NBER WORKING PAPER SERIES

THE VALUE OF BROWNFIELD REMEDIATION

Kevin Haninger
Lala Ma
Christopher Timmins

Working Paper 20296
http://www.nber.org/papers/w20296

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2014

We gratefully acknowledge the U.S. Environmental Protection Agency (EPA) Office of Solid Waste and Emergency Response Center for Program Analysis, the EPA Office of Brownfields and Land Revitalization, and EPA Regional Offices for their support. We thank the EPA for their financial support and for providing access to the data. We would also like to thank seminar participants at Iowa State University and CEnREP Camp Resources for helpful comments. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the EPA or the National Bureau of Economic Research. This work was conducted while Kevin Haninger was an American Association for the Advance of Science Policy Fellow at the EPA. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Value of Brownfield Remediation
Kevin Haninger, Lala Ma, and Christopher Timmins
NBER Working Paper No. 20296
July 2014
JEL No. Q51,R11

ABSTRACT

The U.S. Environmental Protection Agency Brownfields Program awards grants to redevelop contaminated lands known as brownfields. This paper estimates cleanup benefits based on a nationally representative sample of brownfields using a variety of quasi-experimental techniques. To our knowledge, this is the first paper that combines non-public EPA administrative records with high-resolution, high-frequency housing data to estimate the effects of brownfield cleanup across the entire federal Brownfields Program. We find increases in property values accompanying cleanup, ranging from 4.9% to 11.1%; for a welfare interpretation that does not rely on the intertemporal stability of the hedonic price function, a double-difference matching estimator finds even larger effects of up to 32.2%. Our various specifications lead to the common conclusion that Brownfields Program cleanups yield a positive, statistically significant, but highly-localized effect on housing prices.

Kevin Haninger
U.S. Environmental Protection Agency
and AAAS
1200 Pennsylvania Avenue, NW
Room 3417G, EPA West, Mail Code 5101T
Washington, DC 20460
kevin.haninger@gmail.com

Lala Ma
University of Kentucky
Gatton College of Business & Economics
Department of Economics
lala.x.ma@gmail.com

Christopher Timmins
Department of Economics
Duke University
209 Social Sciences Building
P.O. Box 90097
Durham, NC 27708-0097
and NBER
christopher.timmins@duke.edu

1 Introduction

Land revitalization is a beneficial, yet costly, process to undertake. Lands are often contaminated with harmful substances that require expensive procedures to treat. In some cases, toxic waste sites pose a direct threat to human health. In other cases, sites may pose a low risk to nearby residents, but are left unused or under-used until even small amounts of contaminants are removed. Most would agree upon the importance of treating (or at least containing) health hazards at high-risk sites. As for low-risk sites, however, it is far less obvious that the benefits of remediation should exceed the costs. Even though these sites may not be especially toxic, their oftentimes poor aesthetic quality combined with their additional need for special treatment in order to be redeveloped causes the surrounding area to be an undesirable place to live or work. Thus, the benefits of revitalizing these sites include the economic development that would result from making them more productive and attractive. The U.S. Environmental Protection Agency (EPA) has designated these lower-risk sites as brownfields and has aimed to promote their revitalization through grant funding.

This paper uses a slate of quasi-experimental approaches to estimate the benefits of brownfield cleanup by examining its effect on nearby property values. In this respect, the paper draws upon the extensive literature on property-value hedonics to recover homeowner willingness-to-pay for remediation.[1] The value of cleanup, as captured by the value capitalized into nearby housing prices, is suitable for measuring a variety of beneficial effects, which includes the effects on numerous local neighborhood amenities.

Although our paper is not the first to estimate brownfield sites' impact on nearby residential and commercial property values (Ihlanfeldt and Taylor, 2004; Leigh and Coffin, 2005; Linn, 2013), we advance the existing body of work on two important fronts. First, while these previous papers are careful in their investigation of the potential threats from correlated unobservables, all focus on only a subset of brownfields within one or two states. To our knowledge, this is the first paper that combines non-public EPA administrative records with

[1] See Taylor (2003) and Palmquist (2005) for summaries of this literature

high-resolution, high-frequency housing data to estimate the effects of brownfield cleanup across the entire federal Brownfields Program. Next, only under certain conditions that we describe below, can the capitalization of disamenities into local housing markets be given a welfare interpretation (Kuminoff and Pope, 2012). We utilize different sources of variation available in our unique data to estimate cleanup benefits without relying on those assumptions, which makes our estimates particularly useful for cost-benefit analysis.

1.1 Identifying the Effects of Brownfield Remediation

In an ideal research environment, one would randomly select brownfield sites for cleanup and observe the impacts of that cleanup on nearby housing prices. The random selection of sites into the remediation process would guarantee that unobservable determinants of changes in local housing prices would not be correlated with changes induced by remediation, allowing the researcher to cleanly identify the latter. While more common in some areas of research, opportunities for these sorts of experiments are not often available in environmental economics.[2] Indeed, it is the case that the Brownfields Program awards cleanup grants based on a competitive process. The outcome of this process may lead to the award of cleanup funds to locations that differ systematically from locations that do not receive funds. To the extent that we can control for these differences with observable characteristics, they do not present a problem. Data describing sites and the neighborhoods around them are limited, however, so there necessarily will be variables that we cannot control for directly.

We therefore adopt a variety of quasi-experimental approaches to identifying the effect of cleanup on brownfields. The idea of these approaches is to exploit some source of exogenous variation in data so as to approximate that which would result from a truly random experiment. We begin by demonstrating the bias that could result from ignoring unobservable confounders altogether with a cross-sectional specification. In particular, we compare locations with

[2] See Banerjee and Duflo (2009) for a description of the extensive role played by randomized experiments in development economics, and Greenstone and Gayer (2009) for a discussion of the benefits and limitations of quasi-experimental methods for environmental questions.

an untreated brownfield to areas with a remediated brownfield. The problem is that these groups may differ systematically with respect to unobservables that could be correlated with treatment status.

We then demonstrate how even a simple fixed effects specification, which uses changes in a neighborhood's exposure to an unremediated brownfield site, can help solve the problem. In particular, if unobservable differences between houses in the different neighborhoods are constant over time, we can difference that heterogeneity away by looking at changes in exposure status accompanying cleanup activities. Of course, only houses surrounding sites that are remediated experience a change in exposure status, so we must limit our analysis to houses in these neighborhoods.

The problem with the fixed effects specification is that not all unobserved factors will be time-invariant. If brownfield cleanup funds are typically awarded to 'up-and-coming' neighborhoods, the effect of cleanup will be confounded by those other improvements. The opposite would be true if awards were made in an attempt to turn around declining neighborhoods. Fixed effects are unable to deal with these time-varying unobservable factors that are correlated with cleanup activity. This is where we move to techniques traditionally considered 'quasi-experimental.'

First, we consider the 'difference-in-differences' (DID) specification. This approach defines a *treatment* group (e.g., the houses immediately surrounding a brownfield that is treated at some point in time t^*) and a *control* group (e.g., houses nearby to those in the treatment group, so that we can safely assume that other time-varying neighborhood factors will be the same, but far enough away so as to be able to assume that the impact of the brownfield site is negligible). DID then compares the change in prices in the treatment group from houses sold in $t > t^*$ to those sold in $t < t^*$ to a similarly defined change in the control group. The change in prices in the control group, intuitively, controls for any changes in price induced by neighborhood-specific factors aside from brownfield remediation. The remaining effect can therefore be ascribed to the cleanup. Note in addition that, in the process of differencing within the treatment or the control groups, any time-invariant differences between these groups are controlled for as well.

The DID approach to estimation requires a number of non-trivial assumptions. The most important is the ‘common trends’ assumption - in particular, that the change over time in log price in the treatment and control groups would be the same (conditional upon observable covariates) were the treatment group to have remained untreated. In addition to the common trends assumption, the DID specification requires that the equilibrium hedonic price function remain stable over time in order to give estimates a welfare interpretation. The same is also true of the fixed effects specification. We describe this issue in more detail in the following subsection, and use a DID matching estimator that avoids using time variation to deal with it.

1.2 Capitalization v. Marginal Willingness to Pay

The fixed effect and DID approaches to recovering the benefits of site remediation suffer from a similar problem. In particular, each requires an assumption that the hedonic price function, which describes the equilibrium relationship between house attributes (including exposure and treatment status) and price, is stable over time. However, recent work on environmental gentrification, a process whereby changes in the socioeconomic characteristics of a community accompany changes in environmental amenities, suggests there may be substantial neighborhood turnover in response to brownfield redevelopment (Banzhaf and McCormick, 2007; Wolverton, 2002), rendering this assumption questionable. Put differently, with a new local population, the willingness-to-pay for not being exposed to an untreated brownfield site that is revealed by the hedonic price function may be very different after cleanup. Kuminoff and Pope (2012) show that the results of simple fixed effect estimation of the price response to cleanup may therefore fail to identify the MWTP of either those living in proximity to the brownfield before or after cleanup. Instead, it will recover a ‘capitalization’ effect (i.e., the simple response of price to a cleanup, without any additional welfare interpretations). The capitalization effect of a cleanup may be interesting in its own right (e.g., considering implications for property tax revenue collection), but it does not imply a welfare interpretation.

To overcome this problem, we suggest an alternative to using time variation under the traditional DID estimator. In particular, we use a DID nearest-neighbor matching estimator (DD-NNM) that exploits the differences between both treatment and control groups within a neighborhood surrounding a particular site, and the differences between cleaned and uncleaned sites. This method compares similar houses in treatment and control groups around sites that were and were not cleaned, but does not require any comparisons over time. Matching of similar sites relies, in particular, on the state the brownfield is in, the number of previous assessments performed, the type of grant proposal (petroleum or hazardous substances), and Brownfields Program grant scores, which provide a good source of exogenous variation in cleanup status for otherwise similar sites.[3] Furthermore, we match on household characteristics available at the census tract level in each year to control for neighborhood composition. These include race, median income, and number of children, as well as measures of educational attainment, occupation type, spoken languages, and citizenship status. By 'double differencing' in this manner, we are able to cleanly identify a different hedonic price function in each year. In not relying on time variation and an assumption of a stationary hedonic gradient, we are able to interpret our estimates as willingnesses to pay instead of simply capitalization effects.

Together, our fixed effect and quasi-experimental approaches to estimation all lead to a common conclusion - that cleanups conducted under the Brownfields Program yield a large, statistically significant, positive, but highly-localized effect on housing prices.

[3]Applications receiving higher scores are more likely to be funded, but in any particular year a given score may or may not be funded owing to variability in the program's budget - simply put, the program works its way down the list of ranked proposals allocating funds until the budget runs out.

1.3 Outline

This paper is divided into six sections. Section 2 describes the EPA Brownfields Program and cleanup process, paying particular attention to the cleanup grant application and scoring process. Section 3 describes our methodological approach, detailing the different specifications we use to recover estimates of MWTP in the presence of correlated unobservables. Section 4 describes the data, and Section 5 reports estimates from each specification. Section 6 concludes with a brief discussion and cost-benefit calculation.

2 The EPA Brownfields Program

A brownfield is a 'real property, the expansion, redevelopment, or reuse of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant.'[4] Typically, brownfields are lands that were previously used for industrial or commercial purposes and include areas that are contaminated by low concentrations of hazardous substances. These sites are diverse in nature and can range from being old dry cleaning establishments and gas stations to processing plants for materials such as steel, bricks, and asbestos. Generally, brownfields pose lower risk to human health than other types of hazardous waste sites, as they exclude sites listed or proposed for listing on the National Priorities List and sites that are remediated under the Toxic Substances Control Act of 1976. The U.S. Government Accountability Office estimates that there are more than 450,000 brownfields nationwide. In 1995, the U.S. Environmental Protection Agency initiated the Brownfields Program to assist public and private sector organizations in revitalizing brownfields, mainly by providing grant funding. The aim was not only to improve the environment, but also to promote social and economic reinvestment in these unused lands. In 2002, the Small Business Liability Relief and Brownfields Revitalization Act (i.e., the 'Brownfields Law') was signed as an amendment to the Comprehensive Environmental

[4]http://epa.gov/brownfields/. See the EPA's website for further details on the Brownfields Program and a link to public law 107-118 (H.R. 2869), 'Small Business Liability Relief and Brownfields Revitalization Act'.

