The Value of Brownfield Remediation

Kevin Haninger, Lala Ma, Christopher Timmins

Abstract: The US Environmental Protection Agency Brownfields Program awards grants to redevelop contaminated lands known as brownfields. This paper estimates cleanup benefits by combining administrative records for a nationally representative sample of brownfields with high-resolution, high-frequency housing data. With cleanup, we find that property values increase by an average of 5.0% to 11.5%. For a welfare interpretation that does not rely on the intertemporal stability of the hedonic price function, a double-difference matching estimator finds even larger effects of up to 15.2%. Our various specifications lead to the consistent conclusion that Brownfields Program cleanups yield positive, statistically significant, but highly localized effects on housing prices.

JEL Codes: Q51, R11

Keywords: Brownfields, Difference in differences, Nearest neighbor matching, Property value hedonics

LAND REVITALIZATION is a beneficial, yet costly, process to undertake. Lands are often contaminated with harmful substances that require expensive procedures to treat. In some cases, toxic waste sites pose a direct threat to human health. In other cases, sites may pose a low risk to nearby residents but are left unused or underused until even small amounts of contaminants are removed. Most would agree upon the importance

Kevin Haninger is at the US Environmental Protection Agency and American Association for the Advance of Science (AAAS), 1200 Pennsylvania Avenue NW, Washington, DC 20460 (kevin.haninger@gmail.com). Lala Ma is at the University of Kentucky, Department of Economics, Business and Economics Building, Lexington, KY 40506 (lala.ma@uky.edu). Christopher Timmins (corresponding author) is at Duke University, Department of Economics, and National Bureau of Economic Research, Social Sciences Building, PO Box 90097, Durham, NC 27708 (christopher.timmins@duke.edu). We gratefully acknowledge the US Environmental Protection Agency (EPA) Office of Land and Emergency Response Center for Program Analysis, the Office of Brownfields and Land Revitalization (OBLR), and EPA Regional Offices for providing financial support, program expertise, and access to the data. We especially thank Ryan Smith of the OBLR for his partnership. We would also like to thank seminar par-

Received September 26, 2015; Accepted May 23, 2016; Published online January 26, 2017.

JAERE, volume 4, number 1. © 2017 by The Association of Environmental and Resource Economists. All rights reserved. 2333-5955/2017/0401-0006$10.00 http://dx.doi.org/10.1086/689743

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 preceding redevelopment makes the surrounding area 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 US Environmental Protection Agency (EPA) has designated these lower-risk sites as brownfields and has aimed to promote their revitalization through grant funding.

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

Although our paper is not the first to estimate brownfield sites' impacts on nearby residential and commercial property values (Ihlanfeldt and Taylor 2004; Leigh and Coffin 2005; Linn 2013; Savchenko and Braden 2014), we advance the existing body of work on two important fronts. First, while previous papers are careful in their investigation of the potential threats from correlated unobservables, all focus on only a subset of brownfields within one or two states. To our knowledge, ours is the first paper that combines EPA administrative records with high-resolution, high-frequency housing data to estimate the effects of brownfield cleanup across the entire federal Brownfields Program. Next, only under certain conditions that we describe below can the capitalization of disamenities into local housing markets be given a welfare interpretation (Kuminoff and Pope 2014). We utilize different sources of variation available in our unique data to estimate cleanup benefits without relying on those assumptions, making our estimates particularly useful for cost-benefit analysis.

1. IDENTIFYING THE EFFECTS OF BROWNFIELD REMEDIATION

In an ideal research environment, one would randomly select brownfield sites for cleanup and observe the impacts of that cleanup on nearby housing prices. The ran-

---

ticipants at Iowa State University, University of Kentucky, and Center for Environmental and Resource Economic Policy Camp Resources for helpful comments. Any opinions and conclusions expressed herein are ours and do not necessarily represent the views of the EPA or the National Bureau of Economic Research. This work was conducted while Kevin Haninger was an AAAS Policy Fellow at the EPA. All errors are our own.

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

dom selection of sites into the remediation process would guarantee that unobservable determinants of changes in local housing prices would not be correlated with changes induced by remediation, allowing the researcher to cleanly identify the latter. While more common in some areas of research, opportunities for these sorts of experiments are not often available in environmental economics.[2] Indeed, it is the case that the Brownfields Program awards cleanup grants based on a competitive process. The outcome of this process may lead to the award of cleanup funds to locations that differ systematically from locations that do not receive funds. We therefore adopt a variety of quasi-experimental approaches to identifying the effect of cleanup on brownfields, including fixed effects and differences in differences (DID).

