Application No.: BN-2020-006

Attachment "C"

Estimating the Impact of Brownfield Remediation on Housing Property Values

Establishment of the
Cutler Bay Civic And Resiliency Enhancement (CARE) Zone

Brownfield Designation

A RESOLUTION OF THE MAYOR AND TOWN COUNCIL OF THE TOWN OF CUTLER BAY, FLORIDA, MAKING FINDINGS AND DESIGNATING REAL PROPERTY IDENTIFIED BY FOLIO NUMBERS 36-6009-005-0015, 36-6009-005-0010, 36-6009-006-0010, 36-6009-006-0012, and 36-6009-006-0011 AS A BROWNFIELD AREA PURSUANT TO SEC. 376.80(2)(A), FLORIDA STATUTES, FOR PURPOSES OF ENVIRONMENTAL REHABILITATION, JOB CREATION, AND ECONOMIC DEVELOPMENT; PROVIDING AUTHORIZATION; AND PROVIDING FOR AN EFFECTIVE DATE.

DUKE ENVIRONMENTAL ECONOMICS WORKING PAPER SERIES
organized by the
NICHOLAS INSTITUTE FOR ENVIRONMENTAL POLICY SOLUTIONS

Estimating the Impacts of Brownfield Remediation on Housing Property Values

Kevin Haninger[*]
Lala Ma[†]
Christopher Timmins[†‡]

Working Paper EE 12-08
August 2012

[*]U.S. Department of Health and Human Services
[†]Department of Economics, Duke University
[‡]National Bureau of Economic Research

Duke
UNIVERSITY

Duke
NICHOLAS INSTITUTE
FOR ENVIRONMENTAL POLICY SOLUTIONS

Estimating the Impacts of Brownfield Remediation on Housing Property Values

by

Kevin Haninger
U.S. Department of
Health and Human Services

Lala Ma
Duke University
Department of Economics

Christopher Timmins
Duke University
Department of Economics
and NBER

June 2012

Abstract

The U.S. Environmental Protection Agency Brownfields Program provides grants to assess and clean up brownfields – 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." The highly localized nature of brownfields lends itself well to measuring the value of site remediation with property value hedonics. The application of that technique is, however, complicated by the presence of correlated unobservable determinants of housing prices (both time-invariant and those that vary over time). This report uses a variety of quasi-experimental techniques to overcome this problem. The analysis finds evidence of large increases in property values accompanying cleanup, ranging from 5.1% to 12.8%; a double-difference matching estimator that does not rely on the intertemporal stability of the hedonic price function finds even larger effects, implying that evidence of property value increases is consistent with a willingness to pay interpretation.

1. Introduction

Land revitalization is a beneficial, yet costly, process to undertake. Lands are often contaminated with various harmful substances that require expensive procedures to treat. In some instances, toxic waste sites are shown to 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 re-developed 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.

1.1 Identifying the Effects of Brownfield Remediation

This report 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 a good way to measure a variety of beneficial effects, including effects on numerous local neighborhood amenities. Under

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

certain conditions that we describe below, these capitalization effects can be given a welfare interpretation, making them particularly useful for benefit-cost analysis.

In an ideal research environment, one would randomly select brownfields 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,[2] opportunities for these sorts of experiments are not often available in environmental economics. 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 observables, they do not present a problem. Data describing sites and the neighborhoods around them are limited, however, so there are necessarily going to be variables that we cannot control for directly.

We therefore adopt a variety of quasi-experimental approaches to identify the effect of cleanup on brownfields. The strategy of these approaches is to exploit some source of exogenous variation in data that approximates the variation that 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 (i) locations with no brownfield to (ii) locations with an untreated brownfield and (iii) locations with a remediated brownfield. The problem is

---

[2] See Banerjee and Duflo (2009) for a description of the extensive role played by randomized experiments in development economics.

that all three of 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, 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 constant over time. 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., the 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 is negligible).[3] 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 have been the same (conditional upon observable covariates) were the treatment group to have remained untreated. This assumption is not testable.

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 introduce an estimator that deals 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 approach requires an

[3] In practice, "treatment" consists of several stages, including assessment and cleanup activities that we will model explicitly.

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. Given the substantial neighborhood turnover that may occur in response to brownfield redevelopment, this assumption may be questionable. Put differently, with a new local population, the willingness-to-pay for not being exposed to an untreated brownfield that is revealed by the hedonic price function may be very different after cleanup. Kuminoff and Pope (2010) show that the results of simple fixed effect estimation of the price response to cleanup may therefore fail to identify the marginal willingness-to-pay (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 “double difference matching” (DDM) estimator that exploits the differences between both treatment and control groups within a neighborhood surrounding a particular site, and the differences between remediated and unremediated sites. In particular, the DDM method compares similar houses in treatment and control groups around sites that were and were not cleaned up, but does not require any comparisons over time. Matching of similar sites relies, in particular, on the state the brownfield is in and on the Brownfield Program cleanup grant proposal scores, which provide a good source of exogenous variation in cleanup status for otherwise similar sites.[4] By double differencing in this manner (instead of over time), we are able to cleanly identify a different hedonic price function in each year. By 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 Brownfield Program yield a large, statistically significant, positive, but highly-localized effect on housing prices.

1.3 A Note on Localized Externalities

Brownfields, like many other environmental disamenities (Superfund sites, TSDFs, TRI plants) may have very localized impacts on housing prices. As such, recovering these impacts without access to high-resolution data can prove difficult. 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, Mastromonaco, and Timmins 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 site remediation, brownfield cleanups can be relatively quick, leaving a great deal of remaining time over a ten-year period for

[4] Cleanup grant proposals 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.

other events to happen. Second, cleanups under the Brownfield 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 (ACS), 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 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.

1.4 Limitations of the Analysis

Before proceeding, we acknowledge a few limitations of our analysis. First, looking at the price of housing in close proximity to brownfields 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 small 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, Smith, and Timmins 2011).

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 environmental disamenities (Superfund sites, TSDFs, or other toxic waste exposure), we do not expect this to be as much of an issue for brownfields, making property value hedonics a good approach in this context.

1.5 Outline of Report

This report is divided into six sections. Section 2 describes the Brownfields Program, 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 "back-of-the-envelope" benefit-cost analysis.

2. The U.S. 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." [5] 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

[5] http://epa.gov/brownfields/. See the EPA webpage 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."

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. EPA 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 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

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 provide up to $200,000 for a grant recipient to inventory, characterize, assess, and conduct planning and community involvement related to brownfields sites. Job training grants provide funding to recruit mostly unemployed, low-income and minority residents from brownfield-affected areas and to train these individuals to secure full-time jobs in site assessment and cleanup. Cleanup grants provide up to $200,000 to perform cleanup activities at a brownfield contaminated by petroleum and hazardous substances. Finally, revolving loan fund grants provide funding to capitalize a revolving loan fund, which is used to make loans and sub-grants for cleanup activities at brownfields.

Since passage of the Brownfields Law through FY 2011, EPA has competitively awarded 1,479 assessment grants totaling $331.3 million, 143 revolving loan fund grants totaling $167.5 million, 801 cleanup grants totaling $150.7 million, and 121 job training grants totaling $25.2 million.

2.2 Cleanup Grant Applications, Proposal Scoring, and Awards

This paper focuses on the effect of cleanup grants on housing values. As stated above, cleanup grants provide up to $200,000 to perform cleanup activities at a brownfield 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.[6]

If a proposal is not awarded in one year, the applicant can reapply in a subsequent year. 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] This implies that the brownfield 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.

[6] More information on the cleanup grant application process can be found at http://www.epa.gov/brownfields/applicat.htm.

[7] In our final sample after the cuts described in the data section, we are left with 18 properties that reapplied for funding, 11 of which were eventually awarded with cleanup.

This paper uses the hedonic method to model a property's price.[8] For a thorough discussion of the hedonic method, see the reviews by Taylor (2003) and Palmquist (2004). 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 can be measured by observing compensating price differentials in housing markets. In general, the homeowner's MWTP for some desirable attribute (e.g., increased distance from a brownfield) can be read off of the hedonic gradient (i.e., the derivative of the hedonic price function), owing to utility-maximizing homeowners' sorting behavior. Rosen's (1974) seminal paper 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

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

researcher but correlated with the attribute of interest. 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) difference-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 and estimates the effect of treatment by the Brownfields Program by comparing the prices of houses in the vicinity of a treated brownfield to those surrounding an untreated site while controlling for the determinants of the prices of houses that are not exposed to a site at all. Considering the set of all houses in all U.S. counties containing a brownfield from our sample,[9] we estimate the following regression specification:

$$P_i = \beta_0 + \beta_1 BF_i + \beta_2 BF_i \times CLEANUP_i + X_i'\delta + YEAR_i'\gamma + \varepsilon_i \quad (1)$$

where

P_i	log of transaction price of house i
BF_i	1 if house i is exposed to a brownfield (= 0 otherwise)
$CLEANUP_i$	1 if the brownfield that house i is exposed to has been treated under the Brownfields Program[10]
X_i	vector of attributes of house i
$YEAR_i$	vector of dummy variables indicating year in which house i is sold

[9] We describe the sample of brownfields we use for estimation in Section 4.
[10] In practice, we will consider houses in four phases of the remediation process – (i) pre-assessment, (ii) post-assessment but pre-cleanup, (iii) cleanup started, and (iv) post-cleanup.

