change in the counts of sites that are Listed	0.118	***
	(0.013)	
Change in the counts of sites that are Construction Completed	0.131	***
	(0.014)	
Change in the counts of sites that are Deletion	0.147	***
	(0.015)	
Number of observations	59,055	
<u>Test statistics that Superfund milestones have equal coefficients</u>		
Coefficients for Proposal and Listing	0.000	
Coefficients for Listing and Construction Complete	0.009	
Coefficients for Construction Complete and Deletion	0.027	
Notes: The observations are blocks that are contained within the tracts that lie within 3 mile buffers around the 321 sites whose 1982 HRS lie within (16.5, 40.5). In other words, the panel analysis examines blocks within tracts that were examined in the tract panel analysis in Table 3. Counts of sites are measured 2 miles from the block centroid. The panel analysis measures changes between 1990 and 2000. Control variables (in changes) are analogous to those listed in Table 1. Robust standard errors are in parenthesis. Statistically significant at ***1%, **5% and *10%.		

Table 7: Housing Transactions and Site Counts by State and Site Status

State	All		Listed		Deleted		Pre-Proposal	
	Houses	Sites	Houses	Sites	Houses	Sites	Houses	Sites
AZ	27,892	5	27,892	5				
CA	77,376	29	55,263	16			22,113	13
CT	29,504	12	16,090	8			13,414	4
MA	68,231	15	46,865	10			21,366	5
NJ	65,459	62	40,647	37	781	1	24,031	24
NC	1,634	2					1,634	2
OR	13,723	6	5,108	1			8,615	5
RI	36,346	13	18,162	8			18,184	5
TN	21,739	7	4,179	1			17,560	6
WA	78,067	13	44,794	8	1,296	1	31,977	4
Total	419,971	164	259,000	94	2,077	2	158,894	68

Figure 1: Housing Transaction Exposure Measures by Decile (Pooled Sites)