The fixed effects and DID approaches to recovering the benefits of site remediation suffer from a similar problem. In particular, each requires an assumption that the hedonic price function, which describes the equilibrium relationship between house attributes (including exposure) and price, is stable over time. However, recent work on environmental gentrification, a process whereby changes in the socioeconomic characteristics of a community accompany changes in environmental amenities, suggests that there may be substantial neighborhood turnover in response to brownfield redevelopment (Banzhaf and McCormick 2007; Wolverton 2009), rendering this assumption questionable. Put differently, with a new local population, the willingness to pay for not being exposed to an untreated brownfield site that is revealed by the hedonic price function may be very different after cleanup. Kuminoff and Pope (2014) show that the results of simple fixed effects estimation of the price response to cleanup may therefore fail to identify the marginal willingness to pay (MWTP) of those living in proximity to the brownfield either before or after cleanup. Instead, it will recover a "capitalization" effect (i.e., the simple response of price to a cleanup). 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 return a change in welfare. According to Banzhaf (2015), the capitalization estimate recovered from DID does, however, represent a lower bound for the equivalent surplus of an improvement in some amenity. Because this measure may still underestimate benefits by a large margin, to overcome this problem we suggest a DID nearest-neighbor matching estimator (DD-NNM) that does not require any comparisons over time as an alternative to the traditional DID estimator.

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

---

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

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

2. THE EPA BROWNFIELDS PROGRAM

A brownfield is a "property, the expansion, redevelopment, or reuse of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant."[3] Typically, brownfields are lands that were previously used for industrial or commercial purposes and include areas that are contaminated by low concentrations of hazardous substances. These sites are diverse in nature and can range from being old dry cleaning establishments and gas stations to processing plants for materials such as steel, bricks, and asbestos. Generally, brownfields pose a 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 US Government Accountability Office estimates that there are more than 450,000 brownfields nationwide. In 1995, the US Environmental Protection Agency initiated the Brownfields Program to assist public and private sector organizations in revitalizing brownfields, mainly by providing grant funding. The aim was not only to improve the environment but also to promote social and economic reinvestment in these unused lands. In 2002, the Small Business Liability Relief and Brownfields Revitalization Act (i.e., the "Brownfields Law") was signed as an amendment to the Comprehensive Environmental 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.

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

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

This paper focuses on the effect of cleanup grants on housing values. Cleanup grants provide up to $200,000 to perform cleanup activities at a brownfield site contaminated by petroleum or hazardous substances. While the amount of cleanup funding is small in the grand scheme of land redevelopment, grants from the program are perceived to provide liability relief, the lack of which is seen to be a real deterrent to development as a result of lawsuits pursued under CERCLA (Alberini et al. 2005). Entities eligible for cleanup grants include state, local, and most tribal governments; quasi-government entities such as redevelopment agencies; and nonprofit organizations. Eligible entities must have completed a site assessment to determine the contamination to be remediated and have sole ownership of the cleanup site before proposal submission.[4] Due to budgetary limitations, no eligible entity may apply for funding cleanup activities at more than three sites. Cleanup grants require a 20% 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 above requirements represent "threshold criteria" against which cleanup grant proposals are evaluated. Applicants must pass all threshold criteria in order to qualify for funding. Other threshold criteria include community notification, opportunity for public comment prior to proposal submission, and a letter from the appropriate state or tribal environmental authority acknowledging that the applicant plans to apply for federal brownfield assistance.

Conditional upon passing all threshold criteria, the proposal will receive a numerical score from the evaluation panel. Scores are based on several evaluation fields, including community need, project description and feasibility, community involvement and partnerships, and reduction of threats to human health and the environment. Once scored, cleanup grant proposals are ranked from highest to lowest score and then awarded funding in rank order until the program budget has been exhausted.[5] Since passage of the Brownfields Law through FY2015, the EPA has competitively awarded 1,128 cleanup grants totaling approximately $214.9 million.

If a proposal is not awarded in one year, the applicant can reapply in a subsequent year.[6] This implies that the brownfield site could be associated with different proposal

---

4. Historically, many brownfields are owned by local governments due to bankruptcy of a previous business that occupied the site.

5. Guidelines for cleanup grants can be found at http://www.epa.gov/brownfields/applicat.htm. The score and award cutoff may initially seem well suited for identification using regression discontinuity (RD). However, since the main score is the sum of the various subscores for each evaluation field, and each field describes a very different aspect of the applicant, it is unclear what RD would control for in absence of proposal subscores, which we do not have.