Response, Compensation, and Liability Act of 1980 (CERCLA), which established the Superfund Program. The passage of the Brownfields Law formalized EPA policies regarding brownfields and expanded financial and technical assistance for brownfield remediation through the Brownfields Program.

2.1 Brownfield Grants, Applications, Scoring, and Awards

Brownfields grants serve as the foundation of the Brownfields Program and support land revitalization efforts by funding environmental site assessment, cleanup, and job training activities. There are four types of competitive grants that serve specific purposes in the land revitalization process: assessment grants, cleanup grants, job training grants, and revolving loan fund grants.

This paper focuses on the effect of cleanup grants on housing values. Cleanup grants provide up to $200,000 to perform cleanup activities at a brownfield site contaminated by petroleum or hazardous substances. Due to budgetary limitations, no eligible entity may apply for funding cleanup activities at more than three sites. Cleanup grants require a 20 percent cost share in the form of a contribution of money, labor, material, or services for eligible and allowable costs; however, applicants may request a waiver of the cost share requirement based on financial hardship. The performance period for cleanup grants is three years. Cleanup grant proposals are evaluated against both threshold and ranking criteria. Applicants must pass all threshold criteria in order to quality for funding. Threshold criteria include site ownership and eligibility for federal brownfield assistance, community notification and opportunity for public comment prior to proposal submission, and a letter from the appropriate state or tribal environmental authority acknowledging that the applicant plans to apply for federal brownfield assistance.

Conditional upon passing all threshold criteria, the proposal will receive a numerical score from the evaluation panel. Scores are based on several evaluation fields, including community need, project description and feasibility, community involvement and partnerships, and reduction of threats to human health and the environment. Once scored, cleanup grant proposals are ranked from highest to lowest score and then awarded funding in rank order until the program budget has been exhausted.[5] Since passage of the Brownfields Law through FY 2013, EPA has competitively awarded 993 cleanup grants totaling $188.6 million.

If a proposal is not awarded in one year, the applicant can reapply in a subsequent year.[6] This implies that the brownfield site could be associated with different proposal scores and different award statuses. We take the applicant's most recent score and application outcome, assuming that it represents the applicant's best and most knowledgeable proposal effort. More details on how scores are compared across grant years are provided in Section 4.

3 Model and Identification

Since brownfield cleanup activity is not directly traded in markets, a revealed preference approach is used to infer its value from its impact on nearby housing prices. This paper uses the hedonic method to model a property's price.[7] For a thorough discussion of the hedonic method, see the reviews by Taylor (2003) and Palmquist (2005). The hedonic price function is defined as a mapping from the attributes of a house, including the presence of a nearby brownfield, to a price in equilibrium. The implicit price of brownfield exposure may be measured with, for example, the hedonic price gradient with respect to distance.

The hedonic method is based on the idea that homeowners' disutility from living in close proximity to a brownfield site can be measured by observing compensating price differentials in housing markets. In general, the homeowner's marginal willingness to pay (MWTP) for some desirable attribute (e.g., distance from a brownfield site) can be read off of the hedonic

[5]Guidelines for cleanup grants can be found at http://www.epa.gov/brownfields/applicat.htm.

[6]Within the universe of brownfield cleanup proposals, we identified 172 properties that reapplied for funding at least once in the six-year period after the program began, 87 of which was eventually awarded funding.

[7]Assuming that the housing supply is fixed in the short-run, any improvement to a brownfield is assumed to be completely capitalized into price and not in the quantity of housing supplied. Given that the Brownfields Program is relatively recent, we would expect to still be in the 'short-run'. As more time passes, researchers will be able to study whether cleanups have had a discernible impact on new development.

gradient (i.e., the derivative of the hedonic price function), owing to utility-maximizing homeowners' sorting behavior. Rosen's seminal paper (Rosen, 1974) and the literature it sparked describe procedures for recovering the MWTP functions for heterogeneous individuals. Bishop and Timmins (2011) describe many of the difficulties encountered in this exercise - because of these difficulties, the typical approach in the applied hedonics literature has been to ignore this heterogeneity and either recover a function that describes price as a linear function of distance, or one that treats exposure discretely, defining it according to whether a house falls inside a particular distance band drawn around a brownfield. That is the approach we adopt here.

One of the more difficult problems that arises when implementing the hedonic method is the presence of house and neighborhood attributes that are unobserved by the researcher but correlated with exposure. These unobservables have the potential to bias the results of a simple cross-sectional specification. Empirical approaches that are used to deal with this problem include (i) fixed effects, (ii) differences-in-differences, and (iii) matching estimators. We briefly review the econometric theory behind each of these modeling strategies below.

3.1 Cross-Sectional Estimates

The simplest specification ignores any panel variation in the data. Considering all houses in counties that contain brownfields,[8] the most straightforward comparison is between houses in the vicinity of a brownfield site and houses in the county that are not exposed to a site at all. However, houses and neighborhoods near brownfields are likely to be different in unobservable ways from those that are not, and these unobservables may lead to biased estimates. Table I describes the observable attributes of houses located inside a circular buffer of 5 kilometers surrounding brownfields in our sample compared with those not exposed to brownfields in the rest of the county, regardless of cleanup status (but before any cleanup has occurred at sites that are cleaned). A simple inspection of this table suggests several reasons to be concerned

[8]We describe the sample of brownfield sites we use for estimation in Section 4.

about the results of this cross-sectional comparison. In particular, there are statistically and economically significant differences between houses that lie in close proximity (5 kilometers) to a brownfield and those that do not - e.g., houses within 5 kilometers of a brownfield site are less expensive and tend to be older and smaller than those in the rest of the county. These large differences in observables suggest that there may also be differences in unobservable attributes of each of these groups of sites.

An alternative approach limits the analysis to only houses surrounding brownfields within 5 kilometers of sites (both those that have and those that have not been cleaned). By limiting the sample in this way, we narrow the variation in unobservable heterogeneity that might be correlated with brownfield exposure. We estimate the following regression specification:

$$P_i = \beta_0 + \beta_1 CLEANUP_i + X_i'\delta + YEAR_i'\gamma + \epsilon_i \quad (1)$$

where

P_i = log of transaction price of house i

$CLEANUP_i$ = 1 if the brownfield that house i is exposed to has been treated under the Brownfields Program[9]

X_i = vector of attributes of house i

$YEAR_i$ = vector of dummy variables indicating year in which house i is sold

The effect of cleanup is then measured by β_1. The problem here is that $CLEANUP_i$ is still likely to be correlated with ϵ_i. Potential bias arises if brownfields that received treatment were systematically different in unobservable ways from those that did not receive treatment. We might, for example, expect that houses located in close proximity to awarded brownfields may be of lower quality than those located near non-awarded sites. Table II compares houses surrounding cleaned brownfields from our sample to those surrounding brownfields that have not been cleaned. While Table II shows that the size of those differences

[9] In practice, this will be houses exposed to awarded brownfields *after* the brownfield has undergone cleanup.

is dramatically lower than are the differences between areas with and without brownfield sites, evidence of significant differences between houses lying inside a 5 kilometer buffer of sites that are eventually cleaned compared with those that are not eventually cleaned still exist. For subsequent methods, we limit the analysis to only using houses within 5 kilometers of brownfield sites.

3.2 Fixed Effects

The simplest approach to dealing with unobserved house and neighborhood attributes that may be correlated with brownfield remediation is to exploit the variation in panel data to control for time-invariant neighborhood attributes. Suppose P_{itk} measures the natural log of the price of house i located in the neighborhood around brownfield k which transacts in year t. X_{itk} is a vector of attributes of that house,[10] and $CLEANUP_{itk}$ is a dummy variable that takes the value 1 if the brownfield k has completed the cleanup process by period t ($= 0$ otherwise).[11] As in Equation 1, consider only houses that are in close proximity to brownfields.[12] μ_k is a time-invariant attribute associated with the neighborhood around brownfield site k that may or may not be observable by the researcher, and ν_{itk} is a time-varying unobservable attribute associated with the house. Importantly, μ_k may be correlated with $CLEANUP_{itk}$ (i.e., sites that receive cleanup treatment may be in neighborhoods that are systematically different from those that do not receive cleanup).

$$P_{itk} = \beta_0 + \beta_1 CLEANUP_{itk} + X'_{itk}\delta + \mu_k + \nu_{itk} \tag{2}$$

[10]Note that, with Dataquick data, house attributes do not vary over time because only the structural attributes from the most recent property assessment are recorded. We subscript X by k and t simply to indicate the neighborhood in which the house is found and the year in which it transacts.

[11]Housing transactions observed before the start of the cleanup period are given a value of $CLEANUP_{itk} = 0$. In practice, we distinguish between houses sold before and during cleanup from those sold after. We discuss this in detail in Section 4.

[12]We present estimates from using multiple buffers to demonstrate robustness.

Using $(i,k) \in k$ to denote all houses in all years that lie in the neighborhood surrounding site k, we can take the within-neighborhood means of each variable and generate mean-differenced data, denoted by $\tilde{}$'s. Noting that $\mu_{itk} - \overline{\mu_k} = 0$, we can then re-write Equation 2:

$$\tilde{P}_{itk} = \beta_1 CLE\tilde{A}NUP_{itk} + \tilde{X}'_{itk}\delta + \tilde{\nu}_{itk} \tag{3}$$

Estimating this specification therefore controls for any permanent unobservable differences between places that received cleanup treatment and those that did not.

3.3 Difference-in-Differences (DID)

Let P_{itk} be the log of the price of house i in the neighborhood surrounding brownfield site k at time t. At some point in time, site k is cleaned. Considering only houses in the vicinity of brownfields that are cleaned (5 kilometers), and let the treatment group of houses be defined by those that are close enough (i.e. closer than 5 kilometers) to be affected by that cleanup. A specific definition of treatment is discussed in section 3.4, but the intuition is that these houses are particularly close to the brownfield, while there may be other houses in the same local neighborhood that experience the same local public goods but are far enough from the brownfield to not be 'treated' by it. We define this distance below.

The dummy variable $TREAT_{ik}$ is equal to 1 if house i belongs to the treatment group (i.e., is located within some buffer b, less than 5 kilometers, surrounding the brownfield), and it is equal to 0 if it belongs to the control group (i.e., inside 5 kilometers but outside the treatment group). Let $POST_{tk}$ indicate post-treatment, which equals 1 if a house lying within 5 kilometers of site k (in either the treatment or control group) sells after site k is cleaned. The model for the observed log price is then written as

$$P_{itk} = \beta_0 + \beta_1 TREAT_{ik} + \beta_2 POST_{itk} + \pi TREAT_{ik} \times POST_{itk} + u_{itk} \tag{4}$$

where π represents the expected change in log price for the treated group less the expected change in price for the control group. π is equal to:

$$\begin{aligned}\pi = \Big(E\left[P_{i1k}^{1} \mid TREAT_{ik}=1\right] - E\left[P_{i0k}^{0} \mid TREAT_{ik}=1\right]\Big) \\ - \Big(E\left[P_{i1k}^{0} \mid TREAT_{ik}=0\right] - E\left[P_{i0k}^{0} \mid TREAT_{ik}=0\right]\Big)\end{aligned} \tag{5}$$

where the superscripts on price denote the counterfactual cleanup status (=1 if cleaned and 0 otherwise) regardless of actual cleanup status (denoted in the subscript). The main identifying assumption underlying the DID model is that of common trends,

$$\begin{aligned}E\left[P_{i1k}^{0} \mid TREAT_{ik}=1\right] - E\left[P_{i0k}^{0} \mid TREAT_{ik}=1\right] \\ = E\left[P_{i1k}^{0} \mid TREAT_{ik}=0\right] - E\left[P_{i0k}^{0} \mid TREAT_{ik}=0\right]\end{aligned} \tag{6}$$

In the case of brownfields, this assumption implies that, in the absence of cleanup, the potential log prices of properties in the treated group would have followed the same trend as log prices in the control group. Under this assumption, π identifies the Average Treatment Effect on the Treated (ATT). Failing to control for observable covariates may invalidate the common trends assumption. One can easily control for them by extending the regression model used to recover π:

$$P_{itk} = \beta_0 + \beta_1 TREAT_{ik} + \beta_2 POST_{itk} + \pi TREAT_{ik} \times POST_{itk} + X_{ik}'\delta + u_{itk} \tag{7}$$

In practice this regression model can be expanded to include multiple groups and multiple treatment periods. For application to brownfield cleanup, we separate the pre-cleanup time frame into two periods and make all comparisons to prices before cleanup activities begin. This will be elaborated in Section 4.1.