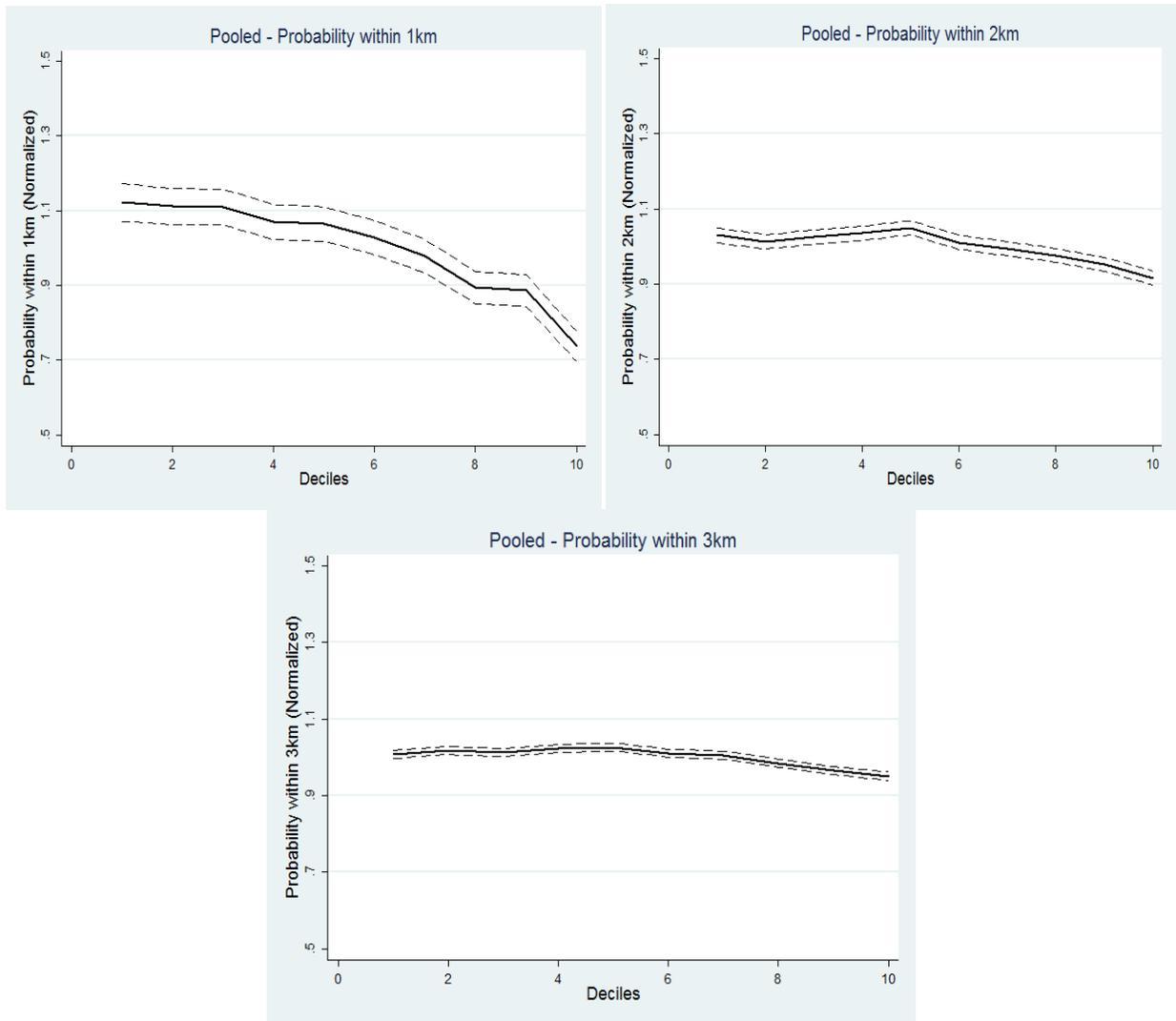


Figure 2: Housing Transaction Exposure Measures by Decile (Pre-Proposal Sites)

