

**Does Hazardous Waste Matter?
Evidence from the Housing Market and the Superfund Program***

Michael Greenstone and Justin Gallagher

September 2005

* We thank Daron Acemoglu, David Autor, Maureen Cropper, Esther Duflo, Dick Eckhaus, Alex Farrell, Don Fullerton, Ted Gayer, Jon Gruber, Jon Guryan, Joe Gyourko, Paul Joskow, Matthew Kahn, David Lee, Jim Poterba, Katherine Probst, Randall Walsh, Rob Williams and seminar participants at ASSA, BYU, UC-Santa Barbara, CEMFI, Colorado, HEC Montreal, Kentucky, LSE, MIT, NBER, Rand, Resources for the Future, SITE, Stanford, Syracuse, Texas, and UCLA for insightful comments. Leila Agha, Brian Goodness, Elizabeth Greenwood, Rose Kontak, William Li, and Jonathan Ursprung provided outstanding research assistance. We especially thank Katherine Probst for generously sharing data on the costs of Superfund clean-ups. Funding from the Center for Integrating Statistical and Environmental Science at the University of Chicago and the Center for Energy and Environmental Policy Research at MIT is gratefully acknowledged. Although the research described herein has been funded wholly or in part by the United States Environmental Protection Agency, through STAR Cooperative Agreement #R-82940201-0 to the University of Chicago, it has not been subjected to the Agency's required peer and policy review and does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

**Does Hazardous Waste Matter?
Evidence from the Housing Market and the Superfund Program**

Abstract

We exploit knowledge of the selection rule that determined the first 400 hundred hazardous waste sites that were to be cleaned up under the Superfund program to estimate individuals' valuations of proximity to hazardous waste sites. In the program's first year, the selection rule was based on a nonlinear function of a continuous and noisy measure of risk, which allows for the implementation of a quasi-experimental regression discontinuity design. The application of this approach to a new data file suggests that nearly two decades after these sites became eligible for federal clean-up, the property values in surrounding areas had increased by an economically small and statistically indistinguishable from zero amount, relative to areas surrounding sites that narrowly missed eligibility for a Superfund clean-up. Notably, the quasi-experimental results contrast with estimates from a conventional least squares approach that finds a substantial positive effect on local housing prices. Overall, the preferred quasi-experimental estimates suggest that the benefits of Superfund clean-ups as measured through the housing market are substantially lower than \$40 million, our best estimate of the mean costs of a Superfund clean-up.

Michael Greenstone
MIT, Department of Economics
50 Memorial Drive, E52-359
Cambridge, MA 02142-1347
and NBER
mgreenst@mit.edu

Justin Gallagher
NBER
1050 Massachusetts Avenue
Cambridge, MA 02138
jgallagh@nber.org

Introduction

The estimation of individuals' valuations of environmental amenities with revealed preference methods has been an active area of research for more than three decades. Numerous theoretical models outlining methods to recover economically well defined measures of willingness to pay have been developed. In principle, these methods can be used in a variety of settings, including housing markets, recreational choices, health outcomes, and the consumption of goods designed to protect individuals against adverse environmentally-induced health outcomes (for reviews see Freeman 2003; Champ, Boyle, and Brown 2003). The application of these approaches, however, is often accompanied by seemingly valid concerns about misspecification that undermine the credibility of any findings. Consequently, many are skeptical that markets can be used to determine individuals' valuation of environmental amenities. Further, the increasing reliance on stated preference techniques to value environmental amenities is surely related to dissatisfaction with the performance of revealed preference techniques.¹

Hazardous waste sites are an example of an environmental disamenity that provokes great public concern. The 1980 Comprehensive Environmental Response, Compensation, and Liability Act, which became known as Superfund, was intended to address this concern. The Act gave the EPA the right to initiate remedial clean-ups at sites where a release or significant threat of a release of a hazardous substance poses an imminent and substantial danger to public welfare and the environment. Since the passage of the Superfund legislation, more than 1,500 sites have been placed on the National Priorities List (NPL), qualifying those sites for the expenditure of federal remediation funds. As of 2000, approximately \$30 billion (2000\$) has been spent on clean-ups, and remediation efforts are incomplete at roughly half of the sites. Despite these expenditures, there has not been a systematic accounting of the benefits of Superfund clean-ups of hazardous waste sites. As a result, Superfund is a controversial program.²

¹ See Hanemann (1994) and Diamond and Hausman (1994) for discussions of stated preference techniques.

² In March 1995 in Congressional testimony, Katherine Probst of Resources for the Future said, "Although the program has been in existence for over 14 years, we still know very little about the benefits of site cleanup or about the associated costs." At the same hearing, John Shanahan of the Heritage Foundation said, "Superfund...is widely regarded as a wasteful and ineffective program in dire need of substantive reform."

This study uses the hedonic method to empirically estimate individuals' valuations of the remediation of hazardous waste sites from the residential housing market. The results can also be used to estimate the benefits of the Superfund program. The appeal of the housing market is that if it is operating correctly, prices will capture the health and aesthetic benefits of clean-ups.

The empirical challenge is that NPL sites, by their very assignment to the NPL, are the most polluted sites in the US. Thus, the evolution of housing prices in the areas proximate to them may not be comparable to the evolution in the rest of the US. To solve this problem, we implement a quasi-experiment based on knowledge of the assignment rule that the EPA used to develop the first NPL.

Nearly 15,000 sites were referred to the EPA as potential NPL sites in 1980-1, but the EPA's budget could only accommodate 400 clean-ups. Consequently, they initially winnowed the list to the 690 worst sites and then developed the Hazardous Ranking System (HRS). The HRS was used to assign each site a score from 0 to 100 based on the risk it posed with 100 being the most dangerous. The EPA placed the 400 sites with HRS scores exceeding 28.5 on the initial NPL in 1983. These sites then became eligible for Superfund remedial clean-ups.

Our empirical strategy exploits this selection process by focusing on the census tracts containing the 690 hazardous waste sites that were finalists for the initial NPL. We compare the evolution of property values in the tracts containing sites with initial HRS scores above 28.5 to those with sites that had HRS scores below the 28.5 threshold. The assumption is that the below 28.5 sites form a valid counterfactual for the evolution of housing prices at sites above the threshold. We also implement a quasi-experimental regression discontinuity design (Cook and Campbell 1979) to focus the comparisons in the "neighborhood" of the cut-off.

These approaches suggest that the presence of a Superfund site is associated with economically small and statistically indistinguishable from zero gains in residential property values. The validity of these results is supported by the finding that observable determinants of housing price appreciation are well balanced among census tracts containing sites with HRS scores above and below 28.5, especially near the cut-off. Further, the result is robust to a wide variety of specification checks. Thus, these

findings suggest that the benefits of Superfund clean-ups as measured through the housing market are substantially lower than \$40 million, our best estimate of the mean costs of a Superfund clean-up.

These estimates of marginal willingness to pay for remediation contrast sharply with those obtained from a more conventional comparison of census tracts with NPL sites to all other tracts in the US. The conventional approach implies property value gains approximately equal to the costs of the clean-ups. The determinants of changes in housing prices, however, differ dramatically between census tracts with NPL sites and the remainder of the country. This finding underscores the importance of our research design as a potential solution to the likely confounding in conventional approaches.

We also examine whether the clean-ups lead to changes in the demographic characteristics and population of the tracts containing NPL sites. The quasi-experimental estimates fail to find significant evidence of sorting in response to the clean-up. For example, the percentage of households that receives public assistance, the percentage black, and the percentage under the age of 6 all appear to be unchanged. There is modest evidence that clean-ups are associated with population increases. These findings contrast with evidence of sorting or migration in response to changes in emissions of toxic air pollutants (Banzhaf and Walsh 2005) and a case study of the listing of four hazardous waste sites on the NPL (Cameron and McConnaha 2005).

In addition to the availability of the quasi-experiment, our approach differs from previous research on the benefits of the Superfund program in two important ways. First, the analysis is conducted with a data file on all Superfund hazardous waste sites, the sites that narrowly missed placement on the initial NPL, and census-tract level housing prices for 1980, 1990, and 2000. The data file also contains a wealth of information about the hazardous waste sites, including the census tract they are located in, HRS scores, the expected and actual costs (where remediation is complete), and the size in acres. This data file is the most comprehensive one ever compiled by the EPA or other researchers on the Superfund program and its effects. The result is that this study is a substantial departure from previous Superfund hedonic studies that examine a single site or handful of sites (Smith and Michael 1990; Kohlhase 1991; Kiel 1995; Gayer Hamilton, and Viscusi 2000 and 2002; McCluskey and Rausser 2003; Messer et al. 2004; Cameron and McConnaha 2005).

Second, we assume that individuals transmit their valuations of the reduction in health risks and aesthetic improvements of future clean-ups through the housing market. This assumption frees us from a reliance on the notoriously poor laboratory estimates of risk to human health associated with the thousands of chemicals present at these sites. Consequently, all welfare calculations are derived from consumers' revealed preferences and not from EPA laboratories and assumptions about the appropriate value of a statistical life.³

The paper proceeds as follows. Section I describes the conceptual framework. Section II provides background on the Superfund program and how its initial implementation may provide the conditions necessary to credibly estimate the benefits of Superfund clean-ups. Section III details the data sources and provides some summary statistics. Sections IV and V review the econometric methods and report on the empirical findings. Section VI interprets the findings and VII concludes the paper.

I. Conceptual Framework and Research Design

This paper's primary goals are to obtain reliable estimates of the benefits of Superfund clean-ups and, more broadly, measures of individuals' valuations of clean local environments. An explicit market for proximity to hazardous waste sites, however, does not exist. This section explains why we believe that current knowledge is insufficient to successfully implement the health effects approach. It also explains why data on the housing market may offer an opportunity to achieve our goals. Specifically, we briefly review hedonic theory, which spells out the assumptions necessary to interpret changes in housing prices as welfare changes, and then explain the econometric identification problems that plague the conventional implementation of the hedonic method.

A. Difficulties with the Health Effects Approach to Valuing Clean-Ups

³ Viscusi and Hamilton (1999) use EPA provided estimates of the probability of cancer cases at a subsample of sites and find that at the median site expenditure the average cost per cancer case averted by the clean-up exceeds \$6 billion. They also find evidence that decisions about the timing of clean-ups of NPL sites are associated with local measures of political activism. Other researchers have found less decisive evidence on the relationship between local community's characteristics and EPA decisions on the timing of clean-ups (Hird 1993, 1994; Zimmerman 1993; Gupta et al., 1995 and 1996).

The basis of the “health effects” approach is the determination of the reductions in rates of morbidity and mortality associated with proximity to a hazardous waste site following a clean-up. These reductions are then multiplied by estimates of willingness to pay to avoid morbidities and fatalities.

The four steps involved in these calculations help to clarify the difficulties with the health effects approach. The first step is to identify each of the chemicals present at a site and the pathway(s) (e.g., air, water, or soil) by which humans come into contact with them. Through tests conducted at the site, the EPA obtains pathway-specific estimates of the concentrations of each chemical before the clean-up. The agency also specifies goals for these concentrations once the clean-up is completed, resulting in expected chemical by pathway specific reductions in concentrations.

The second step is estimating the health benefits of these pollution reductions, which requires the development of pathway-specific estimates of the health risk from each chemical. The difficulty is that more than 65,000 industrial chemicals have been in commercial production since WWII in the U.S. and the human health effects of many of them are unknown. This problem is further complicated by heterogeneity in the health effects across the pathway of exposure.⁴

Third calculating the health benefits, even for chemicals where reliable dose-response evidence is available, requires assumptions about the size of the affected population and the length of exposure through each potential pathway. This task is complicated by the fact that individuals tend to avoid contact with known risks, so residential proximity to a hazardous waste site may provide a poor measure of exposure. Hamilton and Viscusi (1999) underscore the difficulty in developing reliable exposure assumptions and argue that the EPA often uses assumptions that seem unrealistic.

Fourth, monetizing the changes in morbidity and mortality requires estimates of individuals’ willingness to pay (WTP) to avoid these events. There is an extensive literature on the value of a statistical life (see e.g., Viscusi 1993 and Ashenfelter and Greenstone 2004a and 2004b), but estimates of trade-offs between wealth and morbidity are less pervasive. Moreover, the application of available

⁴ For example, swallowing endrin has negative health consequences but inhalation or contact with the skin is believed to be safe. Similarly, swallowing or inhaling (through dust) arsenic is dangerous, but skin contact from dirt or water is relatively harmless.

estimates relies on the potentially controversial assumption that the literature's estimates of WTP are relevant to the affected subpopulation.

Each of the four steps of the health effects approach involves substantial uncertainty. In light of the scientific, empirical, and data quality concerns outlined here, our view is that at present this approach is unlikely to produce credible estimates of the benefits of Superfund clean-ups. Further, by its very nature, this approach cannot account for the potential aesthetic benefits of these clean-ups.

B. The Hedonic Method

As an alternative to the health effects approach, we use the housing market to infer individuals' valuations of clean-ups. Economists have estimated the association between housing prices and environmental amenities at least since Ridker (1967) and Ridker and Henning (1967). However, Rosen (1974) and Freeman (1974) were the first to give this correlation an economic interpretation. In the Rosen formulation, a differentiated good can be described by a vector of its characteristics, $Q = (q_1, q_2, \dots, q_n)$. In the case of a house, these characteristics may include structural attributes (e.g., number of bedrooms), neighborhood public services (e.g., local school quality), and local environmental amenities (e.g., distance from a hazardous waste site). Thus, the price of the i^{th} house can be written as:

$$(1) \quad P_i = P(q_1, q_2, \dots, q_n).$$

The partial derivative of $P(\bullet)$ with respect to the n^{th} characteristic, $\partial P / \partial q_n$, is referred to as the marginal implicit price. It is the marginal price of the n^{th} characteristic implicit in the overall price of the house.

In a competitive market, the locus between housing prices and characteristic, or the hedonic price schedule (HPS), is determined by the equilibrium interactions of consumers and producers.⁵ In the hedonic model, the HPS is formed by tangencies between consumers' bid functions and suppliers' offer functions. The gradient of the implicit price function with respect to distance from a hazardous waste site gives the equilibrium differential that compensates individuals living in close proximity to hazardous waste sites. Locations close to hazardous waste sites must have lower housing prices to attract potential

⁵ See Rosen (1974), Freeman (1993), and Palmquist (1991) for details.

homeowners, so the HPS reveals the price that allocates individuals across locations. In principle, the price differential reflects both individuals valuations of the health risk associated with proximity to a site and the site's damage to a neighborhood's aesthetics. In this respect, the hedonic approach provides a fuller examination of the valuation than an exclusive focus on the health risks would.⁶

At each point on the HPS, the marginal price of a housing characteristic is equal to an individual's marginal willingness to pay (MWTP) for that characteristic and an individual supplier's marginal cost of producing it. Since the HPS reveals the MWTP at a given point, it can be used to infer the welfare effects of a marginal change in a characteristic. The overall slope of the HPS provides a measure of the average MWTP across all consumers. In principle, the hedonic method can also be used to recover individuals' demand or MWTP functions, which is of tremendous practical importance because these functions allow for the calculation of the welfare effects of nonmarginal changes.⁷

The consistent estimation of (1) is the foundation for accurate welfare calculations of both marginal and non-marginal changes. However, consistent estimation may be difficult since it is likely that there are unobserved factors that covary with both distance from a hazardous waste site and housing prices.⁸ Although this possibility cannot be directly tested, it is notable that proximity to a hazardous waste site is associated with a number of important observable predictors of housing prices. For examples, areas with hazardous waste sites tend to have lower population densities and a higher proportion of mobile homes and locations in the Northeast.

Consequently, cross-sectional estimates of the association between housing prices and proximity to a hazardous waste site may be severely biased due to omitted variables. In fact, the cross-sectional estimation of the HPS has exhibited signs of misspecification in a number of other settings, including the relationships between land prices and school quality, total suspended particulates air pollution, and

⁶ The hedonic approach cannot account for aesthetic benefits that accrue to nonresidents that, for example, engage in recreational activities near the site. The health effects approach has this same limitation.

⁷ Rosen (1974) proposed a 2-step approach for estimating the MWTP function, as well as the supply curve. In recent work, Ekeland, Heckman and Nesheim (2004) outline the assumptions necessary to identify the demand (and supply) functions in an additive version of the hedonic model with data from a single market.

⁸ See Halvorsen and Pollakowski (1981) and Cropper et al. (1988) for discussions of misspecification of the HPS due to incorrect choice of functional form for observed covariates.

climate variables (Black 1999; Chay and Greenstone 2005; Deschenes and Greenstone 2004).⁹ Small (1975) recognized the consequences of the misspecification of equation (1) just one year after publication of the Rosen and Freeman papers:

I have entirely avoided...the important question of whether the empirical difficulties, especially correlation between pollution and unmeasured neighborhood characteristics, are so overwhelming as to render the entire method useless. I hope that...future work can proceed to solving these practical problems....The degree of attention devoted to this [problem]...is what will really determine whether the method stands or falls..." [p. 107].

In the intervening years, this problem of misspecification has received little attention from empirical researchers, even though Rosen himself recognized it.¹⁰ One of this paper's aims is to demonstrate that it may be possible to obtain credible hedonic estimates with a quasi-experimental approach.

Although it is tempting to implement the full blown hedonic method to obtain estimates of marginal willingness to pay functions, our focus is limited to the successful estimation of equation (1). We rely on the decennial census files for census tract-level housing prices. In our view, the inability to link individuals across censuses and the long period between observations makes it invalid to interpret changes in house prices across censuses as measures of a fixed set of individuals' willingness to pay. These data are, however, fine for the estimation of the market locus as specified in (1).

II. The Superfund Program and a New Research Design

A. History and Broad Program Goals

Before the regulation of the disposal of hazardous wastes by the Toxic Substances Control and Resource Conservation and Recovery Acts of 1976, industrial firms frequently disposed of wastes by burying them in the ground. Love Canal, NY is perhaps the most infamous example of these disposal practices. Throughout the 1940s and 1950s, this area was a landfill for industrial waste and more than 21,000 tons of chemical wastes were ultimately deposited there. The landfill closed in the early 1950s

⁹ Similar problems arise when estimating compensating wage differentials for job characteristics, such as the risk of injury or death. The regression-adjusted association between wages and many job amenities is weak and often has a counterintuitive sign (Smith 1979; Black and Kneisner 2003).

¹⁰ Rosen (1986) wrote, "It is clear that nothing can be learned about the structure of preferences in a single cross-section..." (p. 658), and "On the empirical side of these questions, the greatest potential for further progress rests in developing more suitable sources of data on the nature of selection and matching..." (p. 688).

and over the next two decades a community developed in that area. In the 1970s, Love Canal residents began to complain of high rates of health problems, such as cancer, birth defects, miscarriages, and skin ailments.¹¹ After New York state investigators found high concentrations of dangerous chemicals in the air and soil, concerns about the safety of this area prompted President Carter to declare a state of emergency that led to the permanent relocation of the 900 residents of this area. The Love Canal incident helped to galvanize support for addressing the legacy of industrial waste. These political pressures led to the creation of the Superfund program in 1980.

The centerpiece of the Superfund program, and this paper's focus, is the long-run remediation of hazardous waste sites.¹² These multi-year remediation efforts aim to reduce permanently the serious, but not imminently life-threatening, dangers caused by hazardous substances. As of 2000, roughly 1,500 sites have been placed on the NPL and thereby chosen for these long-run clean-ups. The next subsection describes the selection process, which forms the basis of our research design.

B. Site Assessment Process

As of 1996, more than 40,000 hazardous waste sites had been referred to the EPA for possible inclusion on the NPL. The EPA follows a multi-step process to determine which of these sites pose the greatest danger to humans and the environment. The assessment process involves determining which hazardous chemicals are present at the site and the overall risk level.

The final step of the assessment process is the application of the Hazardous Ranking System (HRS), which is reserved for the most dangerous sites. The EPA developed the HRS in 1982 as a standardized approach to quantify and compare the human health and environmental risk among sites to identify the ones that pose the greatest threat. The original HRS evaluated the risk for exposure to

¹¹ EPA (2000) claims that 56% of the children born in Love Canal between 1974 and 1978 had birth defects.

¹² The Superfund program also funds immediate removals. These clean-ups are responses to environmental emergencies and are generally short-term actions aimed at diminishing an immediate threat. Examples of such actions including cleaning up waste spilled from containers and constructing fences around dangerous sites. These actions are not intended to remediate the underlying environmental problem and account for a small proportion of Superfund activities. Importantly, they are administered at hazardous waste sites that are and are not on the NPL.

chemical pollutants along three migration ‘pathways’: groundwater, surface water, and air.¹³ The toxicity and concentration of chemicals, the likelihood of exposure and proximity to humans, and the population that could be affected are the major determinants of risk along each pathway. The non-human impact that chemicals may have is considered during the process of evaluating the site but plays a minor role in determining the HRS score.

The HRS produces a score for each site that ranges from 0 to 100, with 100 being the highest level of risk. From 1982-1995, the EPA assigned all hazardous waste sites with a HRS score of 28.5 or greater to the NPL.^{14 15} These sites are the only ones that are eligible for Superfund remedial clean-up. The Data Appendix provides further details on the determination of HRS test scores.