3.4 Defining Treatment and Control Groups

The DID specification allows one to control for two types of unobservables. First, it controls for unobservables that vary by treatment group but not over time. Second, it controls for unobservables that affect outcomes over time but are common to both groups. This motivates the definition of treatment and control groups to identify cleanup impact. This paper follows the strategy employed by Linden and Rockoff (2008), using adjacent neighborhoods around a brownfield to define treatment and control groups to alleviate the problem of group- and time-specific unobservables.[13] That is, houses located within a certain distance of a brownfield are considered to be in the treatment group, while houses located outside of that distance (where the site has no effect regardless of cleanup) are designated as controls. To find that distance, we estimate two functions describing the relationship between price and the distance to the nearest brownfield for all property transactions occurring before and after cleanup. Ideally, the distance at which the difference in the price functions becomes insignificant is the point at which we would define the cutoff between the treatment and control groups.

Specifically, one would expect that prices of properties closer to brownfield sites are impacted more by cleanup than those located far away. Furthermore, at some distance far enough away from the site, cleanup should not influence property prices at all. As the effects of hazardous waste sites such as those on the National Priorities List decrease very quickly with distance from the site (Adler et al., 1982; Kohlhase, 1991; Kiel, 1995), the price shocks that would affect the trend of the treated group would arguably affect that of the control group as well. Ultimately, the common trend assumption is untestable. However, this paper provides graphical evidence in the data section and specification tests in the results section that allow us to better assess its validity.

[13] Linden and Rockoff (2008) estimate the impact of sex offender arrival in Mecklenberg County, North Carolina.

3.5 Difference-in-Differences Nearest Neighbor Matching (DD-NNM)

In the previous two sub-sections, we discussed estimators where the distinction between capitalization and MWTP is a potential issue. While we can take some comfort in the fact that we are typically relying on variation in prices over just a few years (and, hence, the hedonic price function may not have much time to evolve), we propose a strategy that deals explicitly with this problem. In particular, we estimate a separate hedonic price function in each year by exploiting variation in data across treated houses around cleaned and uncleaned sites.

Returning to the specification used to estimate the difference-in-differences model in sub-section 3.3, but allowing all of the parameters of the hedonic price function to vary with time, we index each observation by i (house), t (year) and k (site near to which house i is located). Some of the sites have been cleaned by time t ($CLEANUP_{tk} = 1$) while others have not ($CLEANUP_{tk} = 0$). Note that we include the set of sites that applied for, but were denied funding (i.e., $CLEANUP_{tk} = 0 \; \forall \; t$). Finally, we include a flexible function of house, brownfield, and neighborhood attributes (h). We consider only transactions that occur in a particular year t; we therefore do not need to differentiate between a pre- and post-treatment periods. Instead, we only need to differentiate between sites that have and have not been cleaned:

$$\begin{aligned} P_{itk} = \beta_{0t} + \beta_{1t} TREAT_{ik} + \beta_2 CLEANUP_{tk} + \\ \pi_t TREAT_{ik} \times CLEANUP_{tk} + f(h_{itk}; \theta_t) + u_{itk} \end{aligned} \tag{8}$$

We begin by considering only houses in a particular year t that are inside the treatment buffers of either a cleaned or an uncleaned site. As such, $TREAT_{ik} = 1$ for all houses in this sample,

$$P_{itk} = (\beta_{0t} + \beta_{1t}) + (\beta_2 + \pi_t)\, CLEANUP_{tk} + f(h_{itk}; \theta_t) + u_{itk} \tag{9}$$

Using a nearest-neighbor matching algorithm, we pair each house inside the treatment buffer in each neighborhood with $CLEANUP_{tk} = 1$ with a set of J houses that are as similar as possible in h_{itk} and located inside the treatment buffer of a neighborhood with $CLEANUP_{tk} = 0$.

Specifically, for a particular house i located in the treatment buffer of a cleaned site (price designated by P_{itk}), we find the J = 5 'nearest neighbors' to i, t, k from houses located in the treatment buffer of an uncleaned site (prices denoted by $P_j^{(itk)}$)

$$(\beta_{2t} + \pi_t) = \frac{1}{N_t} \sum_{i=1}^{N_t} \left(P_{itk} - \frac{1}{J} \sum_j P_j^{(itk)} \right) \tag{10}$$

Since the covariates of the 'nearest neighbors' may not match exactly, this will introduce additional bias into the estimated treatment effect (Abadie and Imbens, 2011). Therefore, we follow Abadie and Imbens (2011) in using a bias-corrected matching estimator to account for differences in covariate values between the treated observation and its matched counterpart, which has an additional benefit of producing estimates that are more robust to the number of matches used (Abadie and Imbens, 2011). Recalling that house i is located within the treatment buffer of a cleaned site, let $\mu_0(h_{itk})$ denote its conditional expected price, given attributes h_{itk}, had it been near an uncleaned site,

$$\mu_0(h_{itk}) = E[P_{itk}^0 \mid h_{itk}] \tag{11}$$

To implement the bias correction, we first approximate $\mu_0(h_{itk})$ with a linear model,

$$\hat{\mu}_0(h_{jtk}) = h'_{jtk} \hat{\theta}_{w=0} \tag{12}$$

where the parameters ($\theta_{\mathbf{w=0}}$) are estimated with weighted OLS using only the observations from the matched sample (i.e. the houses near *uncleaned* sites, indexed by j). The weight for a house j is given by the number of times it is used as a match for properties near cleaned sites. We then predict prices for houses near *cleaned* sites, with covariates h_{itk}, using $\hat{\theta}_{w=0}$, estimated based on those near uncleaned sites,

$$\hat{\mu}_0(h_{itk}) = h'_{itk}\hat{\theta}_{\mathbf{w}=\mathbf{0}} \tag{13}$$

Finally, the bias-adjusted estimator replaces the counterfactual from the simple matching estimator, $\frac{1}{J}\sum_j P_j^{(itk)}$, with

$$\frac{1}{J}\sum_j \hat{P}_j^{(itk)} = \frac{1}{J}\sum_j P_j^{(itk)} + \hat{\mu}_0(h_{itk}) - \hat{\mu}_0(h_{jtk}) \tag{14}$$

where the adjustment accounts for the difference in the counterfactual outcome due to covariate differences in the matched observation.

Next, we repeat this process using only those houses transacted in year t that are located outside the treatment buffer (i.e., $TREAT_{ik} = 0$ for all of these houses). Denoting the prices of houses located outside the treatment buffer with a $\tilde{P}_{itk}$, we get:

$$\beta_{2t} = \frac{1}{\tilde{N}_t}\sum_{i=1}^{\tilde{N}_t}\left(\tilde{P}_{itk} - \frac{1}{J}\sum_j \hat{\tilde{P}}_j^{(itk)}\right) \tag{15}$$

As such, we are able to recover an estimate of the treatment effect on the treated for each year t by calculating:

$$\pi_t = \frac{1}{N_t}\sum_{i=1}^{N_t}\left(P_{itk} - \frac{1}{J}\sum_j \hat{P}_j^{(itk)}\right) - \frac{1}{\tilde{N}_t}\sum_{i=1}^{\tilde{N}_t}\left(\tilde{P}_{itk} - \frac{1}{J}\sum_j \hat{\tilde{P}}_j^{(itk)}\right) \tag{16}$$

The success of this strategy, of course, depends upon being able to find high-quality matches for houses in neighborhoods around cleaned sites from the set of houses around sites that have not been cleaned. This is what assures that the unspecified function $f(h_{itk};\theta_t)$ will be differenced away. To do this, we match based on house and brownfield characteristics, restrict matches to be amongst sites in the same state, and eliminate other forms of heterogeneity at the neighborhood level by matching on attributes of the census tract in which sites are located.[14]

4 Data

Brownfields, like many other disamenities (Superfund sites, TSDF's, TRI plants) may have very localized impacts on house prices. As such, it can prove difficult to recover these impacts without access to high-resolution data. Cleanup of a brownfield, for example, may not be perceptible in information about census tract median housing prices, while it may in fact have large impacts on nearby houses. One solution to this problem is to use high-resolution decennial census block-level data (Gamper-Rabindran et al., 2011). That approach, however, introduces two potential problems. First, low-frequency decennial data may confound brownfield cleanup with other unobserved events that occurred at some other time during the same decade. Unlike Superfund remediation, brownfield cleanups can be relatively quick, leaving a great deal of remaining time over a ten-year period for other things to happen. Second, cleanups under the Brownfields Program have all taken place in the last decade, and long-form decennial census data have not been collected since 2000. These data are now collected as part of the American Community Survey, and are available at high geographic resolution only on a 'moving average' basis (e.g., for the period 2005-2009). Given that brownfield cleanup can be initiated and completed relatively quickly, we would not know whether most of the cleanups in our data set occurred before or after the homeowner valuations stated in the 2005-2009 ACS data.

In light of all of these concerns, we employ housing transactions data from Dataquick, Inc. that are both high-resolution (i.e., latitude and longitude) and high-frequency (i.e., day of transaction). This allows us to measure the impact of the cleanup with a great deal of precision, both in space and time. In the following three subsections, we describe the

[14]Brownfield characteristics used for matching include sites' average proposal scores, proposal type, and whether there was a phase II assessment performed. Neighborhood characteristics used include race, median income, and number of children, as well as measures of educational attainment, occupation type, spoken languages, and citizenship status.

data, define our pre- and post-treatment periods, and provide summary statistics along with graphical evidence supporting our identification assumptions.

4.1 Data Description

Administrative data on cleanup grant applicants, proposal scores, and brownfields properties are provided by EPA. The non-public data set includes all brownfields that applied for cleanup grants since passage of the Brownfields Law in 2002 through 2008. The data provide characteristics of brownfields, including the exact location (latitude and longitude),[15] property size (for awarded sites only), and types of grant application (i.e. targeted to treat petroleum sites, sites with hazardous substances, or both). A subset of this applicant pool is awarded a cleanup grant.[16] For the properties that were awarded funding, the data include related award and cleanup progress information. Since funding for brownfields varies each year and is awarded beginning with the highest scoring applicant and working downward until funding runs out, there is not one score cutoff that determines whether a property is cleaned. Moreover, because of changing scoring rules, the raw scores are difficult to compare across competition years. To make scores comparable across years, we standardize the scores to be between 0 and 100 by dividing the raw score by the maximum score in its respective competition year.

Dates of different milestones in the process to remediate the brownfield exist starting from site assessment and ending with cleanup. However, these dates are not always available for all of the awarded and non-awarded sites, so we consider all houses sold before any cleanup activities commence to belong to a period, 'Pre-cleanup.' Next, we define an interim treatment period that starts from the earliest recorded cleanup start date, and ends on the

[15] Available information describes the centroid of the brownfield property, but not property boundaries. This is a common feature in data describing the geographic siting of locally undesirable land uses (i.e., LULU's). Like most of this literature, we use distance from the centroid as a measure of exposure. Obtaining more detailed information that would allow us to measure the distance to a site's boundary would be desirable.

[16] Generally, one brownfield is tied to one cleanup grant, although there are a few cases where a brownfield is tied to multiple grants.

cleanup completion date.[17] We distinguish this interim period as houses sold during this time are not exposed to the full effect of cleanup. Lastly, we define the post-cleanup period during which properties have been fully treated with brownfield cleanup as starting with the cleanup completion date and lasting for the duration of our sample.