Exposure is defined geographically; houses located inside a circular buffer surrounding the brownfield are considered to be "exposed." We return to the determination of the radius of that buffer below. For a particular definition of exposure, we measure the effect of exposure to an unremediated brownfield by β_1, while the effect of cleanup is given by β_2.

The problem with this approach is that both BF_i and $CLEANUP_i$ are likely to be correlated with ε_i. Houses and neighborhoods near brownfields are likely to be different in unobservable ways from those that are not near brownfields, and amongst houses and neighborhoods near brownfields, those that receive cleanup are likely to be different in unobservable ways from those that do not. We might, for example, expect that houses located in close proximity to brownfields (cleaned up or not cleaned up) may be of lower quality than those located elsewhere in the county.

An alternative approach limits the analysis to houses in buffers surrounding brownfields (both those that have and those that have not been cleaned up). For this sample, $BF_i = 1 \ \forall \ i$. By limiting the sample in this way, we narrow the variation in unobservable heterogeneity that might be correlated with brownfield exposure.

$$P_i = \beta_0 + \beta_1 CLEANUP_i + X_i'\delta + YEAR_i'\gamma + \varepsilon_i \tag{2}$$

The effect of cleanup is then measured by β_1. There is still the potential for bias in this specification, which would arise if brownfields that received treatment were systematically different in unobservable ways from those that did not receive treatment.

Table 1 describes the observable attributes of houses surrounding brownfields in our sample compared with those that do not surround those brownfields but are located in the same county, regardless of cleanup status (but before any cleanup has occurred at sites that are cleaned up). Table 2 compares houses surrounding cleaned up brownfields from our sample to those surrounding brownfields that have not been cleaned up. A simple inspection of these tables suggests several reasons to be concerned about the results of a simple cross-sectional analysis. In particular, there are statistically and economically significant differences between houses that lie in close proximity to a brownfield and those that do not – e.g., while they are more expensive on average, houses within 5 kilometers of a brownfield also tend to be older and smaller. These large differences in observables suggest that there may also be differences in unobservable attributes of each of these groups of sites. These unobservables would lead to biased estimates. While Table 2 shows evidence of significant differences between houses lying inside a 5 kilometer buffer of sites that are eventually cleaned up compared with those that are not eventually cleaned up, we note that the size of those differences is dramatically lower than are the differences between houses located near and far from a brownfield.

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 $P_{i,t,k}$ measures the natural log of the price of house *i* located in the neighborhood around brownfield *k* which transacts in year *t*. $X_{i,t,k}$ is a vector of attributes of that house,[11] and $CLEANUP_{i,t,k}$ is a dummy variable that takes the value 1 if the brownfield *k* has completed cleanup by period *t* (= 0 otherwise).[12] As in equation (2), consider only houses that are in close proximity to brownfields (i.e., $BF_i = 1 \;\forall\, i$). μ_k is a time-invariant attribute associated with the neighborhood around brownfield *k* that may or may not be observable by the researcher, and $\upsilon_{i,t,k}$ is a time-varying unobservable attribute associated with the house. Importantly, μ_k may be correlated with $CLEANUP_{i,t,k}$ (i.e., sites that receive cleanup treatment may be in neighborhoods that are systematically different from those that do not receive cleanup).

$$P_{i,t,k} = \beta_0 + \beta_1 CLEANUP_{i,t,k} + X'_{i,t,k}\delta + \mu_k + \upsilon_{i,t,k} \tag{3}$$

Using $(i,t) \in k$ to denote all houses in all years that lie in the neighborhood surrounding brownfield *k*, we can take the within-neighborhood means of each variable:

$$\bar{P}_k = \frac{1}{N_k} \sum_{(i,t)\in k} P_{i,t,k} \tag{4a}$$

$$\overline{CLEANUP}_k = \frac{1}{N_k} \sum_{(i,t)\in k} CLEANUP_{i,t,k} \tag{4b}$$

$$\bar{X}_k = \frac{1}{N_k} \sum_{(i,t)\in k} X_{i,t,k} \tag{4c}$$

[11] Note that, with Dataquick data, house attributes do not vary over time. 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.
[12] Housing transactions observed before the start of the cleanup period are given a value of $CLEANUP_{i,t,k} = 0$, while transactions observed during the cleanup period are dropped.

$$\bar{\mu}_k = \frac{1}{N_k} \sum_{(i,t)\in k} \mu_k = \mu_k \tag{4d}$$

$$\bar{\upsilon}_k = \frac{1}{N_k} \sum_{(i,t)\in k} \upsilon_{i,t,k} \tag{4e}$$

and generate mean-differenced data:

$$\tilde{P}_{i,t,k} = P_{i,t,k} - \bar{P}_k \tag{5a}$$

$$\widetilde{CLEANUP}_{i,t,k} = CLEANUP_{i,t,k} - \overline{CLEANUP}_k \tag{5b}$$

$$\tilde{X}_{i,t,k} = X_{i,t,k} - \bar{X}_k \tag{5c}$$

$$\tilde{\upsilon}_{i,t,k} = \upsilon_{i,t,k} - \bar{\upsilon}_k \tag{5d}$$

Noting that $\mu_k - \bar{\mu}_k = 0$, we can then re-write equation (3):

$$\tilde{P}_{i,t,k} = \beta_1 \widetilde{CLEANUP}_{i,t,k} + \tilde{X}'_{i,t,k}\delta + \tilde{\upsilon}_{i,t,k} \tag{6}$$

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

3.3 Difference-in-Differences (DID)

Let $P_{i,t,k}$ be the log of the price of house *i* in the neighborhood surrounding brownfield *k* at time *t*. At some point in time, brownfield *k* is cleaned up. For now, consider only houses in the vicinity of brownfields that are cleaned up, and let the treatment group of houses be defined by those that are close enough 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 $TREAT_{i,k}$ 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 $POST_{t,k}$ indicate post-treatment, which equals 1 if a house lying within 5 kilometers of brownfield *k* (in either the treatment or control group) sells after brownfield *k* is cleaned up. The model for the observed log price is then written as

$$P_{i,t,k} = \beta_0 + \beta_1 TREAT_{i,k} + \beta_2 POST_{t,k} + \pi\, TREAT_{i,k} \times POST_{t,k} + u_{i,t,k} \qquad (7)$$

where π represents the expected change in log price for the treated group less the expected change in price for the control group. According to the above model, it is equal to

$$\pi = \left\{ E\left[P_{i,t,k} \mid TREAT_{i,k} = 1, POST_{t,k} = 1 \right] - E\left[P_{i,t,k} \mid TREAT_{i,k} = 1, POST_{t,k} = 0 \right] \right\} \\ - \left\{ E\left[P_{i,t,k} \mid TREAT_{i,k} = 0, POST_{t,k} = 1 \right] - E\left[P_{i,t,k} \mid TREAT_{i,k} = 0, POST_{t,k} = 0 \right] \right\} \tag{8}$$

Using the "potential outcomes" notation (Rubin 1974), where $P^0_{i,t,k}$ represents the log of i's potential price if the house does not receive treatment and $P^1_{i,t,k}$ represents the log of i's potential price if it does.

$$\pi = \left\{ E\left[P^1_{i,1,k} \mid TREAT_{i,k} = 1 \right] - E\left[P^0_{i,0,k} \mid TREAT_{i,k} = 1 \right] \right\} \\ - \left\{ E\left[P^0_{i,1,k} \mid TREAT_{i,k} = 0 \right] - E\left[P^0_{i,0,k} \mid TREAT_{i,k} = 0 \right] \right\} \tag{9}$$

The main identifying assumption underlying the DID model is that of common trends, which specifies that

$$\left\{ E\left[P^0_{i,1,k} \mid TREAT_{i,k} = 1 \right] - E\left[P^0_{i,0,k} \mid TREAT_{i,k} = 1 \right] \right\} = \\ \left\{ E\left[P^0_{i,1,k} \mid TREAT_{i,k} = 0 \right] - E\left[P^0_{i,0,k} \mid TREAT_{i,k} = 0 \right] \right\} \tag{10}$$

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). In particular, we can use equation (10) to replace the third and fourth terms in equation (9):

$$\pi = \left\{ E\left[P^{1}_{i,1,k} \mid TREAT_{i,k} = 1 \right] - E\left[P^{0}_{i,0,k} \mid TREAT_{i,k} = 1 \right] \right\}$$
$$- \underbrace{\left\{ E\left[P^{0}_{i,1,k} \mid TREAT_{i,k} = 0 \right] - E\left[P^{0}_{i,0,k} \mid TREAT_{i,k} = 0 \right] \right\}}_{\left\{ E\left[P^{0}_{i,1,k} \mid TREAT_{i,k} = 1 \right] - E\left[P^{0}_{i,0,k} \mid TREAT_{i,k} = 1 \right] \right\}} \tag{11}$$