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

scores and different award statuses. We take the applicant's most recent score and application outcome, assuming that it represents the applicant's best and most knowledgeable proposal effort. More details on how scores are compared across grant years are provided in section 4.

3. MODEL AND IDENTIFICATION

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

The hedonic method is based on the idea that a homeowner's disutility from living in close proximity to a brownfield site can be measured by observing compensating price differentials in housing markets. In general, the homeowner's marginal willingness to pay (MWTP) for some desirable attribute (e.g., distance from a brownfield site) can be read off of the hedonic gradient (i.e., the derivative of the hedonic price function), owing to the sorting behavior of utility-maximizing homeowners. Rosen's seminal paper (Rosen 1974) and the literature it sparked describe procedures for recovering the MWTP functions for heterogeneous individuals. Bishop and Timmins (2011) describe many of the difficulties encountered in this exercise—because of these difficulties, the typical approach in the applied hedonics literature has been to ignore this heterogeneity and recover a function that describes price as either a linear function of distance or as one that treats exposure discretely, defining it according to whether a house falls inside a particular distance band drawn around a brownfield. That is the approach we adopt here.

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

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

to contaminated land that was remediated under the federal Brownfields Program and returned to potentially productive use.[8]

3.1. Cross-Section and Fixed Effects

The simplest specification ignores any panel variation in the data, and compares houses exposed to sites that have been awarded cleanup under the Brownfields program to those that have not. Potential bias arises if brownfields that received treatment were systematically different in unobservable ways from those that did not receive treatment. An inspection of attributes by award status (available in table A2 of the appendix; tables A1–A6 and appendix available online) suggests several reasons to be concerned about the results of a cross-sectional comparison.[9]

A common approach to dealing with unobserved house and neighborhood characteristics that may be correlated with brownfield remediation is to exploit variation in panel data. Considering houses in close proximity (e.g., within 5 kilometers) of awarded sites only, we compare prices of houses sold before the nearby brownfield was cleaned to prices of those sold after to control for any permanent unobservable differences between places that received cleanup treatment and those that did not.[10] We also introduce fixed effects at the brownfield (or house) level to remove time-invariant differences across sites (or houses).[11]

---

8. While cleanup standards oftentimes depend on anticipated future use so that the treatment impact could theoretically include the value of the site after it becomes redeveloped, there is much uncertainty about the specific reuses of a site. For example, only 35% of awarded brownfields in our sample had any form of future use information at the time of the application, suggesting that the specific future uses have not been determined. With regard to sites that have undergone and completed cleanup, to our knowledge there is currently no data on post-cleanup uses.

9. Table A2 compares houses surrounding awarded brownfields that have been cleaned from our sample to those surrounding brownfields that have not been awarded with cleanup. There are statistically and economically significant differences in house attributes by site award status—for example, houses near an awarded site are less expensive, tend to be older and have fewer bathrooms than those near a nonawarded site. These large differences in observables suggest that there may also be differences in unobservable attributes of each of these groups of sites. Attribute comparisons by exposure to a site versus not at all (table A3) reveal even larger differences. Part of the depressed housing prices surrounding areas with brownfield sites may also be attributed to excess supply observed in these markets (Vigdor 2002). For subsequent methods, we thus limit the analysis to only using houses within 5 kilometers of brownfield sites.

10. We present estimates from using multiple distance buffers to demonstrate robustness.

11. No county or state fixed effects were included in brownfield-level fixed effects specifications. Since it is possible that our fixed effects estimates result from a comparison of houses belonging to different states or cities even with brownfield-level fixed effects (e.g., for sites that are located near a jurisdictional boundary), we provide fixed effects estimates that include county-level fixed effects. Estimates with county-level fixed effects do not change our results significantly and are included in table A6 of the appendix.

3.2. Difference-in-Differences (DID)

Let P_{itk} be the log of the price of house i in the neighborhood surrounding brownfield site k at time t. At some point in time, site k is cleaned. Consider only houses in the vicinity of brownfields that are cleaned (5 kilometers), and let the treatment group of houses be defined by those that are close enough (i.e., closer than 5 kilometers) to be affected by that cleanup. A specific definition and estimation of treatment is discussed in section 4.3; the intuition is that these houses are particularly close to the brownfield, while there may be other houses in the same local neighborhood that experience the same local public goods but are far enough from the brownfield to not be "treated" by it. We define this distance below.