The time period dummy variables that will be used in all of the specifications are $Interim_{tk}$ and $Post_{tk}$, which respectively equal to 1 if a house is sold during and after cleanup of the nearby brownfield. For the DID specification, interactions between each of the above time period dummies with the treatment dummy are included. In that specification, the coefficient on $Treat_{tk} \times Post_{tk}$ is the treatment effect on the treated, and should be interpreted with respect to the houses in the pre-cleanup period, which is the omitted group. There are several brownfields where cleanup activities have not begun or are not yet complete. We retain these brownfields for the analysis. As long as the types of brownfields that are awarded earlier in our sample (and importantly the types of neighborhoods in which they are situated) do not differ from those that are awarded later on, this should not affect our estimates.

The second data source comes from Dataquick Information Systems, used under a license agreement with the Duke Department of Economics, which provides housing transactions data. These data contain the history of transactions and characteristics for houses in a large number of U.S. counties. The data include information on the sale of newly constructed houses, re-sales, refinance or equity dealings, timeshare sales, and subdivision sales. The data saves transaction-related information such as price, date and associated loans. For each house in the data set, the attributes are recorded from the most recent tax assessment. The attribute fields are detailed and include characteristics such as the number of bedrooms, bathrooms, square footage, lot size, number of units, and number of stories. The housing assessment data also include the latitude and longitude of each property.

In addition to house-level attributes, we control for county level effective real estate tax (RET) rates (Siniavskaia, 2011), as defined by the percentage of the property value that is paid

[17] Dates on which information are released to the public about cleanup, such as the public announcement of grant awards, are also reasonable to consider.

in taxes every year. The county-level RET rates are calculated using homeowner-reported home values and annual real estate taxes from the Census Bureau's 2005-2009 American Community Survey.[18]

The set of brownfields under consideration are those tied to cleanup grant applications between 2002 and 2008. There are a total of 1383 brownfield applications in the EPA data, 446 of which are awarded cleanup grants and 937 are not. Applicants could reapply for a grant in another year following a rejection. Taking into consideration re-applications, we identified 1178 unique brownfield properties. After removing brownfields with missing or inaccurate longitude and latitude coordinates, we are left with 797 sites (437 awarded and 360 non-awarded). Property locations were individually verified with Google maps and checked to ensure that the background of the reported location corroborated with the information from the grant proposal. Dataquick does not have housing data for all counties in which brownfields are located; therefore, only a subset of the properties that are tied to cleanup grants are included. Out of a total of 797 unique brownfields from the EPA data with geocoordinates, 327 had associated housing transactions data within 5 kilometers of only 1 brownfield site. Of those 327 sites, 197 are awarded with cleanup and 130 are not.[19] Currently, the window of observations used for housing transactions starts in 1998 (four years before the start of the Brownfields Program)[20] and ends in 2012, which is the last available year for housing sales.

Focusing on the housing data, our analysis limits transactions to house sales or re-sales of owner occupied properties. Houses with missing prices, bathrooms, bedrooms, or square footage are dropped. Furthermore, since only housing characteristics from the most recent tax assessment are recorded, any house indicated to have undergone major improvements is dropped, as its attributes may be incorrect for previous transactions. To reduce possible errors in record-keeping and sales anomalies, the analysis excludes houses that sold more than

[18] For details, see Siniavskaia (2011).

[19] This figure is after removing certain locations where house attributes are missing.

[20] The extent of geographic coverage by Dataquick becomes much greater in 1998. Going back further in time would require dropping more brownfield sites for lack of housing data.

once per year or five times in the eleven year window of house sales.[21] Prices are normalized to January 2000 dollars using the monthly, regional All Urban Consumers Housing CPI taken from the Bureau of Labor Statistics. The analysis excludes the 1st and 99th percentile of the observed price distribution.

Knowing the exact locations of all properties allows us to calculate the distance between each house and the nearest brownfield. This is our measure of brownfield 'exposure'. Using Graphical Information Systems (GIS), each property is first matched to the nearest brownfield within a 5 kilometer radius. The distances to those brownfields are then recovered. Houses not within 5 kilometers of any brownfield are dropped. Houses located near multiple brownfields, in which case the effect of cleanup may be hard to measure, are dropped.[22] The treatment and control groups are then defined using houses within this 5 kilometer radius to minimize the threat of any location-specific unobservable differences that may affect price dynamics.

An important note is that the available EPA data describe the set of brownfield sites associated with applications for cleanup grants. This precludes analysis of brownfields that did not apply for funding. Therefore, it is possible that there are brownfields (along with other locally undesirable land uses) in neighborhoods that are not accounted for. Even though the analysis cannot control for these sites, it is unlikely that the status of these brownfields will have changed over the course of our analysis, making them time-invariant unobservables that will be differenced out of our analysis using several of the methods described in the previous section. Moreover, if they do change status over time, our DID estimator will control for this to the extent that they equally affect treatment and control groups.

[21] The former often represent non-arms-length transactions that can sometimes lead to multiple transactions on the same day. The latter (i.e., more than 5 transactions in 11 years) signals that the house may be used as an investment property by a house 'flipper' (Bayer et al. (2011)).

[22] Houses near multiple brownfields are (mean difference in parentheses) less expensive ($3,838), older (12 years), smaller (23 square feet), are more likely to be condominiums (4%) and multifamily homes (6%), and less likely to be single family homes (10%). These differences, though statistically significant, are small and unlikely to make difference in our estimates.

4.2 Graphical Evidence

The next step is to determine the distance at which the control and treatment groups are defined. We begin by estimating a pair of price functions over distance from the nearest brownfield site - one for pre-cleanup transactions and one for post-cleanup transactions. The distance at which the pre-cleanup and post-cleanup price functions converge is where brownfield cleanup no longer impacts house prices; this is ideally where we would define the cutoff between treatment and control groups.

Rather than impose a functional form for the price function, we use a local linear polynomial estimator (Fan and Gijbels, 1996), which is described in detail in the appendix.[23] We make one modification to this procedure to account for the fact that the mix of houses sold before and after cleanup changes with respect to distance. In particular, Figure I describes the average square footage of houses sold at each distance from a brownfield site before and after cleanup. It is clear from this figure that houses sold before cleanup of brownfield sites within approximately 2 kilometers tend to be larger than those sold in that same buffer after cleanup. We therefore control parametrically for house attributes before recovering the non-parametric relationship between house prices and distance in Figure II. Figure II also controls parametrically for year effects to allow for general inflationary trends, and differences in brownfield characteristics including the proposal scores, proposal type, and the number of times the sites are assessed.[24]

Figure II provides evidence in support of the assumption that houses that are 'far' enough from brownfields represent a valid control group. While we find that houses at all distances have higher prices on average after cleanup, we find that this difference narrows outside of 2040 meters. Taking the treatment group to be defined by a 2040 meter buffer, the simple

[23]The bandwidth, determined by inspection, is three times Silverman's Rule of Thumb. For the distance gradient, this is about 308 meters. For the time gradient, it is approximately 381 days. A Gaussian kernel is used for weighting.

[24]All brownfields must undergo Phase I and II site assessments. Under certain circumstances, however, additional testing may be advised by a Licensed Site Professional, and a supplemental site assessment is conducted. Recognizing those sites that demand additional testing may control for differences in the severity of contamination at sites.

DID estimator will compare the average change in prices before and after cleanup inside the buffer with the similarly defined change outside the buffer.

Given the definition of the treatment and control groups, a natural way to check whether the common trend assumption is reasonable is to compare the price trends of the treatment and control groups pre- and post-treatment. If the common trend assumption is valid, then price trends should exhibit a few characteristics. First, if the relationship between price and cleanup is causal, one would expect a significant price increase for treatment houses around the time of cleanup, as opposed to a gradual upward trend in price. This would support the claim that cleanup in fact leads to an increase in prices of houses near brownfields. Second, the price trends of the two groups in the pre-cleanup period should be relatively similar (i.e., common trends before cleanup). Third, in the post-cleanup period, the prices of the control houses should not change significantly, but rather should follow a path similar to that in the pre-treatment period. The latter two characteristics would suggest that price trends for houses near brownfields would have been the same as those far from brownfields had they not been treated with cleanup.

Figure III plots the prices of treatment (i.e., inside 2040 meters) and control houses against time relative to the cleanup date.[25] The trends pre- and post-treatment are similar for the two groups. While both groups exhibit a jump at the point of treatment, suggesting that some of the treatment may spill-out into the control group, the discontinuity for the control group going from pre- to post-cleanup (-0.59%) is smaller than that in the treatment group (6.77%). The differences-in-differences approach measures the jump in the treatment group relative to that in the control group.

[25] As was the case when generating Figure II, we parametrically control for housing attributes, year effects, and brownfield characteristics before non-parametrically estimating price as a function of time relative to the cleanup period.

4.3 Summary Statistics

The tables in this section summarize our brownfields data set on multiple dimensions. This is a useful exercise given that this is the first time a national level data set has been compiled for the Brownfields Program. Table III provides summary statistics for the brownfields in the sample. The table provides statistics for subsets of brownfields by housing data availability in order to examine the representativeness of the sample after data cuts and merges. Columns (1) - (3) and (4) - (6), respectively, summarize characteristics of the subsets of brownfields with and without Dataquick housing data. Tests for the equality of group means for the various attributes across these subsets are provided in columns (7) and (8). Table III suggests that proposal scores are marginally higher for non-funded brownfields in locations with Dataquick data, compared to non-funded brownfields in locations without Dataquick data. The difference is not statistically significant for the set of funded properties. Hazardous substances contamination is more common in the funded brownfields for which we do not have housing data; since Dataquick does not provide data for many rural communities, significant differences may reflect the more common occurrence of certain types of brownfields in more urbanized areas.

Table V provides summary statistics for house attributes by treatment status. Columns (1) - (2) and (3) - (4), respectively, summarize the housing characteristics for the treatment group (within 2040 meters of a brownfield) and the control group (between 2040 meters and 5 kilometers of a brownfield). Columns (5) and (6) test for equality of group means. Although we reject the equality of means for many attributes, we do take comfort in the fact that the differences are far smaller than in Table I, which compares houses within 5 kilometers to houses in the rest of the county. We take Table V as evidence that there are important differences between treatment and control groups that should be accounted for parametrically in the DID specification.

Table VI provides a yearly breakdown of cleanup starts and completions for the brownfields that were awarded cleanup grant funding.[26] Since the Brownfields Law was only recently enacted in 2002, many cleanup completions occur towards the end of the window of observations, which limits the number of post-cleanup transactions we have to work with. Table VIII reports the mean cleanup duration by toxin-found and media of contamination. The average cleanup duration for all brownfields for which we can calculate durations is approximately 444 days with a standard deviation of 451 days. These figures imply that brownfield cleanups are relatively quick (e.g., in comparison to the cleanup of a Superfund site); this requires that we use high-frequency housing data (i.e., daily transactions information) for estimation.

Even with the relatively short average duration of brownfield cleanup, right-censoring (i.e., cleanups that are not completed by the end of our sample) is still an issue - particularly for cleanups begun in later years. Table IX describes the fraction of cleanups initiated in each year that were not completed by 2012.[27] Not surprisingly, cleanups begun later in the sample are less likely to be completed. There is, however, a significant fraction of cleanups with petroleum contamination begun early in the sample that have not been completed by 2012.

4.4 Neighborhood Turnover

As suggestive evidence of neighborhood turnover that could alter the equilibrium hedonic price function over time, we use a subsample of the data from Massachusetts for which we can obtain the income and race of the primary homebuyer to examine changes in race and income distributions before and after cleanup.[28] Tabulations of the data (Table X) show that there are more minorities and low income households in areas close to brownfields (i.e. within the treatment buffer), an outcome that is consistent with many studies in the environmental justice literature.[29] We find that neighborhoods that are cleaned experience a relative increase

[26]There are 2 sites that began cleanup before the 2002 - one for areas with Dataquick coverage, and without. These are likely from pilot programs that receive funding before the formal program began.

[27]There were no cleanups initiated in 2012 from the pool of awarded sites between 2002 and 2008.