Canceling repeated terms yields

$$\pi = \left\{ E\left[P^{1}_{i,1,k} \mid TREAT_{i,k} = 1 \right] - E\left[P^{0}_{i,1,k} \mid TREAT_{i,k} = 1 \right] \right\} \tag{12}$$

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:

$$P_{i,t,k} = \beta_0 + \beta_1 TREAT_{i,k} + \beta_2 POST_{t,k} + \pi\, TREAT_{i,k} \times POST_{t,k} + X^{'}_{i,k}\delta + \varepsilon_{i,t,k} \tag{13}$$

In practice this regression model can be expanded to include multiple groups and multiple treatment periods. For application to brownfield cleanup, there may be various program-related events prior to cleanup that may affect prices. Thus, lumping together all of the prices prior to cleanup may cause the pre-policy treatment and control differences to depend on policy-related factors. We therefore introduce two additional periods and make all comparisons to prices before any remediation-related actions are taken. 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 group (treatment and control) but not over time. Second, it controls for unobservables that affect outcomes over time but are common to both groups. Controlling for both sets of unobservables motivates our definition of the treatment and control groups. One approach might be to define the treatment group as properties near a brownfield that has been cleaned up and the control group as properties near a brownfield that has not been cleaned up. However, if the two brownfields are located in different places, it is likely that the prices of surrounding houses will be subject to unobservables that are not only group-specific, but which change over time. For example, brownfields that are cleaned up might be located in up-and-coming neighborhoods compared to brownfields that are not cleaned up. Over time, the prices of houses near brownfields that are cleaned up would reflect this improvement, compromising the DID identification strategy.

Instead of defining treatment and control groups as above, 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

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

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 located closer to brownfields are impacted more by cleanup than those located farther away. Furthermore, at some distance far enough away from the site, cleanup should not influence property prices at all. Studies show that 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)). This suggests that the treatment and control groups can be defined by the distance at which brownfields begin to have no impact. If this were the case, then price shocks that would affect the trend of one group would arguably affect that of the other 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 the validity of this assumption.

3.5 MWTP v. Capitalization

The intention when running the hedonic specifications described above is to recover an estimate of the MWTP for the amenity in question (here, cleanup of a proximate brownfield). Kuminoff and Pope (2010) note that price function estimates identified using changes in prices and amenities over time formally recover a *capitalization rate* (i.e., the rate at which housing prices increase with the change in the amenity). This may not be the same as the MWTP (i.e., the actual slope of the hedonic price function) either before or after the amenity change. Moreover, it is hard to say a priori which direction the difference between the capitalization effect and the MWTP might go. As long as the hedonic price function is constant over time, there should not be a difference between capitalization and MWTP. One would therefore expect the difference between MWTP and capitalization to be smaller the shorter the time-period is between observations.

To make this point clear, consider the simple example of two hedonic gradients that apply to two different time periods (indexed by t = 1, 2):

$$P_{1,k} = \rho_1 + \theta_1 g_{1,k} + \mu_k + \varepsilon_{1,k} \tag{14}$$

$$P_{2,k} = \rho_2 + \theta_2 g_{2,k} + \mu_k + \varepsilon_{2,k}$$

In this example, $g_{t,k}$ indicates the policy being valued.[14] The MWTP in each period is given by θ_1 and θ_2, respectively. If we were to take the difference between these two equations in order to eliminate the fixed effect, μ_k, we would obtain:

[14] Applied to the question of brownfield remediation, $P_{t,k}$ might refer to the log of the median price in the neighborhood surrounding brownfield k and $g_{t,k}$ would refer to whether that site has been cleaned up by period t. In our estimates, we allow for house-level variation in the transaction price data. For the purposes of illustrating Kuminoff and Pope's point, however, it is simpler to describe the model estimated using only site-level variation.

$$\Delta P_k = (\rho_2 - \rho_1) + (\theta_2 g_{2,k} - \theta_1 g_{1,k}) + \Delta\varepsilon_k \tag{15}$$

However, estimating this equation requires the (stronger than usual) assumption that both $g_{1,k}$ and $g_{2,k}$ are uncorrelated with $\Delta\varepsilon_k$. As such, we typically assume $\theta_1 = \theta_2 = \phi$ and instead estimate

$$\Delta P_k = \psi + \phi \Delta g_k + \Delta\varepsilon_k \tag{16}$$

where $\psi = \rho_2 - \rho_1$. To see how this may yield biased estimates of both θ_1 and θ_2, note the following:

$$\phi = \frac{B - A}{g_2 - g_1} \qquad A = \rho_1 + \theta_1 g_1 \qquad B = \rho_2 + \theta_2 g_2 \tag{17}$$

Therefore,

$$\phi = \frac{\theta_2 g_2 - \theta_1 g_1}{g_2 - g_1} + \frac{\rho_2 - \rho_1}{g_2 - g_1} \tag{18}$$

$$\phi = \frac{\theta_2 g_2 - \theta_1 g_1}{g_2 - g_1} + \left(\frac{\theta_2 g_1}{g_2 - g_1} - \frac{\theta_2 g_1}{g_2 - g_1} \right) + \frac{\rho_2 - \rho_1}{g_2 - g_1} \tag{19}$$

$$\phi = \frac{g_1(\theta_2 - \theta_1)}{g_2 - g_1} + \theta_2 + \frac{\rho_2 - \rho_1}{g_2 - g_1} \tag{20}$$

It is therefore easy to see that, if $\theta_1 = \theta_2$, ϕ will recover the common MWTP estimate. However, if this is not the case, there is no reason why ϕ even has to lie inside the range defined by θ_1 and θ_2.

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 the following sub-section. In particular, we describe a technique that allows us to estimate a separate hedonic price function in each year by exploiting variation in data across treated houses near remediated and unremediated brownfields.

3.6 Double Difference Matching

We begin this sub-section by 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. Furthermore, we index each observation by *i* (house), *t* (year) and *k* (brownfield near to which house *i* is located). Some of the brownfields have been cleaned up by time *t* ($CLEANUP_{t,k} = 1$) while others have not ($CLEANUP_{t,k} = 0$). Note that we include the set of brownfields that applied for but were denied funding (i.e., $CLEANUP_{t,k} = 0 \quad \forall\, t$). Finally, we include a flexible function of house 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 up:

$$P_{i,t,k} = \beta_{0,t} + \beta_{1,t} TREAT_{i,k} + \beta_2 CLEANUP_{t,k} + \qquad (21)$$

$$\pi_t \, TREAT_{i,k} \times CLEANUP_{t,k} + f\left(h_{i,t,k};\theta_t\right) + u_{i,t,k}$$

We begin by considering only houses in a particular year *t* that are inside the treatment buffers of either a remediated or an unremediated brownfield. As such, $TREAT_{i,k} = 1$ for all houses in this sample.

$$P_{i,t,k} = (\beta_{0,t} + \beta_{1,t}) + (\beta_{2,t} + \pi_t) CLEANUP_{t,k} + f\left(h_{i,t,k};\theta_t\right) + u_{i,t,k} \qquad (22)$$

Using a nearest-neighbor matching algorithm, we pair each house inside the treatment buffer in each neighborhood with $CLEANUP_{t,k} = 1$ with a set of J houses that are as similar as possible in $h_{i,t,k}$ and located inside the treatment buffer of a neighborhood with $CLEANUP_{t,k} = 0$. We also match on the brownfields' cleanup grant proposal scores and restrict matches to be between brownfields in the same state.

Specifically, for a particular house *i* located in the treatment buffer of a cleaned up brownfield (price designated by $P_{i,t,k}$), we find the J = 10 "nearest neighbors" to *i,t,k* (prices denoted by $P_j^{(i,t,k)}$).

$$(\beta_{2,t} + \pi_t) = \frac{1}{N_t} \sum_{i=1}^{N_t} \left(P_{i,t,k} - \frac{1}{J} \sum_j P_j^{(i,t,k)} \right) \qquad (23)$$

Next, we repeat this process using only those houses transacted in year t that are located outside the treatment buffer in neighborhoods surrounding sites that were not cleaned (i.e., $TREAT_{i,k} = 0$ for all of these houses). Denoting the prices of houses located outside the treatment buffer with a %, we get:

$$\beta_{2,t} = \frac{1}{\tilde{N}_t} \sum_{i=1}^{\tilde{N}_t} \left(\tilde{P}_{i,t,k} - \frac{1}{J} \sum_j \tilde{P}_j^{(i,t,k)} \right) \tag{24}$$

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

$$\pi_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \left(P_{i,t,k} - \frac{1}{J} \sum_j P_j^{(i,t,k)} \right) - \frac{1}{\tilde{N}_t} \sum_{i=1}^{\tilde{N}_t} \left(\tilde{P}_{i,t,k} - \frac{1}{J} \sum_j \tilde{P}_j^{(i,t,k)} \right) \tag{25}$$

The success of this strategy, of course, depends upon being able to find high-quality matches for houses in neighborhoods around cleaned up brownfields from the set of houses around brownfields that have not been cleaned up. This is what assures that the unspecified function $f\left(h_{i,t,k};\theta_t\right)$ will be differenced away. By matching based on proposal score and restricting matches to be amongst sites in the same state, moreover, we eliminate other forms of heterogeneity at the neighborhood level.