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

$$P_{itk} = \beta_0 + \beta_1 TREAT_{ik} + \beta_2 POST_{itk} + \pi TREAT_{ik} \times POST_{itk} + u_{itk}, \quad (1)$$

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

$$\begin{aligned} \pi = {} & (E[P^1_{i1k}|TREAT_{ik} = 1] - E[P^0_{i0k}|TREAT_{ik} = 1]) \\ & - (E[P^0_{i1k}|TREAT_{ik} = 0] - E[P^0_{i0k}|TREAT_{ik} = 0]), \end{aligned} \quad (2)$$

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

$$\begin{aligned} & E[P^0_{i1k}|TREAT_{ik} = 1] - E[P^0_{i0k}|TREAT_{ik} = 1] \\ & = E[P^0_{i1k}|TREAT_{ik} = 0] - E[P^0_{i0k}|TREAT_{ik} = 0]. \end{aligned} \quad (3)$$

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

$$P_{itk} = \beta_0 + \beta_1 TREAT_{ik} + \beta_2 POST_{itk} + \pi TREAT_{ik} \times POST_{itk} + X'_{itk}\delta + u_{itk}. \quad (4)$$

In practice this regression model can be expanded to include multiple groups and multiple treatment periods. For our application to brownfield cleanup, we separate the period before cleanup completion into two periods, the period before cleanup commences and an interim period when cleanup takes place, and make all comparisons to prices before cleanup activities begin.

3.3. Difference-in-Differences Nearest Neighbor Matching (DD-NNM)

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

Returning to the specification used to estimate the difference-in-differences model in section 3.2, but allowing all of the parameters of the hedonic price function to vary with time, we index each observation by i (house), t (year), and k (site near to which house i is located). Some of the sites have been cleaned under the Brownfields Program by time t ($CLEANUP_{itk} = 1$) while others have not ($CLEANUP_{itk} = 0$). Note that we include the set of houses near sites that applied for but were denied funding (i.e., $CLEANUP_{itk} = 0 \quad \forall t$).[12] Finally, we include a flexible function of house, brownfield, and neighborhood attributes (X). We consider only transactions that occur in a particular year t; we therefore do not need to differentiate between pre- and post-treatment periods. Instead, we only need to differentiate between sites that have and have not been cleaned:

$$\begin{aligned} P_{itk} = {} & \beta_{0t} + \beta_{1t} TREAT_{ik} + \beta_{2t} CLEANUP_{itk} \\ & + \pi_t TREAT_{ik} \times CLEANUP_{itk} + f(X_{itk}; \theta_t) + u_{itk}. \end{aligned} \tag{5}$$

Our DD-NNM method proceeds in two stages. We describe each stage below and highlight some advantages that DD-NNM has over the DID estimator at the end of this section.

12. While these sites may not have been cleaned under the federal Brownfield Program, they could, in theory, experience cleanup. We discuss the implications of this on our results in section 4.1.

DD-NNM: Stage 1

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

$$P_{itk} = (\beta_{0t} + \beta_{1t}) + (\beta_{2t} + \pi_t)CLEANUP_{itk} + f(X_{itk}; \theta_t) + u_{itk}. \tag{6}$$

Using a nearest-neighbor matching algorithm, we pair each house inside the treatment buffer in each neighborhood with $CLEANUP_{itk} = 1$ with a set of J houses that are as similar as possible in the covariate space and located inside the treatment buffer of a neighborhood with $CLEANUP_{itk} = 0$.[13]

Specifically, for a particular house i located in the treatment buffer of a cleaned site (price designated by P_{itk}), we find the $J = 5$ "nearest neighbors" to i, t, k from houses located in the treatment buffer of an uncleaned site (prices denoted by $P_j^{(itk)}$). We use these matches to construct the counterfactual outcome for each house i had it not been cleaned, and save individual treatment effects for stage 2 into a vector $\boldsymbol{P}_t^{In}$ of length N_t. The treatment effect averaged over all of the houses near awarded sites is given by

$$(\beta_{2t} + \pi_t) = \frac{1}{N_t}\sum_{i=1}^{N_t}\left(P_{itk} - \frac{1}{J}\sum_j P_j^{(itk)}\right). \tag{7}$$

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

$$\beta_{2t} = \frac{1}{\widetilde{N}_t}\sum_{i=1}^{\widetilde{N}_t}\left(\widetilde{P}_{itk} - \frac{1}{J}\sum_j \widetilde{P}_j^{(itk)}\right). \tag{8}$$

We similarly save the individual treatment effects into a vector $\boldsymbol{P}_t^{Out}$ of length $\widetilde{N}_t$.