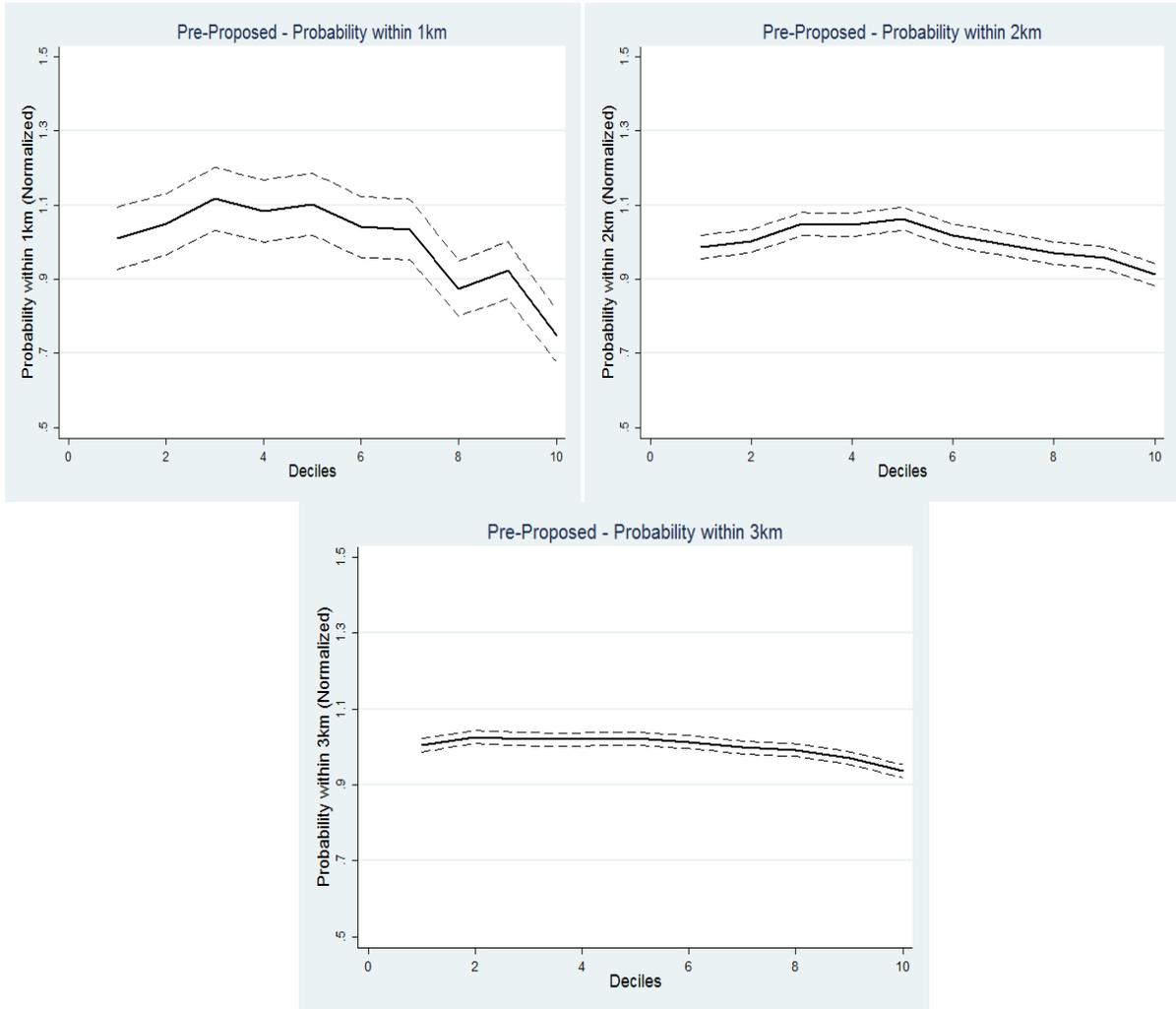


Figure 3: Housing Transaction Exposure Measures by Decile (Listed Sites)