4. Data

Our analysis is based on two main sources of data. In the following three subsections, we describe the data, define our pre- and post-treatment periods, and provide summary statistics along with graphical evidence supporting our identification assumptions.

4.1. Data Description

Data on brownfield properties are provided by the U.S. EPA.[15] The data set includes administrative records of all brownfields that applied for cleanup grant funding in fiscal years 2003 through 2008, which represent the first six years of the cleanup grant competition following enactment of the Brownfields Law. The data provide characteristics of the brownfields, including the exact location (latitude and longitude),[16] property size, and types of toxic materials present.

A subset of applicants is awarded cleanup grants based on their proposal scores and the program budget. Since cleanup grant funding for brownfields varies each year and is awarded from the highest scoring applicant until funding runs out, there is not a common score cutoff across all competition years that determines whether a property will receive funding . Moreover, because of changing scoring rules, the raw scores are difficult to compare across years. To make scores comparable across years, we

---

[15] As part of this study, EPA contacted all grant recipients and provided them with a summary of the most current information EPA had concerning assessment, cleanup, and redevelopment at the properties associated with their grants. The letters requested that grant recipients submit updated information about these properties, if necessary, and provided them with instructions for submitting this information to EPA.

[16] 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. 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, although this is less of a concern with brownfields because of their relatively small size.

standardize the scores to be between 0 and 100 by dividing the raw score by the maximum possible score in its respective competition year.

Especially relevant to this study are the dates of different milestones in the process to remediate the brownfield, starting from site assessment and ending with a completed cleanup. The data track the dates of assessment completions, planning activities, initiations and completions of interim cleanup objectives (for funded brownfields), and cleanup completions (for funded brownfields). Our analysis requires that we define and control for different periods over which the brownfield is evaluated and remediated. Many of the dates we are given are internal milestones for those executing the cleanup, and may not matter much to (potential) property owners in nearby neighborhoods. We do consider events that may observably alter (for the property owner) the amount of brownfield exposure, and define four periods as follows. First, we define two pre-treatment periods. All brownfield properties that applied for cleanup grant funding must have recently performed (or at least be in the process of performing) a phase II environmental site assessment. The phase II assessment is a process by which a licensed environmental professional inventories site contamination. In addition, assessment results are communicated to the public. Since the phase II assessment process involves intrusive sampling and is observable to homeowners (as opposed to phase I assessments, which are based on reviewing site records and visual inspection), the prices of properties sold after the phase II assessment but before any cleanup has started may reflect responses to assessment outcomes. Assessments could simultaneously signal to property owners a higher chance that the brownfield may be cleaned up, and so prices of properties sold during this period could also be affected by these expectations. We therefore separate the period before any cleanup has begun into a pre-assessment period for any house sold before the phase II assessment complete date, and a post-assessment period for those houses sold after the phase II assessment date but before cleanup has begun. The aim is to contain any effects related to cleanup or its possibility and establish a good comparison group for the effect of cleanup by using pre-assessment prices. Next, we define an interim treatment period that starts from the earliest recorded cleanup start date, and ends on the cleanup completion date.[17] We distinguish this interim period since 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 that starts at the cleanup completion date and lasts the duration of our sample.

The time period dummy variables that will be used in all of the specifications are $POST-ASSESS_{t,k}$, $INTERIM_{t,k}$, and $POST_{t,k}$, which respectively equal 1 if a house is sold after assessment, during cleanup, 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 $POST_{t,k} \times BF_{t,k}$ is the treatment effect on the treated, and should be interpreted with respect to the houses in the pre-assessment period, which is the omitted time period. There are several brownfields where cleanup activities have not begun or are not yet complete. Brownfields with incomplete cleanup activities or missing dates are excluded from the analysis.

[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.

The second data source comes from housing transactions data provided by Dataquick Information Systems, used under a license agreement with the Duke Department of Economics. 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) rate (Siniavskaia, 2011), as defined by the percentage of the property value that is paid in taxes every year. The county-level RET rates are calculated using homeowner-reported home values and annual real estate taxes from the U.S. Census Bureau's 2005-2009 American Community Survey.[18]

The set of brownfields under consideration are those that applied for cleanup grant funding in fiscal years 2003 through 2008. There are a total of 1,383 brownfield cleanup grant applications in the EPA data, 446 of which were awarded funding and 937 which were not. Applicants could reapply for a cleanup grant in a subsequent year following a rejection. Taking into consideration re-applications, we identified 1,178 unique brownfield properties. 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

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

cleanup grant proposals are included. Out of a total of 1,178 unique brownfields from the EPA data, 584 had associated housing transactions data within 5 kilometers. Currently, the window of observations used for housing transactions starts in 1998 (four years before passage of the Brownfields Law)[19] and ends in 2009, which is the last available year for housing sales. Since brownfield properties with completions after 2009 will not have post-treatment data, these brownfields are not included in the analysis. After removing brownfields with incomplete cleanup, those with missing or miscoded dates, we are left with a final sample of 110 brownfields, 66 that were awarded cleanup grant funding and 44 that were not.[20]

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. Houses with a negative age, calculated as year sold minus year built, were removed as well. 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 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

[19] The extent of geographic coverage by Dataquick became much greater in 1998. Going back further in time would require dropping more brownfields for lack of housing data.

[20] After removing houses located near multiple brownfields (i.e. brownfields that are located in proximity to other sites), we are left with 279 brownfields. There are 72 brownfields with incomplete cleanup by 2010, and 19 brownfields with incomplete cleanup by 2009. There are 78 brownfields that are missing assessments dates (62 for non-awarded sites and 18 for awarded sites). After removing these sites, we are left with 110 brownfields, 66 of which are funded and 44 of which are not.

[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]

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 a graphical information system (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 may be located near multiple brownfields, in which case the effect of cleanup may be hard to measure. The treatment and control groups are then defined using houses within this 5 kilometer radius. Even though the houses outside of 5 kilometers will not be used in the estimation, it is of interest to compare differences between houses close (within 5 kilometers) to brownfields and houses located in the rest of the county (in addition to comparing treatment and control houses within 5 kilometers) in order to motivate the employed definition of treatment. We define both treatment and control groups to be contained in a small area around brownfields (5 kilometers) 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 brownfields 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.

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 – one for pre-assessment transactions and one for post-cleanup transactions. The distance at which the pre-assessment 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.[22] 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 1 describes the average square footage of houses sold at each distance from a brownfield before and after cleanup. It is clear from this figure that houses sold before cleanup within 1 kilometer of brownfields 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 2. Figure 2 also controls parametrically for year effects to allow for general inflationary trends, differences in county-level real estate tax rates, as well as brownfield

[22] The bandwidth, determined by inspection, is three times Silverman's Rule of Thumb. For the distance gradient, this is about 366 meters. For the time gradient, it is approximately 247 days. A Gaussian kernel is used for weighting.

characteristics including the cleanup grant proposal scores and the number of times the sites were assessed.[23]

Figure 2 provides evidence in support of the assumption that houses that are "far" enough from the brownfield 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 1 kilometer. Taking the treatment group to be defined by a 1 kilometer buffer, the simple DID estimator will compare the average change in prices before assessment and after cleanup inside the buffer with the similarly defined change outside the buffer. We demonstrate the sensitivity of some of our results to the assumed buffer size in the following section.

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 leads to an increase in prices of houses near brownfields. Second, the price trends of the two groups in the pre-assessment 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

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

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 3 plots the prices of treatment (i.e., inside 1 kilometer) and control houses against time relative to the cleanup date.[24] 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-assessment to post-cleanup ($9,202) is significantly smaller than that in the treatment group ($45,545). The difference-in-differences approach measures the jump in the treatment group relative to that in the control group – a conservative approach.

4.3 Summary Statistics

Table 3 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 valid 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 3 suggests that mean proposal scores for non-funded brownfields in locations with Dataquick data are marginally higher than mean proposal scores for non-funded brownfields in locations without Dataquick data. The analogous difference in mean proposal scores is not statistically significant for the set of funded properties. Hazardous

[24] As was the case when generating Figure 2, we parametrically control for housing attributes, year effects, RET rates, and brownfield characteristics before non-parametrically estimating price as a function of time relative to the cleanup period.

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 4 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 1 kilometer of a brownfield) and the control group (between 1 kilometer 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 1, which compares houses within 5 kilometers of a brownfield to houses in the rest of the county. We take Table 4 as evidence that there are important differences between treatment and control groups that should be accounted for in the DID specification.