C. Clean-Up of NPL Sites

Once a site is placed on the NPL, it generally takes many years until clean-up is complete. The first step is a further study of the extent of the environmental problem and how best to remedy it. This assessment leads to the publication of a Record of Decision (ROD) that outlines the clean-up actions that are planned for the site. In our primary sample, the median time between NPL listing and the release of the ROD is roughly 4 years. The next step is the initiation of clean-up activities; the median time between NPL placement and initiation of clean-up is between 6 and 7 years.

The site receives the “construction complete” designation once the physical construction of all clean-up remedies is complete, the immediate threats to health have been removed, and the long-run threats are “under control.” In our primary sample, the median number of years between NPL placement

¹³ In 1990, the EPA revised the HRS test so that it also considers soil as an additional pathway.

¹⁴ In 1980 every state received the right to place one site on the NPL without the site having to score at or above 28.5 on the HRS test. As of 2003, 38 states have used their exception. It is unknown whether these sites would have received a HRS score above 28.5.

¹⁵ In 1995 the criteria for placement on the NPL were altered so that a site must have a HRS score greater than 28.5 and the governor of the state in which the site is located must approve the placement. There are currently a number of potential NPL sites with HRS scores greater than 28.5 that have not been proposed for NPL placement due to known state political opposition. We do not know the precise number of these sites because our Freedom of Information Act request for information about these sites was denied by the EPA.

and the application of the “construction complete” designation is 12 years. On average, it takes about one more year for the EPA to officially delete the site from the NPL.

D. 1982 HRS Scores as the Basis of a New Research Design

This paper’s goal is to obtain reliable estimates of the benefits of the Superfund program and, more generally, local residents’ willingness to pay for the clean-up of hazardous waste sites. The empirical difficulty is that NPL sites are the most polluted in the US, so the evolution of housing prices near these sites may not be comparable to that of the remainder of the US, even conditional on observable covariates. To avoid confounding the effects of the clean-ups with unobserved variables, it is necessary to develop a valid counterfactual for the evolution of property values at Love Canal and the other Superfund sites in the absence of those sites’ placement on the NPL and eventual clean-up.

A feature of the initial NPL assignment process that has not been noted previously by researchers may provide a credible solution to the likely omitted variables problem. In the first year after the legislation’s passage in 1980, 14,697 sites were referred to the EPA and investigated as potential candidates for remedial action. Through the assessment process, the EPA winnowed this list to the 690 most dangerous sites. The Superfund legislation directed the EPA to develop a NPL of “at least” 400 sites (Section 105(8)(B) of CERCLA). Budgetary considerations caused the EPA to set a goal of placing exactly 400 sites on the NPL.

The EPA developed the HRS to determine scientifically the 400 out of the 690 sites most worthy of clean-ups. Pressured to initiate the clean-ups quickly, the EPA developed the HRS in about a year. The HRS test was applied to the 690 worst sites, and their scores were ordered from highest to lowest. A score of 28.5 divided number 400 from number 401, so the initial NPL published in September 1983 was limited to sites with HRS scores exceeding 28.5.¹⁶

¹⁶ Exactly 400 of the sites on the initial NPL had HRS scores exceeding 28.5. The original Superfund legislation gave each state the right to place one site on the NPL without going through the usual evaluation process. Six of these “state priority sites” were included on the original NPL released in 1983. Thus, the original list contained the 400 sites with HRS scores exceeding 28.5 and the 6 state exceptions. See the Data Appendix for further details.

The central role of the HRS score provides a compelling basis for a research design that compares outcomes at sites with initial scores above and below the 28.5 cut-off for at least three reasons. First, it is unlikely that sites' HRS scores were manipulated to affect their placement on the NPL, because the 28.5 threshold was established after the testing of the 690 sites was completed. The HRS scores therefore reflected the EPA's assessment of the human health and environmental risks associated with the site and were not based on the expected costs or benefits of clean-up.

Second, the HRS scores were noisy measures of risk, so it is possible that true risks were similar above and below the threshold. This noisiness was a consequence of the scientific uncertainty about the health consequences of many of the tens of thousands of chemicals present at these sites.¹⁷ Further, the threshold was not selected based on evidence that HRS scores below 28.5 sites posed little risk to health. In fact, the Federal Register specifically reported that the "EPA has not made a determination that sites scoring less than 28.50 do not present a significant risk to human health, welfare, or the environment" (Federal Register 1984, Section IV) and that a more informative test would require "greater time and funds" (EPA 1982).¹⁸ Together, the HRS scores' noisiness and the budgetary motivation for the 28.5 cutoff are attractive features of this research design.

Third, the selection rule that determined the placement on the NPL is a highly nonlinear function of the HRS score. This naturally lends itself to a comparison of outcomes at sites "near" the 28.5 cut-off. If the unobservables are similar around the regulatory threshold, then a comparison of these sites will control for all omitted factors correlated with the outcomes. This test has the features of a quasi-experimental regression-discontinuity design (Cook and Campbell 1979).

¹⁷ A recent summary of Superfund's history makes this point. "At the inception of EPA's Superfund program, there was much to be learned about industrial wastes and their potential for causing public health problems. Before this problem could be addressed on the program level, the types of wastes most often found at sites needed to be determined, and their health effects studied. Identifying and quantifying risks to health and the environment for the extremely broad range of conditions, chemicals, and threats at uncontrolled hazardous wastes sites posed formidable problems. Many of these problems stemmed from the lack of information concerning the toxicities of the over 65,000 different industrial chemicals listed as having been in commercial production since 1945" (EPA 2000, p. 3-2).

¹⁸ One way to measure the crude nature of the initial HRS test is by the detail of the guidelines used for determining the HRS score. The guidelines used to develop the initial HRS sites were collected in a 30 page manual. Today, the analogous manual is more than 500 pages.

An additional feature of the analysis is that an initial score above 28.5 is highly correlated with eventual NPL status but is not a perfect predictor of it. This is because some sites were rescored, with the later scores determining whether they ended up on the NPL.¹⁹ The subsequent analysis uses an indicator variable for whether a site's initial (i.e., 1982) HRS score was above 28.5 as an instrumental variable for whether a site was on the NPL in 2000 (and 1990). We use this approach rather than a simple comparison of NPL and non-NPL sites, because it purges the variation in NPL status that is due to political influence, which may reflect the expected benefits of the clean-up.

Furthermore, the research design of comparing sites with HRS scores "near" the 28.5 is unlikely to be valid for sites that received an initial HRS score after 1982. This is because once the 28.5 cut-off was set, the HRS testers were encouraged to minimize testing costs and simply determine whether a site exceeded the threshold. Consequently, testers generally stop scoring pathways once enough pathways are scored to produce a score above the threshold. When only some of the pathways are scored, the full HRS score is unknown and the quasi-experimental regression discontinuity design is inappropriate.

E. What Questions Can Be Answered?

Our primary outcome of interest is the median housing value in census tracts near hazardous waste sites. In a well-functioning market, the value of a house equals the present discounted value of the stream of services it supplies into the infinite future. In light of the practical realities of the long period of time between a site's initial listing on the NPL and eventual clean-up and the decennial measures of housing prices, this subsection clarifies the differences between the theoretically correct parameters of interest and the estimable parameters.

Define R as the monetary value of the stream of services provided by a house over a period of time (e.g., a year), or the rental rate. We assume that R is a function of an index that measures

¹⁹ As an example, 144 sites with initial scores above 28.5 were rescored and this led to 7 sites receiving revised scores below the cut-off. Further, complaints by citizens and others led to rescoring at a number of sites below the cut-off. Although there has been substantial research on the question of which sites on the NPL are cleaned-up first (see, e.g., Viscusi and Hamilton 1991 and Sigman 2001), we are unaware of any research on the determinants of a site being rescored.

individuals' perception of the desirability of living near a hazardous waste site. We denote this index as H and assume that it is a function of the expected health risks associated with living in this location and any aesthetic disamenities. It is natural to assume that $\partial R / \partial H < 0$, because, for example, individuals' willingness to pay for renting a house is decreasing in the health risk associated with residing in it.

Now, consider how H changes for residents of a site throughout the different stages of the Superfund process. Specifically,

(2) $H_0 =$ Index Before Superfund Program Initiated

$H_1 =$ Index After Site Placed on the NPL

$H_2 =$ Index Once ROD Published/Clean-Up is Initiated

$H_3 =$ Index Once "Construction Complete" or Deleted from NPL

It seems reasonable to presume that $H_0 > H_3$ so that $R(H_3) > R(H_0)$ because the clean-up reduces the health risks and increases the aesthetic value of proximity to the site. It is not evident whether H_1 and H_2 are greater than, less than, or equal to H_0 . This ordering depends on how H evolves during the clean-up process. It is frequently argued that the announcement that a site is eligible for Superfund remediation causes H to increase, because placement on the NPL may cause residents to revise their expectation of the health risk upwards.²⁰

We can now write the constant dollar price of a house (measured after NPL listing) that is in the vicinity of a hazardous waste site, with a HRS score exceeding 28.5:

$$(3) P^{HRS > 28.5} = \sum_{t=0}^{\infty} 1(H_t = H_1) \delta^t R(H_1) + 1(H_t = H_2) \delta^t R(H_2) + 1(H_t = H_3) \delta^t R(H_3).$$

In this equation, the indicator variables $1(\cdot)$ equal 1 when the enclosed statement is true in period t and δ is a discount factor based on the rate of time preferences. The equation demonstrates that upon placement on the NPL, $P^{HRS > 28.5}$ reflects the expected evolution of H throughout the clean-up process.²¹ The key

²⁰ McCluskey and Rausser (2003) and Messer, Schulze, Hackett, Cameron, and McClelland (2004) provide evidence that prices immediately decline after the announcement that a local site has been placed on the NPL. The intuition is that residents knew that proximity to the site was undesirable but did not know that it was one of the most dangerous sites in the country.

²¹ The stigma hypothesis states that even after remediation individuals will assume incorrectly that properties near Superfund sites still have an elevated health risk. Thus, there is a permanent negative effect on property values. See Harris (1999) for a review of the stigma literature.

implication of equation (3) is that $P^{HRS>28.5}$ varies with the stage of the Superfund clean-up at the time that it is observed. For example, it is higher if measured when $H_t = H_3$ than when $H_t = H_1$, because the years of relatively low rental rates have passed.

The constant dollar price of a house located near a hazardous waste site with a HRS score below 28.5 is:

$$(4) P^{HRS<28.5} = \sum_{t=0}^{\infty} \delta^t R(H_0).$$

We assume that H is unchanged for the sites that narrowly missed being placed on the NPL due to HRS scores below 28.5. If this assumption is valid, then $P_t^{HRS<28.5}$ is identical in all periods.

At least two policy-relevant questions are of interest. First, how much are local residents' willing to pay for the listing of a local hazardous waste site on the NPL? This is the ideal measure of the welfare consequences of a Superfund clean-up. In principle, it can be measured as:

$$(5) t_{WTP \text{ for Superfund}} = [P^{HRS>28.5} | H_t = H_1] - P^{HRS<28.5}.$$

It is theoretically correct to measure $P^{HRS>28.5}$ at the instant that the site is placed on the NPL to account for the Superfund program's full effect on the present discounted stream of housing services at that site. Notice, the sign of $t_{Superfund}$ is ambiguous and depends on the time until clean-up, the discount rate, and the change in H at each stage of the clean-up. In practice, our estimates of $t_{WTP \text{ for Superfund}}$ are likely to be biased upwards relative to the ideal because we can only observe $[P^{HRS>28.5} | 1990 \text{ or } 2000]$ when many of the low rental rate years where $H_t = H_1$ have passed.

Second, how does the market value the clean-up of a hazardous waste site? This is represented by:

$$(6) t_{Clean-Up} = [P^{HRS>28.5} | H = H_3] - P^{HRS<28.5},$$

which is the difference in the value of the property after remediation is completed and the average value of sites that narrowly miss placement on the NPL. This is a measure of how much local governments should pay for a clean-up. Numerous sites from the initial NPL list were cleaned up by 2000, so it is feasible to estimate $[P^{HRS>28.5} | H = H_3]$ with data from that census year. It is important to note that $t_{Clean-Up}$ is not a welfare measure since by 2000 the composition of consumers is likely to have changed.

III. Data Sources and Summary Statistics

A. Data Sources

To implement the analysis, we constructed the most comprehensive data file ever compiled on the Superfund program. The data file contains detailed information on all sites placed on the NPL by 2000, as well as the hazardous waste sites with initial HRS scores below 28.5. We also compiled housing price and demographic information for the areas surrounding these sites.

The housing price, housing characteristics, and demographic data come from Geolytics's *Neighborhood Change Database*, which includes information from the 1970, 1980, 1990, and 2000 Censuses. It provides a panel data set of census tracts based on 2000 census tract boundaries, which are drawn so that they include approximately 4,000 people in 2000.²² We restrict the analysis to the 48,556 out of the 65,443 2000 census tracts that have non-missing housing price data in 1980 (before the Superfund legislation was passed), 1990, and 2000.

We spent considerable effort collecting precise location data (e.g., longitude and latitude) for each of the NPL sites and hazardous waste sites with initial HRS scores below 28.5. This information was used to place these sites in unique census tracts. We also used GIS software to identify the census tracts that neighbor the tracts with the sites. We define neighbors in two ways: tracts that share a border and tracts that fall within distance rings (e.g., 1 mile) of the site. The Data Appendix provides further details on these issues.

The 1982 HRS composite scores are a crucial component of the analysis. We collected these scores for the 690 hazardous waste sites considered for placement on the initial NPL from various issues of the *Federal Register* (1982). We also obtained the groundwater, surface water, and air pathway scores from the same source.

²² Census tracts are the smallest geographic unit that can be matched across the 1980, 1990, and 2000 Censuses. See the Data Appendix for a description of how 1980 and 1990 census tracts were adjusted to fit 2000 census tract boundaries.

We collected a number of other variables for the NPL sites. Various issues of the *Federal Register* were used to determine the dates of NPL listing. The EPA provided us with a data file that reports the dates of release of the ROD, initiation of clean-up, completion of construction, and deletion from the NPL for sites that achieved these milestones. We also collected data on the expected costs of clean-up before remediation was initiated and our estimates of the actual costs for the sites that reached the construction complete stage.²³ The RODs also provided information on the size (measured in acres) of the hazardous waste sites. See the Data Appendix for more details on these variables and our sources.

B. Summary Statistics

The analysis is conducted with two data samples. We refer to the first as the “All NPL Sample,” and it includes the 1,398 hazardous waste sites in the 50 US states and the District of Columbia placed on the NPL by January 1, 2000. The second is labeled the “1982 HRS Sample,” and it is comprised of the hazardous waste sites tested for inclusion on the initial NPL, regardless of their eventual placement on the NPL.

Table 1 presents some summary statistics on the hazardous waste sites in these samples. The entries in column (1) are from the All NPL Sample and are limited to those in a census tract for which there is non-missing housing price data in 1980, 1990, and 2000. After these sample restrictions, there are 985 sites, which is more than 70% of the sites placed on the NPL by 2000. Columns (2) and (3) report data from the 1982 HRS Sample. The column (2) entries are based on the 487 sites that we were able to place in a census tract with complete housing price data. Column (3) reports on the remaining 189 sites with certain census tract placement but incomplete housing price data.

Panel A reports on the timing of the placement of sites on the NPL. Column (1) reveals that about 75% of all NPL sites received this designation in the 1980s. Together, columns (2) and (3) demonstrate that 443 of the 676 sites in the 1982 HRS sample eventually were placed on the NPL. This

²³ We measure actual cost as the sum of government outlays and estimates of the costs of the remediation that were paid for by non-governmental responsible parties. The actual outlays by these parties are unavailable to the public and the EPA, so these parties’ estimated outlays come from EPA engineers’ estimates of the costs of completing the required remediation. See the Data Appendix for further details.

number exceeds the 400 sites that Congress set as an explicit goal, because, as we have discussed, some sites with initial scores below 28.5 were rescored and then received scores above the threshold. Most of this rescoring occurred in the 1986-1989 period. Panel B provides mean HRS scores, conditioned on scores above and below 28.5. Notably, the means are similar across the columns.

Panel C reports on the size of the hazardous waste sites measured in acres. This variable is only available for NPL sites since it is derived from the RODs. In the three columns, the median site size ranges between 25 and 35 acres. The substantially larger mean results from just a few sites. The modest size of most sites suggests that any expected effects on property values may be confined to relatively small geographic areas.²⁴ In the subsequent analysis, we will separately test for effects on property values in the census tracts that contain the sites and tracts that neighbor these tracts. We also report results on property values for tracts with circles of varying radii (e.g., 1 mile, 2 miles, 3 miles and 4 miles) around the sites.

Panel D provides evidence on the amount of time required for the completion of clean-ups. The median time until different clean-up milestones are achieved is reported, rather than the mean, because many sites have not reached all of the milestones yet. As an example, only 16 of the NPL sites in column (2) received either the construction complete or deleted designation by 1990. Thus, when we measure the effect of NPL status on 1990 housing prices, this effect will almost entirely be driven by sites where remediation activities are unfinished. By 2000, the number of sites in the construction complete/deleted category had increased dramatically to 198. In column (1), the numbers of sites that were construction complete by 2000 (1990) is 478 (22).

Panel E reports the expected costs of clean-up for NPL sites. This information was obtained from the sites' RODs and provides a measure of the expected costs (2000 \$'s) of the clean-up before any remedial activities have begun.²⁵ These costs include all costs expected to be incurred during the active clean-up phase, as well as the expected costs during the operation and maintenance phase that is

²⁴ In the subsequent analysis, we explore whether the effect of the clean-ups differs in census tracts where a relatively high fraction of households relies on well water. In principle, sites in these tracts may affect a larger share of houses.

²⁵ All monetary figures are reported in 2000 \$'s, unless otherwise noted.

subsequent to the assignment of the construction complete designation.

In the All NPL Sample, the estimated cost data is available for 753 of the 985 NPL sites. The mean and median expected costs of clean-up are \$28.3 million and \$11 million. The larger mean reflects the high cost of a few clean-ups—for example, the 95th percentile expected cost is \$89.6 million. In the 1982 HRS Sample in column (2), the analogous figures are \$27.5 million and \$15.0 million. Conditional on construction complete status, the mean cost is \$20.6 million among these sites.

The final panel reports estimated and actual costs for the subsample of construction complete sites where both cost measures are available. To the best of our knowledge, the estimated and actual cost data have not been brought together before. The conventional wisdom is that the actual costs greatly exceed the estimated costs of clean-up, and this table provides the first opportunity to test this view. The data appear to support the conventional wisdom as the mean actual costs are 35%-55% higher than the mean expected costs across the three columns. The findings are similar for median costs.

A comparison of columns (2) and (3) across the panels reveals that the sites with and without complete housing price data are similar on a number of dimensions. For example, the mean HRS scores conditional on scoring above and below 28.5 are remarkably similar. Further, the median size and various cost variables are comparable in the two columns. Consequently, it seems reasonable to conclude that the sites without complete housing price data are similar to the column (2) sites, suggesting the subsequent results may be representative for the entire 1982 HRS Sample.

We now graphically summarize some other features of the two samples. Figure 1 displays the geographic distribution of the 985 hazardous waste sites in the All NPL Sample. There are NPL sites in 45 of the 48 continental states, demonstrating that Superfund is genuinely a national program. The highest proportion of sites, however, is in the Northeast and Midwest (i.e., the “Rust Belt”) that reflects the historical concentration of heavy industry in these regions.

Figures 2A and 2B present the geographic distribution of the 1982 HRS sample. Figure 2A displays the distribution of sites with initial 1982 HRS scores exceeding 28.5, while those with scores below this threshold are depicted in 2B. The sites in both categories are spread throughout the United States, but the below 28.5 sites are in fewer states. For example, there are not any below 28.5 sites in

Minnesota, Florida, and Delaware. The unequal distribution of sites across the country in these two groups is a potential problem for identification in the presence of the local shocks that are a major feature of the housing market. To mitigate concerns about these shocks, we will estimate models that include state fixed effects for changes in housing prices.

Figure 3 reports the distribution of HRS scores among the 487 sites in the 1982 HRS Sample. The figure is a histogram where the bins are 4 HRS points wide. The distribution looks approximately normal, with the modal bin covering the 36.5-40.5 range. Importantly, 227 sites have HRS scores between 16.5 and 40.5. This set is centered on the regulatory threshold of 28.5 that determines placement on the NPL and constitutes the regression discontinuity sample that we exploit in the subsequent analysis.

Figure 4 plots the mean estimated costs of remediation by 4-unit intervals, along with the fraction of sites in each interval with non-missing cost data. The vertical line denotes the 28.5 threshold. The non-zero mean costs below the threshold are calculated from the sites that received a score greater than 28.5 upon rescaling and later made it onto the NPL. The estimated costs of remediation appear to be increasing in the HRS score. This finding suggests that despite its widely acknowledged noisiness, the 1982 HRS scores may be informative about relative risks. In the neighborhood of 28.5, however, estimated costs are roughly constant, providing some evidence that risks are roughly equal among the sites in the regression discontinuity sample.