[28]Housing transactions in Dataquick are merged to data from the Home Mortgage Disclosure Act (HMDA) based on the Census tract, loan amount, and lender name of each property transaction. For details, see Bayer et al. (2011).

[29]See Been et al. (1997); Pastor et al. (2001); Baden and Coursey (2002); Wolverton (2002)

of 2.19% in the share of white homeowners compared to nearby neighborhoods unaffected by the presence of brownfield sites (Table XI, Panel A).[30] With regards to income, Table XI, Panel B shows increases in the middle and lower class (second and third quintiles) following cleanup. Although we cannot explain these shifts in race and income,[31] we can be more certain that dynamic forces that alter the socioeconomic makeup of communities are at play, which can signify preference shifts at the communities of interest, further motivating the use of our DD-NNM estimator.

5 Empirical Results

5.1 Cross-Sectional Estimates

Table XII reports the results of our cross-sectional specification described in Equation 1, where we restrict the comparison to be between houses that are in the vicinity of brownfields - some of which have been cleaned, others of which have not. We find that the value of cleanup is negative at -11.3%. The counterintuitive sign of this effect may be a result of omitted variables bias if cleanup grants are targeted towards struggling neighborhoods. Table XII suggests that unobservable neighborhood attributes may be correlated with their cleanup status, necessitating a different empirical approach.

5.2 Fixed Effect Estimates

Next, we use the fixed effects specification described in Equation 3, which controls for time-invariant unobservables associated with neighborhoods. These unobservables can be the source of bias that leads to the counterintuitive results found in the cross-sectional specifications. The fixed effects specification uses all houses in a buffer; we consider buffers

[30] The fall in the share white homeowners in neighborhoods surrounding brownfields is smaller than the same decrease in white homeowners in nearby neighborhoods far from brownfields.

[31] The literature on gentrification also finds mixed evidence on the direction of demographic and income change following improvements in environmental amenities.

of 1000, 2000, 3000, and 5000 meters to demonstrate robustness. We also include controls for year fixed effects, house attributes, and the real estate tax rate. The results of the fixed effects specification, described in Table XIII, differ strikingly from the cross-sectional results, with increases in house prices from cleanup that range between 6.24% and 11.1%, depending on the size of the buffer.[32]

5.3 Difference-in-Differences Estimates (DID)

While it is able to deal with time-invariant unobservable neighborhood attributes, the fixed effects specification described in Table 3 does nothing to control for time-varying unobservables that may be correlated with brownfield cleanup. Estimates would still be biased if, for example, cleanup were systematically directed towards locations that were considered bad neighborhoods, but were improving in unobservable ways. The DID approach overcomes this problem with the 'common trends' assumption - namely, that the change over time in unobservables in the control group is the same as it would have been in the treatment group in the absence of treatment. By assigning the control group to be houses in the same neighborhood as those in the treatment group, but far enough away from the site to not be impacted by cleanup, we try to satisfy this assumption and obtain estimates that account for any time-varying unobservables that are common to both the treatment and control groups. Moreover, by differencing over time, the DID approach also controls for time-invariant unobservables just as the fixed effects specification did.

As described in Section 3, the average treatment effect on the treated is measured by the coefficient on the interaction of the indicators for a house being in the treatment group ($Treat$) and its transaction occurring after the cleanup has been completed ($Post$). These estimates can be found in the fifth row of Table XIV. With only year fixed effects and brownfield-level controls, we find a treatment effect of 5.85% using the preferred buffer size of 2040 meters. In a specification that includes year fixed effects, house-level and brownfield-

[32]Only the fixed effect estimate using a buffer at 2000 meters is significant using cluster-robust standard errors.

level controls, and controls for the real estate tax, this effect increases to 7.17%. Further introducing brownfield fixed effects decreases this effect to 4.93%, which is significant at 10% after clustering standard errors at the brownfield level.[33] It is worth noting that the coefficient on the time dummy variables, $Interim$ and $Post$, are both positive and significant across all specifications, which suggest that at the community level, broad neighborhood improvements are simultaneously being made over time. This reinforces the importance of controlling for time-varying unobservables through the use of a control group in order to avoid overstating the impact of cleanup. The estimate for the cleanup interim interaction, $Interim \times BF$, additionally reveals that in spite of large housing price increases during the interim of cleanup, the actual cleanup process might slightly depress housing values compared to the houses in the same area that are located outside of the exposure buffer, suggesting that the cleanup effort, though on average fairly quick, can be disruptive. Lastly, we re-estimate the model with false cleanup periods 180, 365, 720 and 1095 days prior to the real cleanup period (Table XV). Notice that the coefficients on the treatment effect become insignificant, providing some evidence that the current specification successfully controls for temporal and spatial confounders.

5.4 Difference-in-Differences Nearest Neighbor Matching Estimates (DD-NNM)

Both the fixed effects and DID approaches rely on the strong assumption that the hedonic price function remains stable over time. If cleanup activities initiate neighborhood turnover, the identities of those living in close proximity to the site may change, and with them, marginal willingness to pay may change as well. In fact, Kuminoff and Pope (2012) demonstrate that estimates of the hedonic price function may provide no information about MWTP. As such, one needs a method that both controls for unobservables that may be correlated with cleanup activities while not relying on variation in cleanup status over time. The difference

[33] See Bertrand et al. (2004) for the importance of clustering standard errors.

in differences nearest neighbor matching estimator described in Section 3 is designed to do this by controlling for brownfield, house, and neighborhood characteristics.

Estimates of the average treatment effect on the treated (π) are recovered without using time variation by taking the difference between two sets of parameter estimates - one derived by comparing houses inside the treatment buffer of cleaned sites to houses inside the treatment buffers of uncleaned sites ($\beta_2 + \pi$), and the other derived by comparing houses in the control groups of cleaned sites to houses in the control groups of uncleaned sites (β_2). We cannot consider results for the years 2004-2007, since yearly data on tract-level attributes, which are used to control for neighborhood composition, are only available starting in the year 2008.[34] Additionally, since our aim is to estimate the benefits from cleaning up brownfield sites, we consider limiting the post-cleanup period to end at most 5 years after the cleanup completion date, as brownfield cleanup can trigger other types of neighborhood redevelopment activities that are not directly tied to the cleanup itself.

Panel A of Table XVI first describes these estimates for our preferred buffer size of 2040 meters using J = 5 matches. Estimates and standard errors are based on Abadie and Imbens (2006). The matching estimates for houses outside the treatment buffer are negative and significant across all years, ranging from -14.0% to -25.0% for J=5 matches, which is consistent with the idea that cleanup targets neighborhoods that are worse off. Several of the within-buffer estimates are insignificant, as the sample sizes have been reduced by a fair amount after limiting the data to post-cleanup sales by year. Differencing the inside and outside estimators, the DD-NNM estimator finds cleanup effects of 12.4% to 32.2%. Estimates are fairly robust to using matches of different sizes,[35] and to limiting the post-cleanup sales to occur within 5 years of cleanup.[36] The estimates that remain most stable regardless of number of matches or timeframe after cleanup is for the years 2008-2010, which find effects ranging from 12.4% to 23.3% for our main specification with J=5 matches and no limit on the

[34] This data is available through SimplyMap, a web-based mapping application accessible with a license agreement with Duke University.

[35] See Panel B for estimates using J=10 matches

[36] See Table XVII for estimates that limit post-sales timeframe for J=5 and J=10 matches

post-cleanup timeframe. The estimated treatment effect for the year 2010 is also significant for both the inside- and outside- buffer matches across all specifications, with the smallest finding an effect of 14.2% after limiting the post-cleanup sales to 5 years and using J=5 matches, and the largest finding an effect of 21.0% without limiting cleanup timeframe and using J=10 matches.

These results suggest that we can indeed interpret our results as implying a positive and significant willingness to pay for brownfield remediation (i.e., a welfare interpretation). The fixed effects estimator, which estimates an 11.1% increase in housing values, is smaller than the smallest of the DD-NNM estimator. Compared with the results of the fixed effects and DID specifications, these larger estimates suggest that changes in the price function over time may have indeed had the effect of reducing the estimated MWTP. Hence, caution must be exercised for assuming time-constant hedonic price functions in policy evaluation if the policy under consideration induces large enough changes such that the population considered before treatment is inherently different from the one after.

6 Conclusion

6.1 Cost-Benefit Analysis

Finally, we can address the simple question, 'is brownfield remediation worth it?' In answering this question, we take a conservative approach. First, we take our most conservative estimate of the cleanup effect - the difference-in-differences estimate based on a 2040 meter treatment buffer (4.93%), rather than the larger estimates generated by the fixed effect and DD-NNM specifications. Next, we take a conservative estimate of the value of housing that sold inside the treatment buffer prior to cleanup. Ideally, we would like to measure the total value of all housing units inside each buffer prior to the start of cleanup, but we do not observe every house sell during that pre-cleanup period. Rather than try to impute values for houses that we do not see transact during that period, we take the conservative approach of aggregating the value of only the houses that do sell in the five years prior to the start of cleanup inside the treatment buffer. We are able to construct this aggregate value for 51 of the brownfields - $4,052,267,776. Multiplying by 4.93% yields an estimate of the aggregate increase in housing value owing to cleanup of $199,776,800. This represents an average benefit value of $3,917,192 per site, with a median of $2,117,982. Figure IV plots the distribution of benefits across sites. The Northeast Midwest Institute (NEMW) estimates an average cost of brownfield cleanup to be $602,000 based on cleanup data provided by the EPA (Paull, 2008). Although the smallest of our benefits estimate is below the estimated cleanup costs (17 of 51 brownfields have estimated benefits less than $600,000), the benefits for the majority of the cleaned sites still far exceed the cost.[37] Furthermore, brownfield remediation should easily pass a cost-benefit test if we considered all the properties located inside the treatment buffer, a larger treatment buffer, or one of our larger treatment effect estimates.

6.2 Discussion

The EPA Brownfields Program provides grants to assess and cleanup properties the 'expansion, re-development, or re-use of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant.' In this paper, we quantify the benefits associated with these remediation activities using property value hedonic techniques. As is typically the case in property value hedonic applications, omitted neighborhood attributes have the potential to bias these estimates. Indeed, our evidence suggests that neighborhoods that successfully clean brownfields under the program may be worse in other unobserved dimensions. As such, we offer a slate of quasi-experimental approaches to overcome this problem, including simple neighborhood fixed effects, a difference-in-differences approach that relies on a treatment and control group defined by geographic proximity, and a difference-in-differences nearest neighbor matching estimator that exploits the advantages of our treatment and control group definitions while not requiring that the hedonic price function remain

[37] Although it is beyond the scope of this paper as we do not have data on planned use, it would be interesting to see whether estimated benefits are systematically different depending on planned future use.

stable over time. Furthermore, our paper offers the added benefit of external validity given our unique data, as it is the first to use a nationally representative sample of brownfield sites considered by the EPA federal Brownfields Program.

Before concluding, we acknowledge a few limitations of our analysis. First, looking at the price of housing in close proximity to brownfield sites will not capture equilibrium effects that are realized elsewhere in the urban area - i.e., cleanup of brownfields may have impacts on local labor markets and on particular housing markets far from the brownfield in question. We will fail to capture these effects to the extent that they appear in other parts of the city. Given the size of a typical brownfield (relative to the size of an urban area), this may not be much of a practical issue. Still, we do note that new methods (i.e., estimable sorting models) may be able to deal with these sorts of concerns (Kuminoff et al., 2013).

Second, our approach will also not capture health benefits from remediation that people are not aware of (and, hence, are not reflected in house purchase decisions and transactions prices). In contrast to other nuisances (Superfund sites, TSDF's, or other toxic waste exposure), we do not expect this to be as much of an issue for brownfield sites, making property value hedonics a good approach in this context.