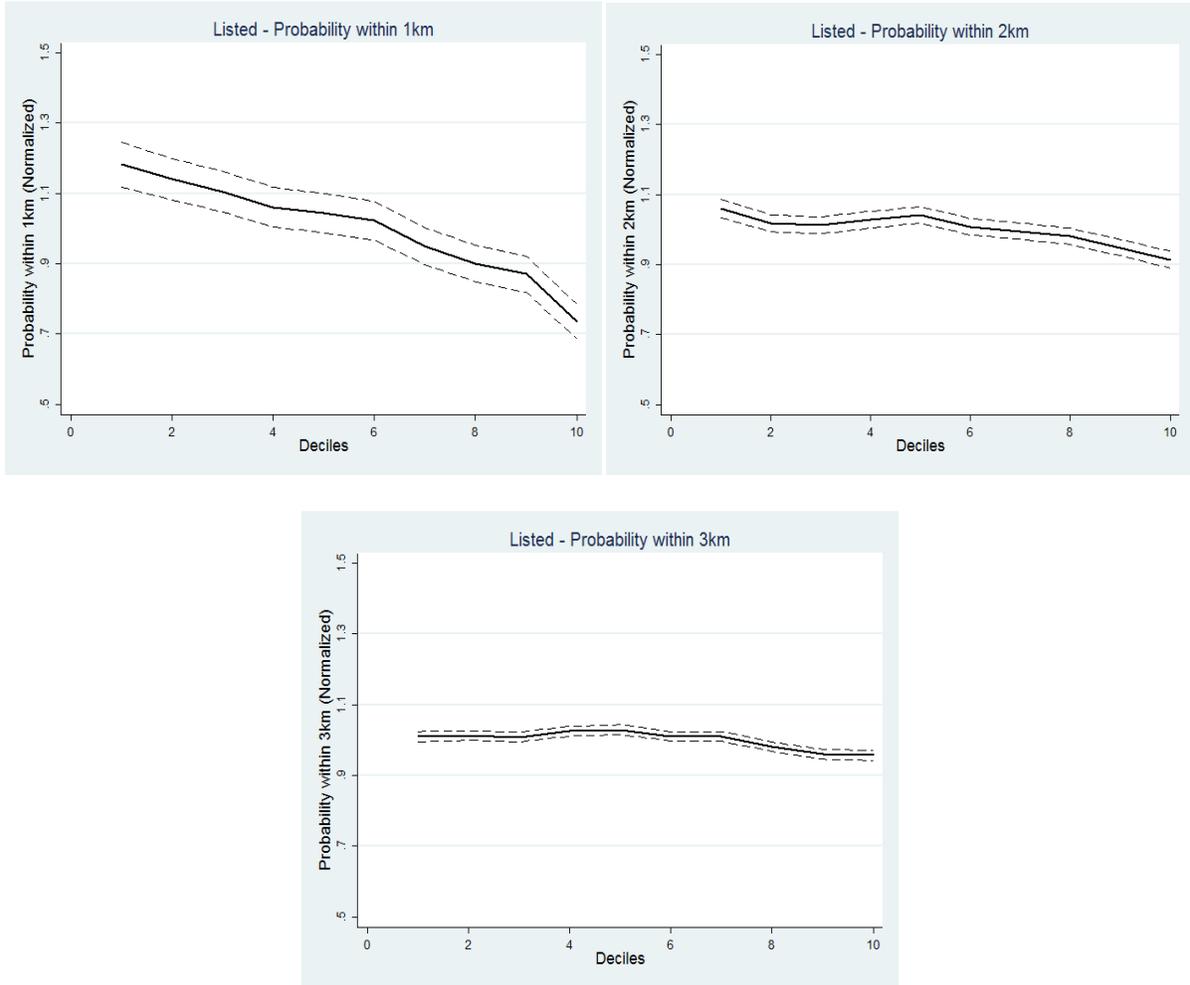


Figure 4: Housing Transaction Exposure Measures by Decile (Deleted Sites)