The success of this strategy, of course, depends upon being able to find high-quality matches for houses in neighborhoods around awarded sites that have been cleaned from the set of houses around sites that have not been awarded with cleanup. This is what assures that the unspecified function $f(X_{itk}; \theta_t)$ will be differenced away. To do this, we match based on house and brownfield characteristics, restrict matches to be among sites in the same state, and eliminate other forms of heterogeneity at the neighborhood level by matching on attributes of the census tract in which sites are located. Match variables are described in section 4.1.

13. The Mahalanobis metric is used to measure the distance between two sets of covariates. The number of matches used may exceed J in the case of ties.

Before proceeding to stage 2, we take an additional step to correct for potential bias in our first-stage estimates from inexact covariates matches from the set of nearest neighbors (Abadie and Imbens 2011). We use the bias-corrected matching estimator from Abadie and Imbens (2011) to account for differences in covariate values between the treated observation and its matched counterparts. Furthermore, this has an additional benefit of producing estimates that are more robust to the number of matches used (Abadie and Imbens 2011).

Take a house i that is located within the treatment buffer of a cleaned site, and let $\mu_0(X_{itk})$ denote its conditional expected price, given attributes X_{itk}, had it been near an uncleaned site,

$$\mu_0(X_{itk}) = E[P^0_{itk}|X_{itk}]. \tag{9}$$

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

$$\hat{\mu}_0(X_{jtk}) = X'_{jtk}\hat{\theta}_{w=0}, \tag{10}$$

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

$$\hat{\mu}_0(X_{itk}) = X'_{itk}\hat{\theta}_{w=0}. \tag{11}$$

Finally, the bias-adjusted estimate for an individual i replaces the counterfactual from the simple matching estimator, $(1/J)\Sigma_j P_j^{(itk)}$, with $(1/J)\Sigma_j \hat{P}_j^{(itk)}$

$$\begin{aligned} P^{In,bcm}_{i,t} &= P_{itk} - \frac{1}{J}\sum_j \hat{P}_j^{(itk)} \\ &= P_{itk} - \left(\frac{1}{J}\sum_j P_j^{(itk)} + \hat{\mu}_0(X_{itk}) - \hat{\mu}_0(X_{jtk})\right), \end{aligned} \tag{12}$$

where the adjustment accounts for the difference in the counterfactual outcome due to covariate differences in the matched observation. The bias-corrected estimator is used to recover treatment effects for houses near ($P^{In,bcm}_t$) and far ($P^{Out,bcm}_t$) from awarded brownfields. We then stack these bias-corrected estimates into a vector $\boldsymbol{P}^{bcm}_t$ of length $N_t + \widetilde{N}_t$.

DD-NNM: Stage 2

We are able to recover an estimate of the average treatment effect on the treated for each year t by simply differencing the average estimates from stage 1 in the following manner,

$$\pi_t = \frac{1}{N_t}\sum_{i=1}^{N_t}\left(P_{itk} - \frac{1}{J}\sum_j \hat{P}_j^{(itk)}\right) - \frac{1}{\widetilde{N}_t}\sum_{i=1}^{\widetilde{N}_t}\left(\widetilde{P}_{itk} - \frac{1}{J}\sum_j \widetilde{\widehat{P}}_j^{(itk)}\right). \tag{13}$$

However, even after limiting houses to the immediate vicinity of brownfields, there may still be reason to believe that systematic differences exist between houses near and far from sites.[14] Therefore, it will be important to control for observable differences in characteristics. As such, we estimate the treatment effect on the treated, π_t, by regressing $\boldsymbol{P}_t^{bcm}$ on a dummy variable for whether the house is located inside of the buffer while controlling for the same set of characteristics used in matching and restricting comparisons to be within brownfield sites (i.e., including site fixed effects).

The DD-NNM necessarily means comparing different geographic areas, a trade-off we make that yields a change in cleanup status without using temporal variation. Although it is true that this method replaces one assumption about time invariant price gradients with another, namely, that gradients are the same over space, we argue that this is superior to DID for two reasons. First, cross-sectional comparisons are better able to control for re-sorting in populations in response to cleanup. How much better (as compared to comparing across time) may be context specific. In section 4, we provide empirical evidence that, at least in our context, temporal comparisons from DID potentially leave changes in race distributions from cleanup uncontrolled, whereas most of these differences in race are removed with cross-sectional comparisons as used in DD-NNM. Second, the DD-NNM estimator does more than DID to avoid a violation in the stability assumption by using strict matching (e.g., only looking within the same state and time period) in conjunction with the bias correction. While there are versions of DID that more flexibly control for differences across observable attributes (see Heckman et al. [1998] for an example), our evidence from section 4 gives us more confidence that matching and bias adjustment do a better job of controlling for observable and, we hope, unobservable, differences.