IV. Econometric Methods

A. Least Squares Estimation with Data from the Entire U.S.

Here, we discuss the econometric models that we use to estimate the relationship between housing prices and NPL listing. We begin with the following system of equations:

$$(7) \quad y_{c2000} = \theta \mathbf{1}(\text{NPL}_{c2000}) + \mathbf{X}_{c1980} \beta + \varepsilon_{c2000}, \quad \varepsilon_{c2000} = \alpha_c + u_{c2000}$$

$$(8) \quad \mathbf{1}(\text{NPL}_{c2000}) = \mathbf{X}_{c1980}' \Pi + \eta_{c2000}, \quad \eta_{c2000} = \lambda_c + v_{c2000},$$

where y_{c2000} is the log of the median property value in census tract c in 2000. The indicator variable $\mathbf{1}(\text{NPL}_{c2000})$ equals 1 only for observations from census tracts that contain a hazardous waste site that has been placed on the NPL by 2000. Thus, this variable takes on a value of 1 for any of the

Superfund sites in column (1) of Table 1, not just those that were on the initial NPL list. The vector X_{c1980} includes determinants of housing prices measured in 1980, and ε_{c2000} and η_{c2000} are the unobservable determinants of housing prices and NPL status, respectively. We are also interested in the effect of NPL status in 1990, and the year 1990 versions of these equations are directly analogous.

A few features of the X vector are noteworthy. First, we restrict this vector to 1980 values of these variables to avoid confounding the effect of NPL status with “post-treatment” changes in these variables that may be due to NPL status. Second, we include the 1980 value of the dependent variable, y_{c80} , to adjust for permanent differences in housing prices across tracts and the possibility of mean reversion in housing prices.

Third in many applications of Rosen’s model, the vector of controls, denoted by X , is limited to housing and neighborhood characteristics (e.g., number of bedrooms, school quality, and air quality). Income and other similar variables are generally excluded on the grounds that they are “demand shifters” and are needed to obtain consistent estimates of the MWTP function. The exclusion restriction is invalid however if individuals treat wealthy neighbors as an amenity (or disamenity). In the subsequent analysis, we are agnostic about which variables belong in the X vector and report estimates that are adjusted for different combinations of the variables available in the Census data.

The coefficient θ is the ‘true’ effect of NPL status on 2000 property values after controlling for 1980 property values and the other covariates. For consistent estimation, the least squares estimator of θ requires $E[\varepsilon_{c2000}\eta_{c2000}] = 0$. If there are unobserved permanent (α_c and λ_c) or transitory (u_{c2000} and v_{c2000}) factors that covary with both NPL status and housing prices, then this estimator will be biased. In order to account for transitory factors, we report the results from specifications that include a full set of state fixed effects to account for local shocks.

Ultimately, the approach laid out in equations (7) and (8) relies on a comparison of NPL sites to the rest of the country. The validity of this approach rests on the assumption that linear adjustment can control for all determinants of housing prices between census tracts with and without a NPL site, besides the effect of the presence of a NPL site.

B. A Quasi-Experimental Approach with the 1982 HRS Sample

Here, we discuss an alternative identification strategy that has two key differences with the one described above. First, we limit the sample to the subset of census tracts containing the 487 sites that were considered for placement on the initial NPL and have complete housing price data. Thus, all observations are from census tracts with hazardous waste sites that were initially judged to be among the nation's most dangerous by the EPA. If, for example, the β 's differ across tracts with and without hazardous waste sites or there are differential trends in housing prices in tracts with and without these sites, then this approach is more likely to produce consistent estimates. Second, we use an instrumental variables strategy to account for the possibility of endogenous rescaling of sites.

More formally, we replace equation (8) with:

$$(9) \quad 1(\text{NPL}_{c2000}) = X_{c1980}'\Pi + \delta 1(\text{HRS}_{c82} > 28.5) + \eta_{c2000}, \quad \eta_{c2000} = \lambda_c + v_{c2000},$$

where $1(\text{HRS}_{c82} > 28.5)$ is an indicator function that equals 1 for census tracts with a site that exceeds the 28.5 threshold, based on their HRS score from before the threshold was known. This approach exploits the variation in NPL status that is due to the site's 1982 HRS score.

For the IV estimator (θ_{IV}) to provide a consistent estimate of the HPS gradient, the instrumental variable must affect the probability of NPL listing without having a direct effect on housing prices. The next section will demonstrate that the first condition clearly holds. The second condition requires that the unobserved determinants of 2000 housing prices are orthogonal to the portion of the nonlinear function of the 1982 HRS score that is not explained by X_{c1980} . In the simplest case, the IV estimator is consistent if $E[1(\text{HRS}_{c82} > 28.5) \varepsilon_{c2000}] = 0$. We implement the IV estimator on the data from the 487 sites with nonmissing housing price data to obtain θ_{IV} .

We also calculate IV estimates another way that allows for the possibility that $E[1(\text{HRS}_{c82} > 28.5) \varepsilon_{c2000}] \neq 0$ over the entire sample. In particular, we exploit the regression discontinuity design implicit in the $1(\bullet)$ function that determines NPL eligibility in two separate ways. In the first approach, we include a quadratic in the 1982 HRS score in X_{c80} to partial out any correlation between residual housing prices and the indicator for a HRS score exceeding 28.5. Consequently, this approach does not require that the determinants of housing prices are constant across the range of HRS scores. Instead, it relies on the

arguably plausible assumption that residual determinants of housing price growth do not change discontinuously at HRS scores just above 28.5.

The second regression discontinuity approach involves implementing our IV estimator on the sample of 227 census tracts with sites that have 1982 HRS scores greater than 16.5 and less than 40.5. Here, the identifying assumption is that all else is held equal in the “neighborhood” of the regulatory. More formally, it is $E[1(\text{HRS}_{c82} > 28.5) \varepsilon_{c2000} | 16.5 < 1982 \text{ HRS} < 40.5] = 0$. Both of these regression discontinuity approaches are demanding of the data.

V. Empirical Results

A. Balancing of Observable Covariates

This subsection examines the quality of the comparisons that underlie the subsequent least squares and quasi-experimental instrumental variables estimates of the effect of NPL status on housing price growth. We begin by examining whether NPL status and the $1(\text{HRS}_{c82} > 28.5)$ instrumental variable are orthogonal to the *observable* predictors of housing prices. Formal tests for the presence of omitted variables bias are as always unavailable, but it seems reasonable to presume that research designs that meet this criterion may suffer from smaller omitted variables bias. First, designs that balance the observable covariates may be more likely to balance the unobservables (Altonji, Elder, and Taber 2000). Second if the observables are balanced across NPL status or values of the instrument, then consistent inference does not depend on functional form assumptions on the relations between the observable confounders and housing prices. Estimators that misspecify these functional forms (e.g., linear regression adjustment when the conditional expectations function is nonlinear) will be biased.

Table 2 shows the association of NPL status and $1(\text{HRS}_{c82} > 28.5)$ with potential determinants of housing price growth measured in 1980. Column (1) reports the means of the variables listed in the row headings in the 985 census tracts with NPL hazardous waste sites and complete housing price data. Column (2) displays the means in the 41,989 census tracts that neither contain a NPL site nor share a border with a tract containing one. Columns (3) and (4) report on the means in the 181 and 306 census tracts with hazardous waste sites with 1982 HRS scores below and above the 28.5 threshold, respectively.

Columns (5) and (6) repeat this exercise for the 90 and 137 tracts below and above the regulatory threshold in the regression discontinuity sample. The remaining columns report p-values from tests that the means in pairs of the columns are equal. P-values less than 0.01 are denoted in bold.

Column (7) compares the means in columns (1) and (2) to explore the possibility of confounding in the least square approach. The entries indicate that 1980 housing prices are more than 20% higher in the rest of the US. Among the potential determinants of housing prices, the hypothesis of equal means can be rejected at the 1% level for 9 of the 11 demographic and economic variables and 7 of 11 of the housing characteristics. The tracts with NPL sites are less densely populated (e.g., due to locations in industrial parts of urban areas and rural areas) and have lower household incomes. One variable that seems to capture this is that the fraction of the housing stock that is comprised of mobile homes is more than 80% greater (0.0862 versus 0.0473) in these tracts. Additionally, the fraction of Blacks and Hispanics is greater in tracts without NPL sites, which undermines “environmental justice” claims in this context. Due to the confounding of NPL status with these other determinants of housing price growth, it may be reasonable to assume that the estimation of equation (7) will produce biased estimates of the effect of NPL status.

Column (8) compares the tracts with hazardous wastes that have 1982 HRS scores below and above the 28.5 regulatory threshold. It is immediately evident that by narrowing the focus to these tracts, the differences in the potential determinants of housing prices are greatly mitigated. For example, the population density and percentage of mobile homes are well balanced. One important difference that remains is that the mean housing price in 1980 is roughly 16% higher in tracts with HRS scores above 28.5. Overall, the entries suggest that the above and below 28.5 comparison reduces the confounding of NPL status.

Column (9) repeats this analysis in the regression discontinuity sample. Notably, the difference in 1980 housing prices is reduced to 10% and is no longer statistically significant at conventional levels (p-value = 0.084). With respect to the other determinants of 2000 housing prices, the findings are remarkable in that the hypothesis of equal means cannot be rejected at the 1% level for any of these variables. The differences in the means are substantially reduced for many of the variables, so this result

is not simply due to the smaller sample (and larger sampling errors). This finding suggests that the regression discontinuity approach may provide the most reliable results of the effect of NPL status on housing price growth.

There is not enough room to present the results here, but there are substantial differences in the geographic distribution of sites across states in both the above and below 28.5 (i.e., columns 3 and 4) and regression discontinuity (i.e., columns 5 and 6) comparisons. It is likely that this is due to the small samples in each state. This is a salient issue, because there were dramatic differences in state-level trends in housing prices in the 1980s and 1990s. Consequently, we will emphasize specifications that include state fixed effects in most of the subsequent models.

B. Least Squares Estimates from the All NPL Sample

Table 3 presents the first large-scale effort to test the effect of Superfund clean-ups on property values. Specifically, it reports the regression results from fitting 4 versions of equation (7) for 1990 and 2000 housing prices. The sample size is 42,974 in all regressions. In Panel A, 746 (985) observations in 1990 (2000) are from census tracts that contain a hazardous waste site that had been on the NPL at any time prior to 1990 (2000).²⁶ In addition to tracts with incomplete housing price data from 1980 – 2000, tracts that share a border with tracts containing NPL sites (unless they also contain a NPL site) are dropped from the sample since Superfund clean-ups may affect housing prices in neighboring tracts.

This possibility is investigated in Panel B. Here, the observation from each tract with a NPL site is replaced with the average of all variables across the census tracts that share a border with the tract containing the NPL site and have complete housing price data. The intent in this panel is to test whether the NPL designation affects housing prices in adjacent tracts. The individual observations on these neighboring tracts are excluded from the sample as in Panel A. The rest of the sample is comprised of

²⁶ For tracts that include multiple NPL sites, the observation from that tract is repeated in the data file for each site. Thus, there would be 2 identical observations for a tract that contains two NPL sites. The 746 (985) NPL sites in 1990 (2000) are located in 682 (892) individual tracts.

observations on the 41,989 tracts with complete housing price data that neither have a NPL site nor are adjacent to a tract with a NPL site.

The dependent variables are underlined in the first column. The entries report the coefficient and heteroskedastic-consistent standard error on the NPL indicator. All specifications control for the natural log of the mean housing price in 1980, so the reported parameter should be interpreted as the growth in housing prices in tracts with a NPL site (or its neighbors), relative to tracts without NPL sites. The exact covariates in each specification are noted in the row headings at the bottom of the table.

The Panel A results show that this least squares approach finds a positive association between NPL status and housing price increases. Specifically, the estimates in the first row indicate that median housing prices grew by 9.8% to 16.4% (measured in ln points) more in tracts with a NPL site between 1980 and 1990. All of these estimates would easily be judged statistically significant by conventional criteria. The column (4) estimate of 9.8% is the most reliable one, because it is adjusted for all unobserved state-level determinants of price growth. Recall, Table 1 demonstrated that remediation activities had been initiated at less than half of these sites and only 22 sites were construction complete by 1990. Thus, these findings appear to contradict the popular “stigma” hypothesis (see, e.g., Harris 1999) that a site’s placement on the NPL causes housing prices to decline. We investigate this issue further below.

The next row repeats this exercise for the period from 1980 to 2000. Here, the estimated effect of the presence of a NPL site within a tract’s boundaries is associated with a 4.0% to 7.3% increase in the growth of house prices depending on the specification. Notably, the standard errors on the 2000 estimates are about half the size of the 1990 standard errors.

Panel B explores the growth of housing prices in the census tracts that are adjacent to the tracts containing the NPL sites. All of the estimates are statistically different from zero and imply that the placement of a site on the NPL is associated with a positive effect on the growth of housing prices in

neighboring tracts. The column (4) specification indicates 10.8% and 8.8% gains by 1990 and 2000, respectively.²⁷

It is possible to use the Table 3 estimates to do a crude cost-benefit calculation. The 1980 mean of aggregate property values in tracts with NPL sites is roughly \$77 million. The multiplication of this value by the 2000 own tract, column (4) point estimate of 0.066 implies that housing prices were roughly \$5 million higher in these tracts in 2000 due to the NPL designation. The 1980 mean of aggregate property values in adjacent tracts is \$638 million and a similar calculation suggests that the NPL designation led to an increase in property values of roughly \$56 million. Setting aside the legal costs associated with collecting the funds from private parties and deadweight loss of public funds, the \$61 million increase in property values appears larger than our estimate of the mean actual costs of clean-ups of about \$40 million.²⁸

Before drawing any definitive conclusions or policy implications, however, it is worth emphasizing that three features of the evidence presented so far suggest that the Table 3 estimates may be unreliable. First, Table 2 demonstrated that NPL status is confounded by many variables. Second, 7 of the 8 adjacent tract point estimates exceed the own census tract estimates. In light of these tracts' small size (recall, the median size of a site is 29 acres), it seems unlikely that the NPL status would have a larger effect in the adjacent tracts. Third, all of the 2000 own tract and adjacent point estimates are smaller than the 1990 point estimates from the same specifications. This finding is puzzling, because remediation was complete at only 22 of the NPL sites by 1990. In contrast, the clean-ups had been completed at nearly half of the sites by 2000.

²⁷ The subsequent sections are focused on the subset of hazardous waste sites that were tested in 1982 for inclusion on the initial NPL. Consequently, it is relevant to contrast the subsequent estimates with those from the Table 3 approach. When the set of tracts with NPL sites is limited to those that were tested in 1982, the column 4 specification in Table 3 produces point estimates (standard errors) of 0.063 (0.025) and 0.057 (0.013) for 1990 and 2000. The analogous 1990 and 2000 adjacent census tract results are 0.080 (0.016) and 0.073 (0.012). Overall, these estimates are quite similar to those in Table 3 for the full set of NPL sites.

²⁸ Our estimate of the mean actual costs of clean-ups are calculated in the following manner. Panel F of Table 1 indicates that among fully remediated sites the actual costs of clean-up exceed the expected costs by roughly 40%. We then multiply the mean expected costs of clean (calculated among all sites) of \$28.3 million by 1.4, which gives our estimated mean actual cost of \$40 million.

The next section begins our presentation of the results from the 1982 HRS score quasi-experiment. This research design seems to mitigate many of the concerns about omitted variables bias that appear to plague the least squares approach.

C. Is the $1(\text{HRS}_{c82} > 28.5)$ a Valid Instrumental Variable for NPL Status?

This section presents evidence on the first-stage relationship between the $1(\text{HRS}_{c82} > 28.5)$ indicator and NPL status, as well as some suggestive evidence on the validity of the exclusion restriction. Figure 5 plots the bivariate relation between the probability of 1990 (Panel A) and 2000 (Panel B) NPL status and the 1982 HRS score among the 487 sites in the 1982 HRS sample. The plots are done separately for sites above and below the 28.5 threshold and come from the estimation of nonparametric regressions that use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5.²⁹ Thus, they represent a moving average of the probability of NPL status across 1982 HRS scores. The data points represent the mean probabilities in the same 4-unit intervals of the HRS score as in Figures 3 and 4.

The figures present dramatic evidence that a HRS score above 28.5 is a powerful predictor of NPL status in 1990 and 2000. Virtually all sites with initial scores greater than 28.5 are on the NPL. The figure also reveals that some sites below 28.5 made it on to the NPL (due to rescoring) and that this probability is increasing in the initial HRS score and over time.

Panel A of Table 4 reports on the statistical analog to these figures from the estimation of linear probability versions of equation (9) for NPL status in 1990 and 2000. This is a test of the “first-stage” relationship in our two-stage least squares approach. In the first five columns, the sample is comprised of the full 1982 HRS Sample. The controls in the first four columns are identical to those in the four specifications in Table 3. In the fifth column, the 1982 HRS score and its square are added to the column (4) specification so that the estimates are adjusted for a smooth function of the “running” variable. Thus, the specification tests for a discontinuity in the probability of NPL status at the regulatory threshold of

²⁹ The smoothed scatterplots are qualitatively similar with a rectangular weighting function (i.e., equal weighting) and alternative bandwidths.

28.5. In column (6), the controls are the same as in column (4), and the sample is restricted to the regression discontinuity sample that is comprised of the 227 sites with 1982 HRS scores between 16.5 and 40.5. These specifications and samples are repeated throughout the remainder of paper.

The results confirm the visual impression that a 1982 HRS score above 28.5 increases the probability that a site is placed on the NPL. The point estimates imply a higher probability ranging between 61% and 86%, depending on the year and specification. Overall, these findings reveal a powerful first-stage relationship.

Panel B of Table 4 presents an informal test of the validity of the research design. The table reports the coefficient and standard error on $1(\text{HRS}_{\text{c82}} > 28.5)$ from five regressions. In the first row of the panel, the dependent variable is the natural log of 1980 mean housing prices in the tract with the hazardous waste site. In the second row, the dependent variable and the covariates for each observation are calculated as the mean across the census tracts that share a border with the tract containing the hazardous waste site. The specifications are identical to those in the upper panel (except, of course, they do not control for 1980 prices). Thus, these regressions test for differential 1980 housing prices above and below the threshold after adjustment for observed covariates. Residual housing prices may be an important predictor of the growth in housing prices, so evidence of significant differences would undermine the validity of this research design.

After adjustment for the housing characteristics, the point estimate on the above 28.5 indicator is small both economically and statistically in the own census tract regressions. In columns (3) – (6), the estimates range from -0.010 to 0.001 and the largest t-statistic is 0.55. The adjacent census tract results suggest that a HRS score exceeding 28.5 is virtually uncorrelated with adjusted 1980 housing prices.

Figure 6 allows for a graphical investigation of this test. It plots the results from nonparametric regressions between 1980 residual housing prices after adjustment for the column (4) covariates (except the indicator for a HRS score above 28.5) and the 1982 HRS scores. These plots are done separately for sites above and below the threshold and again use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5. The qualitative findings are unaffected by reasonable alternatives for the weighting

function and the bandwidth. In Panel A, the 1982 HRS sample is used, whereas Panel B employs the regression discontinuity sample.³⁰

There is little evidence of a relationship between the 1982 HRS score and 1980 residual housing prices in either figure. Importantly, there are only modest differences in residual housing prices just above and below the regulatory threshold. Along with the findings in Tables 2 and 4, these plots support the validity of this paper's quasi-experimental approach.

D. Quasi-Experimental Estimates of NPL Status on Housing Prices

Table 5 presents instrumental variables estimates of the effect of NPL status on housing prices in 1990 and 2000. In Panel A the observations are from the census tracts containing the 487 hazardous waste sites in the 1982 HRS sample. In Panel B each observation is comprised of the average of all variables across tracts that share a border with these tracts. The 6 combinations of samples and control variables are identical to those in Panel A of Table 4.

The Panel A results suggest that a site's placement on the NPL has little impact on the growth of property values in its own census tract, relative to tracts with sites that narrowly missed placement on the NPL. The 1990 and 2000 point estimates range from -0.6% to 6.6%. However, they all have associated t-statistics less than 2.

The most reliable specifications are presented in columns (4) – (6). The column (4) specification controls for state fixed effects in housing price growth and produces smaller point estimates than the identical specification in Table 3. However, the imprecision of the estimates means that the 95% confidence intervals of these estimates overlap. The regression discontinuity specifications in columns (5) and (6) may be the least subject to omitted variables bias concerns, so it is notable that they produce

³⁰ Figure 6 is intended to provide a qualitative graphical exploration of the regression results. The relationship between housing prices and 1982 HRS scores cannot be exactly inferred from this graph, because the HRS score has not been adjusted for the column (4) covariates. This same caveat applies to the subsequent Figures 7 and 8. However, the meaningfulness of this graph is supported by Table 2's finding that the covariates are well balanced among sites with 1982 HRS scores above and below the regulatory threshold, especially in the regression discontinuity sample.

the smallest point estimates. These two specifications are demanding of the data and this is reflected in the larger standard errors.

Panel B presents the adjacent tract results. The column (4) specification produces economically small point estimates that are estimated precisely. Specifically, the 1990 and 2000 point estimates (standard errors) are -0.006 (0.027) and 0.012 (0.022), respectively. The regression discontinuity estimates are also small and the null of zero cannot be rejected at anything approaching conventional levels. Overall, these quasi-experimental results suggest that a site's placement on the NPL does not affect housing prices in neighboring census tracts either positively or negatively. In this respect they contrast sharply with the adjacent tract results from the least squares approach in Table 3.