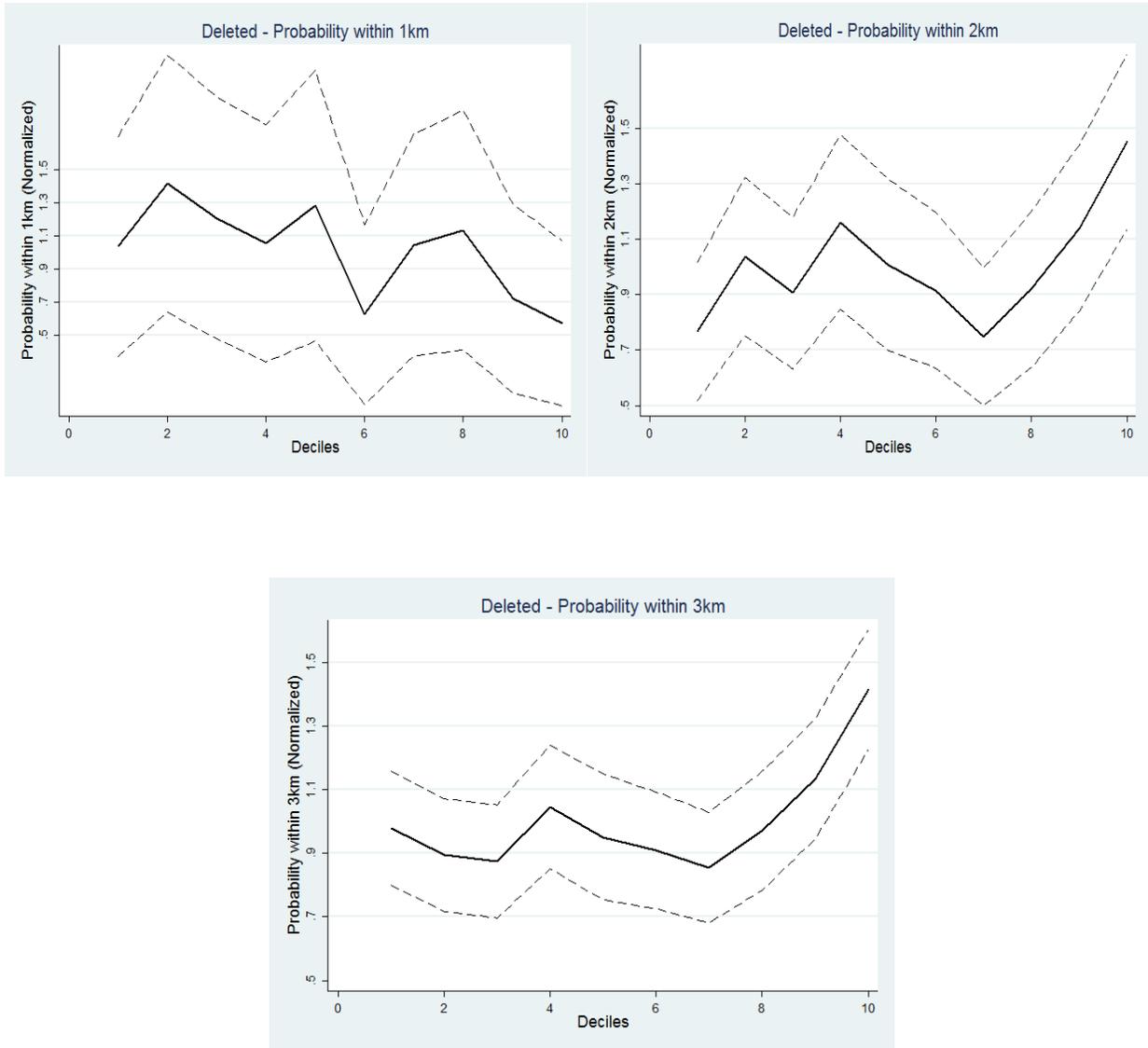
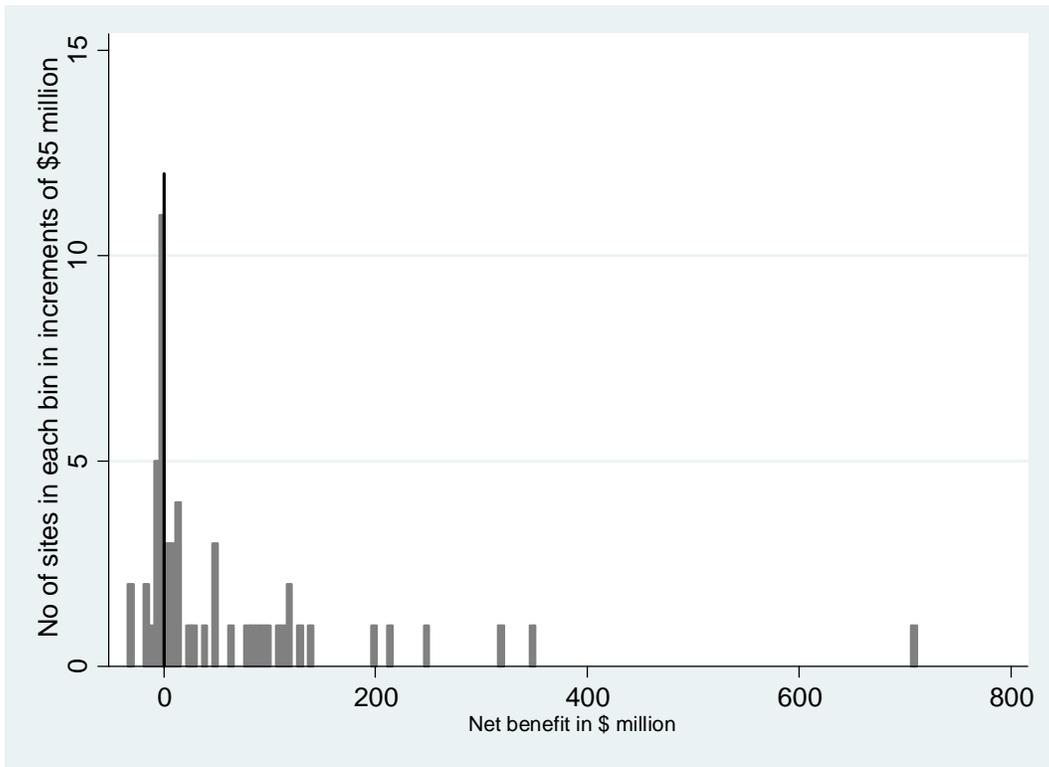
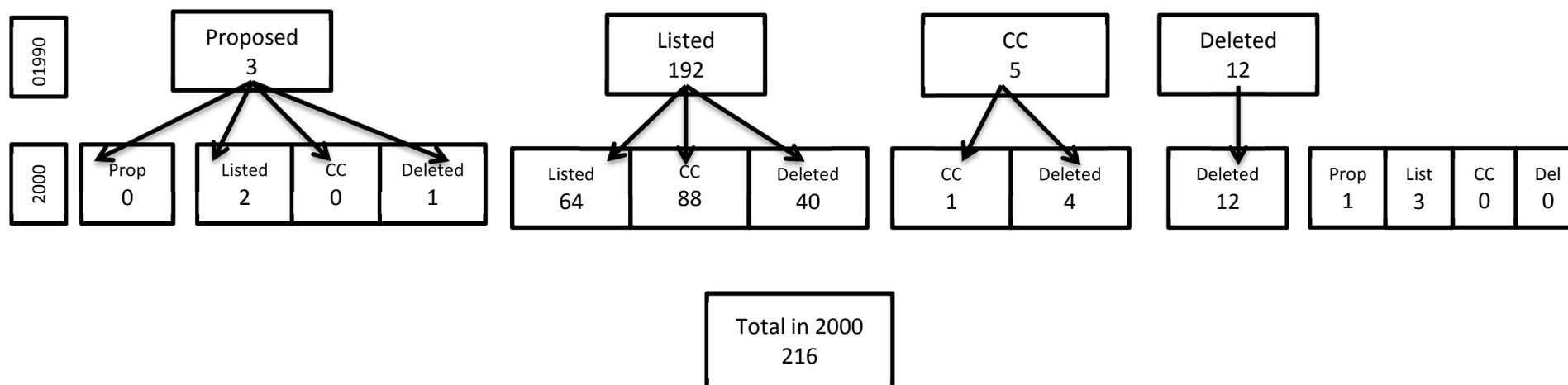