In light of these limitations, the alternative specifications explored yield a consistent conclusion - averaging over the experiences at a nationally representative sample of brownfield properties, cleanup leads to housing price increases between 4.9% and 32.2%. Taking the most conservative estimate of the value of an average site cleanup, we find that it indeed passes cost-benefit analysis by an order of magnitude based on the expenditures from the Brownfields Program. Moreover, our estimate using a difference-in-differences matching estimator without time variation is consistent with a willingness to pay (i.e., welfare) interpretation, not simply a capitalization effect. Although only one part of the larger EPA Brownfields Program, cleanup of brownfield sites alone yields large increases to nearby housing values and, given the DD-NNM results, has unambiguously positive welfare impacts on communities nearby.

Appendix

Local Polynomial Modeling of the Hedonic Price Gradient

Let $(X_0^1, \ldots, X_0^j, \ldots, X_0^k)$ be a set of k equally-spaced focal points on the support of the variable defining distance from brownfield. Using k focal points divides the support of distance into $k+1$ intervals of length

$$l = \frac{dist_{max} - dist_{min}}{k+1}$$

where $X_0^j = dist_{min} + l \times j$ for $j = 1, 2, \ldots k$. We fit a linear function for each focal point:

$$P_i \mid X_0^j = a + b \cdot dist_i + \epsilon_i$$

where P_i is the price for house i and X_0^j is distance. The covariate and the focal points used in the kernel weight are normalized to have mean 0 and standard deviation 1. The problem is to minimize the following weighted sum of squared residuals,

$$\sum_{i=1}^{n} \left(P_i - [a + b(dist_i - X_0^j)]\right)^2 \cdot K_h\left(\frac{dist_i - X_0^j}{\hat{\sigma}}\right)$$

where $K_h(\cdot)$ is a Gaussian kernel; i.e. $K_h(z) = \frac{1}{h}K_h(\frac{z}{h}) = \frac{1}{h}\phi(\frac{z}{h})$, and $\hat{\sigma}$ is the estimated standard deviation of the covariate, X_i . The smoothing parameter h is chosen according to three times Silverman's Rule of Thumb, which states:

$$h = \frac{1.06\hat{\sigma}}{n^{1/5}}$$

Comparing the price gradients with respect to distance pre- and post- treatment, the estimates find that the difference becomes close to 0 at a distance from the brownfield of about 2 kilometers. Price gradients with respect to time are estimated similarly where the X variable is instead the days relative to cleanup initiation and completion.

References

Abadie, A. and G. W. Imbens (2006). Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica 74*(1), 235–267.

Abadie, A. and G. W. Imbens (2011). Bias-corrected Matching Estimators for Average Treatment Effects. *Journal of Business & Economic Statistics 29*(1).

Adler, K. J., R. C. Anderson, Z. Cook, R. C. Dower, A. R. Ferguson, and M. J. Vickers (1982). The Benefits of Regulating Hazardous Waste Disposal: Land Values as an Estimator. *Washington, DC: Public Interest Economics Center*.

Baden, B. M. and D. L. Coursey (2002). The Locality of Waste Sites Within The City of Chicago: A Demographic, Social, and Economic Analysis. *Resource and Energy Economics 24*(1), 53–93.

Banerjee, A. and E. Duflo (2009). The Experimental Approach to Development Economics. *Annual Review of Economics 74*(1), 151–178.

Banzhaf, S. and E. McCormick (2007). Moving Beyond Cleanup: Identifying the Crucibles of Environmental Gentrification.

Bayer, P., C. Geissler, and J. Roberts (2011). Speculators and Middlemen: The Role of Flippers in the Housing Market. *NBER Working Paper 16784*.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2011). A Dynamic Model of Demand for Houses and Neighborhoods.

Been, V. et al. (1997). Coming to the Nuisance or Going to the Barrios? A Longitudinal Analysis of Environmental Justice Claims. *Ecology LQ 24*, 1–887.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics 119*(1).

Bishop, K. and C. Timmins (2011). Hedonic Prices and Implicit Markets: Estimating Marginal Willingness to Pay for Differentiated Products Without Instrumental Variables. *NBER Working Paper w17611*.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability*, Volume 66. CRC Press.

Gamper-Rabindran, S., R. Mastromonaco, and C. Timmins (2011). Valuing the Benefits of Superfund Site Remediation: Three Approaches to Measuring Localized Externalities. *NBER Working Paper w16655*.

Greenstone, M. and T. Gayer (2009). Quasi-Experimental and Experimental Approaches to Environmental Economics. *Journal of Environmental Economics and Management 57*(1), 21–44.

Ihlanfeldt, K. R. and L. O. Taylor (2004). Externality Effects of Small-Scale Hazardous Waste Sites: Evidence from Urban Commercial Property Markets. *Journal of environmental economics and management 47*(1), 117–139.

Kiel, K. A. (1995). Measuring the Impact of the Discovery and Cleaning of Identified Hazardous Waste Sites on House Values. *Land Economics*, 428–435.

Kohlhase, J. E. (1991). The Impact of Toxic Waste Sites on Housing Values. *Journal of urban Economics 30*(1), 1–26.

Kuminoff, N. V. and J. C. Pope (2012). Do 'Capitalization Effects' for Public Goods Reveal the Public's Willingness to Pay? *Forthcoming in International Economic Review*.

Kuminoff, N. V., V. K. Smith, and C. Timmins (2013). The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets. *Journal of Economic Literature 51*(4), 1007–1062.

Leigh, N. G. and S. L. Coffin (2005). Modeling the Relationship Among Brownfields, Property Values, and Community Revitalization. *Housing Policy Debate 16*(2), 257–280.

Linden, L. and J. E. Rockoff (2008). Estimates of the Impact of Crime Risk on Property Values from Megan's Laws. *The American Economic Review*, 1103–1127.

Linn, J. (2013). The Effect of Voluntary Brownfields Programs on Nearby Property Values: Evidence from Illinois. *Journal of Urban Economics 78*, 1–18.

Palmquist, R. B. (2005). Property value models. *Handbook of environmental economics 2*, 763–819.

Pastor, M., J. Sadd, and J. Hipp (2001). Which Came First? Toxic Facilities, Minority Move-in, and Environmental Justice. *Journal of Urban Affairs 23*(1), 1–21.

Paull, E. (2008). The Environmental and Economic Impacts of Brownfields Redevelopment.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy 82*(1), 34–55.

Siniavskaia, N. (2011). Property Tax Rates by County and City.

Taylor, L. O. (2003). The hedonic method. *A Primer on Nonmarket Valuation 3*, 331.

Wolverton, A. (2002). Does Race Matter? An Examination of a Polluting Plant's Location Decision. *US Environmental Protection Agency, National Center for Environmental Economics, mimeograph*.

Tables

Table I: House Attributes By Within 5km Versus Rest Of County

	Within 5km		Rest of County		Equality of Means	
Attributes	Mean	St. Dev.	Mean	St. Dev.	t-stat	Reject?
Price	230,426.57	1,377,013.13	263,210.91	1,266,378.88	37.04	Y
Real Estate Rate (County)	10.94	4.78	9.90	4.98	-307.13	Y
Age	55.78	34.85	31.62	27.98	-1,141.92	Y
Square Footage	1,495.73	6,482.25	1,814.20	19,809.96	26.33	Y
Bathrooms	1.70	6.85	1.96	2.09	98.56	Y
Bedrooms	2.30	1.98	2.18	1.82	-88.12	Y
Sold in Year Built	0.03	0.16	0.07	0.25	251.08	Y
Condominium	0.19	0.39	0.16	0.37	-96.81	Y
Multifamily	0.09	0.28	0.02	0.13	-549.31	Y
Single Family	0.71	0.45	0.81	0.39	340.42	Y
Mobile	0.00	0.05	0.01	0.09	113.32	Y
Misc.	0.01	0.08	0.00	0.06	-45.89	Y
Obs.	2,769,158		9,288,332			

Note: Compares all houses within 5000m of a brownfield (funded or unfunded) before cleanup to houses located outside 5000m in the rest of the county.

Table II: House Attributes By Whether Brownfield Is Funded Or Unfunded

	Funded Brownfields		Unfunded Brownfields		Equality of Means	
Attributes	Mean	St. Dev.	Mean	St. Dev.	t-stat	Reject?
Price	197,596.68	140,223.30	228,864.38	153,341.36	82.56	Y
Real Estate Rate (County)	10.51	4.01	9.74	4.86	-66.55	Y
Age	46.47	32.54	45.76	30.95	-8.76	Y
Square Footage	1,560.11	690.50	1,551.65	652.17	-4.97	Y
Bathrooms	1.87	0.81	1.99	0.83	56.55	Y
Bedrooms	3.01	1.05	3.03	1.01	6.42	Y
Sold in Year Built	0.04	0.20	0.03	0.18	-18.85	Y
Condominium	0.15	0.35	0.16	0.37	15.28	Y
Multifamily	0.05	0.22	0.04	0.19	-22.74	Y
Single Family	0.80	0.40	0.79	0.40	-6.48	Y
Mobile	0.00	0.04	0.00	0.06	13.32	Y
Misc.	0.00	0.03	0.00	0.06	17.12	Y
Obs.	250,809		395,756			

Note: Sample includes all houses located within 5000m of a brownfield (funded or unfunded). For funded brownfields, attributes are taken from houses selling before cleanup.

Table III: Brownfield Attributes By Availability Of Housing Data

Variable	With Dataquick Data Mean	St. Dev.	Obs.	Without Dataquick Data Mean	St. Dev.	Obs.	t-state	Reject?
Funded and Unfunded								
Petroleum	0.27	0.45	401	0.17	0.38	777	-4.02	Y
Hazardous Substances	0.75	0.43	401	0.85	0.35	777	4.26	Y
Proposal Score (std.)	78.38	11.88	401	76.03	11.38	512	-3.04	Y
Funded Only								
Petroleum	0.30	0.46	239	0.18	0.39	201	-2.76	Y
Hazardous Substances	0.73	0.45	239	0.85	0.36	201	3.00	Y
Proposal Score (std.)	84.54	4.71	239	84.85	4.68	201	0.69	N
Property Size (acres)	10.83	31.29	239	12.91	44.53	197	0.57	N
Ready for Reuse	0.43	0.50	239	0.44	0.50	201	0.25	N

Table IV: Housing Attributes By Proximity To Number Of Sites

	Near 1 Brownfield		Near Multiple Brownfields			
Attributes	Mean	St. Dev.	Mean	St. Dev.	t-stat	Reject?
Price	218,479.00	153,031.80	214,641.00	150,705.70	-17.97	Y
Real Estate Rate (County)	10.18	4.64	10.20	4.31	2.48	Y
Age	46.78	31.49	58.85	34.22	259.07	Y
Square Footage	1,565.92	684.63	1,542.84	757.77	-22.53	Y
Bathrooms	1.95	0.84	1.83	0.93	-92.87	Y
Bedrooms	3.04	1.05	3.02	1.34	-14.00	Y
Sold in Year Built	0.04	0.19	0.03	0.16	-43.59	Y
Condominium	0.15	0.35	0.19	0.39	83.12	Y
Multifamily	0.05	0.21	0.10	0.31	153.83	Y
Single Family	0.80	0.40	0.70	0.46	-167.85	Y
Mobile	0.00	0.05	0.00	0.03	-27.75	Y
Misc.	0.00	0.05	0.00	0.06	24.55	Y
Obs.	876,693		1,186,858			

Table V: Housing Attributes By Treatment Status (Determined By Buffer)

	Treat ($\leq$ 2040m)		Control (>2040m)			
Attributes	Mean	St. Dev.	Mean	St. Dev.	t-stat	Reject?
Price	192,415.71	130,258.47	198,729.03	142,283.50	8.65	Y
Real Estate Rate (County)	9.96	3.62	10.64	4.08	32.61	Y
Age	52.94	36.05	45.05	31.54	-46.76	Y
Square Footage	1,559.46	689.81	1,560.25	690.65	0.22	N
Bathrooms	1.83	0.78	1.88	0.81	11.59	Y
Bedrooms	3.10	1.19	3.00	1.01	-18.18	Y
Sold in Year Built	0.04	0.19	0.05	0.21	8.36	Y
Condominium	0.13	0.34	0.15	0.36	8.53	Y
Multifamily	0.09	0.29	0.04	0.20	-43.92	Y
Single Family	0.77	0.42	0.81	0.39	16.00	Y
Mobile	0.00	0.03	0.00	0.04	4.01	Y
Misc.	0.00	0.04	0.00	0.03	-1.77	N
Obs.	44,539		206,270			

Note: All Houses located within 5km of a brownfield (awarded only). Attributes are taken from houses selling before cleanup.