Figures 7 and 8 provide an opportunity to better understand the source of these regression results. They are constructed identically to Figure 6 and plot the nonparametric regressions of 2000 residual housing prices (after adjustment for the column (4) covariates) against the 1982 HRS Score in the 1982 HRS and Regression Discontinuity samples. The figures are for the own and adjacent census tract housing prices, respectively, and are the graphical version of the reduced form relationship between housing price growth and the instrument.

These graphs confirm that there is little association between 2000 residual housing prices and the 1982 HRS score. A comparison of the plots at the regulatory threshold is of especial interest in light of the large jump in the probability of placement on the NPL there. In all four of these plots the moving averages from the left and right are virtually equal at the threshold.

One potentially troublesome issue with the approach thus far is the arbitrary nature of census tract boundaries with respect to the location of the hazardous waste sites. Although there is no a priori reason to believe that it is likely, it is possible that the houses affected by the clean-ups fall predominantly outside the sites' tracts and the adjacent tracts. To explore this possibility, we obtained the exact longitude and latitude of all the hazardous waste sites in the 1982 HRS sample and used GIS software to draw circles around them with radii of 1, 2, and 3 miles. We then calculated the means of all variables as the weighted average across the portion of tracts that fall within the relevant circle around the hazardous waste sites. The weight is the fraction of the tract's area that falls within the relevant circle multiplied by

the tract's 1980 population. To place the size of these circles in some context, the 1980 aggregate values of the housing stock in the three circles are roughly \$300, \$755, and \$1,410 million (2000 \$'s).

Appendix Table 1 reports results from the application of the 6 instrumental variables specifications used in Table 5 to this data set. None of the 36 point estimates presented would be judged to be statistically different from zero by conventional criteria. Thus, the conclusions from this approach are very similar to those from Table 5—sites' receipt of the NPL designation appears to have little effect, either positive or negative, on the growth of housing prices among houses in close proximity to those sites.³¹

Table 6 presents the results from a series of other specification checks, all of which return to using census tracts, rather than circles around a site, as the unit of observation. For conciseness, we only present estimates from the column (1) specification and the three most robust specifications (i.e., columns 4, 5, and 6) from Table 5. We also limit the table to the 2000 own census tract results. We focus on the 2000 own census tract data, because it is more likely that we will find an effect with them since more clean-ups are completed in 2000 and the data come from the houses in closest proximity to the sites. Consequently, a continued failure to find an effect of NPL status will further undermine the view that NPL status affects local property values.

Each of the specification checks is reported in a separate panel. The HRS score is a nonlinear function of the ground water, surface water, and air migration pathway scores, so we add these three pathway scores to X_{c1980} in equations (7) and (9). Panel A reports the results from this alternative method to achieve identification from the nonlinearity in the selection rule. In Panel B, we add the 1970 values of the covariates (including the natural log of 1970 housing prices) to the Table 5 specification to adjust for mean reversion or pre-existing trends. The 1970 data is available for a subset of counties, so the sample sizes decline to 353 in columns (1) – (3) and 157 in (4). Panel C adds the 2000 value of the covariates.

³¹ A limitation of the GIS approach in this setting is that address level data on housing prices and the covariates is unavailable. Consequently, we assign a census tract's average to the portion of the tract that falls within the circle, which is equivalent to assuming that there is no heterogeneity in housing prices or other variables within a tract. This is a strong assumption and underscores that the publicly available census tract data is not ideal for the circle approach. Another limitation of the circle approach is that it is not possible to implement the GIS circles with data from the entire country as there is no clear way to center the circles in areas without a hazardous waste site.

These variables are potentially endogenous, because they may be affected by the clean-ups. To isolate the effect of NPL status on land values (rather than housing values), however, it may be appropriate to adjust for these variables.

The results in the first three panels of Table 6 continue to suggest that the NPL designation has little effect on the growth of property values between 1980 and 2000. The adjustment for the individual pathway scores generally reduces the point estimates and causes the standard errors to increase. The addition of the 1970 and 2000 covariates also tends to reduce the point estimates. All 12 of these estimates have an associated t-statistic less than 1.1.

Panel D. tests whether the effect of the NPL designation differs in census tracts with a population density exceeding 1,588 per square mile, which is the top quartile among tracts in this sample. The intuition is that the price response to clean-ups may be greater in higher density tracts where opportunities to build new houses may be more restricted. Specifically, the specification now includes an indicator for these tracts and the additional variable of interest, which is the interaction of this indicator and the indicator for 2000 NPL status. The latter variable is treated as endogenous and instrumented with the interaction of indicators for tracts with a population density exceeding 1,588 and sites with a HRS score exceeding 28.5. Two of the estimates of the interaction are positive and two are negative. None of them would be judged to statistically differ from zero by conventional criteria.³²

Panel E follows a similar approach, except that its focus is on tracts where it is most likely that the hazardous waste site has contaminated local drinking water supplies. These tracts are defined as those where the site's groundwater pathway score exceeds 28.5 and the tract is in the top half of the distribution of mean household well water usage (i.e., more than 11.8% of households use well water). It seems reasonable to suppose that tracts where households rely on contaminated water may place a higher value on clean-ups. Here, the point estimates are in the relatively narrow range of -0.007 to 0.024, although individually they are poorly determined.³³

³² The results are qualitatively similar when we test whether the effect of the NPL designation differs in tracts located in counties in the top quartile of county population density.

³³ Of the 122 tracts with a population density exceeding 1,588, 72 have a HRS score exceeding 28.5 and 79 have sites placed on the NPL by 2000. Among the 204 tracts that qualify as potentially having contaminated drinking

E. Quasi-Experimental Estimates of Stages of Superfund Clean-Ups on Rental Rates

We now turn to using the 1990 and 2000 ln median rental rates as the outcome variable. Rental units account for roughly 20% of all housing units and generally differ on observable characteristics from owner occupied homes, so the results may not be indicative of the overall housing market. This outcome's appeal is that rental rates are a flow measure, so we can abstract from the problem that individuals' expectations about the length of time until the completion of the clean-up are unknown. Consequently, we can test whether individuals' valuation of the local housing services varies at different stages of Superfund clean-ups.

Table 7 presents separate two stage least squares estimates of the effect of the different stages of the remediation process on 1990 and 2000 rental rates. The three endogenous variables are separate indicator variables for tracts that contain a site that by 1990/2000: is on the NPL but a ROD has not been issued; has a ROD and/or clean-up has been initiated but not completed; or, has been designated construction complete or deleted from the NPL. Importantly, these categories are mutually exclusive, so each of the tracts with a NPL site only helps to identify one of the indicators. The three instruments are the interaction of the $1(\text{HRS}_{c82} > 28.5)$ variable with the indicators for these three stages of the NPL process. The table reports the point estimates and their standard errors, along with the p-value from a F-test that the three point estimates are equal. The number of sites in each category and the mean HRS score is also listed in brackets.

The sites in the "NPL Only" category have been on the NPL for 7 (17) years by 1990 (2000), but the EPA had not developed a remediation plan yet. It seems reasonable to presume that individuals have had enough time to use the NPL listing to update their expectation of the risk associated with proximity to the site. In both the 1990 and 2000 regressions, three of the four point estimates are positive for this group. The large standard errors, especially in the regression discontinuity specifications, make definitive conclusions unjustified. However, the results in this table (along with results from analogous housing

water, 164 have HRS scores exceeding 28.5 and 170 have sites that had been placed on the NPL by 2000.

price regressions³⁴) fail to lend much credence to the conventional wisdom that a site's placement on the NPL leads to a reduction in the value of housing services near the site.

The 2000 results for the "Construction Complete or NPL Deletion" category are noteworthy. The point estimates from the more reliable specifications range from -0.020 to 0.046 and zero cannot be rejected for any of them. This finding is telling, because these sites have been fully remediated and yet there seems to be little effect on the growth of rental rates.³⁵ We also report the results from F-tests of whether the three parameter estimates are equal. Further, the null hypothesis of equality of the three parameters cannot be rejected at conventional levels in any of the specifications. The imprecision of the estimates helps to explain this finding, but their rank ordering is also surprising. In three of the four 2000 specifications, the point estimate on the "NPL Only" indicator exceeds the "Construction Complete or NPL Deletion" one.

A possible explanation for the small housing price effects in Tables 5 and 6 is that gains near sites where clean-ups are complete are obscured by declines at sites where the clean-ups are not complete or have not been initiated. The findings in this table fail to support this explanation. Overall, these results lend further support to the view that Superfund clean-ups had small effects on individuals' valuations of local housing services.

F. Quasi-Experimental Estimates of the Effect of NPL Status on Total Population and Demographics?

Table 8 estimates IV models for a series of housing supply, population, and demographic variables measured in 2000. We continue to focus on the tracts containing the hazardous waste sites. The intent is to determine whether individuals sort in response to changes in local environmental amenities, or specifically the Superfund clean-ups. We report the parameter estimate and standard errors on the

³⁴ We also estimated a version of these models for housing prices. The column (2) NPL Only parameter estimates (standard errors) for 1990 and 2000 are 0.008 (0.066) and 0.062 (0.059), respectively.

³⁵ The point estimates (standard errors) from the columns (2), (3), and (4) specifications in the 2000 housing price analogs to the Table 7 regressions are 0.064 (0.029), 0.019 (0.060), and 0.042 (0.042). Thus, these regressions also fail to provide much evidence that Superfund clean-ups have a substantial impact on housing prices.

dummy for NPL status by 2000. As in the earlier analysis, this variable is instrumented with an indicator for a HRS score above 28.5.

The first specification simply adjusts for the 1980 value of the dependent variable. The column (2) specification adds state fixed effects. The column (3) specification adds a quadratic in the 1982 HRS score and in column (4) the smaller regression discontinuity sample is used. To keep the interpretation simple, the estimates are not adjusted for the wide range of housing characteristics, economic conditions and demographics as was done in the price and rental rate regressions.

Panel A tests for an association between NPL status and the total number of housing units and population in tracts containing the sites. The results provide mixed evidence on total housing units. There is modest evidence of a relative increase in population in the tracts with NPL sites. The estimates in the regression discontinuity specifications are poorly determined, but the point estimates are similar to those in the other specifications. Figure 9 provides an opportunity to examine the population results graphically. It is evident that tracts with HRS scores exceeding 28.5 have higher population growth but there is only weak evidence of a discontinuous jump in population growth at the regulatory threshold.

Panels B and C examine whether the NPL designation is associated with sorting or migration that leads to changes in the income and demographic characteristics of local residents. The first two specifications suggest that household income increased more in NPL tracts and that the fraction of households receiving public assistance declined. The regression discontinuity specifications, however, are less supportive of this conclusion. It appears that the fraction of Blacks or Hispanics did not change in a meaningful way. Although the clean-up of Superfund sites is intended to reduce environmental risks for local residents, especially children, the fraction of the population under the age of 6 is unchanged.

VI. Interpretation

This paper's primary finding is that the NPL designation and eventual clean-up of hazardous waste sites have small effects on the value of housing services in the areas proximate to these sites. In light of the substantial resources devoted to these clean-ups and the claims of large health benefits, this

finding may be considered surprising. We review a number of potential explanations here and assess their credibility.

First, the quasi-experimental estimates are downward biased. Our view is that the wealth of evidence presented here contradicts this view. The quasi-experiment relies on an arbitrary and discontinuous rule that was unknown until sites' initial HRS scores were in the books. Additionally, the observable predictors of housing price growth are well balanced in the tracts with sites that have scores above and below the regulatory threshold. Further, we have presented numerous regression specifications and none of them suggest that individuals living in close proximity to these sites place a large valuation on these clean-ups. Although it remains a possibility that the results are downward biased by an unobserved form of heterogeneity, there is little in the way of concrete evidence to support this view.

Second, a popular theory is that sites are permanently stigmatized by placement on the NPL. A straightforward implication of this theory is that the value of housing services should decline upon a site's placement on the NPL. The results on rental rates in Table 7 contradict this implication and, in turn, the validity of this explanation.

Third, there is heterogeneity in individuals' valuations of the risks associated with proximity to hazardous waste sites and the individuals that choose to live near these sites have a low willingness to pay to avoid the associated health risks. This could be the case either because of differences in income or heterogeneity in preferences. In this case, the paper's estimate would be uninformative about the population average willingness to pay. It is important to note, however, that from a policy perspective the crucial parameter is the willingness to pay of the population that lives near these sites and this is the parameter that the paper has tried to estimate.

Fourth, individuals have imperfect information. This imperfect information could take the form of ignorance about proximity to sites or unawareness of the reduction in health risks associated with clean-up. It is unlikely that individuals are unaware of the sites, because local newspapers often provide extensive coverage when a site is placed on the NPL (or even considered for placement) and/or when remediation is completed. Further, many states (e.g., New Jersey) require home sellers to disclose whether there are hazardous waste sites in close proximity to a residence.

It seems reasonable to presume that individuals do not perfectly understand the health benefits of Superfund clean-ups. As discussed above, scientists are actively studying the health consequences of exposure to the thousands of toxics present at Superfund sites. However, laboratory evidence on the human (rather than mouse or rat) health effects of these toxics may not be available any time soon since scientific review boards reasonably place many safeguards on the study of human health. Put another way, the scientific uncertainty surrounding the toxics at many Superfund sites makes it difficult for individuals to correctly assess the health benefits of clean-ups.³⁶

In our view, these last two explanations are the most plausible ones for the paper's finding that Superfund clean-ups have little effect on the valuation of local housing services. The former explanation would suggest that society devotes too many resources to these clean-ups and that it may be prudent to pursue less expensive options. The latter explanation suggests that the benefits of these clean-ups may be substantially larger than the application of this quasi-experiment to the housing market has shown. To make rational decisions about the value of Superfund clean-ups, it is important to develop a better understanding of their health benefits. This is an important area for future research.

VII. Conclusions

This paper has exploited knowledge of the selection rule that determined the first 400 hundred hazardous waste sites that were to be cleaned up under the Superfund program to estimate individuals' valuations of proximity to hazardous waste sites. In the program's first year, the selection rule was based on a nonlinear function of a continuous and noisy measure of risk, which allows for the implementation of a quasi-experimental regression discontinuity design. The application of this approach to a new data set suggests that nearly two decades after these sites became eligible for federal clean-up, the property values

³⁶ One caveat to this line of reasoning is that although there is uncertainty about the magnitude of the benefits of clean-ups, it is not evident why individuals would incorrectly assume they are small rather than large. It is possible that some individuals have difficulty assessing low probability events (e.g., a small increase in the probability of contracting cancer in the distant future) and just assume that the clean-ups induce no change in risk. However, if this is a mistake, it is unclear why all individuals would make it. Markets allow for heterogeneity in beliefs and presumably the individuals that believe that clean-ups have large health benefits should purchase houses from the individuals that believe the opposite.

in surrounding areas had increased by an economically small and statistically indistinguishable from zero amount, relative to areas surrounding sites that narrowly missed eligibility for a Superfund clean-up. Interestingly, the quasi-experimental results contrast with estimates from a conventional approach of linear adjustment that finds a substantial positive effect on local housing prices. Overall, the preferred quasi-experimental estimates suggest that the benefits of Superfund clean-ups as measured through the housing market are substantially lower than \$40 million, our best estimate of the mean costs of a Superfund clean-up.

DATA APPENDIX

I. Assignment of HRS Scores

The HRS test scores each pathway from 0 to 100, where higher scores indicate greater risk.³⁷ Each pathway score is capped at 100.³⁸ The individual pathway scores are calculated using a method that considers characteristics of the site as being included in one of three categories: waste characteristics, likelihood of release, and target characteristics. The final pathway score is a multiplicative function of the scores in these three categories. The logic is, for example, that if twice as many people are thought to be affected via a pathway then the pathway score should be twice as large.

The final HRS score is calculated using the following equation:

$$(1) \text{ HRS Score} = [(S_{\text{gw}}^2 + S_{\text{sw}}^2 + S_{\text{a}}^2) / 3]^{1/2},$$

where S_{gw} , S_{sw} , and S_{a} , denote the ground water migration, surface water migration, and air migration pathway scores, respectively. As equation (1) indicates, the final score is the square root of the average of the squared individual pathway scores. It is evident that the effect of an individual pathway on the total HRS score is proportional to the pathway score.

It is important to note that HRS scores can't be interpreted as strict cardinal measures of risk. A number of EPA studies have tested how well the HRS represents the underlying risk levels based on cancer and non-cancer risks.³⁹ The EPA has concluded that the HRS test is an ordinal test but that sites with scores within 2 points of each pose roughly comparable risks to human health (EPA 1991).⁴⁰

II. Primary Samples

We have two primary samples. The first sample includes sites that were placed on the National Priority List (NPL) before January 1, 2000. There are 1,436 sites in this sample. The second sample is all sites that were tested between 1980 and 1982 for inclusion on the initial National Priority List announced on September 8, 1983.

A. All NPL Sample

The all NPL sample only includes National Priority List sites located in US states and does not include sites that were proposed for but not listed on the NPL before January 1, 2000. As noted in the text, we use census tract data from the 1980, 1990, and 2000 year US Census reports. Although there are NPL sites located in US territories such as Puerto Rico, we do not include these in the sample because the same census data are not available for US territories. Further we only include sites in the sample that were listed on the NPL before January 1, 2000 to ensure that site listing occurred before any data collection for the 2000 census.

B. 1982 HRS Sample

The second sample consists of sites tested for inclusion on the initial NPL published on September 8, 1983. 690 sites were tested for inclusion on this list. As noted in the text, not all sites tested between 1980 and 1982 were placed on the first NPL list due to initial HRS scores below 28.5.

³⁷ See the EPA's *Hazard Ranking System Guidance Manual* for further details on the determination of the HRS score.

³⁸ The capping of individual pathways and of attributes within each pathway is one limiting characteristic of the test. There is a maximum value for most scores within each pathway category. Also, if the final pathway score is greater than 100 then this score is reduced to 100. The capping of individual pathways creates a loss of precision of the test since all pathway scores of 100 have the same effect on the final HRS score but may represent different magnitudes of risk.

³⁹ See Brody (1998) for a list of EPA studies that have examined this issue.

⁴⁰ The EPA states that the HRS test should not be viewed as a measure of "absolute risk", but that "the HRS does distinguish relative risks among sites and does identify sites that appear to present a significant risk to public health, welfare, or the environment" (Federal Register 1984).

Additionally, 12 sites proposed for the NPL on December 30, 1982 were not listed on the first NPL which was issued on September 8, 1983. Specifically, 418 sites were proposed for the NPL, but 406 sites were listed. The difference between the proposed list and the final list is due mostly to the rescoring of sites. The EPA received 343 comments on 217 sites (all of which were proposed NPL sites) that led to score changes in 156 sites. Revised scores for 5 of these sites fell below 28.5. These sites were dropped from the proposed list. Also not included on the 1983 NPL are 7 more sites. These 7 sites were considered “still under consideration” and did not have final rescores available as of September 8, 1983.

Here is a detailed explanation of the difference between the 1982 proposed list and the first NPL issued in 1983:

- (1) Included on the 1982 proposed list and not on 1983 final list
 - a. Sites with a revised HRS score below 28.5:
 1. Crittenden County Landfill (Marion, AR)
 2. Flynn Lumber (Caldwell, ID)
 3. Parrot Road (New Haven, IN)
 4. Phillips Chemical (Beatrice, NE)
 5. Van Dale Junkyard (Marietta, OH)
 - b. Sites “still under consideration”:
 1. Clare Water Supply (Clare, MI)
 2. Electravoice (Buchanan, MI)
 3. Littlefield Township Dump (Oden, MI)
 4. Whitehall Wells (Whitehall, MI)
 5. Kingman Airport Industrial Area (Kingman, AZ)
 6. Airco (Calvert City, KY)
 7. Bayou Sorrel (Bayou Sorrel, LA)
 - c. State priority sites that were dropped:
 1. Plastifax (Gulfport, MS)
 - d. Sites cleaned up by the responsible party before the 1983 NPL:
 1. Gratiot Co Golf Course (St. Louis, MI)
 - e. Sites split into two separate sites:
 1. Vestal Water Supply (Vestal, NY)
- (2) Included on the 1983 final list but not on the 1982 proposed list
 - a. Two separate sites, formally Vestal Water Supply:
 1. Vestal 1-1 (Vestal, NY)
 2. Vestal 4-2 (Vestal, NY)
 - b. Site identified and tested after the 1982 proposed list:
 1. Times Beach (Times Beach, MO)

Note that 5 of the 7 “still under consideration” sites (Airco, Bayou, Clare, Electravoice, and Whitehall Wells) were later added to the NPL. All five sites had score changes (3 revised upward, 2 revised downward). Two sites (Littlefield, Kingman) were never listed on the NPL. These sites would have had scores that dropped below 28.5. For consistency, we included the score changes for the 5 sites that were later placed on the NPL under the 1983 score variable in the dataset. However, as described above, these scores were not actually released along with the other score changes in 1983.

Changes to site status for the sites in (1)c-(1)e, (2)a, and (2)b above did affect our sample. Gratiot Co Golf Course (1)d was remediated before publication of the final NPL and therefore dropped from our sample. The original Vestal Water Supply (1)e split into 2 sites, with Vestal 4-2 retaining all of the original attributes of the site. We therefore considered Vestal 4-2 as a continuation of the original site. Vestal 1-1 is not included in our sample as there is no 1982 score associated with this site. Likewise Times Beach (2)b is not included in our sample since there is no 1982 score. Plastifax (1)c received a 1982 score that would not have qualified the site for remediation. The site remains in the sample as would any other site that scored below 28.5.