4. DATA

Brownfields, like many other disamenities (Superfund sites, TSDFs, TRI plants) may have very localized impacts on house prices. As such, it can prove difficult to recover these impacts without access to high-resolution data. Cleanup of a brownfield, for example, may not be perceptible in information about census tract median housing prices, while it may in fact have large impacts on nearby houses. One solution to this problem is to use high-resolution decennial census block-level data (Gamper-Rabindran, Mastromonaco, and Timmins 2011). That approach, however, introduces potential problems. Low-frequency decennial data may confound brownfield cleanup with other unobserved events that occurred at some other time during the same decade. Unlike

14. This will be evident from group mean comparisons in section 4.3

Superfund remediation, brownfield cleanups can be relatively quick, leaving a great deal of remaining time over a 10-year period for other things to happen.

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. In the following four subsections, we describe the data, estimate the treatment buffer, provide summary statistics (based on our definition of treatment), and describe neighborhood turnover in brownfield neighborhoods that additionally motivate the concern for time-varying hedonic price functions.

4.1. Data Description

EPA provided nonpublic administrative data on all cleanup grant applicants and proposal scores since passage of the Brownfields Law in 2002 through 2008, as well as public data on the subset of brownfields that were awarded funding. The data provide characteristics of brownfields, including the exact location (latitude and longitude),[15] property size (for awarded sites only), dates of brownfield assessments, and types of grant application (i.e., targeted to treat petroleum sites, sites with hazardous substances, or both). For the properties that were awarded funding,[16] the data include related award and cleanup progress information. Since funding for brownfields varies each year and is awarded beginning with the highest scoring applicant and working downward until funding runs out, there is not one score cutoff that determines whether a property is cleaned. Moreover, because of changing scoring rules, the raw scores are difficult to compare across competition years. To make scores comparable across years, we standardize the scores to be between 0 and 100 by dividing the raw score by the maximum score in its respective competition year. Where site-level fixed effects are not included, the following brownfield characteristics are used as controls: the standardized proposal score, and indicators for whether a site is assessed twice, a petroleum grant applicant, and/or a hazardous substances grant applicant.

There are a total of 1,383 brownfield applications, 446 of which are awarded cleanup grants and 937 are not. Applicants could reapply for a grant in another year following a rejection. Taking into consideration reapplications, we identified 1,178 unique brownfield properties. After removing brownfields with missing or inaccurate longitude

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

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

and latitude coordinates, we are left with 797 sites (437 awarded and 360 nonawarded). Property locations were individually verified with Google maps and checked to ensure that the background of the reported location corroborated with the information from the grant proposal. Figure 1 plots the brownfield sites in our sample against states in the continental United States.

Dates of different milestones in the process to remediate the brownfield start with site assessment and end with cleanup. We consider all houses sold before any cleanup activities commence to belong to a period "pre-cleanup." Next, we define an interim treatment period that starts from the earliest recorded cleanup start date and ends on the cleanup completion date.[17] We distinguish this interim period to control for expectations of future development that potentially affect prices before cleanup is complete.[18] Last, we define the post-cleanup period during which properties have been fully treated with brownfield cleanup as starting with the cleanup completion date and lasting for the duration of our sample.

The time period dummy variables that will be used in all of the specifications are $INTERIM_{itk}$ and $POST_{itk}$, which respectively equal 1 if a house is sold during and after cleanup of the nearby brownfield. For the DID specification, interactions between each of the above time period dummies with the treatment dummy are included. In that specification, the coefficient on $TREAT_{ik} \times POST_{itk}$ is the treatment effect on the treated and should be interpreted with respect to the houses in the pre-cleanup period, which is the omitted group.

The second data source comes from Dataquick Information Systems, used under a license agreement with the Duke University Department of Economics, which provides housing transactions data. These data contain the history of transactions and characteristics for houses (including exact location) in a large number of US counties. Dataquick does not have housing data for all counties; therefore, only a subset of the properties that are tied to cleanup grants are included. Our final sample of houses near brownfield sites covers 176 counties across 38 states. The data consist of information on the sale of newly constructed houses, re-sales, refinance or equity dealings, timeshare sales, and subdivision sales. The data save transaction-related information such as price, date, and associated loans, as well as structural characteristics recorded from the most recent tax assessment. The attribute fields used as controls for differences in structural characteristics include age, number of bathrooms, bedrooms, square footage, and indicator variables for selling in the year built, being categorized as a condo,

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.