Table VI: Timeline of Brownfield Start And Completion Frequencies

	With Dataquick Data		Without Dataquick Data	
	Starts	Completions	Starts	Completions
2000	1			
2001			1	
2002	1		1	
2003	4		2	
2004	23	6	17	5
2005	37	12	23	13
2006	35	35	36	18
2007	23	24	28	26
2008	30	17	34	27
2009	30	23	17	33
2010	8	22	8	14
2011	1		1	2

Table VII: Brownfield Properties By Proposal Fiscal Year

Proposal Year	Hazardous Substances (percent)	Petroleum (percent)	Property Size (acres)	Soil Contamination (percent)	Groundwater Contamination (percent)
2003	0.52	0.52	10.41	0.81	0.45
2004	0.83	0.22	7.97	0.88	0.48
2005	0.79	0.23	6.06	0.75	0.56
2006	0.85	0.17	18.82	0.87	0.39
2007	0.87	0.14	16.56	0.77	0.49
2008	0.81	0.23	9.88	0.78	0.52

Note: 'Hazardous Substances' and 'Petroleum' refer to the type of grant application.

Table VIII: Brownfield Cleanup Duration (In Days) By Contaminant

Contaminant Funding Type	Mean	St. Dev.	Obs.
Petroleum only	444.08	468.75	60
Hazardous Substances only	442.72	449.64	210
Contaminant Found			
Controlled Substances	741.90	645.86	10
Asbestos	493.62	476.18	86
PCBs	489.58	468.92	45
VOCs	501.88	464.46	108
Lead	445.65	415.13	156
Other Metals	438.97	436.78	117
PAHs	448.07	436.83	117
Other	495.85	520.38	75
Unknown	383.00	513.36	2
Media of Contamination	Mean	sd	N
Soil	464.06	450.36	234
Air	329.33	289.09	12
Surface Water	356.00	282.92	21
Groundwater	520.44	499.20	126
Drinking Water	634.00		1
Sediments	422.45	441.28	11
Unknown	456.00	359.05	7

Table IX: Fraction of Proposals Initiated In Each Year (Column) That Did Not Complete Cleanup By 2012

Contaminant Funding Type	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Petroleum only		0.0000	0.1765	0.0952	0.2857	0.3333	0.4444	0.6000		
Hazardous Substances only	0.0000	0.0000	0.0455	0.0541	0.1818	0.1905	0.2222	0.4865	0.4000	1.0000
		0.0000	0.0000	0.0000	0.0000		0.0000		1.0000	
Contaminant Found	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Controlled Substances			0.0000	0.0000	0.0000	0.0000		0.0000		
Asbestos	0.0000	0.0000	0.0769	0.0714	0.0000	0.1364	0.3684	0.5714	0.5000	
PCBs		0.0000	0.1667	0.0833	0.1000	0.2500	0.4444	0.5000	0.4286	1.0000
VOCs		0.0000	0.1429	0.1429	0.2667	0.3000	0.2286	0.5556	0.5000	1.0000
Lead		0.0000	0.0500	0.0345	0.1250	0.1563	0.2250	0.5517	0.4167	1.0000
Other Metals		0.0000	0.1538	0.0400	0.2258	0.2609	0.2963	0.5000	0.3636	1.0000
PAHs	0.0000	0.0000	0.0769	0.0800	0.1818	0.2000	0.2647	0.6800	0.4286	1.0000
Other	0.0000	0.0000	0.0000	0.0000	0.1333	0.2857	0.0000	0.3000	0.3333	
Unknown			0.0000				0.0000			
Media of Contamination	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Soil	0.0000	0.0000	0.1143	0.0600	0.2131	0.2222	0.2857	0.5610	0.4000	1.0000
Air			0.2500	0.0000	0.0000			0.0000		
Surface Water			0.2500	0.1667	0.1429	0.0000	0.0000	0.0000	0.5000	
Groundwater		0.0000	0.0909	0.1071	0.3030	0.3182	0.2895	0.6154	0.5455	1.0000
Drinking Water				0.0000		1.0000	1.0000			
Sediments			0.3333	0.2500	0.5000	0.0000	0.7500	0.5000	0.5000	
Unknown			0.0000	0.0000		0.3333	0.0000	0.5000	1.0000	

Note: No cleanups were initiated in 2012 that submitted applications for the 2002-2008 fiscal years.

Table X: Spatial Distribution Of Income Before Cleanup (In %)

	Race		
	Inside Buffer	Outside Buffer	Diff (in) - Diff (out)
White	88.67	89.36	-0.69%
Asian/PI	3.78	4.19	-0.41%
Black	4.4	2.8	1.6%
Hispanic	2.22	2.85	-0.63%
Other	0.92	0.79	0.13%

	Income Quintiles†		
	Inside Buffer	Outside Buffer	Diff (in) - Diff (out)
1	25.99	20.14	5.85 %
2	18.96	19.13	-0.17%
3	18.98	20.23	-1.25%
4	18.98	20.19	-1.21%
5	17.09	20.31	-3.22%

† Quintiles based on distribution of household incomes of all houses sold before cleanup.

Table XI: Change In Race And Income Distribution After Cleanup

Panel A: Race							
	Inside Buffer			Outside Buffer			Diff (in)
	Post = 0	Post = 1	Diff (In)	Post = 0	Post = 1	Diff (Out)	- Diff (out)
White	88.67	88.52	-0.15	89.36	87.02	-2.34	2.19%
Asian/PI	3.78	1.66	-2.12	4.19	2.68	-1.51	-0.61%
Black	4.4	5.74	1.34	2.8	5.13	2.33	-0.99%
Hispanic	2.22	3.93	1.71	2.85	5.01	2.16	-0.45%
Other	0.92	0.15	-0.77	0.79	0.16	-0.63	-0.14%

Panel B: Income							
	Inside Buffer			Outside Buffer			Diff (in)
	Post = 0	Post = 1	Diff (In)	Post = 0	Post = 1	Diff (Out)	- Diff (out)
1	20.72	20.09	-0.63	20.14	20.36	0.22	-0.85%
2	20.24	22.96	2.72	20.43	22.72	2.29	0.43%
3	19.36	26.13	6.77	19.93	21.14	1.21	5.56%
4	19.69	17.82	-1.87	19.56	20.2	0.64	-2.51%
5	19.99	12.99	-7	19.95	15.58	-4.37	-2.63%

Table XII: Cross-Sectional Specification

Comparison Of Houses Near Cleaned Versus Not Cleaned Brownfields (Within 5km)

VARIABLES	All within 5km
Cleanedup	-0.113***
	(0.00275)
Constant	11.52***
	(0.140)
Obs.	469,928
R-squared	0.481
Controls	
Year Fixed Effects	x
Brownfield Characteristics	x
House Controls	x
BF Fixed Effects	

Note: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample used includes only houses inside buffer 5 km around a funded brownfield that has been cleaned or an unfunded brownfield. Cleanedup = 1 if the house is near a funded site has been cleaned.

Table XIII: Fixed Effects

VARIABLES	b= 1000	b= 2000	b= 3000	b= 5000
Interim × BF	0.164***	0.223***	0.223***	0.209***
	(0.0573)	(0.0527)	(0.0525)	(0.0446)
Post × BF	0.0817	0.112*	0.0978	0.0626
	(0.0718)	(0.0644)	(0.0631)	(0.0521)
Constant	10.58***	10.57***	10.80***	11.39***
	(0.214)	(0.184)	(0.139)	(0.152)
Observations	18,686	64,652	136,480	370,910
Number of Brownfields	0.601	0.612	0.636	0.643
Controls				
Year Fixed Effects	x	x	x	x
Brownfield Characteristics				
House Controls	x	x	x	x
BF Fixed Effects	x	x	x	x

Note: Cluster-Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Sample used includes only houses (i) around awarded brownfields, and (ii) inside buffer b in meters. Post = 1 if transaction occurs after nearby brownfield is cleaned. Interim = 1 if transaction occurs during cleanup.

Table XIV: Differences-In-Differences (b=2040m)

VARIABLES	(1)	(2)	(3)
Treat	-0.0487***	-0.0497***	-0.0464***
	(0.004)	(0.003)	(0.017)
Interim	0.237***	0.113***	0.104***
	(0.004)	(0.003)	(0.039)
Post	0.149***	0.000984	0.0747**
	(0.004)	(0.003)	(0.037)
Interim × Treat	-0.0784***	-0.0370***	-0.0178
	(0.009)	(0.007)	(0.023)
Post × Treat	0.0585***	0.0717***	0.0493*
	(0.007)	(0.006)	(0.029)
Constant	13.97***	13.40***	11.23***
	(0.022)	(0.199)	(0.168)
Obs.	370,910	370,910	370,910
R-squared	0.087	0.471	0.380
Number of Brownfields			197
Controls			
Year Fixed Effects	X	X	X
Brownfield Characteristics	X	X	
House Controls		X	X
BF Fixed Effects			X

Note: Cluster-Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Treatment buffer = 2040m. Sample used includes only houses (i) around awarded brownfields, and (ii) inside 5km buffer. Treat = 1 if house is within b buffer in meters. Post = 1 if transaction occurs after nearby brownfield is cleaned. Interim = 1 if transaction occurs during cleanup.

Table XV: DID Estimates With Cleanup Moved D Days Before Actual Date (Falsification Test)

VARIABLES	D=180	D=365	D=730	D=1095
Treat	-0.0478***	-0.0470***	-0.0458***	-0.0476***
	(0.017)	(0.017)	(0.017)	(0.017)
Post	0.0845***	0.0731**	0.0348	-0.00779
	(0.031)	(0.031)	(0.030)	(0.029)
Post × Treat	0.0256	0.0214	0.0162	0.0179
	(0.022)	(0.022)	(0.022)	(0.021)
Constant	11.25***	11.25***	11.26***	11.27***
	(0.168)	(0.168)	(0.166)	(0.165)
Observations	370,910	370,910	370,910	370,910
R-squared	0.380	0.379	0.378	0.377

Note: Cluster-Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. As it is unclear there is an 'Interim' period with the falsified dates, we do not separately identify cleanup interim effects. Controls used for falsification tests are the same as those use in Column 3 of Table XIV.

Table XVI: Difference-In-Differences Nearest Neighbor Matching Estimator (b=2040)

Panel A: 5 Matches

	Inside Treatment Buffer				Outside Treatment Buffer				Average Treatment Effect on the Treated
B=2040	# Matches	Est.	S.E.	Obs.	# Matches	Est.	S.E.	Obs.	
Y2008	5	-0.0248	(0.0400)	4,610	5	-0.149***	(0.0264)	17,855	12.42%
Y2009	5	0.0926**	(0.0385)	5,282	5	-0.140***	(0.0266)	20,716	23.26%
Y2010	5	-0.0741**	(0.0366)	4,471	5	-0.231***	(0.0249)	17,529	15.69%
Y2011	5	0.0719*	(0.0394)	4,108	5	-0.250***	(0.0322)	16,209	32.19%
Y2012	5	0.0645	(0.0424)	3,470	5	-0.188***	(0.0301)	13,535	25.25%

Panel B: 10 Matches

	Inside Treatment Buffer				Outside Treatment Buffer				Average Treatment Effect on the Treated
B=2040	# Matches	Est.	S.E.	Obs.	# Matches	Est.	S.E.	Obs.	
Y2008	10	0.00687	(0.0313)	4,610	10	-0.109***	(0.0199)	17,855	11.59%
Y2009	10	0.0510	(0.0342)	5,282	10	-0.140***	(0.0201)	20,716	19.10%
Y2010	10	-0.0597*	(0.0326)	4,471	10	-0.270***	(0.0213)	17,529	21.03%
Y2011	10	0.0605*	(0.0324)	4,108	10	-0.267***	(0.0268)	16,209	32.75%
Y2012	10	-0.0208	(0.0356)	3,470	10	-0.192***	(0.0238)	13,535	17.12%

Note: Standard errors in parentheses and calculated according Abadie and Imbens (2006). *** p<0.01, ** p<0.05, * p<0.1.