Finally, as discussed in the text, we use the 1982 HRS score as an instrument for NPL status. Therefore, the score changes do not effect how we treat each site provided the site received an initial score for the 1982 proposed list.

III. Site Size Variable

The size of site is taken directly from EPA documentation on the site. Note that there are two sources for the "actual physical size" of a superfund site. Both sources are from the EPA's on-line CERCLIS system. One is extracted from the Fact Sheet and the other from the Site Narrative. They differ for numerous sites and in these cases we used the average of the two sources. If only one was available we used that one.

Note that is sometimes the case that the site size provided in CERCLIS refers to the area of source of the contamination and not the size of the site. There are relatively few sites that are described this way. To maintain consistency in how we interpret a site's size we excluded these data from our primary data file and indicated the 'actual-size' as missing. Further, there are some sites for which there is no size data available in CERCLIS. It is possible that we may be able to fill in size data for some of these sites using the original HRS scoring sheets. We have requested many of these sheets from the EPA docket center via a Freedom of Information Request.

Finally, sometimes the source of contamination is described as being just one part of the entire site. For example, the description for superfund site "NJD980505424" says that there is a 57 acre landfill on 144 acres of property. For this site and others described similarly, we considered the physical size of the site to be 57 acres.

IV. Note on Construction Complete Site Data

The Construction Complete variable was created by the EPA to gauge progress on remediating National Priority List sites. The EPA first began using this designation in 1990. However, it is evident from the remedial action site histories and a construction complete data file we received from the EPA that construction complete definitions were retroactively assigned to pre-1990 sites.

There are 4 sites in our samples which have been deleted but for which no Construction Complete dates are provided. The site ids for these sites are ARD059636456, KSD007241656, MND000686071, and WID990829475. These sites were all deleted between 1990 and 2000 and since construction complete is a precursor for deletion we assign each site a 1 for con_comp, the variable that tracks whether the site was construction complete before year 2000. Based on Table 1 summary statistics, sites are—on average—listed as Construction Complete approximately 1 year before deletion. We assign proxy Construction Complete dates for these 4 sites that are 1 year before they were deleted. We then use these to assign values to the 1990 and 2000 Construction Complete dummy variables.

V. Notes on No Further Action Rods and Total Rods

A. Number of Operating Units (OUs), Operating Units with Expected Cost Data, and Proportion of Operating Units with Cost Data

Data on the number of operating units at each site and the number with associated expected costs were collected from CERCLIS for each National Priority List site. These variables are used as one measure to determine whether all expected costs have been incurred at each site. Important to note, however, is that there at least 3 reasons why the number of OU's at a site may not correspond to the number of RODs for the site:

1. If remedial action at the site is still in progress and there are OU's that haven't been specifically addressed by a ROD. This is the initial reason that the data were collected. We wanted to be sure that we were capturing all expected costs (i.e. from each part of the site).
2. Occasionally a site will be Construction Complete or Deleted without having a specific ROD for each anticipated OU. The most likely explanation is that remedial action for the

other OU's or the removal activity for another OU or for that OU addressed the problem.

3. Sometimes there is more than one ROD for the same OU. This is different from having an amended ROD. For an amended ROD, the old remedial action in the 1st ROD is changed. In the cases where there are 2 RODs for the same OU, there is additional remedial activity on that OU.

B. No Further Action RODs and No Further Action Sites

A list of No Further Action (NFA) RODs by Operating Unit was provided to us by Kate Probst at Resources for the Future. A NFA ROD is a ROD that is developed with the anticipation that an Operating Unit would require remedial action, but where the EPA determined that no remedial action was still needed.

One implication of sites with NFA RODs is the possibility that there is never any remedial action cleanup at some National Priority List sites. This would occur if all of the RODs at a particular site were NFA RODs. In such a case, a site could be deleted from the National Priority List without ever having a remedial cleanup under the Superfund program.

We cross-checked the list of No Further Action RODs provided to us by Resources for the Future with the EPA Cerclis database. The purpose of this was to verify that (1) these RODs were NFA RODs and (2) to determine the total number of RODs at each site. If all of the RODs for a site were NFA RODs then we defined the site as being a 'No Further Action Site'. It is important to note that given considerations (a)-(c) above that the number of RODs at a site may not correspond to the number of operating units. For NFA RODs we are specifically interested in the total number of RODs and not the total number of OU's.

Finally, just because a ROD is a NFA ROD it doesn't imply that there was no cleanup activity at the OU associated with the ROD. Occasionally an operating unit will have more than one ROD. Also, it is often the case that there were separate removal activities or remedial actions carried out under a consent agreement before the ROD was published. In such cases cleanup work was done at the site and more specifically at the part of the site defined under that OU. There were significant costs associated with these activities at many of the sites. Many of the NFA RODs actually detailed costs in the millions and included operating and maintenance activities similar to those of non-NFA RODs.

Whether or not 'No Further Action Sites' would have received the same cleanup considerations as sites that never made it on the NPL is a relevant question for our analysis. If the answer is "No" then these sites may be more similar to the other NPL sites that did have formal remedial action plans. Below is an example from a ROD Abstract for a NFA ROD that highlights this question.

For site: Varsol Spill, FLD980602346

Only 1 OU and would be considered a NFA Site

"The selected remedial action for phase IV Includes adding air stripping to the existing water treatment system in the study area and operating the Miami Springs and Preston Municipal Wells for the dual purpose of providing potable water and recovering contaminated water from the aquifer. Total capital cost for the selected remedial alternative is estimated to be \$5,268,000 with O&M approximately \$334,400 per year."

VI. Measures of Expected and Actual Remediation Costs

We collected data on the expected and actual costs of remediation at each Superfund site in our samples. Here, we describe the differences in these measures of costs and how they were calculated.

A. Expected (Estimated) Costs

The expected cost data is taken directly from the first ROD for each site (note that the EPA refers to these as estimated costs). Each ROD evaluates possible remedial action plans and selects one that

satisfies all relevant national and state requirements for human health and the natural environment. RODs are issued for NPL sites only, so expected costs are unavailable for sites that fail to make it onto the NPL.

Estimated costs include both the remedial action cost and where available the discounted operations and management cost for the selected remedy. The projected time period for these operation and management costs is usually 20-30 years. All estimated costs are adjusted for year 2000 \$'s using the Consumer Price Index.

Many sites have multiple "operating units" or completely separate sections of the site with different Records of Decision. We include estimated costs from each "operating unit" that has a separate Record of Decision. Savannah Silver Site is the site with the greatest number of operating units included in our sample with at least 73. Many of these operating units do not yet have a published Record of Decision with an estimated cost. The vast majority of sites—approximately 90%—have 3 or less operating units.

Note that the Savannah Silver Site highlights a limitation of the expected cost data. Many sites listed on the National Priority List have Records of Decision and expected costs available for some, but not all of the operating units at the site. To guard against the possibility of under-estimating the expected costs at a site we emphasize expected cost data from those sites that are construction complete. It is clear that all Records of Decision would be published for these sites.

Occasionally sites or "operating units" at a site have updated Records of Decision with new estimated costs. These updates are not included as part of the expected costs we present in this paper. Thus, the interpretation of the expected costs in this paper is that they are a projected total cost of site remediation before remedial cleanup action begins at the site. We did calculate expected costs for sites that included all updates from subsequent Records of Decision. Approximately one quarter of the sites have amended cost estimates. These updated costs, on average, are remarkably similar to the expected costs that only include initial cost estimates. For sites with non-missing data in our 1982-3 sample the mean expected costs for the 1st Record of Decision only and all relevant Records of Decision, conditional on construction complete, are 20.6 and 20.3 million respectively. For sites with non-missing data in the all NPL sample these estimates are 15.5 and 14.8 million.

B. Actual Costs

The actual cost data presented in this paper is our best effort to calculate the actual amount spent on remedial action at each site by the EPA, state governments, and responsible parties. As will be explained in greater detail below, the actual cost data comes from 2 sources. The first source is a complete history of all EPA costs summarized by year and site. These data are from the IFMS database and were provided to us by the financial/accounting department at the federal EPA office. The second source is a document called *Enforcement 3*, also obtained from the accounting department of the national EPA, which estimates all potential responsible party (i.e. private party) costs for each National Priority List site. These potential responsible party (PRP) costs are estimates by EPA engineers of remedial action expenses paid directly by companies and individuals. These costs are not reimbursements to the EPA or another party for work that was already completed. Note that private companies are not required to disclose the actual amount of money spent on remediation efforts. The actual cost data used in this paper is the sum of the EPA actual costs and the PRP estimated costs.

Before explaining in greater detail the data sources used, we should note that we explored the use of two other data sources for actual cost, but we were uncomfortable with the quality of these data. The first source was a data file sent to us by the National EPA office that reportedly included all available actual cost data on National Priority List sites. However, on inspection of this file there were many cases of sites with actual cost amounts of 1, 0, and negative dollar amounts respectively. Our hypothesis is that these data include money reimbursed to the EPA by states and potential responsible parties. This could account for the negative and zero dollar amounts for sites that clearly had remedial action. We are uncertain as to what might explain the arbitrarily low dollar amounts (1, 2, etc.) other than data error. The second source of data we explored using is the "actual cost" figures listed for some National Priority List

sites on the EPA's Superfund website (CERCLIS). On inspection of these cost figures we again found 1, 0, and negative dollar amounts.

Apart from the obvious concerns with the other potential actual cost data sources there are several advantages of the data provided to us by the financial office of the EPA. First, the costs are all listed by site by year. This allows us to adjust all cost expenditures to year 2000 \$'s. Second, the EPA actual cost data include both 'direct' and 'indirect' costs for each site. 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. Third, by including EPA estimates for additional Potential Responsible Party costs we have a more complete accounting of the total costs to remediate each site.

A challenge regarding the actual cost data is how to interpret potential state costs. Initial Superfund legislation required that state governments pay for at least 10% of the remedial costs for sites located in their state along with all of the operations and management costs. The Federal EPA does not track state costs. Conversations with federal EPA personnel have indicated that it is often the case that the Federal EPA pays for the work done at the sites and that the states then reimburse the EPA. This interpretation would be consistent with the fact that the EPA actual cost data file tracks operations and management costs—costs technically supposed to be covered by the states. However, it is likely that there are additional state costs that should be included as part of a state's total actual cost of remediation. It is entirely possible that the actual cost figures presented in this paper under-represent the real actual cost of remediation by approximately 10%. We are currently attempting to contact individual states so as to obtain complete state cost information on all of the sites in our samples. We are also trying to obtain figures for state reimbursement costs from the EPA.

VII. Placement of Hazardous Waste Sites in 2000 Census Tracts

The census tract is used as a unit analysis because it is the smallest aggregation of data that is available in the 1980, 1990 and 2000 US Census. As noted in the text, year 2000 census tract boundaries are fixed so that the size and location of the census tract is the same for the 1980 and 1990 census data. The fixed census tract data were provided by Geolytics, a private company. Information on how the 1980 and 1990 census tracts were adjusted to fit the 2000 census tract boundaries can be found on their website at: www.geolytics.com. An outline of their approach is as follows.

Geolytics mapped 1990 census tracts into 2000 census tracts using block level data. "The basic methodology," the company writes, "was to use the smaller blocks to determine the population-weighted proportion of a 1990 tract that was later redefined as part of a 2000 tract."⁴¹ A 1990 street coverage file was used to weight populations of 1990 blocks included in 2000 census tracts when the 1990 blocks were split among multiple census tracts. The assumption is that local streets and roads served as a proxy for where populations were located. Block level data for 1980 were unavailable. This complicated the mapping of 1980 tracts into 1990 tracts. However, the correspondence between 1980 tracts and 1990 blocks is "very good". As such "splitting a 1980 tract into 1990 tracts had to be done spatially, meaning based solely on the 1990 block to 1980 tract correspondence."⁴²

There are 2 types of hazardous waste sites in our sample—those that were eventually listed on the National Priority List (NPL) and those that have never been listed on the NPL. We placed both types of hazardous waste sites in our sample in a single census tract. The remainder of this section describes the separate procedures we used to determine the year 2000 census tract of location for NPL and non-NPL hazardous waste sites.

⁴¹ Appendix J: Description of Tract Remapping Methodology of Geolytics Data Users' Guide for Neighborhood Change Database (1970-2000), page J3.

⁴² Ibid, page J4.

For the NPL sites, latitude and longitude coordinates are available on the EPA summary page (CERCLIS site page). These coordinates were spot checked against their addresses and site descriptive information. GIS Arc Map software was then used to place these sites in a single census tract.

It is more difficult to place the hazardous waste sites that have never been on the NPL in a single census tract. Our first attempt was to place these sites using a comprehensive file provided to us by the EPA that contained latitude and longitude coordinates for non-NPL sites. However, upon inspection of this file we found numerous errors. Many of our sample sites were placed in different cities, counties, states, or zip codes from the EPA address descriptions provided in CERCLIS and the Federal Register.

In light of the unreliable latitude and longitude data we have used several methods to place these sites. Those sites with complete street address information were placed using a program that converts street addresses to latitude and longitude coordinates. These coordinates were then placed in a census tract using GIS Arc Map software.

Those non-NPL sites with missing or incomplete addresses were the most difficult sites to place. We requested original Hazardous Ranking System (HRS) documents on all of these sites from the Federal Register. The HRS documents are the first comprehensive documents prepared for each site. These documents often contain more detailed descriptive information on the sites. Some HRS documents also contain maps showing a site's location in the surrounding community. Many of these sites could be placed by hand using the more detailed descriptive and location information contained in the HRS documents and an electronic US Census map.

We called regional and state EPA officials regarding all non-NPL sites for which we were not able to place with certainty using either CERCLIS information or the HRS scoring packages. For most of these sites we were able to locate someone with 1st hand knowledge of the site or who was able to provide us with either a map of the site, a more complete address, or more accurate latitude longitude coordinates. Ultimately, we succeeded in obtaining longitude and latitude data for all but 4 of these sites. However, we were able to place these sites in census tracts.

VIII. Neighbor Samples

Each superfund site in our sample is placed in a single census tract. We have been able to exactly place all but 4 of the sites in the 1982 HRS Sample. We use two approaches to define the set of houses outside the sites' tract that may be affected by the clean-up. We refer to this set as "neighbors."

The first approach defines the neighbors as all census tracts that share a border with the tract that contains the site. GIS software was used to find each primary census tract and extract the identity of its adjacent neighbors. In the 1982-3 sample the maximum number of neighboring census tracts is 21 and the median is 7. The population of each adjacent census tract was used to weight the housing price, housing characteristics, and demographic variables for each tract when calculating the mean adjacent neighbor values.

The second approach defines neighbors based on distance 'rings' around the exact location of the site. GIS software is used to draw a 'ring' around the point representing the site (generally the center of the site, but sometimes the point used as a street address). For example, in the 1 mile sample, a GIS program uses a 1 mile radius to draw a circle around the site at a distance of 1 mile. Data from all census tracts that fall within this 'ring' are used in calculating the mean housing values and housing and demographic characteristics for all housing within 1 mile of the site. Each census tract is weighted by the product of the population and the portion of the total area of each census tract that falls within the 'ring.' The maximum number of census tracts included in the 1 mile ring for a site is 37 and the median is 3. For the 2 mile ring the maximum number of neighbor sites is 80, with a median of 5.

REFERENCES

- Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2000. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." NBER Working Paper No. 7831.
- Ashenfelter, Orley and Michael Greenstone. 2004a. "Estimating the Value of a Statistical Life: The Importance of Omitted Variables and Publication Bias." *American Economic Review*, Vol. 94.
- Ashenfelter, Orley and Michael Greenstone. 2004b. "Using Mandated Speed Limits to Measure the Value of a Statistical Life." *Journal of Political Economy*, Vol. 112.
- Banzhaf, H. Spencer and Randall P. Walsh. 2005. "Do People Vote with Their Feet: An Empirical Test of Environmental Gentrification." Mimeograph.
- Black, Sandra. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education," *Quarterly Journal of Economics*. Vol. 114.
- Cameron, Trudy Ann and Ian T. McConnaha. 2005. "Evidence of Environmental Migration." Mimeograph.
- Chay, Kenneth Y., and Michael Greenstone. 2005. "Does Air Quality Matter? Evidence from the Housing Market." *Journal of Political Economy*, Vol. 113.
- Cook, Thomas D., and Donald T. Campbell. 1979. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Boston, MA: Houghton Mifflin..
- Deschenes, Olivier and Michael Greenstone. 2004. "The Economic Impacts of Climate Change: Evidence from Agricultural Profits and Random Fluctuations in Weather." NBER Working Paper No. 10663.
- Ekeland, Ivar, James J. Heckman, and Lars Nesheim (2004): "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, Vol. 112.
- Environmental Protection Agency. 2000. Superfund: 20 Years of Protecting Human Health and the Environment. Available at: <http://www.epa.gov/superfund/action/20years/index.htm>.
- Farrell, Alex. 2004. "A Partial Benefit-Cost Test of NPL Site Remediation: Benefit Transfer with Meta-Analytic Hedonic Data." University of California, Berkeley, Mimeograph.
- Federal Register Rules and Regulations. September 21, 1984. Section IV "Process for Establishing the National Priorities List." Vol. 49, No. 185, pp. 37070-37082.
- Federal Register Rules and Regulations. DATE TO COME. *National Oil and Hazardous Substance Contingency Plan; The National Priorities List; Amendment*.
- Freeman, A Myrick, III. 1974. "On Estimating Air Pollution Control Benefits from Land Value Studies." *Journal of Environmental Economics and Management*. 1(1): pp. 74-83.
- Gayer, Ted, James T. Hamilton, and W. Kip Viscusi. 2000. "Private Values of Risk Tradeoffs at Superfund Sites: Housing Market Evidence on Learning About Risk." *The Review of Economics and Statistics*. 82(3): pp. 439-51.

Gayer, Ted, James T. Hamilton, and W. Kip Viscusi. 2002. "The Market Value of Reducing Cancer Risk: Hedonic Housing Prices with Changing Information." *Southern Economic Journal*. 69(2): pp. 266-289.

Hamilton, James T. and W. Kip Viscusi. 1999. Calculating Risks: The Spatial and Political Dimensions of Hazardous Waste Policy. Cambridge, Mass: The MIT Press.

Freeman III, A. Myrick. 1993. The Measurement of Environmental and Resource Values. Washington, D.C.: Resources for the Future.

Champ, Patricia A., Kevin J. Boyle and Thomas C. Brown (Eds.). 2003. A Primer on Nonmarket Valuation. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Harris, John D. 1999. "Property Values, Stigma, and Superfund." Environmental Protection Agency. Available at: <http://www.epa.gov/superfund/programs/recycle/property.htm>

Kiel, Katherine A. 2001. "Estimating the Economic Benefits of Cleaning Up Superfund Sites: The Case of Woburn, Massachusetts." *The Journal of Real Estate Finance & Economics*. 22 (2): pp. 163-84.

McCluskey, Jill J. and Gordon C. Rausser. 2003. "Hazardous Waste Sites and Housing Appreciation Rates." *Journal of Environmental Economics and Management*, 45: 166-176.

Messer, Kent, William Schulze, Katherine Hackett, Trudy Cameron, and Gary McClelland. 2004. "Stigma: The Psychology and Economics of Superfund." Working Paper.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*. 82: 34-55.

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

Viscusi, W. Kip. 1993. "The Value of Risks to Life and Health." *Journal of Economic Literature* 31: 1912-1946.

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

REFERENCES INCOMPLETE MORE TO COME

Table 1: Summary Statistics on the Superfund Program

	All NPL Sites w/ non-Missing House Price Data (1)	1982 HRS Sites w/ non-Missing House Price Data (2)	1982 HRS Sites w/ Missing House Price Data (3)
Number of Sites	985	487	189
1982 HRS Score Above 28.5	-----	306	95
<u>A. Timing of Placement on NPL</u>			
Total	985	332	111
# 1981-1985	406	312	97
# 1986-1989	340	14	9
# 1990-1994	166	4	3
# 1995-1999	73	2	2
<u>B. HRS Information</u>			
Mean Scores HRS \geq 28.5	41.89	44.47	43.23
Mean Scores HRS < 28.5	-----	15.54	16.50
<u>C. Size of Site (in acres)</u>			
Number of sites with size data	920	310	97
Mean (Median)	1,186 (29)	334 (25)	10,507 (35)
Maximum	195,200	42,560	405,760
<u>D. Stages of Clean-Up for NPL Sites</u>			
<u>Median Years from NPL Listing Until:</u>			
ROD Issued	-----	4.3	4.3
Clean-Up Initiated	-----	5.8	6.8
Construction Complete	-----	12.1	11.5
Deleted from NPL	-----	12.8	12.5
<u>1990 Status Among Sites NPL by 1990</u>			
NPL Only	394	100	31
ROD Issued or Clean-up Initiated	335	210	68
Construction Complete or Deleted	22	16	7
<u>2000 Status Among Sites NPL by 2000</u>			
NPL Only	137	15	3
ROD Issued or Clean-up Initiated	370	119	33
Construction Complete or Deleted	478	198	75
<u>E. Expected Costs of Remediation (Millions of 2000 \$'s)</u>			
# Sites with Nonmissing Costs	753	293	95
Mean (Median)	\$28.3 (\$11.0)	\$27.5 (\$15.0)	\$29.6 (\$11.5)
95 th Percentile	\$89.6	\$95.3	\$146.0
<u>F. Actual and Expected Costs Conditional on Construction Complete (Millions of 2000 \$'s)</u>			
Sites w/ Both Costs Nonmissing	477	203	69
Mean (Median) Expected Costs	\$15.5 (\$7.8)	\$20.6 (\$9.7)	\$17.3 (\$7.3)
Mean (Median) Actual Costs	\$21.6 (\$11.6)	\$32.0 (\$16.2)	\$23.3 (\$8.9)

Notes: All dollar figures are in 2000 \$'s. Column (1) includes information for sites placed on the NPL before 12/31/99. The EPA's 1st Record of Decision for each "operating unit" at a site is the source of the estimated cost information.