Figure 5: Distribution of net benefits from Superfund cleanup at each of the 55 sites that have been deleted from the NPL



ONLINE APPENDIX

Figure A1: Progression of the 321 sites with HRS scores in the (16.5, 40.5) interval³²



³² The remaining 105 sites do not register any Superfund milestones in 1990 or in 2000.

Table A1: Price Intervals for Counts of Owner Occupied Housing
in thousands of dollars in the 1990 and 2000 Decennial Census

# Categories	Year	
	1990	2000
1	<\$15	<\$10
2	\$15-20	\$10-15
3	\$20-25	\$15-20
4	\$25-30	\$20-25
5	\$30-35	\$25-30
6	\$35-40	\$30-35
7	\$40-45	\$35-40
8	\$45-50	\$40-50
9	\$50-60	\$50-60
10	\$60-75	\$60-70
11	\$75-100	\$70-80
12	\$100-125	\$80-90
13	\$125-150	\$90-100
14	\$150-175	\$100-125
15	\$175-200	\$125-150
16	\$200-250	\$150-175
17	\$250-300	\$175-200
18	\$300-400	\$200-250
19	\$400-500	\$250-300
20	>\$500	\$300-400
21		\$400-\$500
22		\$500-750
23		\$750-1 million
24		>\$1 million

Notes: We calculate the discrete distribution of house values within each tract using counts of houses with owners' stated values in various intervals from the Decennial Census. These values, with additional steps, are used to calculate the log price of the owner-occupied housing units for the θ th percentile in the within-tract housing values distribution.

Table A2: Panel analysis of tracts within 3 mile buffers around 299 sites (i.e., restricted to tracts that are exposed to only 1 site)									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Percentiles	10	20	30	40	50	60	70	80	90
Dependent variable: Δ Log price of owner occupied housing units at the θ th percentile									
Δ Proposal	-0.053 (0.110)	-0.044 (0.106)	-0.028 (0.101)	-0.020 (0.099)	-0.011 (0.099)	-0.002 (0.104)	0.003 (0.108)	0.008 (0.105)	0.034 (0.097)
Δ Listing	0.201*** (0.072)	0.203*** (0.070)	0.181*** (0.066)	0.170*** (0.064)	0.157** (0.064)	0.144** (0.066)	0.133* (0.069)	0.116* (0.068)	0.081 (0.066)
Δ Construction	0.252*** (0.074)	0.269*** (0.072)	0.252*** (0.068)	0.236*** (0.066)	0.218*** (0.066)	0.207*** (0.068)	0.192*** (0.070)	0.168** (0.069)	0.118* (0.068)
Δ Deletion	0.327*** (0.077)	0.317*** (0.075)	0.306*** (0.071)	0.293*** (0.069)	0.278*** (0.069)	0.265*** (0.071)	0.238*** (0.074)	0.223*** (0.073)	0.211*** (0.071)
Observations	2993	2993	2993	2993	2993	2993	2993	2993	2993
R-squared	0.234	0.271	0.266	0.249	0.248	0.267	0.278	0.277	0.261

Notes: Columns (1) to (9) represent 9 different regressions of changes in log of housing prices at the θ th percentiles on the change in exposure to Superfund milestones. Exposure to deletion is measured using the ratio of the area of the of the tract that falls in the 3 mile buffer of deleted sites to the total area of the tract. The change in exposure to deletion is measured between 1990 and 2000. The change in other Superfund milestones are defined analogously. The control variables (in changes between 1990 and 2000) are listed in Table 1. The regression is weighted using the number of owner-occupied housing units. No obs=2,993. Robust standard errors in parentheses. Estimates are statistically significant at ***1%, **5% and *10%.

Detailed Comparison with Greenstone and Gallagher (2008)

A detailed comparison our study and Greenstone and Gallagher (2008) (hereafter GG) is summarized in Table A3. We describe some of these differences below.