Table 5 provides a yearly breakdown of cleanup starts and completions for the brownfields that were awarded cleanup grant funding. 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 6 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 456 days with a standard deviation of 411 days. Removing the nineteen brownfields with cleanup completions in 2010 (since the housing data only covers transactions up to 2009) gives an average cleanup duration of 410 days with a standard deviation 364 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 7 describes the fraction of cleanups initiated in each year that were not completed by 2009. Not surprisingly, cleanups begun later in time 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 2009 as well.

5. Empirical Results

5.1 Cross-Sectional Estimates

Table 8 reports the results of our simple cross-sectional specification described in equation (1). We find that being near a brownfield that has been cleaned up yields prices that are consistently lower than houses that are not near sites (by -7.7% to -12.2%, depending upon the buffer size), and lower than values of houses that are near brownfields in any other state of cleanup activity. The counterintuitive sign of this effect may result from omitted variables bias if cleaned up brownfields tend to be located near other (unobservable) undesirable land uses. The coefficients for BF, Post-assessment × BF, and Interim × BF, conversely, indicate that prices are generally higher if a house is located near a brownfield in one of these cleanup stages compared to being located far from a brownfield.

Table 9 reports the results of equation (2), where we restrict the comparison to be between houses that are in the vicinity of brownfields – some of which have been cleaned up, and others of which have not. As in Table 8, we find that the value of cleanup is negative, ranging between -19.9% and -22.9%, depending upon the size of the buffer. We also find that even before cleanup, houses near brownfields awarded cleanup grants have systematically lower prices than their non-awarded counterparts, alluding to the possibility that brownfields located in worse off neighborhoods are more likely to be awarded cleanup grants. Together, Tables 8 and 9 suggest that unobservable neighborhood attributes may be correlated with the presence of brownfields and with their cleanup status, necessitating a different empirical approach.

5.2 Fixed Effect Estimates

Next, we use the fixed effects specification described in equation (6), 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 discussed above. The fixed effects specification uses all houses in a buffer; we consider our preferred buffer size of 1000 meters, along with buffers of 1500 meters and 2000 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 effect specification, described in Table 10, differ strikingly from the cross-sectional results, with cleanup resulting in a 7.7% to 11.1% increase in housing prices. Using a 1 kilometer buffer, the effect is 9.3%.

5.3 Difference-in-Differences Estimates

While it is able to deal with time-invariant unobservable neighborhood attributes, the fixed effects specification described in Table 10 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 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 (BF) and its transaction occurring after the cleanup has been completed (POST). These estimates can be found in the seventh row of Table 11. With only year fixed effects and brownfield-level controls, we find a large treatment effect of 25.5% using the preferred buffer size of 1000 meters. In a specification that includes year fixed effects, house-level and brownfield-level controls, and controls for the real estate tax, this effect falls to 4.87%. Further introducing brownfield fixed effects increases this effect to 5.12%. In the specification with brownfield fixed effects, the estimates for the post-assessment and interim period interactions additionally reveal impacts of other events that occur prior to cleanup being completed. We find that upon the completion of a phase II assessment, which confirms environmental contamination at the brownfield, prices near brownfields will fall by 4.8% compared to their pre-assessment levels. The commencement of cleanup subsequently causes a price change in the opposite direction, leading to an average increase in price of 9.9%. The change in price after an assessment is completed may be attributed to the information release that accompanies it, as one of the main objectives of the assessment is to keep the nearby community informed about environmental contamination. A fall in price could be the result of assessment information drawing attention to brownfield contamination or informing nearby residents about hazards of which they were previously unaware. The fact that cleanup interim prices for the treated group are positive is unsurprising, as households may react to the neighborhood becoming observably better when cleanup begins. However, it is slightly puzzling why interim prices (relative to pre-assessment prices) are even higher than prices after cleanup has finished. A possibility could be that the properties that best capture the benefits of remediation sell quickly, and are immediately purchased before cleanup even finishes. Another possibility could be that information released during this period leads to speculative behavior in the market that inflates the prices before the end result is observable.

5.4 Double Difference Matching

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 brownfield may change, and with them, marginal willingness to pay may change as well. In fact, Kuminoff and Pope (2010) 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 time variation. The double difference matching estimator described in Section 3 is designed to do this.

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 up brownfields to houses inside the treatment buffers of sites that have not been cleaned up $(\beta_2 + \pi)$, and the other derived by comparing houses in the control groups of cleaned up sites to houses in the control groups of sites that have not been cleaned up (β_2). Table 12 describes these estimates for our preferred buffer size of 1000 meters using J = 10 matches. Estimates and standard errors are based on Abadie and Imbens (2006).

We do not consider results for 2004, since there is just one brownfield with completed cleanup on which to base the estimates. For 2005, results based on outside-buffer comparisons are not statistically significant, but results are significant for both comparisons in both 2006 and 2007. In particular, we find cleanup effects of 10.0% and 12.8% in these years, respectively. 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). 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.

5.5. Robustness Check

We check the robustness of our results by varying the baseline and treatment periods. If neighborhoods near brownfields are simply trending differently from those far from brownfields, then artificially changing the cleanup completion date should not affect the treatment estimate (i.e., if this were the case, one would still find significant treatment effects after moving the cleanup completion date earlier from the actual date). In Table 13, falsified cleanup completion dates are constructed by moving the cleanup date 1, 2, and 3 years before assessment is completed. Since the previous specifications allude to policy-related effects in the post-assessment and interim time periods, the falsified dates should be moved into the pre-assessment period. Under the falsified post-treatment periods, the treatment effects are all smaller in magnitude and insignificant, which implies that neighborhoods near brownfields are not simply trending differently from those far from brownfields.

6. Benefit-Cost Analysis

Finally, we can address the simple question, “are brownfield cleanups 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 1 kilometer treatment buffer (5.12%), rather than the larger estimates generated by the fixed effect and double difference matching 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 5 years prior to the start of cleanup.[25] We are able to construct this aggregate value for twenty of the brownfields – $812,870,090. Multiplying by 5.12% yields an estimate of the aggregate increase in housing value owing to cleanup of $41,618,949. This represents a benefit of aggregate value per brownfield of $2,080,947. If the $200,000 EPA cleanup grant represented just 1/10th of the total cleanup cost, brownfield remediation would still pass a benefit-cost analysis. This result would be even stronger if we considered all of the properties located inside the treatment buffer, a larger treatment buffer, or one of our larger treatment effect estimates.

7. Discussion

The U.S. 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 report, 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 cleanup brownfields under the program may be worse in other unobserved dimensions.

[25] If a house sells more than once during this period, we use only the last of the transactions prices.

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 "double difference matching" estimator that exploits the advantages of our treatment and control group definitions while not requiring that the hedonic price function remain stable over time. These alternative specifications yield a consistent conclusion – averaging over the experiences at a large number of brownfields, cleanup leads to housing price increases between 5.1% and 12.8%. Moreover, the latter number is consistent with a willingness to pay (i.e., welfare) interpretation, not simply a capitalization effect. Taking the most conservative estimate of the value of an average site cleanup, we find that it indeed passes benefit-cost analysis by an order of magnitude.

While these results constitute strong evidence of the value of brownfield remediation, there are a number of additional dimensions in which future research could extend our analysis. In terms of data, we were only able to identify brownfield locations as points in space. For large or oddly shaped brownfields, this could have important implications for how we defined our treatment and control groups. Most likely, it would have lead us to include in our treatment buffer areas that were, in fact, inside the brownfield property itself, and to include in the control group some houses that have actually been treated. This would lead us to understate the value of a cleanup.

Second, we hope to soon add more recent data on housing transactions, allowing us to include in the analysis information about brownfields that have been more recently cleaned up. If the attributes of brownfields awarded cleanup grants have changed over time (e.g., if more valuable cleanups were completed first), this could have important implications for our estimates.

Finally, we are in the process of exploring the important role played by information provision – in particular, the information contained in phase II assessments – in house price dynamics both before and after cleanup. That this information has an impact is clear from the estimates provided in this report; how and when this information is released can therefore have important implications for the values we attribute to cleanup activities.

References

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

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, D.C., Public Interest Economics Center, for U.S. Environmental Protection Agency.

Banerjee, A. V. and E. Duflo (2009). "The Experimental Approach to Development Economics." *Annual Review of Economics*. 1:151-178.

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

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 17611.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modeling and its Applications.* Chapman and Hall (eds.). Monographs on Statistics and Applied Probability, No. 66.

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 16655.

Kiel, K. A. (1995). "Measuring the Impact of the Discovery and Cleaning of Identified Hazardous Waste Sites on Home Values." *Land Economics*. 74:428-435.

Kohlhase, J.E. (1991). "The Impact of Toxic Waste Sites on Housing Values." *Journal of Urban Economics*. 30: 1-26.

Kuminoff, N. V. and J. C. Pope (2010). "Rosen Bias" in the Willingness to Pay for Nonmarket Goods."

Linden, L. and J. E. Rockoff (2008). "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *American Economic Review*. 98:1103-1127.

Palmquist, R. B. (2005). "Property Value Models." In *The Handbook of Environmental Economics*. Vol. 2, ed. K. Goran-Maler and J. R. Vincent. Elsevier.

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

Rubin, D.B. (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology,* Vol. 66(5):688-701.