Table 2: Mean Census Tract Characteristics by Categories of the 1982 HRS Score

	NPL Site by 2000 (1)	No NPL Site by 2000 (2)	HRS < 28.5 (3)	HRS > 28.5 (4)	HRS > 16.5 & < 28.5 (5)	HRS > 28.5 & < 40.5 (6)	P-Value (1) vs. (2) (7)	P-Value (3) vs. (4) (8)	P-Value (5) vs. (6) (9)
# Census Tracts	985	41989	181	306	90	137	42974		
<u>Superfund Clean-up Activities</u>									
Ever NPL by 1990	0.7576	-----	0.1271	0.9902	0.2222	0.9854	-----	0.000	0.000
Ever NPL by 2000	1.0000	-----	0.1602	0.9902	0.2667	0.9854	-----	0.000	0.000
<u>Housing Prices</u>									
1980 Mean	58,058	69,904	45,027	52,137	46,135	50,648	0.000	0.000	0.084
1990 Median	100,102	99,552	80,185	96,752	84,461	91,611	0.839	0.005	0.433
2000 Median	139,066	151,712	115,479	135,436	117,528	123,503	0.000	0.001	0.449
<u>Neighbor Housing Prices</u>									
1980 Mean	61,124	-----	49,324	62,711	48,735	54,078	-----	0.038	0.047
1990 Median	101,011	-----	85,476	98,743	85,793	88,743	-----	0.009	0.653
2000 Median	143,805	-----	122,337	140,841	122,776	131,125	-----	0.002	0.277
<u>Demographics & Economic Characteristics</u>									
Population Density	1,406	5,786	1,670	1,157	1,361	1,151	0.000	0.067	0.570
% Black	0.0914	0.1207	0.1126	0.0713	0.0819	0.0844	0.000	0.037	0.926
% Hispanic	0.0515	0.0739	0.0443	0.0424	0.0309	0.0300	0.000	0.841	0.928
% Under 18	0.2940	0.2779	0.2932	0.2936	0.2885	0.2934	0.000	0.958	0.568
% Female Head HH	0.1617	0.1934	0.1879	0.1576	0.1639	0.1664	0.000	0.017	0.862
% Same House 5 Yrs Ago	0.5442	0.5127	0.6025	0.5623	0.5854	0.5655	0.000	0.001	0.244
% > 25 No HS Diploma	0.3425	0.3144	0.4053	0.3429	0.3881	0.3533	0.000	0.000	0.060
% > 25 BA or Better	0.1390	0.1767	0.1003	0.1377	0.1092	0.1343	0.000	0.000	0.036
% < Poverty Line	0.1056	0.1141	0.1139	0.1005	0.1072	0.1115	0.003	0.109	0.716
% Public Assistance	0.0736	0.0773	0.0885	0.0745	0.0805	0.0755	0.083	0.041	0.578
Household Income	20,344	21,526	19,635	20,869	19,812	20,301	0.000	0.013	0.486
<u>Housing Characteristics</u>									
Total Housing Units	1,392	1,350	1,357	1,353	1,367	1,319	0.040	0.951	0.575
% Owner Occupied	0.6821	0.6125	0.6792	0.6800	0.6942	0.6730	0.000	0.959	0.344
% 0-2 Bedrooms	0.4480	0.4722	0.4691	0.4439	0.4671	0.4496	0.000	0.103	0.417
% 3-4 Bedrooms	0.5246	0.5015	0.5098	0.5284	0.5089	0.5199	0.000	0.209	0.586
% Built Last 5 Years	0.1434	0.1543	0.1185	0.1404	0.1366	0.1397	0.006	0.050	0.844
% Built Last 10 Years	0.2833	0.2874	0.2370	0.2814	0.2673	0.2758	0.501	0.012	0.723
% No Air Conditioning	0.4904	0.4220	0.5058	0.4800	0.5157	0.5103	0.000	0.253	0.870

% with Zero Full Baths	0.0254	0.0229	0.0315	0.0259	0.0339	0.0290	0.011	0.089	0.386
% Units Detached	0.8764	0.8773	0.8585	0.8908	0.8545	0.8897	0.868	0.050	0.107
% Units Attached	0.0374	0.0754	0.0603	0.0307	0.0511	0.0317	0.000	0.040	0.297
% Mobile Homes	0.0862	0.0473	0.0813	0.0785	0.0944	0.0787	0.000	0.792	0.285

Notes: Columns (1) - (6) report the means of the variables listed in the row headings across the groups of census tracts listed at the top of the columns. In all of these columns, the sample restriction that the census tract must have nonmissing house price data in 1980, 1990, and 2000 is added. Columns (7)-(9) report the p-values from tests that the means in different sets of the subsamples are equal. The Panel title "Neighbor Housing Price" reports the mean housing prices in all tracts that share a border with the tract containing the hazardous waste site. All other entries in the table refer to characteristics of the tract where the site is located. P-values less than .01 are denoted in bold. For the air conditioning and bath questions, the numerator is year round housing units and the denominator is all housing units. For all other variables in the "Housing Characteristics" category, the denominator is all housing units.

Table 3: Estimates of the Association Between the Presence of a NPL Hazardous Waste Site and the ln of Median Census Tract Housing Prices, 1990 and 2000

	(1)	(2)	(3)	(4)
<u>A. Own Census Tract</u>				
<u>ln (1990 Median Price)</u>				
1(NPL Status by 1990)	0.123 (.023)	0.155 (0.021)	0.164 (0.020)	0.098 (0.018)
<u>ln (2000 Median Price)</u>				
1(NPL Status by 2000)	0.040 (0.012)	0.046 (0.011)	0.073 (0.010)	0.066 (0.009)
<u>B. Adjacent Census Tracts</u>				
<u>ln (1990 Median Price)</u>				
1(NPL Status by 1990)	0.157 (0.018)	0.149 (0.016)	0.173 (0.014)	0.108 (0.011)
<u>ln (2000 Median Price)</u>				
1(NPL Status by 2000)	0.041 (0.011)	0.061 (0.010)	0.092 (0.009)	0.088 (0.008)
1980 Prices	Yes	Yes	Yes	Yes
1980 Housing Char's	No	Yes	Yes	Yes
1980 Economic Conditions	No	No	Yes	Yes
1980 Demographics	No	No	Yes	Yes
State Fixed Effects	No	No	No	Yes

Notes: The table reports results from 16 separate regressions where the unit of observation is a census tract. The sample size is 42,974 in all regressions. In Panel A, 746 (985) observations in 1990 (2000) are from census tracts that contain a hazardous waste site that had been on the NPL at any time prior to observation on housing prices. In Panel B, the observations from tracts containing the NPL sites are replaced with observations comprised of averages of all variables across the census tracts with complete housing price data that share a border with the tracts containing the sites. The dependent variables are underlined in the first column. The entries report the coefficient and heteroskedastic-consistent standard error (in parentheses) on the NPL indicator. The controls are listed in the row headings at the bottom of the table. See the text and Data Appendix for further details.

Table 4: Estimates of the First-Stage Relationship and an Informal Validity Test

	(1)	(2)	(3)	(4)	(5)	(6)
<u>A. First Stage:</u>						
<u>Association Between Instrument and NPL Status</u>						
<u>1(NPL Status by 1990)</u>						
1(1982 HRS Score > 28.5)	0.864 (0.026)	0.859 (0.027)	0.849 (0.028)	0.843 (0.031)	0.670 (0.063)	0.728 (0.057)
<u>1(NPL Status by 2000)</u>						
1(1982 HRS Score > 28.5)	0.827 (0.029)	0.822 (0.029)	0.813 (0.030)	0.801 (0.034)	0.608 (0.066)	0.687 (0.060)
<u>B. Informal Validity Test:</u>						
<u>Association Between Instrument and ln (1980 Mean House Price)</u>						
<u>Own Census Tract</u>						
1(1982 HRS Score > 28.5)	-----	0.035 (0.027)	0.001 (0.021)	-0.010 (0.018)	-0.004 (0.029)	-0.008 (0.021)
<u>Adjacent Census Tracts</u>						
1(1982 HRS Score > 28.5)	-----	0.039 (0.031)	0.016 (0.026)	0.010 (0.022)	-0.003 (0.039)	0.023 (0.026)
1980 Prices (Panel A only)	Yes	Yes	Yes	Yes	Yes	Yes
1980 Housing Char's	No	Yes	Yes	Yes	Yes	Yes
1980 Economic Conditions	No	No	Yes	Yes	Yes	Yes
1980 Demographics	No	No	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	No	Yes	Yes	Yes
Quadratic in 1982 HRS Score	No	No	No	No	Yes	No
Regression Discontinuity Sample	No	No	No	No	No	Yes

Notes: The table reports results from 22 separate regressions. The sample size is 487 in all regressions in columns (1)-(5), which is the number of sites that received 1982 HRS scores and are located in census tracts with non-missing housing price data in 1980, 1990, and 2000. The regression discontinuity sample in column (6) limits the sample to tracts (or neighbors of tracts) with sites with initial HRS scores between 16.5 and 40.5 and totals 227 census tracts. In Panel A, the dependent variables are indicators for whether the hazardous waste sites had been placed on the NPL by 1990 and 2000, respectively. In Panel B, the dependent variables are the ln of 1980 mean housing prices in the tract with the hazardous waste site and the ln of the mean of the mean house price across the census tracts that share a border with the tract containing the NPL site, respectively. The dependent variables are underlined in the first column. The table reports the regression coefficient and heteroskedastic consistent standard error (in parentheses) associated with the indicator variable for whether the hazardous waste site received a 1982 HRS score exceeding 28.5. The controls are listed in the row headings at the bottom of the table. See the text and Data Appendix for further details.

Table 5: Instrumental Variables Estimates of the Effect of NPL Status on House Prices

	(1)	(2)	(3)	(4)	(5)	(6)
<u>A. Own Census Tract</u>						
<u>1990</u>						
1(NPL Status by 1990)	0.034 (0.062)	0.066 (0.057)	0.064 (0.054)	0.020 (0.054)	-0.006 (0.122)	-0.006 (0.078)
<u>2000</u>						
1(NPL Status by 2000)	0.037 (0.035)	0.043 (0.031)	0.056 (0.029)	0.047 (0.027)	0.007 (0.063)	0.027 (0.038)
<u>B. Adjacent Census Tracts</u>						
<u>1990</u>						
1(NPL Status by 1990)	0.068 (0.048)	0.060 (0.039)	0.012 (0.033)	-0.006 (0.027)	-0.005 (0.059)	-0.020 (0.044)
<u>2000</u>						
1(NPL Status by 2000)	0.068 (0.035)	0.019 (0.029)	0.014 (0.025)	0.012 (0.022)	0.021 (0.054)	0.011 (0.035)
1980 Ln House Price	Yes	Yes	Yes	Yes	Yes	Yes
1980 Housing Char's	No	Yes	Yes	Yes	Yes	Yes
1980 Economic Conditions	No	No	Yes	Yes	Yes	Yes
1980 Demographics	No	No	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	No	Yes	Yes	Yes
Quadratic in 1982 HRS Score	No	No	No	No	Yes	No
Regression Discontinuity Sample	No	No	No	No	No	Yes

Notes: The entries report the results from 24 separate regressions, where a census tract is the unit of observation. The samples sizes are 487 in columns (1) through (5) and 227 in column (6). The 1990 and 2000 values of the ln (median house price) are the dependent variables. The variable of interest is an indicator for NPL status and this variable is instrumented with an indicator for whether the tract had a hazardous waste site with a 1982 HRS score exceeding 28.5. The entries are the regression coefficient and heteroskedastic consistent standard errors (in parentheses) associated with the NPL indicator. See the notes to Table 4, the text and the Data Appendix for further details.

Table 6: Further Instrumental Variables Estimates of the Effect of NPL Status on 2000 House Prices

	(1)	(2)	(3)	(4)
<u>A. Add Controls for Ground Water, Surface Water, and Air Migration Pathway Scores</u>				
1(NPL Status by 2000)	-0.010 (0.053)	-0.022 (0.042)	0.013 (0.063)	0.000 (0.068)
<u>B. Add Controls for 1970 Covariates</u>				
1(NPL Status by 2000)	0.037 (0.037)	0.028 (0.033)	-0.036 (0.072)	0.037 (0.056)
<u>C. Add Controls for 2000 Covariates</u>				
1(NPL Status by 2000)	0.037 (0.035)	0.022 (0.022)	-0.044 (0.056)	0.010 (0.033)
<u>D. Does Effect of NPL Status Differ in Tracts in Top Quartile of Population Density?</u>				
1(NPL Status by 2000)	0.078 (0.042)	0.029 (0.032)	-0.001 (0.066)	0.045 (0.048)
1(2000 NPL)*1(Top Quartile Density)	-0.062 (0.077)	0.064 (0.056)	0.060 (0.055)	-0.084 (0.097)
<u>E. Does Effect of NPL Status Differ in Tracts where the Site has Substantially Polluted Ground Water and Households Rely on Well Water?</u>				
1(NPL Status by 2000)	0.019 (0.047)	0.017 (0.035)	-0.002 (0.079)	0.018 (0.053)
1(2000 NPL)*1(Possible Contaminated Water Supply)	-0.007 (0.065)	0.016 (0.057)	0.024 (0.063)	0.011 (0.078)
1980 Ln House Price	Yes	Yes	Yes	Yes
1980 House Char's, Econ Conditions, & Demog's	No	Yes	Yes	Yes
State Fixed Effects	No	Yes	Yes	Yes
Quadratic in 1982 HRS Score	No	No	Yes	No
Regression Discontinuity Sample	No	No	No	Yes

Notes: Each panel reports parameter estimates and standard errors from 4 separate regressions for 2000 housing prices. The panels differ from the Table 5 specifications in the following ways: A. adds controls for the individual pathway scores that are used to calculate the HRS score; B. and C. add controls for the 1970 and 2000 values of the covariates, respectively; D. allows the effect of NPL status to differ in census tracts in the top quartile of population density; and E. allows the effect of NPL status to differ in census tracts where the site has substantially polluted groundwater (i.e., the groundwater pathway score exceeds 28.5) and well water is a significant source of water for households (i.e., the tract is in the top half of well water usage, which is defined as more than 11.8% of households). The sample sizes in Panel B are 353 (cols 1-3) and 157 (col 4) and otherwise are as in Table 5. See the Notes to Table 5 and the text for further details.

Table 7: Two-Stage Least Squares Estimates of Stages of Superfund Clean-ups on Rental Rates

	(1)	(2)	(3)	(4)
<u>Ln (1990 Median Rental Rate)</u>				
1(NPL Only)	0.102	0.019	-0.025	0.037
[107 Sites, Mean HRS = 40.5]	(0.063)	(0.061)	(0.104)	(0.112)
1(ROD & Incomplete Remediation)	0.037	-0.010	-0.062	-0.066
[203 Sites, Mean HRS = 44.8]	(0.048)	(0.050)	(0.095)	(0.074)
1(Const Complete or NPL Deletion)	0.176	0.086	0.051	0.045
[16 Sites, Mean HRS = 35.3]	(0.111)	(0.105)	(0.131)	(0.136)
P-Value from F-Test of Equality	0.25	0.54	0.43	0.42
<u>Ln (2000 Median Rental Rate)</u>				
1(NPL Only)	0.235	0.078	-0.006	0.020
[18 Sites, Mean HRS = 36.6]	(0.107)	(0.059)	(0.083)	(0.109)
1(ROD & Incomplete Remediation)	0.182	0.008	-0.084	-0.137
[116 Sites, Mean HRS = 44.1]	(0.045)	(0.036)	(0.068)	(0.058)
1(Const Complete or NPL Deletion)	0.128	0.046	-0.042	-0.020
[198 Sites, Mean HRS = 42.1]	(0.039)	(0.032)	(0.062)	(0.049)
P-Value from F-Test of Equality	0.28	0.31	0.24	0.09
1980 Prices/Rental Rate	Yes	Yes	Yes	Yes
1980 Housing Char's	No	Yes	Yes	Yes
1980 Economic Conditions	No	Yes	Yes	Yes
1980 Demographics	No	Yes	Yes	Yes
State Fixed Effects	No	Yes	Yes	Yes
Quadratic in 1982 HRS Score	No	No	Yes	No
Regression Discontinuity Sample	No	No	No	Yes

Notes: Here, the indicator variable for NPL status has been replaced by three independent indicator variables. They are equal to 1 for sites that by 1990 (2000) were placed on the NPL but no ROD had been issued, issued a ROD but remediation was incomplete, and were “construction complete” or deleted from the NPL, respectively. The instruments are the interactions of the indicator for a 1982 HRS score above 28.5 and these three independent indicators. The table reports the instrumental variables parameter estimates and heteroskedastic consistent standard errors for the three indicators of clean-up status. The table also reports the p-value associated with a F-test that the three parameters are equal. The sample sizes are 487 in columns (1) - (3) and 227 in column (6). See the text and notes to Table 5 for further details.

Table 8: IV Estimates of 2000 NPL Status on 2000 Population, Wealth, and Demographics

	(1)	(2)	(3)	(4)
<u>A. Housing Units and Population Outcomes</u>				
<u>Total Housing Units</u>	238	213	-154	64
[1980 Mean: 1,354; 2000 -1980 Mean: 463]	(69)	(70)	(186)	(120)
<u>Population</u>	614	491	221	403
[1980 Mean: 3,756; 2000 -1980 Mean: 891]	(188)	(190)	(497)	(333)
<u>B. Residents' Wealth Outcomes</u>				
<u>Household Income</u>	4,923	3,736	-4,358	-90
[1980 Mean: 20,410; 2000 – 1980 Mean: 36,471]	(1,757)	(1,951)	(4,151)	(2,186)
<u>% Public Assistance</u>	-0.016	-0.013	0.012	0.001
[1980 Mean: 0.080; 2000 -1980 Mean: -0.003]	(0.005)	(0.006)	(0.013)	(0.008)
<u>% BA or Better</u>	0.014	0.011	-0.000	0.001
[1980 Mean: 0.124; 2000 -1980 Mean: 0.083]	(0.011)	(0.012)	(0.026)	(0.016)
<u>C. Demographic/Environmental Justice Outcomes</u>				
<u>% Black</u>	-0.022	-0.026	0.040	0.000
[1980 Mean: 0.087; 2000 -1980 Mean: 0.026]	(0.012)	(0.014)	(0.031)	(0.018)
<u>% Hispanic</u>	-0.001	-0.011	-0.018	-0.011
[1980 Mean: 0.043; 2000 -1980 Mean: 0.036]	(0.008)	(0.009)	(0.021)	(0.014)
<u>% Population Under Age 6</u>	-0.004	-0.002	0.001	0.001
[1980 Mean: 0.086; 2000 -1980 Mean: -0.018]	(0.002)	(0.002)	(0.006)	(0.003)
1980 Dependent Variable	Yes	Yes	Yes	Yes
State Fixed Effects	No	Yes	Yes	Yes
Quadratic in 1982 HRS Score	No	No	No	No
Reg Discontinuity Sample	No	No	Yes	Yes

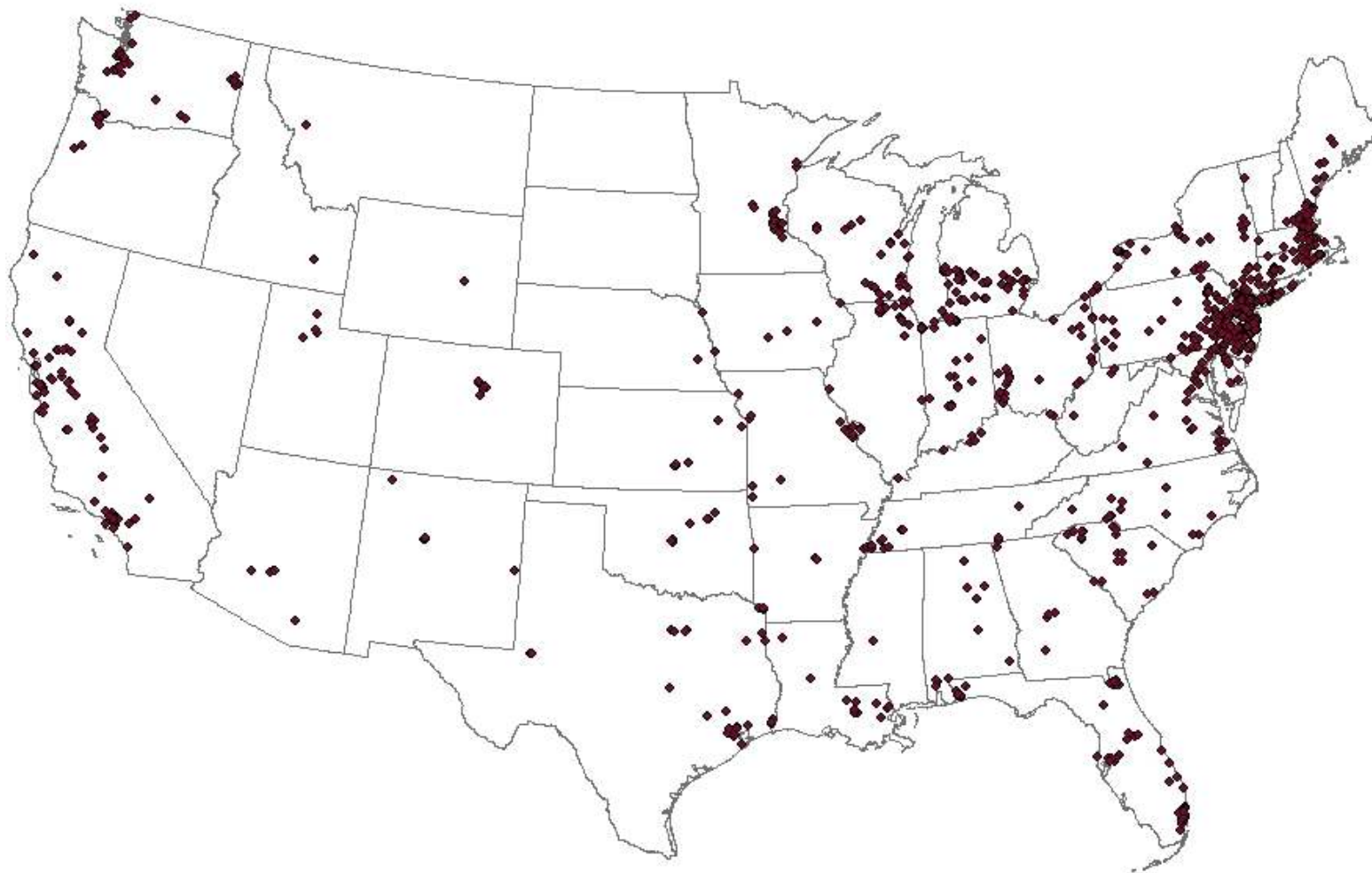
Notes: The entries report the results from 32 separate instrumental variables regressions. The dependent variables are the underlined entries in the first column and are measured in 2000. The table reports the parameter and heteroskedastic consistent standard error associated with an indicator variable that equals 1 for observations from tracts with a hazardous waste site that was placed on the NPL by 2000. The indicator is instrumented with an indicator for whether the hazardous waste site had a 1982 HRS score exceeding 28.5. There are 487 observations in the columns (1) – (3) specifications and 227 in column (4). The entries in brackets are the mean of the dependent variable in 1980 and the mean of the difference between the 2000 and 1980 values. See the text and Notes to Table 5 for further details.