Table XVII: Difference-In-Differences Nearest Neighbor Matching Estimator (b=2040, 5 Years After Cleanup)

Panel A: 5 Matches

	Inside Treatment Buffer				Outside Treatment Buffer				Average Treatment Effect
B=2040	# Matches	Est.	S.E.	Obs.	# Matches	Est.	S.E.	Obs.	on the Treated
Y2008	5	-0.0248	(0.0400)	4,610	5	-0.149***	(0.0264)	17,855	12.42%
Y2009	5	0.0942**	(0.0385)	5,281	5	-0.140***	(0.0266)	20,713	23.42%
Y2010	5	-0.139***	(0.0419)	4,069	5	-0.281***	(0.0265)	16,209	14.20%
Y2011	5	-0.0834*	(0.0439)	3,354	5	-0.472***	(0.0394)	13,290	38.86%
Y2012	5	-0.00393	(0.0680)	2,610	5	-0.175***	(0.0255)	10,205	17.11%

Panel B: 10 Matches

	Inside Treatment Buffer				Outside Treatment Buffer				Average Treatment Effect
B=2040	# Matches	Est.	S.E.	Obs.	# Matches	Est.	S.E.	Obs.	on the Treated
Y2008	10	0.00687	(0.0313)	4,610	10	-0.109***	(0.0199)	17,855	11.59%
Y2009	10	0.0176	(0.0358)	5,281	10	-0.139***	(0.0201)	20,713	15.66%
Y2010	10	-0.131***	(0.0368)	4,069	10	-0.323***	(0.0227)	16,209	19.20%
Y2011	10	-0.111***	(0.0377)	3,354	10	-0.477***	(0.0351)	13,290	36.60%
Y2012	10	-0.0912	(0.0578)	2,610	10	-0.219***	(0.0211)	10,205	12.78%

Note: Standard errors in parentheses and calculated according Abadie and Imbens (2006). *** p<0.01, ** p<0.05, * p<0.1.

Figures

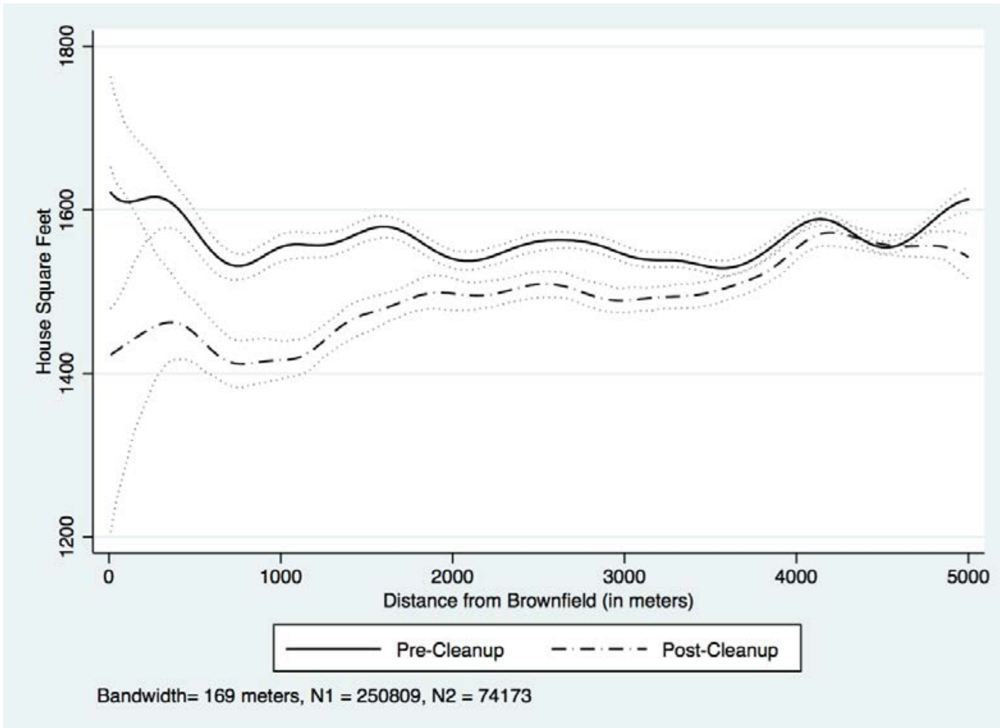

Figure I: Average Square Footage Of Houses Transacted By Distance From Brownfield Before V. After Remediation With 99% Confidence Intervals

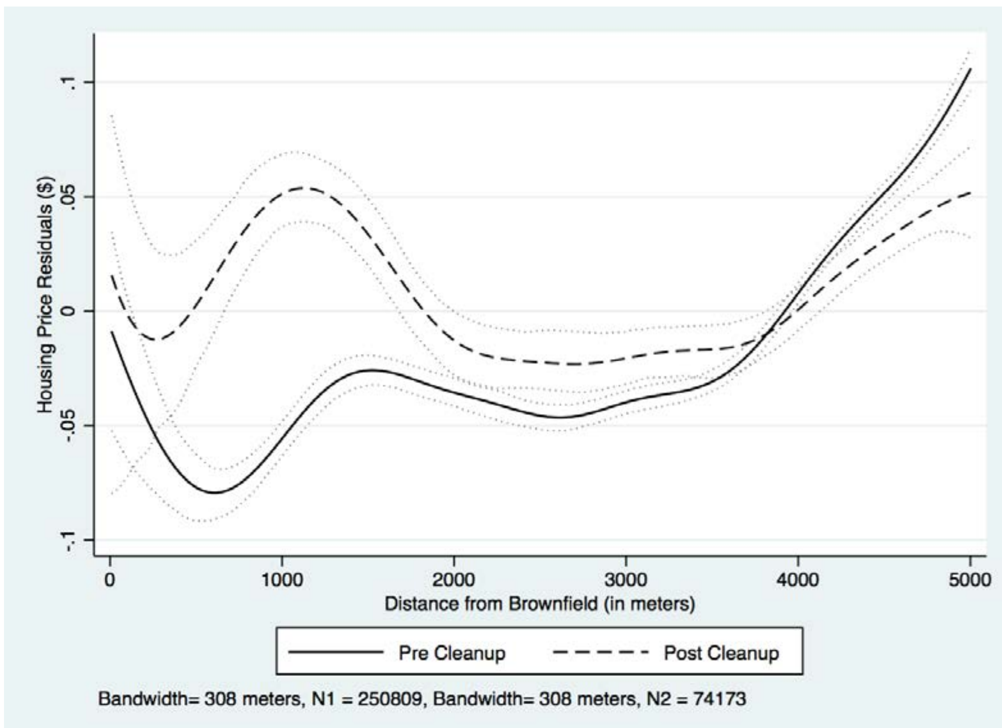

Figure II: Non-Parametric Price Function Estimates Before And After Remediation With 99% Confidence Intervals

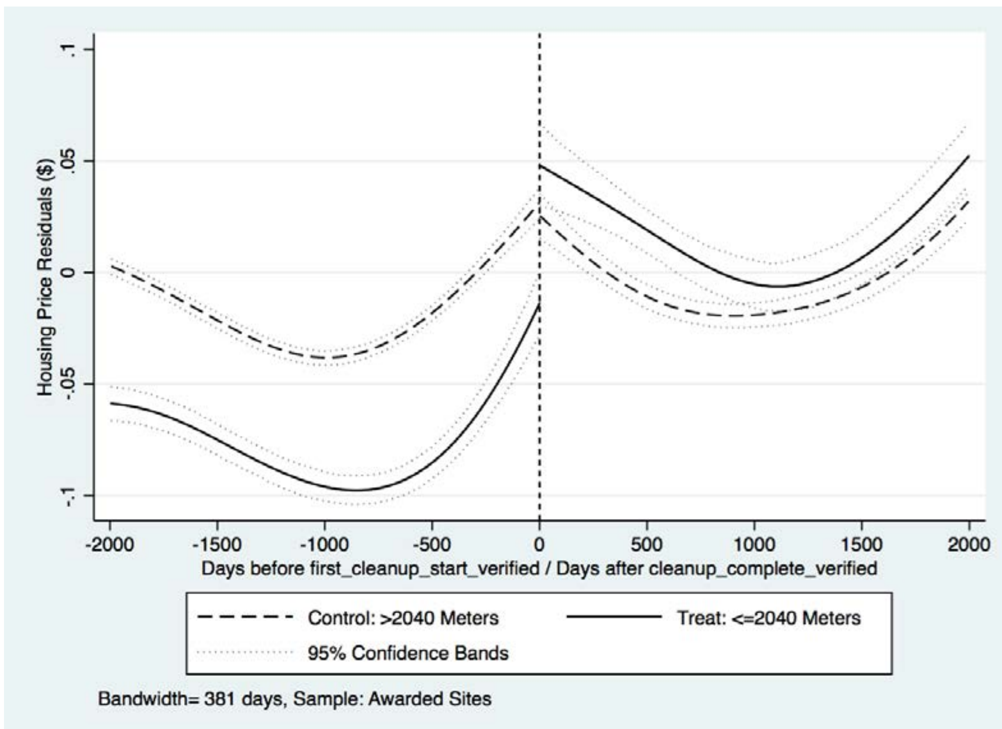

Figure III: Non-Parametric Price Function Estimates Relative To Cleanup Period For Treatment And Control Houses With 95% Confidence Intervals

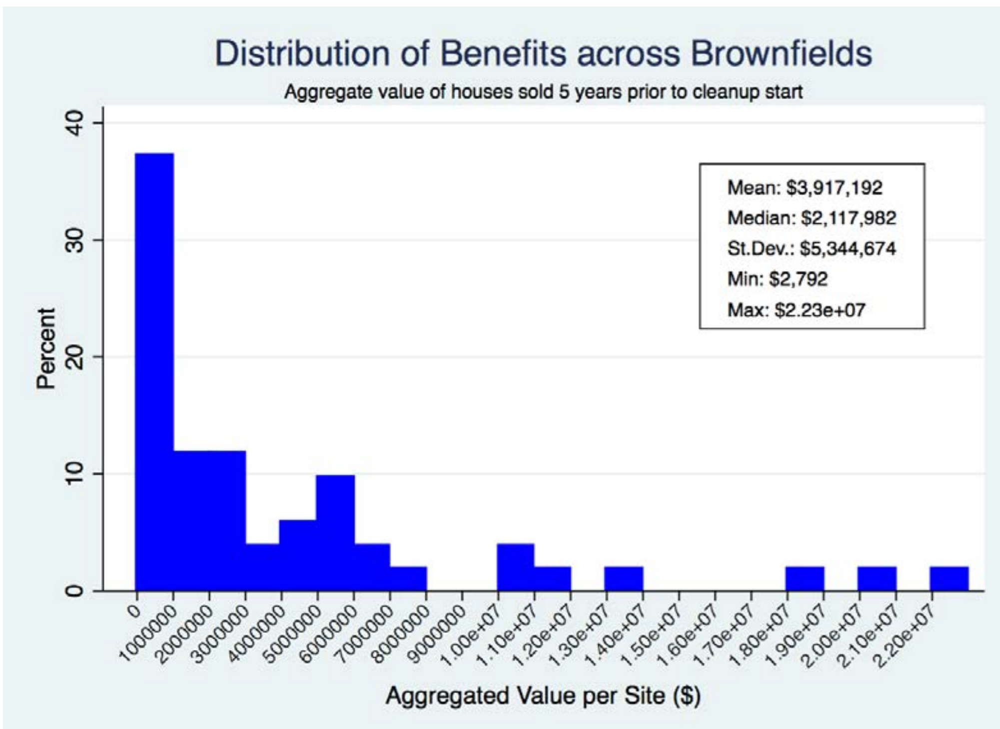

Figure IV: Distribution Of Benefits