Siniavskaia, N. (2011). "Property Tax Rates by County and City." Housing Economics Online, National Association of Home Builders, 2011.

Taylor, L.O. (2003). "The Hedonic Method." In *A Primer on Nonmarket Valuation of the Environment*. , ed. T. Brown P. Champ and K. Boyle. Dordrecht: Kluwer Academic Publishers.

Table 1: House Attributes by within 5 kilometers of a Brownfield versus Rest of County

	Within 5 kilometers		Rest of County		Equality of Means	
Attributes	Mean	Std Dev	Mean	Std Dev	t-stat	Reject?
Price	213,045.341	156,602.063	261,899.797	195,058.313	167.867	Y
Real Estate Rate (County)	10.285	4.479	9.584	4.457	-103.952	Y
Real Estate Rate (Lowest)	5.289	3.053	4.444	2.882	-193.352	Y
Real Estate Rate (Highest)	16.929	6.311	16.168	6.390	-78.799	Y
Age	44.837	32.762	31.893	29.952	-284.029	Y
Square Footage	1,591.919	3,260.815	1,766.912	2,862.458	40.037	Y
Bathrooms	1.911	0.870	2.176	1.035	170.783	Y
Bedrooms	3.033	1.184	3.116	1.203	45.544	Y
Sold in Year Built	0.056	0.229	0.097	0.296	93.611	Y
Condominium	0.183	0.386	0.193	0.394	16.679	Y
Multifamily	0.051	0.220	0.029	0.168	-85.724	Y
Single Family	0.747	0.435	0.761	0.427	20.871	Y
Mobile	0.002	0.049	0.004	0.060	12.933	Y
Misc.	0.016	0.127	0.012	0.111	-23.841	Y
Observations	471,236		6,195,879			

Compares all houses within 5 kilometers of a brownfield (funded or unfunded) before cleanup, to houses located outside 5 kilometers in the rest of the county. The set of brownfields used are those that have completed cleanup by 2009 and are not missing relevant dates.

Table 2: House Attributes by Whether Brownfield is Funded or Unfunded

	Funded Brownfields		Unfunded Brownfields		Equality of Means	
Attributes	Mean	Std Dev	Mean	Std Dev	t-stat	Reject?
Price	160,424.449	103,029.820	151,258.297	126,119.031	-3.553	Y
Real Estate Rate (County)	9.510	3.007	11.557	3.308	29.057	Y
Real Estate Rate (Lowest)	5.393	2.382	7.035	2.514	30.157	Y
Real Estate Rate (Highest)	15.053	2.777	16.108	4.052	13.448	Y
Age	56.556	35.975	61.451	33.693	6.358	Y
Square Footage	1,509.686	666.417	1,536.891	609.069	1.932	N
Bathrooms	1.727	0.766	1.744	0.769	1.013	N
Bedrooms	3.080	1.207	3.169	1.153	3.440	Y
Sold in Year Built	0.012	0.110	0.023	0.151	3.772	Y
Condominium	0.167	0.373	0.073	0.261	-13.379	Y
Multifamily	0.129	0.335	0.058	0.234	-11.285	Y
Single Family	0.699	0.459	0.752	0.432	5.406	Y
Mobile	0.001	0.029	0.011	0.105	5.811	Y
Misc.	0.004	0.062	0.105	0.306	19.686	Y
Observations	3,677		4,556			

Sample includes all houses located within 1 kilometer of a brownfield (funded or unfunded), around the set of brownfields that have completed cleanup by 2009 and are not missing relevant dates. For funded brownfields, attributes are taken from houses selling before cleanup.

Table 3: Brownfield Attributes by Availability of Housing Data

	With Dataquick Data			Without Dataquick Data				
Variable	N	Mean	Std Dev	N	Mean	Std Dev	t-stat	Reject?
Funded and Unfunded								
Petroleum	110	0.245455	0.432326	267	0.183521	0.38782	1.362318	N
Hazardous substances	110	0.772727	0.420988	267	0.842697	0.36477	-1.61684	N
Proposal Score (std.)	110	78.25292	10.9014	186	75.41293	11.21944	2.126671	Y
Funded Only								
Petroleum	66	0.272727	0.448775	83	0.180723	0.387128	1.342575	N
Hazardous substances	66	0.727273	0.448775	83	0.879518	0.327503	-2.39242	Y
Proposal Score (std.)	66	84.41715	4.360875	83	83.51705	4.648741	1.206457	N
Property Size (in acres)	66	7.213939	15.18553	83	13.48012	32.62142	-1.44062	N
Property Ready for Reuse	66	0.69697	0.46309	83	0.722892	0.450291	-0.34469	N

Sample includes brownfields with completed cleanup by 2009 for the funded brownfields, and are not missing relevant assessment, start, and finish cleanup dates.

Table 4: Housing Attributes by Treatment Status (Determined by 1 kilometer Buffer)

	Treat (<= 1 kilometer)		Control (> 1 kilometer)			
Attributes	Mean	Std Dev	Mean	Std Dev	t-stat	Reject?
Price	155,352.056	116,456.898	185,567.734	151,878.125	17.731	Y
Age	59.265	34.814	42.287	29.415	-50.235	Y
Square Footage	1,524.741	635.426	1,507.387	624.359	-2.445	Y
Bathrooms	1.736	0.768	1.811	0.784	8.355	Y
Bedrooms	3.129	1.178	2.946	0.949	-16.759	Y
Sold in Year Built	0.018	0.135	0.056	0.230	14.634	Y
Condominium	0.115	0.319	0.136	0.343	5.362	Y
Multifamily	0.090	0.286	0.033	0.180	-26.504	Y
Single Family	0.729	0.445	0.817	0.387	19.889	Y
Mobile	0.007	0.081	0.006	0.076	-0.874	N
Misc.	0.060	0.237	0.008	0.089	-44.039	Y
Observations	8,233		132,702			

All houses located within 5 kilometers of a brownfield (awarded only), and attributes are taken from houses selling before cleanup.

Table 5: Timeline of Brownfield Start and Completion Frequencies

	With Dataquick Data		Without Dataquick Data	
	Starts	Completions	Starts	Completions
2002	1			
2003	3			
2004	9	5	7	
2005	15	4	17	4
2006	17	17	20	19
2007	11	16	14	15
2008	14	10	23	20
2009	13	14	6	23
2010	2	19	7	11
2011				2
Total	85	85	94	94

Sample includes brownfields with completed cleanup by 2009 for the funded brownfields, and are not missing relevant assessment, start, and finish cleanup dates.

Table 6: Brownfield Cleanup Duration (in days) by Contaminant

Contaminant Funding Type	Mean	Std Dev	N
Petroleum only	437.888889	329.630432	18
Hazardous Substances only	400.291667	379.246948	48
Contaminant Found	Mean	Std Dev	N
Controlled substances	259	0	2
Asbestos	368.142857	370.979126	14
PCBs	470.857143	414.078186	14
VOC	456.923077	384.136322	26
Lead	389.25	419.474335	32
Other metals	453	455.446411	22
PAHs	403.28125	414.563049	32
Other contaminants	455.117647	314.407166	17
Controlled substances	20		1
Media of Contamination	Mean	Std Dev	N
Soil	422.383333	371.124725	60
Air	376.5	447.598602	2
Surface water	172.333333	140.147537	3
Ground water	542.826087	396.993652	23
Drinking water			0
Sediments	23		1
Unknown media	340.5	453.255432	2

Table 7: Fraction of Cleanups Initiated in Each Year (column) that did not Complete Cleanup by 2009

Contaminant Funding Type	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Petroleum only		0.000	0.600	0.200	0.375	0.500	0.600	0.667
Hazardous Substances only	0.000	0.000	0.000	0.143	0.200	0.385	0.650	0.789
Contaminant Found	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Controlled substances				0.000				
Asbestos		0.000	0.000	0.200	0.250	0.500	0.750	0.750
PCBs			0.000	0.143	0.000	0.400	0.750	
VOC		0.000	0.200	0.429	0.231	0.444	0.667	0.714
Lead		0.000	0.000	0.091	0.125	0.250	0.688	0.938
Other metals		0.000	0.125	0.125	0.222	1.000	0.636	0.909
PAHs	0.000	0.000	0.000	0.273	0.167	0.455	0.688	0.800
Other contaminants		0.000	0.000	0.167	0.500	0.600	0.500	0.800
Media of Contamination	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**
Soil	0.000	0.000	0.200	0.182	0.273	0.400	0.667	0.773
Air			0.500					0.000
Surface water			1.000	0.333	0.000		1.000	
Ground water		0.000	0.250	0.300	0.364	0.571	0.765	0.833
Drinking water				0.000		1.000	1.000	
Sediments			0.500		0.500		1.000	1.000
Unknown media						0.500	0.000	1.000

Table 8: Cross-Sectional Specification (Comparison between Houses Inside v. Outside Buffer (b) and Cleaned v. Not Cleaned Brownfields)

VARIABLES	b = 1000m	b = 1500m	b = 2000m
BF	0.0174**	0.0202***	0.0366***
	(0.00867)	(0.00647)	(0.00528)
Post-assess × BF	0.0546***	0.0842***	0.103***
	(0.0138)	(0.00998)	(0.00785)
Interim × BF	0.0675*	0.0522**	0.0273
	(0.0374)	(0.0266)	(0.0193)
Post × BF (= CLEANUP)	-0.0768***	-0.0965***	-0.122***
	(0.0264)	(0.0189)	(0.0140)
Constant	10.65***	10.63***	10.62***
	(0.0949)	(0.0949)	(0.0948)
Observations	143,169	143,169	143,169
R-squared	0.430	0.431	0.432

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Treatment buffer = b. Sample includes only houses (i) inside b buffer (in meters) around funded or unfunded brownfield, (ii) around brownfields with cleanup completed by 2009. Data pooled over years. BF = 1 if house is located within buffer b. Post-assess × BF, Interim × BF, and Post × BF = 1 if nearby brownfield is, respectively, assessed, in the process of cleanup, and cleaned at time of transaction.