Appendix Table 1: Instrumental Variables Estimates of the Effect of NPL Status on House Prices

	(1)	(2)	(3)	(4)	(5)	(6)
<u>A. Houses within 1 Mile of Hazardous Waste Site</u>						
<u>1990</u>						
1(NPL Status by 1990)	0.035 (0.053)	0.086 (0.048)	0.069 (0.044)	0.010 (0.042)	0.101 (0.087)	0.028 (0.070)
<u>2000</u>						
1(NPL Status by 2000)	0.046 (0.034)	0.046 (0.029)	0.047 (0.027)	0.028 (0.025)	0.065 (0.060)	0.021 (0.036)
<u>B. Houses within 2 Miles of Hazardous Waste Site</u>						
<u>1990</u>						
1(NPL Status by 1990)	0.027 (0.048)	0.056 (0.043)	0.047 (0.039)	-0.009 (0.034)	0.007 (0.070)	-0.037 (0.057)
<u>2000</u>						
1(NPL Status by 2000)	0.023 (0.032)	0.005 (0.029)	0.020 (0.026)	0.007 (0.023)	0.015 (0.052)	-0.007 (0.033)
<u>C. Houses within 3 Miles of Hazardous Waste Site</u>						
<u>1990</u>						
1(NPL Status by 1990)	0.063 (0.049)	0.060 (0.041)	0.038 (0.036)	-0.018 (0.029)	-0.033 (0.061)	-0.037 (0.049)
<u>2000</u>						
1(NPL Status by 2000)	0.054 (0.038)	0.024 (0.031)	0.026 (0.026)	0.001 (0.021)	-0.032 (0.051)	-0.007 (0.034)
1980 Ln House Price	Yes	Yes	Yes	Yes	Yes	Yes
1980 House, Econ & Demog Vars	No	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	No	No	No	Yes	Yes	Yes
Quadratic in 1982 HRS Score	No	No	No	No	Yes	No
Regression Discontinuity Sample	No	No	No	No	No	Yes

Notes: The entries report the parameter estimate and heteroskedastic consistent standard error on NPL indicators from 48 separate regressions. Here, the dependent variable and all controls are calculated as the weighted average across the portion of tracts that fall within circles of radius 1 mile (Panel A), 2 miles (Panel B), and 3 miles (Panel C) around the hazardous waste sites with a 1982 HRS score. The weight is the portion of the census tract that falls within the relevant circle multiplied by the tract's 1980 population. The sample sizes are 483 (columns 1-5) and 226 (column 6) in all Panels. The 1980 aggregate values of the housing stock in the three circles are roughly \$300, \$755, and \$1,410 million (2000 \$'s). See the text for further details.

Figure 1: Geographic Distribution of NPL Hazardous Waste Sites in the All NPL Sample



Notes: The All NPL sample is comprised of the 985 hazardous waste sites assigned to the NPL by January 1, 2000 that we placed in a census tract with nonmissing housing price data in 1980, 1990, and 2000.

Figure 2: Geographic Distribution of Hazardous Waste Sites in the 1982 HRS Sample

A. Sites with 1982 HRS Scores Exceeding 28.5

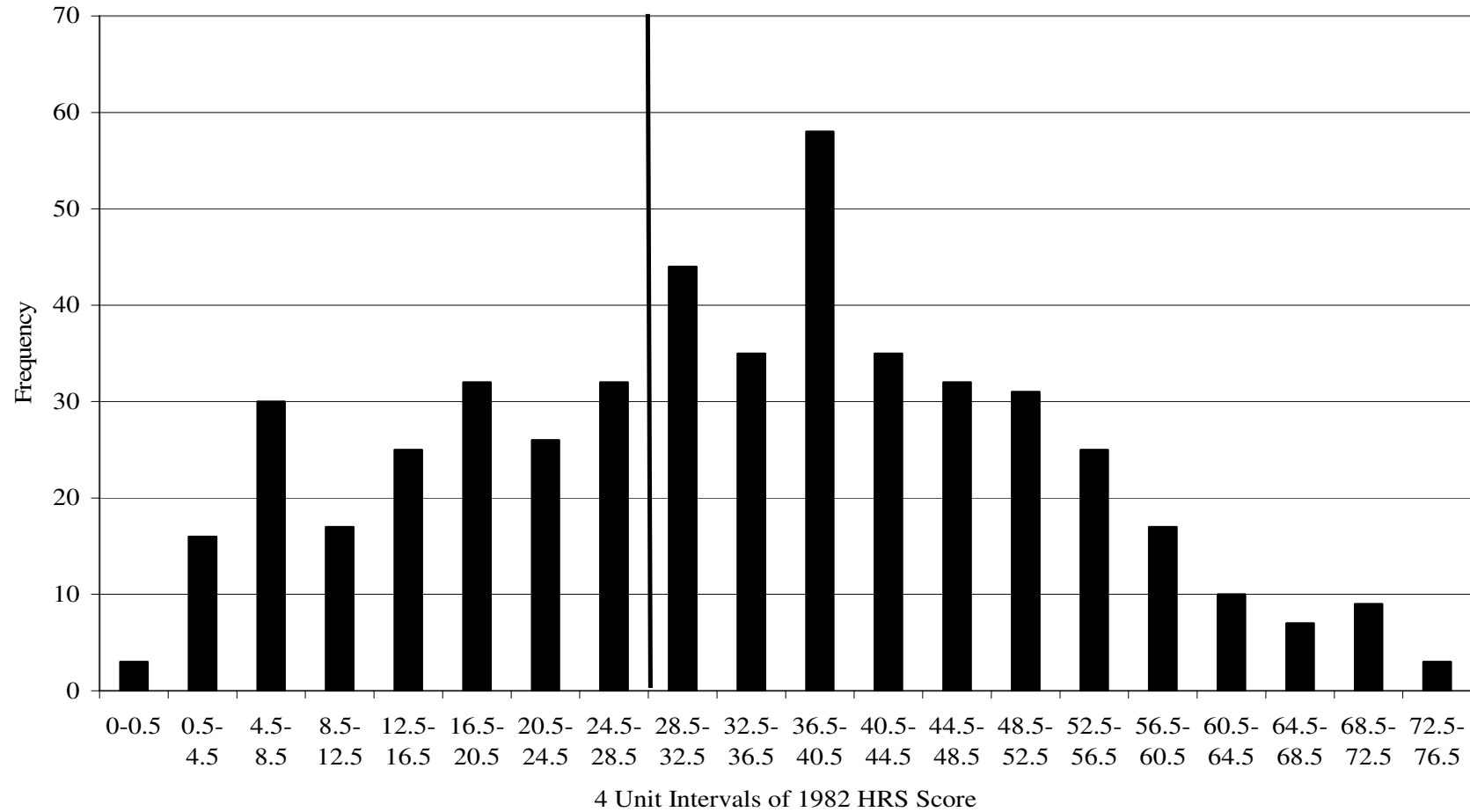


B. Sites with 1982 HRS Scores Below 28.5



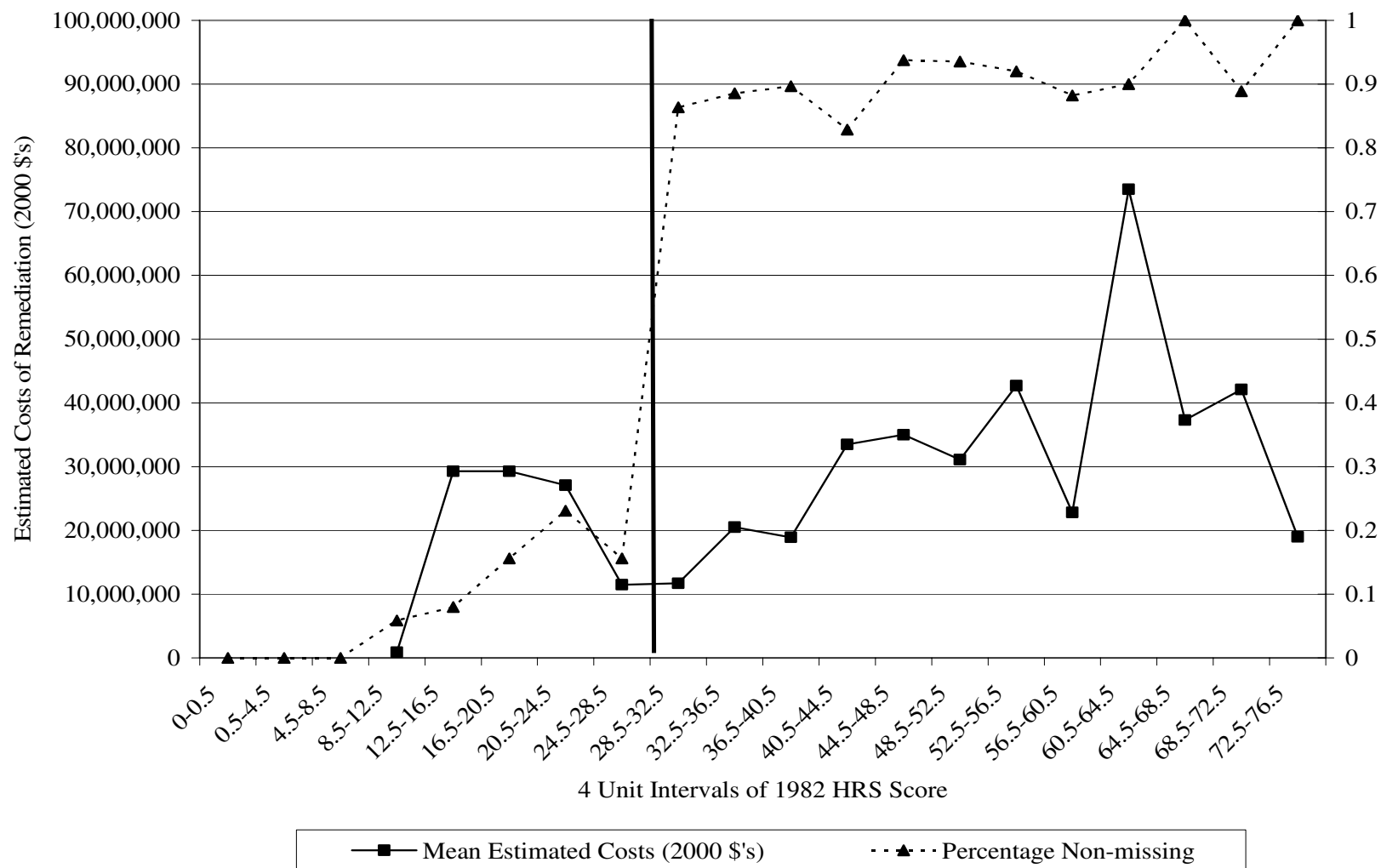
Notes: The 1982 HRS Sample is comprised of the 487 hazardous waste sites that were placed in a census tract with nonmissing housing price data in 1980, 1990, and 2000. 306 (181) of these sites had 1982 HRS scores above (below) 28.5.

Figure 3: Distribution of 1982 HRS Scores



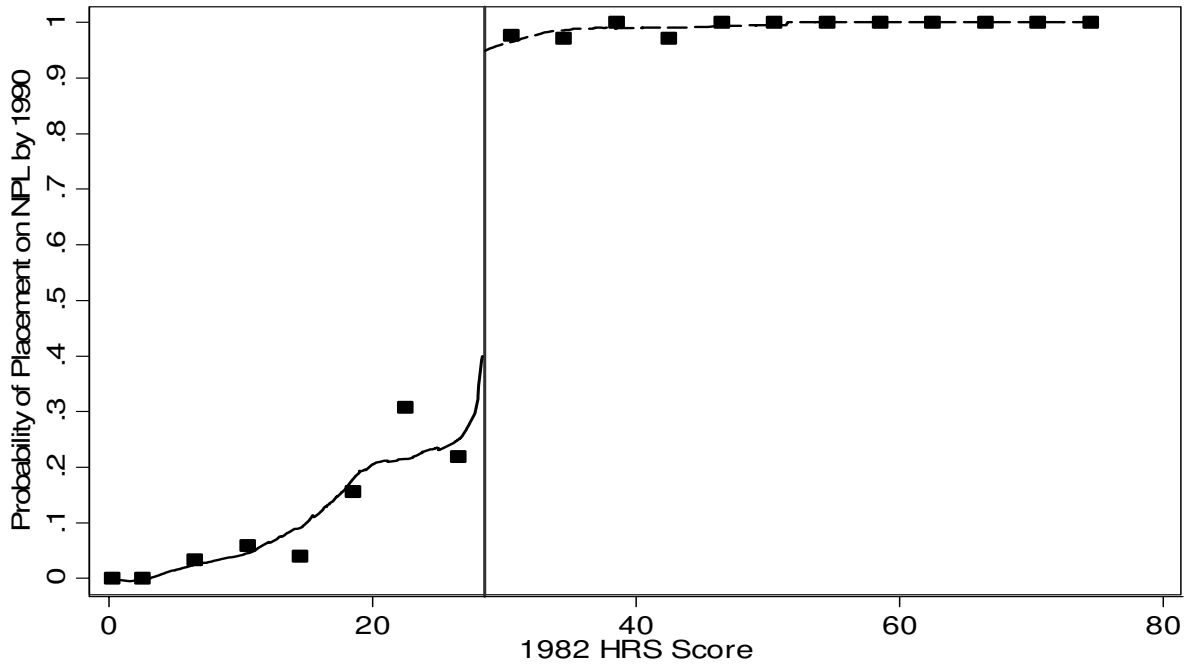
Notes: The figure displays the distribution of 1982 HRS scores among the 487 hazardous waste sites that were tested for placement on the NPL after the passage of the Superfund legislation but before the announcement of the first NPL in 1983. The 188 sites with missing housing data in 1980, 1990, or 2000 are not included in the subsequent analysis and hence are excluded from this figure. The vertical line at 28.5 represents the cut-off that determined eligibility for placement on the NPL.

Figure 4: Estimated Costs of Remediation from Initial Record of Decision, by 4 Unit Intervals of the 1982 HRS Score

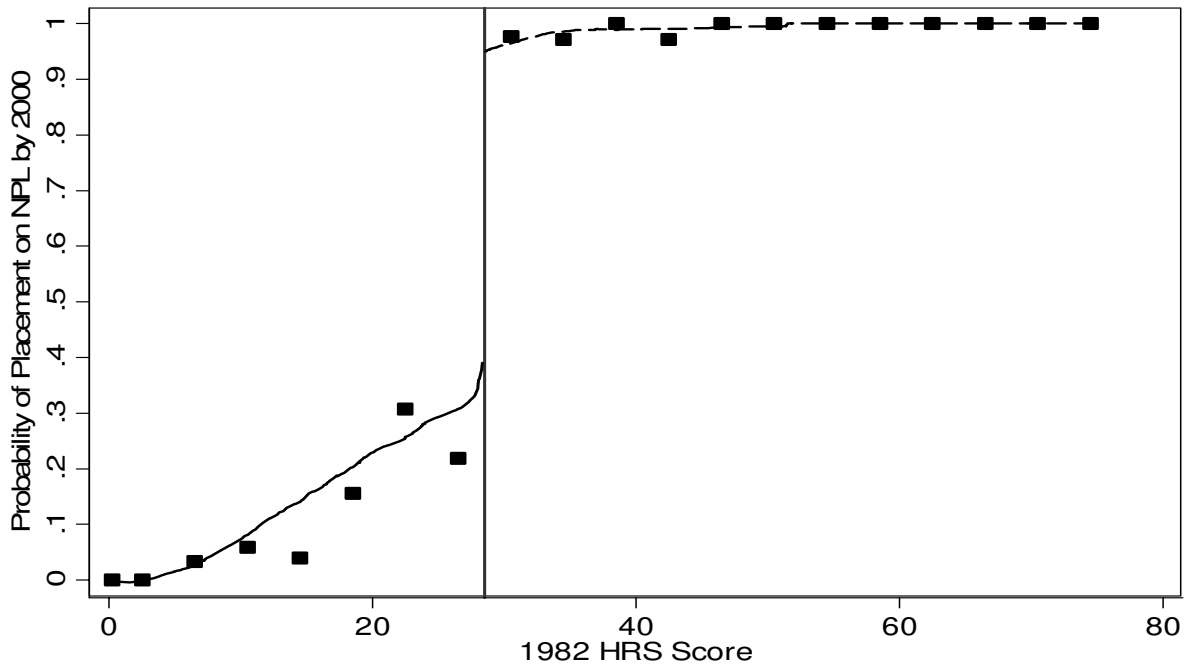


Notes: The figure plots the mean estimated costs (full line) for complete remediation by 4-unit intervals of the sites' 1982 HRS score. The estimated costs are derived from each site's first ROD. The dotted line represents the fraction of sites in each 4-unit interval with non-missing cost data. The vertical line denotes the 28.5 threshold. See the text and Data Appendix for further details.