(1) Unit of Observation

Our primary analysis examines tracts surrounding sites with HRS 1982 as the observations. In other words, each tract that falls into a 3 mile buffer drawn around each of these sites serves as an observations in our study. In contrast, GG's analysis treats the site as the unit of observation, using some combination of attributes of either (i) own-tract, (ii) adjacent tracts, or (iii) tracts inside 2-mile and 3 mile buffers. The number of observations in GG's analysis therefore corresponds to the number of sites in the sample.

(2) Sites

GG makes the persuasive argument that the comparison should be made among sites that are similar except in their receipt of treatment. The first comparison is among the 690 sites which received 1982 HRS scores – 400 are listed on the NPL, and 290 missed being listed on the NPL. The second comparison is among 332 sites with HRS scores that fall just above and below the 28.5 regulatory cutoff, i.e., (16.5, 40.5). The second comparison narrows the differences across sites other than the Superfund treatment.

Our study examines 331 out of the 332 sites whose HRS scores fall within in (16.5, 40.5) (we drop one site with missing geocoordinates). GG examine 227 out of the 332 sites, dropping 95 sites because of missing 1980 covariates. The detection of appreciation in our tract-analysis and the non-detection in the GG analysis does not stem from the differences in the composition

of sites examined in these two studies. In Table A4, we re-estimate our panel model using the tracts that fall in the 3 mile buffer around these 227 sites. Results indicate that cleanup raised housing values by 25.7% at the 10th percentile, 21.9% at the median and 20.3% at the 90th percentile of tract-level housing values. These results are similar to our earlier results of the tracts that fall in the 3 mile buffer around the 331 sites (Table 3), i.e., cleanup raised housing values by 24.5% at the 10% percentile, 20.3% at the median, and 18.6% at the 90th percentile of tract-level housing values.

The GG model's inability to detect statistically significant estimates of the impact of cleanup may result from their treatment of their observations. In their analysis of tracts in buffers surrounding a site, GG aggregate the observations from the collection of tracts surrounding a site to one observation per site. This reduces the information available to estimate the impact of cleanup.

(3) Model Specification

As described in section 4.1, we begin with the standard hedonic specification that relates tract- or block-level housing values with contemporaneous tract or block attributes, and then differences across the 2000 and 1990 specifications in order to control for time invariant unobservables at the tract or block level. In contrast, GG's regression model relates 2000 tract median housing values to 1980 tract characteristics and 1980 tract median values; they argue that 1980 tract attributes are correlated with the 2000 attributes, but are pre-determined with respect to Superfund site status. Deriving the GG regression specification from 1990 and 2000 hedonic price functions, however, we find that the resulting regression error will be correlated with the

key variables appearing on the right-hand-side of the regression, biasing parameter estimates. To illustrate this, begin with a hedonic price function for each year:³³

$$y_{c2000} = \theta 1(NPL_{c2000} = 1) + X'_{c2000}\beta + u_{c2000} \quad (\text{A1})$$

$$y_{c1980} = \theta 1(NPL_{c1980} = 1) + X'_{c1980}\beta + u_{c1980}$$

Noting that $NPL_{c1980} = 0$, we can multiply the second equation by φ and then difference these two equations to obtain:

$$y_{c2000} - \varphi y_{c1980} = \theta 1(NPL_{c2000} = 1) + (X_{c2000} - \varphi X_{c1980})'\beta + (u_{c2000} - \varphi u_{c1980}) \quad (\text{A2})$$

In order to convert this into the equation estimated by GG, we need to (i) add φy_{c1980} to both sides of the equation, and (ii) move $X'_{c2000}\beta$ into the regression error term:

$$y_{c2000} = \theta 1(NPL_{c2000} = 1) - X'_{c1980}\psi + \varphi y_{c1980} + \underbrace{(X'_{c2000}\beta + u_{c2000} - \varphi u_{c1980})}_{\varepsilon} \quad (\text{A3})$$

where $\psi = \varphi\beta$. Note that the resulting regression error (ε) will contain the unobserved determinants of the 1980 median house value (φu_{c1980}), so that the 1980 median (or mean) housing value will naturally be correlated with it, based on the original hedonic specification.