Table 9: Cross-Sectional Specification (Comparison between Houses Inside Buffer (b) Near Cleaned v. Not Cleaned Brownfields.)

VARIABLES	b = 1000m	b = 1500m	b = 2000m
Post-assess × BF	-0.0715***	-0.0129	0.0468***
	(0.0165)	(0.0121)	(0.00980)
Interim × BF	-0.0105	-0.0299	-0.0437**
	(0.0361)	(0.0263)	(0.0200)
Post × BF	-0.199***	-0.228***	-0.229***
	(0.0304)	(0.0220)	(0.0175)
Constant	11.99***	11.36***	10.95***
	(0.0950)	(0.0680)	(0.0547)
Observations	8,746	17,006	28,530
R-squared	0.452	0.462	0.497

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Treatment buffer = b. Sample includes only houses (i) inside b buffer (in meters) around funded or unfunded brownfield, (ii) around brownfields with cleanup completed by 2009. Data pooled over years. BF = 1 if house is located within buffer b. Post-assess × BF, Interim × BF, and Post × BF = 1 if nearby brownfield is, respectively, assessed, in the process of cleanup, and cleaned at time of transaction.

Table 10: Fixed Effects

VARIABLES	b = 1000m	b = 1500m	b = 2000m
Post-assess × BF	-0.0571**	-0.0604***	-0.0329**
	(0.0267)	(0.0203)	(0.0160)
Interim × BF	0.0870**	0.0997***	0.128***
	(0.0389)	(0.0292)	(0.0227)
Post × BF	0.0929**	0.0771**	0.111***
	(0.0415)	(0.0315)	(0.0252)
Observations	4,227	8,116	14,631
R-squared	0.999	0.999	0.998
Controls			
Year Fixed Effects	×	×	×
House Controls	×	×	×
Real Estate Tax (RET) rate*	×	×	×

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Treatment buffer = b. Sample includes only houses (i) around awarded brownfields, (ii) inside b buffer (in meters), (iii) around brownfields with cleanup completed by 2009. BF = 1 if house is located within buffer b. Post-assess × BF, Interim × BF, and Post × BF = 1 if nearby brownfield is, respectively, assessed, in the process of cleanup, and cleaned at time of transaction.

Table 11: Difference-in-Differences (b = 1000 meters)

VARIABLES	(1)	(2)	(3)
BF	0.0356**	0.0548***	-0.0620***
	(0.015)	(0.011)	(0.010)
Post-assess	-0.0688***	0.0603***	0.0254***
	(0.008)	(0.006)	(0.006)
Interim	0.391***	0.275***	0.203***
	(0.012)	(0.009)	(0.009)
Post	0.107***	0.143***	0.187***
	(0.012)	(0.009)	(0.010)
Post-assess × BF	-0.0766***	-0.122***	-0.0480***
	(0.028)	(0.021)	(0.018)
Interim × BF	0.155***	0.0642*	0.0988***
	(0.048)	(0.036)	(0.032)
Post × BF	0.255***	0.0487*	0.0512**
	(0.034)	(0.026)	(0.023)
Assessed twice	-0.454***	-0.417***	
	(0.006)	(0.005)	
Proposal score (std.)	0.0125***	0.0222***	
	(0.001)	(0.001)	
Constant	10.87***	10.03***	9.314***
	(0.072)	(0.106)	(0.463)
Observations	94,596	94,596	94,596
R-squared	0.109	0.497	0.606
Controls			
Year Fixed Effects	×	×	×
House Controls		×	×
Real Estate Tax Controls		×	×
Brownfield Fixed Effects			×

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Treatment buffer = 1000m. Sample includes only houses (i) around awarded brownfields and (ii) around brownfields with cleanup completed by 2009. BF = 1 if house is located within 1000m. Post-assess × BF, Interim × BF, and Post × BF = 1 if nearby brownfield is, respectively, assessed, in the process of cleanup, and cleaned at time of transaction.

Table 12: Double-Difference Matching Estimator (b = 1000 meters, 10 Matches)

	Inside Treatment Buffer					Outside Treatment Buffer					Average Treatment Effect on the Treated
	# Matches		SE	N(Treat)	N(Control)	# Matches		SE	N(Treat)	N(Control)	
2005	10	0.152***	(0.0481)	30	275	10	-0.0456	(0.0527)	688	2265	0.1976
2006	10	0.180	(0.110)	145	150	10	0.0796*	(0.0466)	2560	1534	0.1004
2007	10	1.048***	(0.137)	178	29	10	0.920***	(0.0611)	3805	539	0.128

No estimates for 2008 or 2009, since there is no pre-assessment data in those years, as all houses have assessments completed by 2008. No estimate for 2004, because there is post-cleanup data in that year.

Table 13: Treatment Estimates under False Policy Complete Dates

VARIABLES	-1 year	-2 years	-3 years
BF	-0.0662***	-0.0714***	-0.0808***
	(0.011)	(0.013)	(0.014)
Post	0.0581***	0.0594***	0.0762***
	(0.006)	(0.006)	(0.006)
Post × BF	0.0101	0.0166	0.0267
	(0.015)	(0.016)	(0.017)
Constant	9.823***	9.825***	9.842***
	(0.464)	(0.464)	(0.464)
Observations	94,596	94,596	94,596
R-squared	0.603	0.603	0.603
Controls			
Year Fixed Effects	×	×	×
Brownfield Characteristics			
House Controls	×	×	×
Brownfield Fixed Effects	×	×	×

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Treatment buffer = 1000 meters. Falsified cleanup completion dates are moved 1, 2, and 3 years before assessment is complete

Figure 1: Average Square Footage of Houses Transacted by Distance from Brownfield
Before v. After Remediation With 1% Confidence Intervals

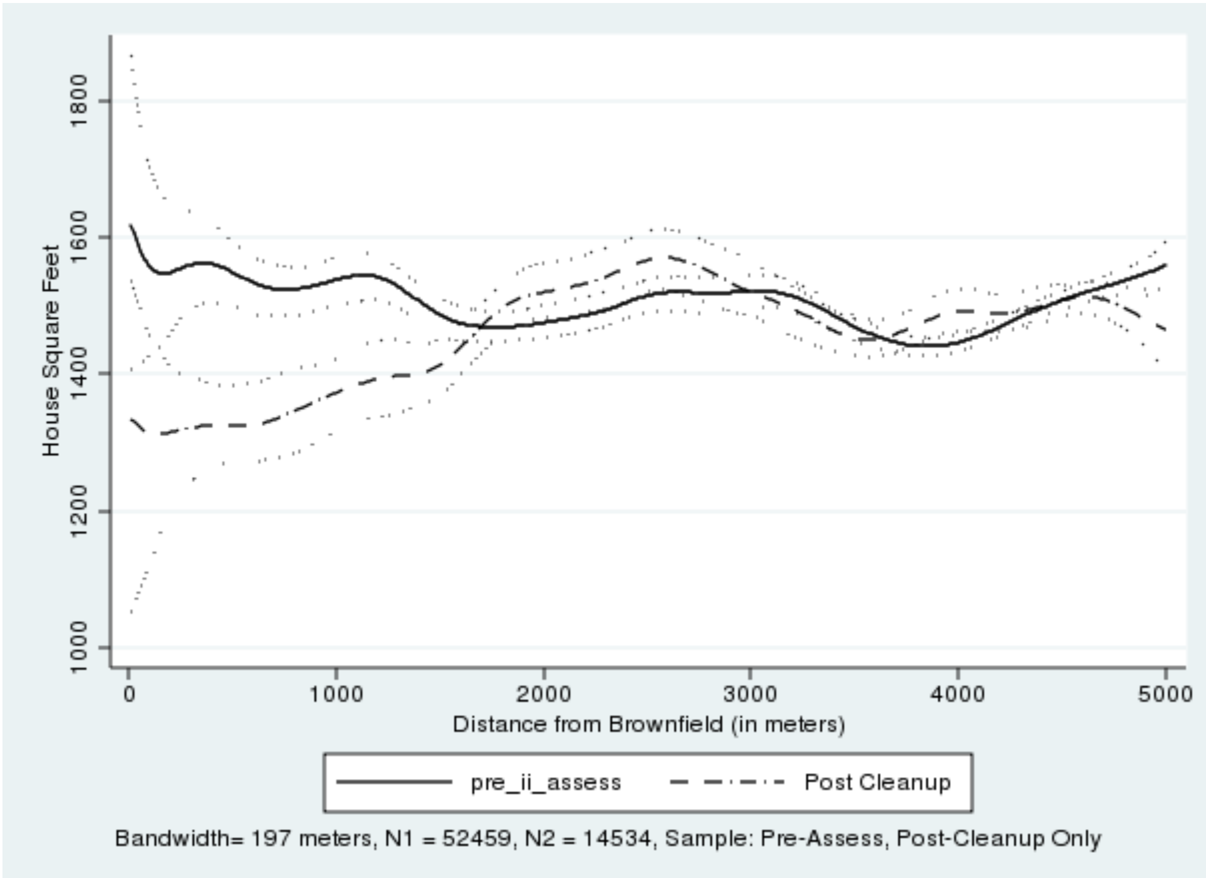


Figure 2: Non-Parametric Price Function Estimates (Price v. Distance to Nearest Brownfield) Before and After Remediation With 1% Confidence Intervals