18. It is possible for speculatory behavior to begin even before this interim period, especially since cleanup grant applications are usually due in the fall before the award announcement in the following spring, after which the cleanup performance period begins. Our robustness checks in section 4 suggest that most anticipatory behavior occurs during the interim period.

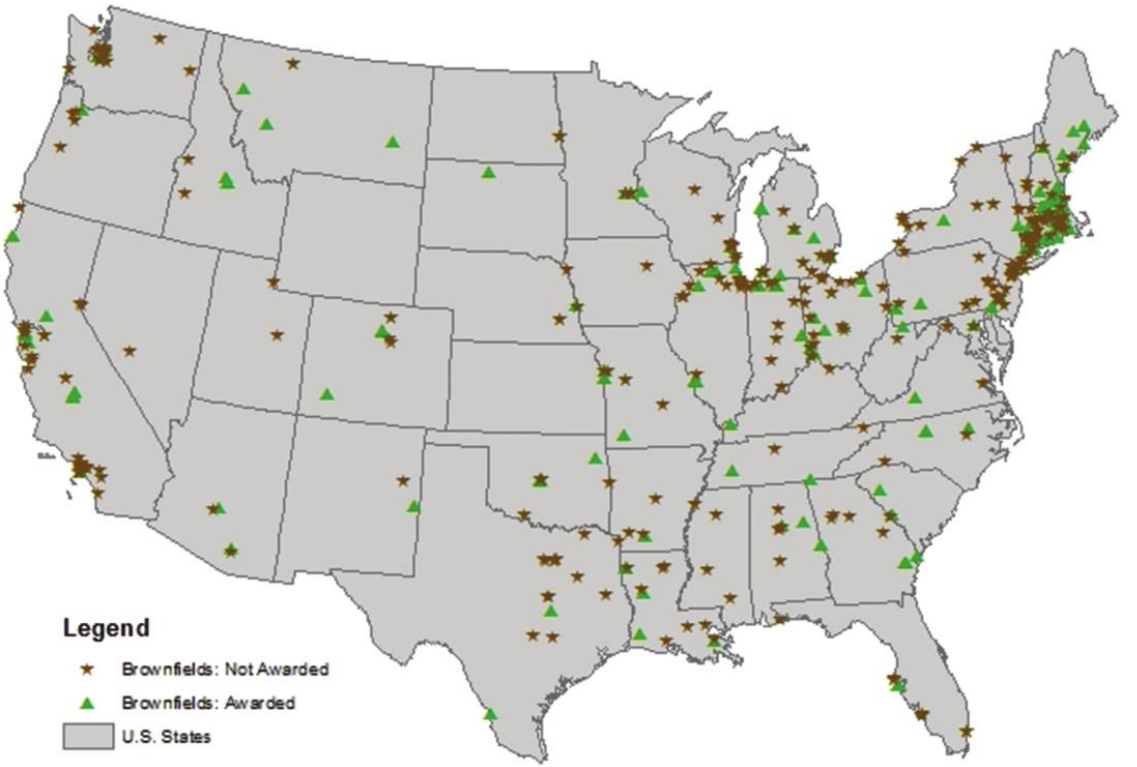

Figure 1. Brownfield sites in the continental United States from the set of cleanup applicants between 2003 and 2008

multifamily, mobile, or an unknown/miscellaneous house type, and year of build for the years 1800s, 1900–1950, 1951–2000, 2001–2005, 2006–2010, and after 2010.

In addition to house-level attributes, we control for county-level effective real estate tax (RET) rates (Siniavskaia 2011), as defined by the percentage of the property value that is paid in taxes every year. The county-level RET rates are calculated using homeowner-reported home values and annual real estate taxes from the Census Bureau's 2005–9 American Community Survey.