Because year 2000 covariates are relegated to the regression error, 1980 covariates (X_{c1980}) will likely be correlated with this component of the error term as well. Finally, the main variable of

³³ For this illustration, we use GG's notation, where y_{c2000} refers to the median housing value in census tract c in year 2000, $1(NPL_{c2000} = 1)$ is an indicator taking the value 1 if that tract was exposed to either a listed or deleted site in that year, and X'_{c2000} is a vector of tract attributes).

interest (year 2000 NPL status) will be correlated with the regression error term if it is correlated with year 2000 covariates, which we might also suspect to be the case. GG will therefore need to rely on their regression discontinuity and IV approaches to eliminate any bias stemming from this correlation.

1. Research Question	Measure the effect of deletion relative to the pre-proposal stage.	Measure the effect of listing versus non-listing.
2. Treatment examined	Deletion from the NPL, which signals the end of cleanup.	Listing on the NPL, i.e. an intermediate Superfund milestone.
3. Outcome variable	Distribution of housing values within the tract.	Median tract-level housing values.
4. Unit of observation	Tracts which fall within the buffer around sites.	Sites. Housing and other attributes are (i) tracts on which sites are located, or (ii) average attribute of tracts that overlap with the buffer around the site.
5. Sites in the Regression Discontinuity (RD) sample	321 sites Restrict to 322 sites that received $16.5 \leq 1982\text{-HRS} \leq 40.5$ Drop 1 site without geocoordinates.	227 sites Restrict to 322 sites that received $16.5 \leq 1982\text{-HRS} \leq 40.5$ Drop 95 sites due to missing 1980 covariates.
6. Affected neighborhoods	Buffer of 3 miles and 2 miles.	Buffer of 3 miles and 5 miles.
7. Measure of exposure to treatment	Share of tract that overlap with buffer surrounding NPL sites i.e., whether proposed, listed, or deleted.	
8. Estimation model	Panel analysis (1990 and 2000) to measure separately the milestones of proposal, listing and deletion.	Instrumental variable 1982 HRS serves as the instrumental variable for "Listing on the NPL in 2000".
9. Model specification	Begin with the standard hedonic model relating housing prices to concurrent exposure and other house and neighborhood attributes. Then difference the 2000 and 1990 hedonic equations.	Model relates 2000 prices with Listing on the NPL in 2000 and 1980 covariates, and 1982 HRS serves as the instrument.

Table A4: Panel analysis of tracts within 3 mile buffers of 227 sites (i.e. sites in GG's sample)									
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Percentiles	10	20	30	40	50	60	70	80	90
Dependent variable: Log price of owner occupied housing units at the θ th percentile									
Δ Proposal	-0.128 (0.079)	-0.123 (0.077)	-0.116 (0.075)	-0.109 (0.075)	-0.101 (0.075)	-0.090 (0.076)	-0.071 (0.077)	-0.058 (0.075)	-0.029 (0.070)
Δ Listing	0.095** (0.039)	0.076** (0.037)	0.064* (0.035)	0.059* (0.035)	0.057* (0.034)	0.054 (0.034)	0.059* (0.034)	0.049 (0.032)	0.037 (0.033)
Δ Construction complete	0.154*** (0.039)	0.153*** (0.037)	0.145*** (0.036)	0.136*** (0.036)	0.132*** (0.036)	0.131*** (0.035)	0.135*** (0.034)	0.121*** (0.032)	0.092*** (0.033)
Δ Deletion	0.257*** (0.043)	0.232*** (0.042)	0.228*** (0.040)	0.220*** (0.040)	0.219*** (0.040)	0.219*** (0.039)	0.212*** (0.040)	0.201*** (0.038)	0.203*** (0.037)
R-squared	0.241	0.280	0.277	0.258	0.256	0.275	0.285	0.284	0.264
Notes: Columns (1) to (9) represent 9 different regressions of changes in log of housing prices at the θ th percentiles on the change in exposure to Superfund milestones. Exposure to deletion is measured using the ratio of the area of the of the tract that falls in the 3 mile buffer of deleted sites to the total area of the tract. The change in exposure to deletion is measured between 1990 and 2000. The change in other Superfund milestones are defined analogously. The control variables (in changes between 1990 and 2000) are listed in Table 1. The regression is weighted using the number of owner-occupied housing units. No obs=3,123. Robust standard errors in parentheses. Estimates are statistically significant at ***1%, **5% and *10%.									

Note that here are more tracts in the 3 mile buffer around the 227 sites in Table A4 than in the 3 mile buffer around the 299 sites in Table A2. On average, the 227 sites have more tracts per site, most likely because these sites are more likely to be located in more urban areas.