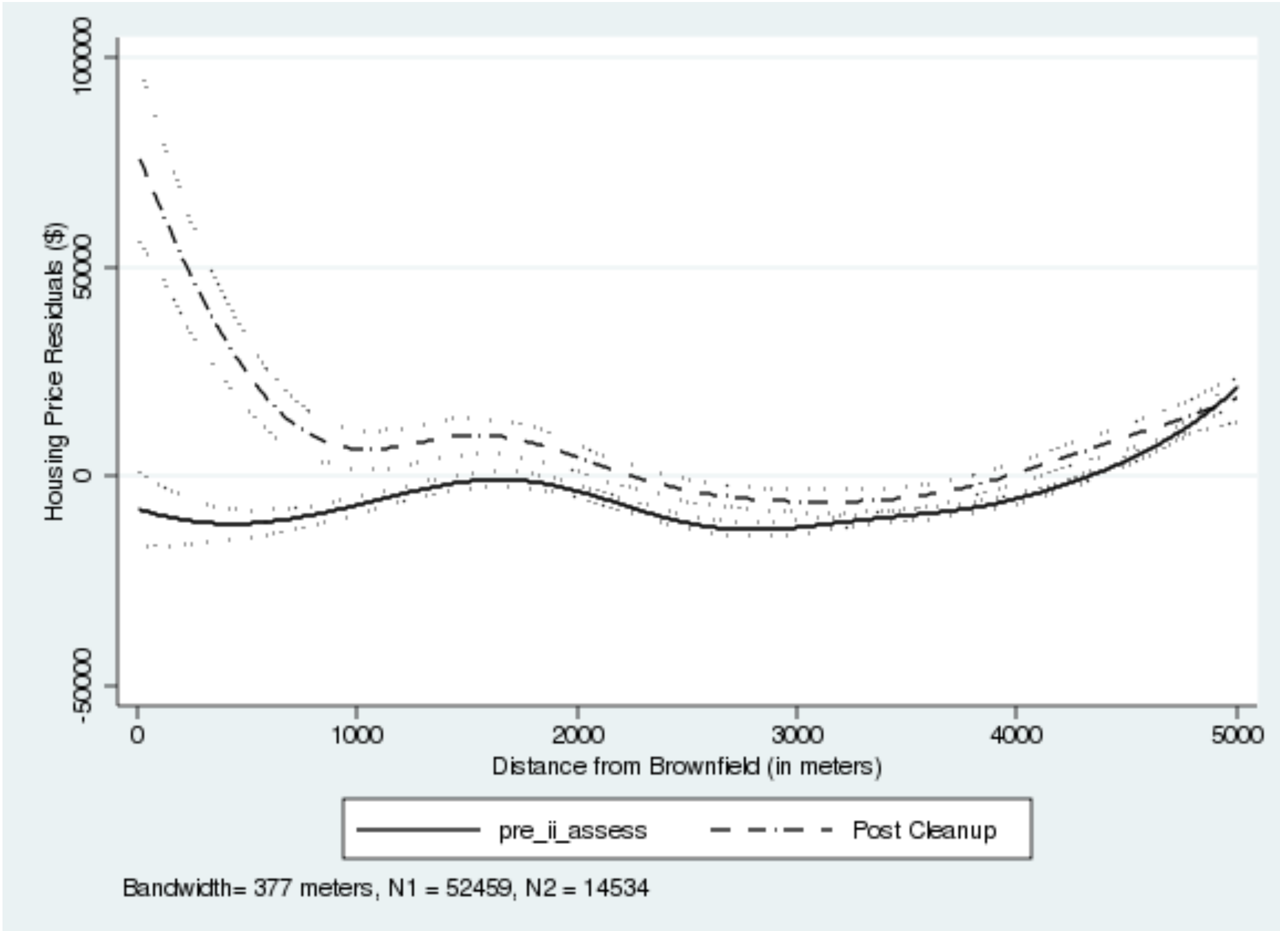
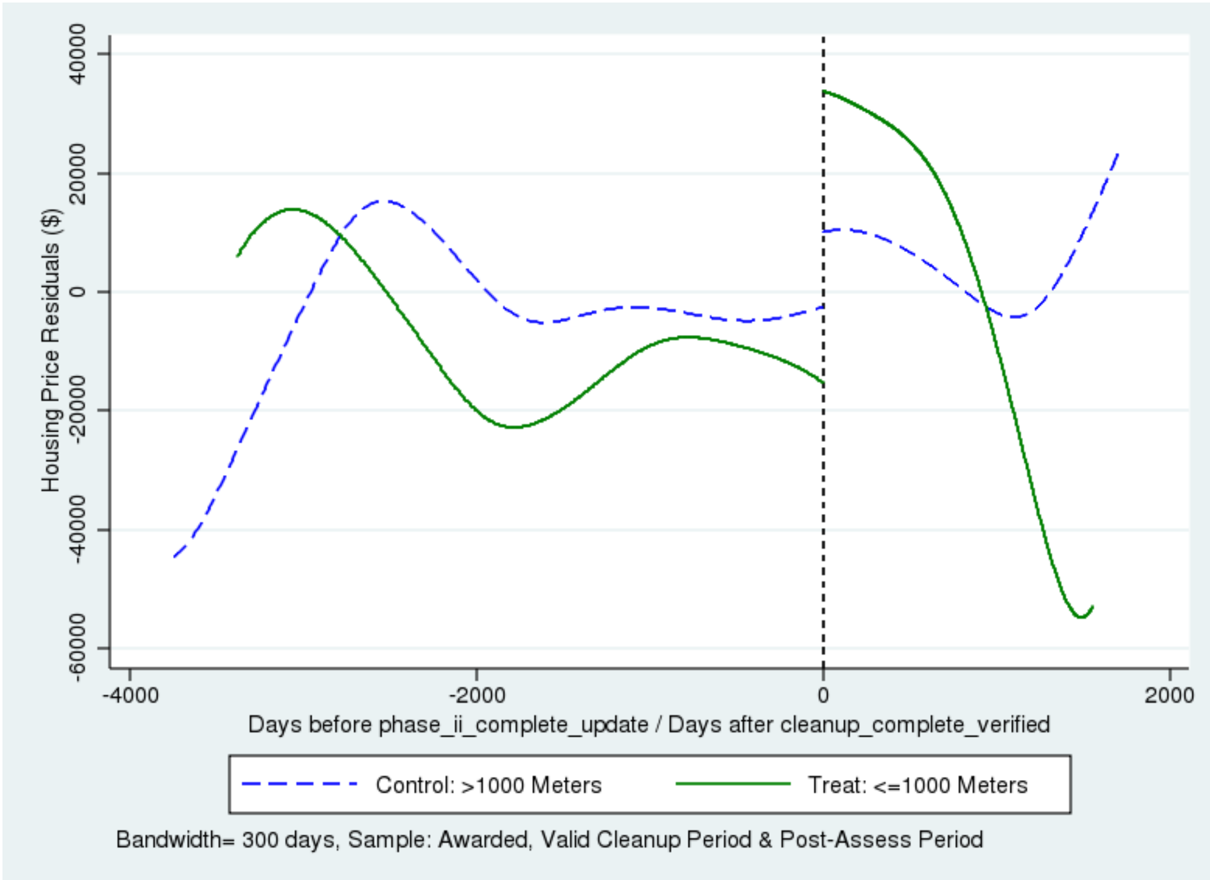


Figure 3: Average Price Relative to Cleanup Period in Cleanup and Control



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:

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

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,

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

where $K_h(g)$ is a Gaussian kernel; i.e. $K_h\left(z\right) = \frac{1}{h}K_h\left(\frac{z}{h}\right) = \frac{1}{h}\phi\left(\frac{z}{h}\right)$, 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.