Figure 5: Probability of Placement on the NPL by 1982 HRS Score
 A. NPL Status by 1990

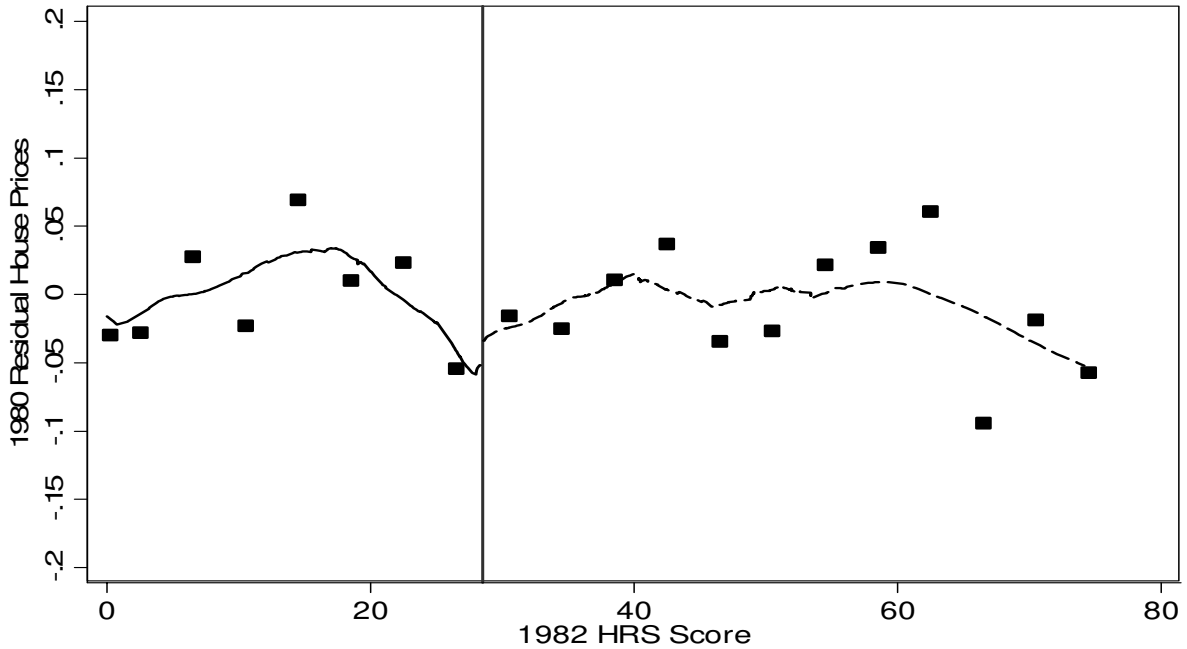


B. NPL Status by 2000

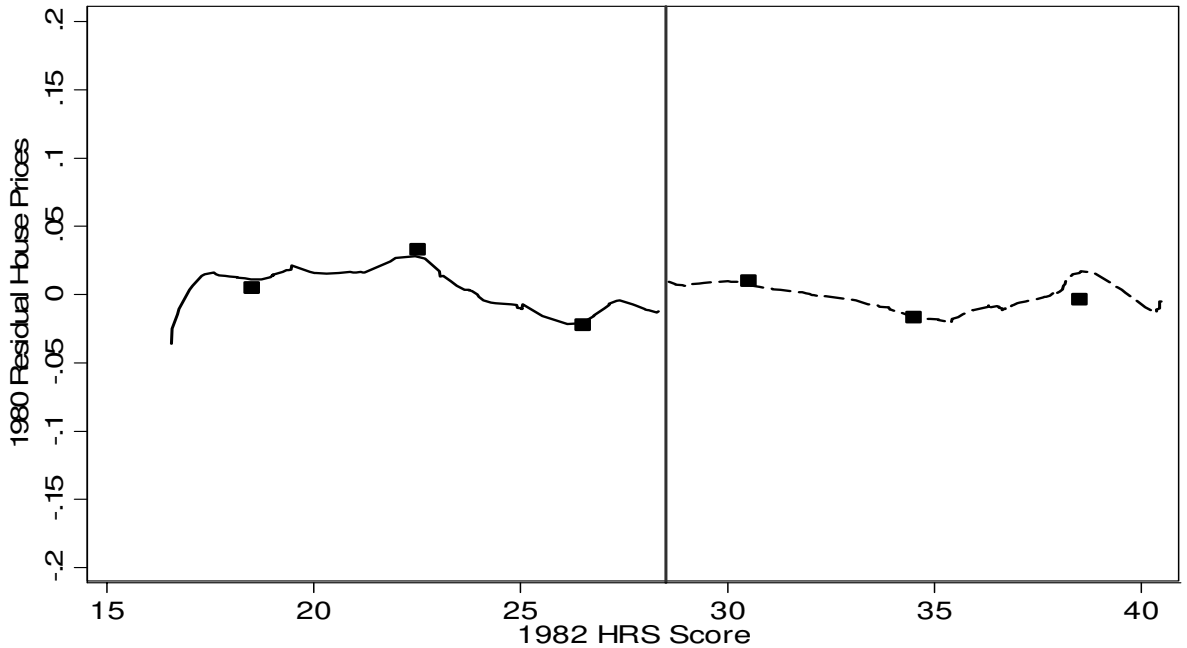


Notes: Each figure plots the bivariate relation between the probability of 1990 (Panel A) and 2000 (Panel B) NPL status and the 1982 HRS score among the 487 sites in the 1982 HRS sample. The plots are done separately for sites above (dashed line) and below (full line) the 28.5 threshold and come from the estimation of nonparametric regressions that use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5. The data points present the mean probabilities in the same 4-unit intervals of the HRS score as in Figures 3 and 4. See the text for further details.

Figure 6: 1980 Residual House Prices After Adjustment for Column 4 Covariates
 A. Full 1982 HRS Sample

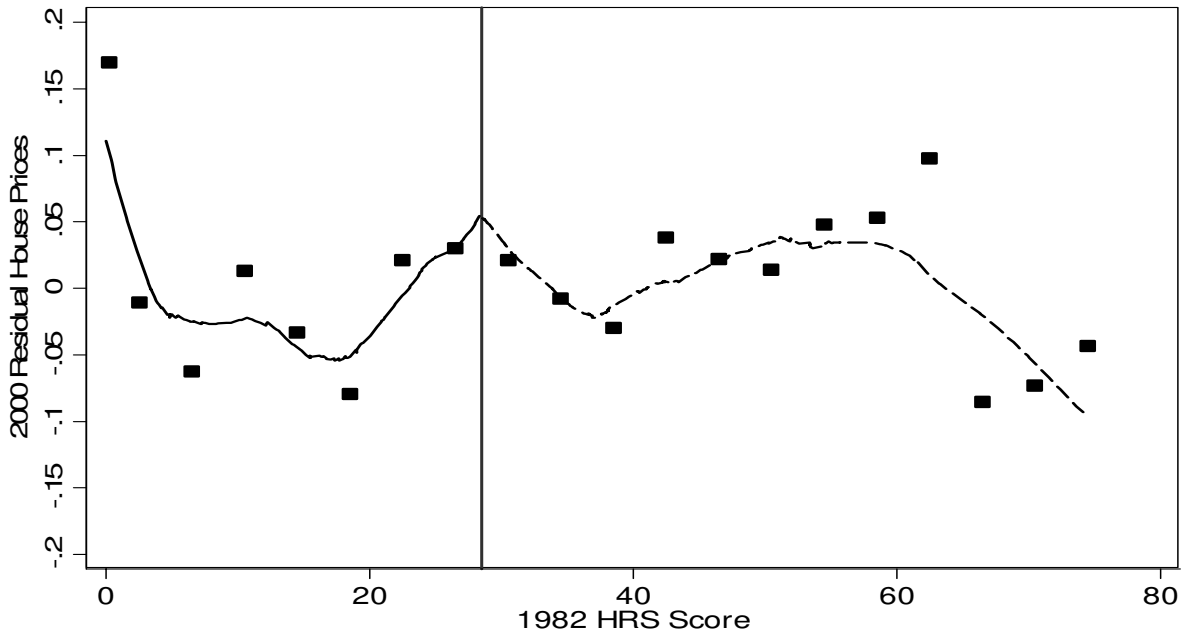


B. Regression Discontinuity Sample

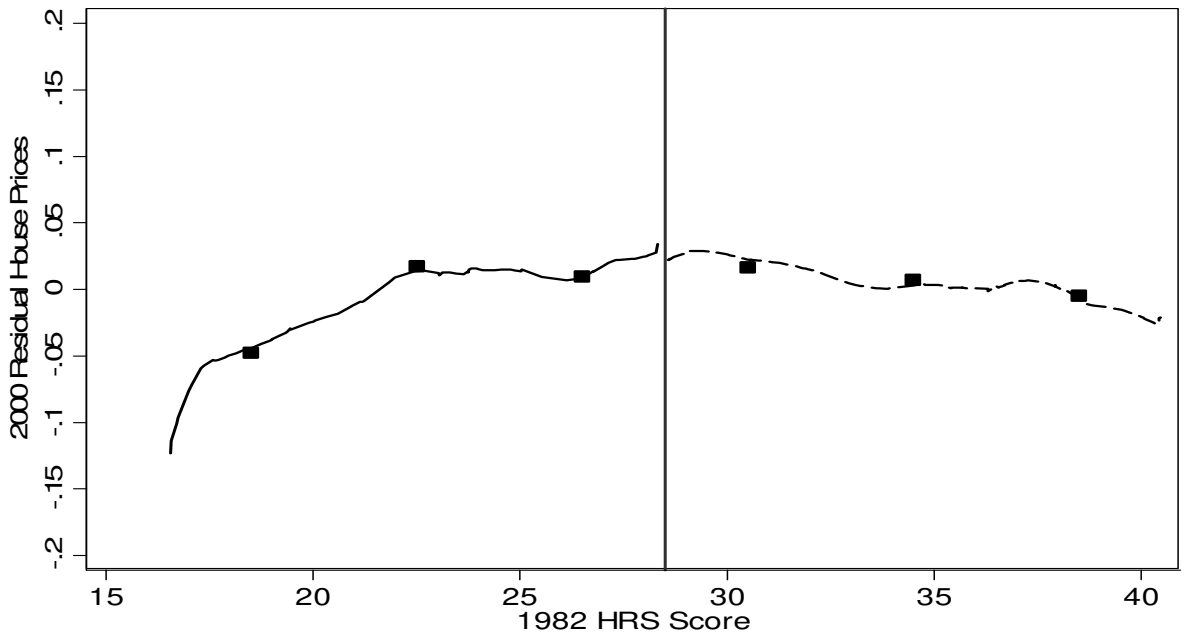


Notes: Each figure plots the results from nonparametric regressions between 1980 residual housing prices after adjustment for the covariates in the column (4) specification of Table 4 (except the indicator for a HRS score above 28.5) and the 1982 HRS scores. The nonparametric regressions use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5. These plots are done separately for sites above (dashed line) and below (full line) the 28.5 threshold. The All NPL sample is used in Panel A, while Panel B uses the regression discontinuity sample. The data points present the mean probabilities in the same 4-unit intervals of the HRS score as in Figures 3 and 4. See the text for further details.

Figure 7: Own Tract 2000 Residual House Prices After Adjustment for Column 4 Covariates
 A. Full 1982 HRS Sample

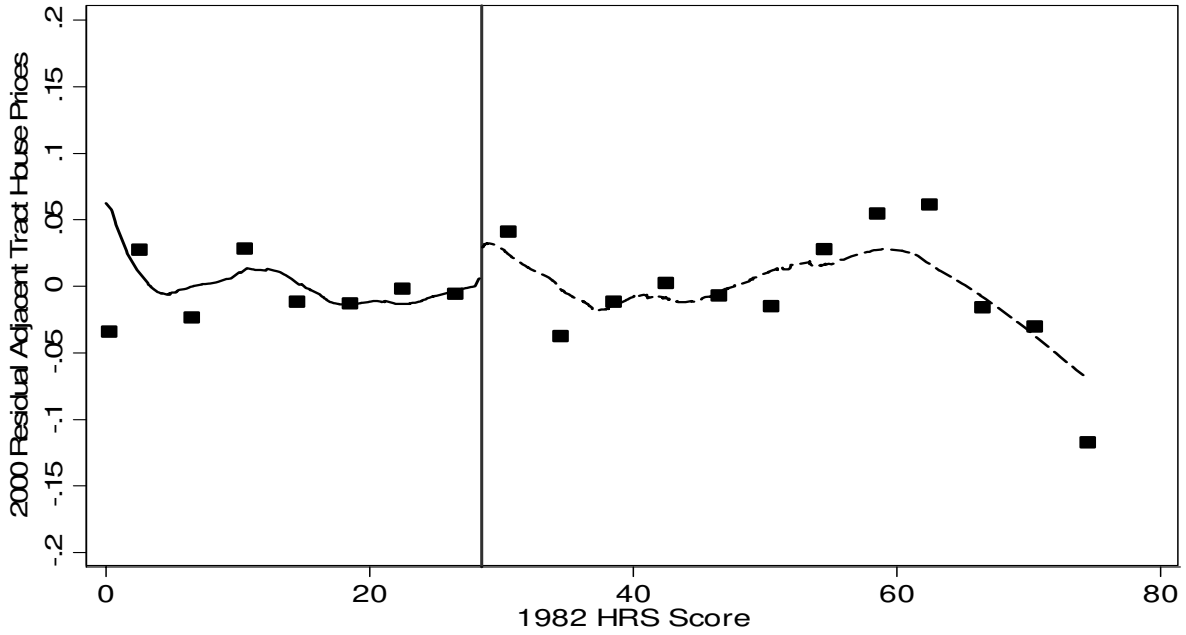


B. Regression Discontinuity Sample

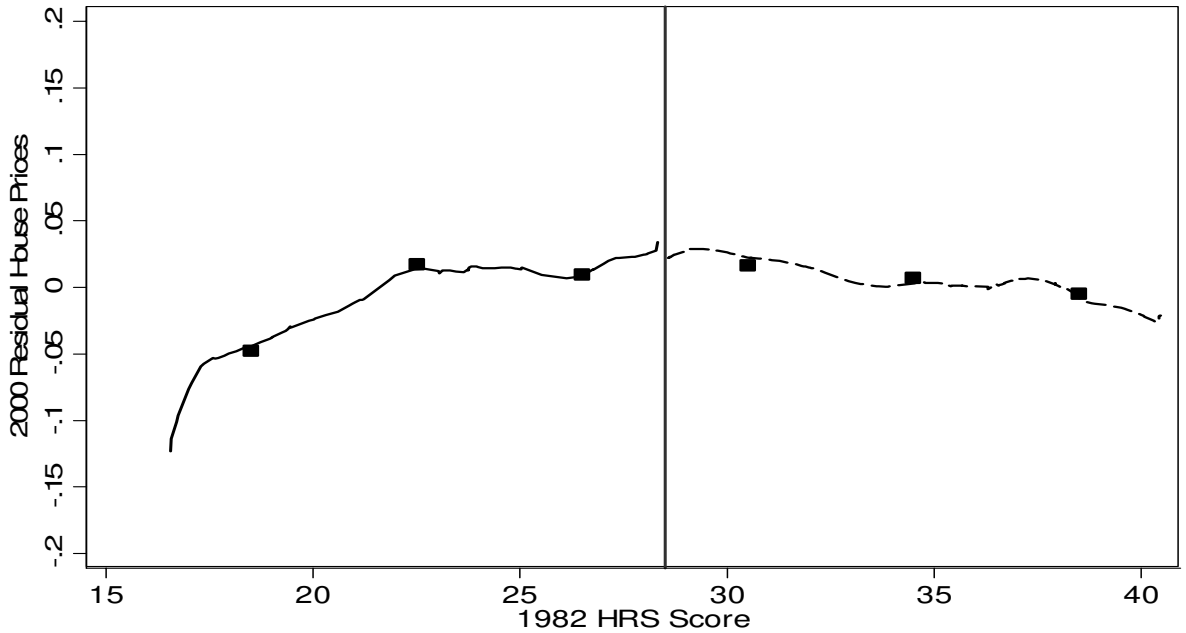


Notes: Each figure plots the results from nonparametric regressions between own tract 2000 residual housing prices after adjustment for the covariates in the column (4) specification of Table 5 (except the indicator for a HRS score above 28.5) and the 1982 HRS scores. The nonparametric regressions use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5. These plots are done separately for sites above (dashed line) and below (full line) the 28.5 threshold. The All NPL sample is used in Panel A, while Panel B uses the regression discontinuity sample. The data points present the mean probabilities in the same 4-unit intervals of the HRS score as in Figures 3 and 4. See the text for further details.

Figure 8: Adjacent Tracts 2000 Residual House Prices After Adjustment for Column 4 Covariates
 A. Full 1982 HRS Sample

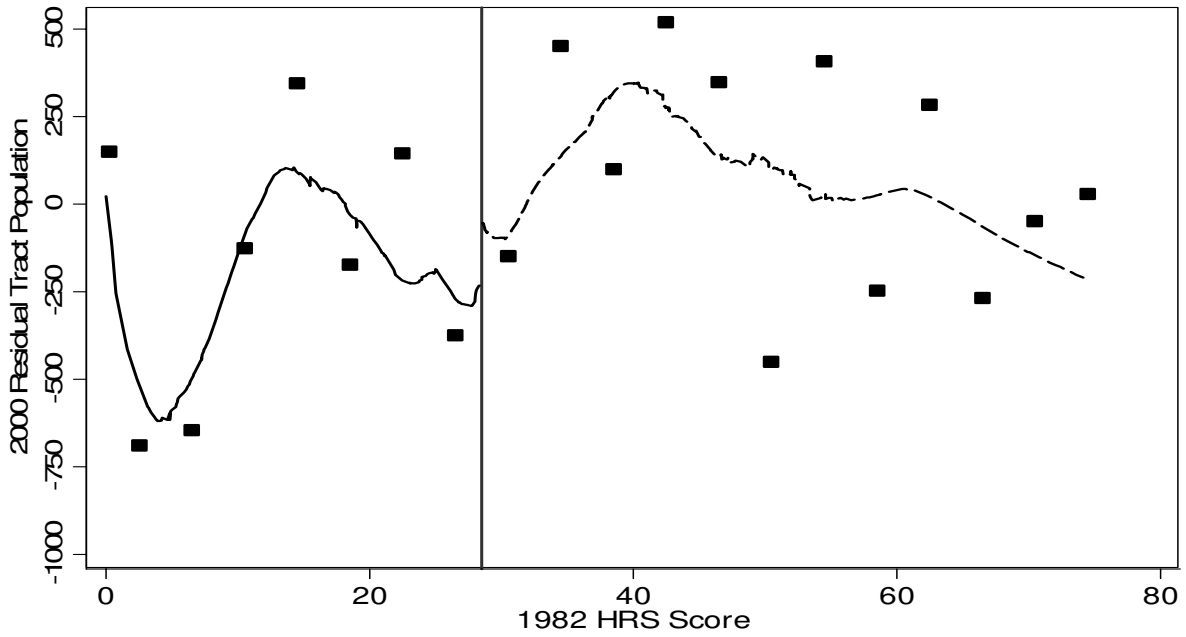


B. Regression Discontinuity Sample

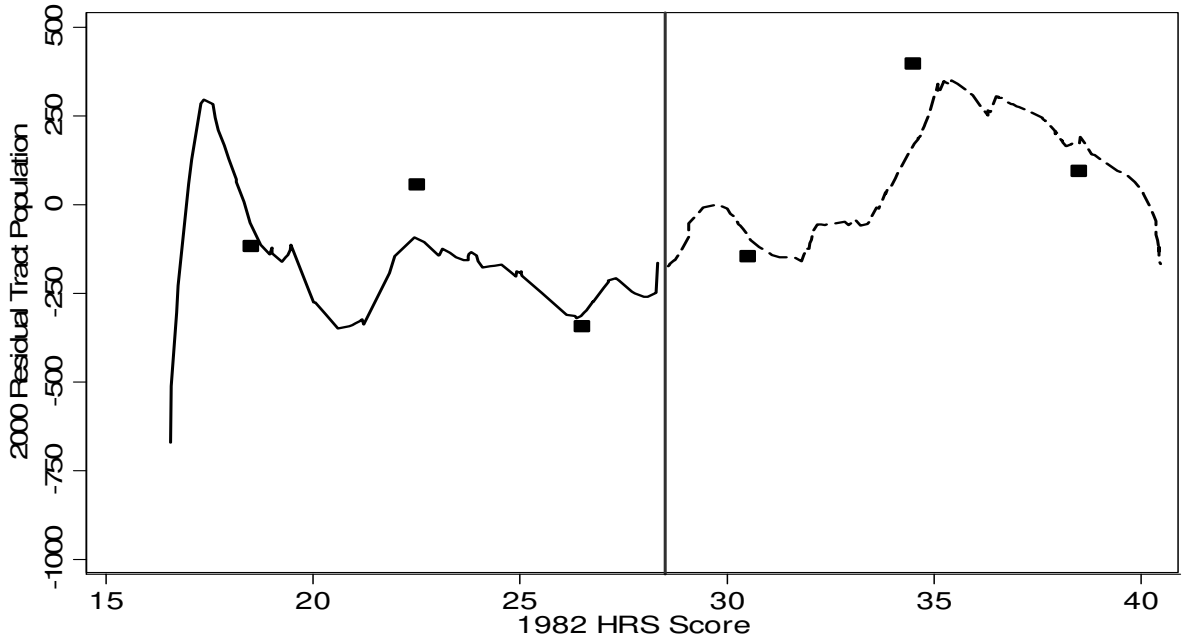


Notes: Each figure plots the results from nonparametric regressions between adjacent tract 2000 residual housing prices after adjustment for the covariates in the column (4) specification of Table 5 (except the indicator for a HRS score above 28.5) and the 1982 HRS scores. The nonparametric regressions use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5. These plots are done separately for sites above (dashed line) and below (full line) the 28.5 threshold. The All NPL sample is used in Panel A, while Panel B uses the regression discontinuity sample. The data points present the mean probabilities in the same 4-unit intervals of the HRS score as in Figures 3 and 4. See the text for further details.

Figure 9: 2000 Residual Own Tract Population After Adjustment for Column 2 Covariates
 A. Full 1982 HRS Sample



B. Regression Discontinuity Sample



Notes: Each figure plots the results from nonparametric regressions between 2000 residual own tract population after adjustment for the covariates in the column (4) specification of Table 8 (except the indicator for a HRS score above 28.5) and the 1982 HRS scores. The nonparametric regressions use Cleveland's (1979) tricube weighting function and a bandwidth of 0.5. These plots are done separately for sites above (dashed line) and below (full line) the 28.5 threshold. The All NPL sample is used in Panel A, while Panel B uses the regression discontinuity sample. The data points present the mean probabilities in the same 4-unit intervals of the HRS score as in Figures 3 and 4. See the text for further details.