Our analysis limits transactions to house sales of owner-occupied properties since the transaction price corresponds to the owner's willingness to pay (WTP), which may be significantly different from that of a renter occupying the house. Houses with missing prices, bathrooms, bedrooms, or square footage are dropped. Furthermore, since only housing characteristics from the most recent tax assessment are recorded, any house indicated to have undergone major improvements is dropped, as its attributes may be incorrect for previous transactions. This is done in order to reduce measurement error and to avoid an upward bias in our welfare estimates if unobserved upgrades are correlated with cleanup. Any WTP for renovations that are bundled with cleanup should be separated out from the WTP to clean a brownfield site. 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 11-year window of house sales.[19] Prices are normalized to December 2000 dollars using the monthly, regional All Urban Consumers Housing CPI taken from the Bureau of Labor Statistics. The analysis excludes the 1st and 99th percentile of the observed price distribution. The window of observations used for housing transactions starts in 1998 (4 years before the start of the Brownfields Program)[20] and ends in 2012, which is the last available year for housing sales.

Neighborhood characteristics used for matching are available through SimplyMap, a web-based data application accessible with a license agreement with Duke University. SimplyMap provides yearly data on tract-level attributes that are only available starting in the year 2008. The specific variables used include percentages for the following: race (white, black, Asian, Hispanic), blue collar workers, educational attainment (high school and college), families with children under 18 years of age, native language (English, Spanish, Asian/Pacific Island, and Indo-European), and US citizens. We also use annual, tract-level median income. We associate these neighborhood characteristics

19. 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 five transactions in 11 years) signals that the house may be used as an investment property by a house "flipper" (Bayer, Geissler, and Roberts 2011).

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

to individual housing transactions based on the year of sale and tract in which the house is located.

Knowing the exact locations of all properties allows us to calculate the distance between each house and the nearest brownfield that applied for cleanup. This is our measure of brownfield "exposure." Using Graphical Information Systems (GIS), each property is first matched to the nearest brownfield within a 5-kilometer radius. The distances to those brownfields are then recovered. Houses not within 5 kilometers of a brownfield that applied for cleanup are dropped. Houses located near multiple brownfields, in which case the effect of cleanup may be hard to measure, are dropped. Out of a total of 797 unique brownfields from the EPA data with geocoordinates, 595 have available Dataquick data within 5 kilometers, 430 of which remain after removing houses that are within 5 kilometers of multiple sites. After cleaning the housing transactions data (e.g., removing missing attributes), this further limits the number of sites to 327, 197 of which are awarded with cleanup and 130 of which are not.[21] While restrictive, this ensures a clean exposure definition given that sites have different cleanup time lines. The treatment and control groups are then defined using houses within this 5-kilometer radius to minimize the threat of any location-specific unobservable differences that may affect price dynamics.

Some comments are in order regarding the sample as well as the additional sample cuts. The main types of selection into the sample are selection into Brownfields Program application, selection based on Dataquick availability, and selection for the sake of a "cleaner" exposure definition that removes houses near clustered brownfield sites. We take these different types of selection separately.

We begin with a sample of brownfields that are associated with EPA applications for cleanup and that also have housing data available for the county in which the site is located. The former precludes analysis of brownfields that did not apply for funding. There are two concerns with this. First, 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 estimate cleanup impacts for these sites, as long as the status of these brownfields does not change over the course of our analysis, they are 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, under certain identifying assumptions, both the DID and DD-NNM estimators are set up to control for this. The DID estimator does so to the extent that unobserved changes in other nearby brownfields that did not apply for cleanup equally affect treatment and control groups over time. For example, suppose cleanup of a particular brownfield site in our sample spurs cleanup at a nearby site that did not apply for brownfield funding. As long as the impact of this other site applies to both the treat-

21. Dataquick has missing house attributes for certain counties altogether.

ment (near) and control (far) houses, the unobserved, time-varying impact of this omitted brownfield will be differenced out. The DD-NNM estimator, in limiting to cross-sectional variation only, deals with this as it will be absent of any unobserved, time-varying impacts.[22]

The second concern with omitting brownfields that did not apply is related to the external validity of our estimates. Selection into application is very difficult to assess since we do not observe the sites that do not apply, preventing us from comparing our sample of brownfields with the universe of brownfield sites in the nation. Previous papers that study brownfield remediation, while often using a broader sample of brownfields taken from state registries, also contain selected samples of sites that voluntarily participate in state remediation or certification programs. A rough comparison of mean house attributes in our sample to that available from three recent papers (Mihaescu and vom Hofe 2012; Linn 2013; Savchenko and Braden 2014) does reveal some differences in house attributes but does not seem show a clear pattern of selection.[23]

That housing data are not available for all locations containing brownfield sites will also affect the type of brownfields for which we are recovering cleanup impacts. In particular, Dataquick does not have data for many rural communities. According to a survey of cities and towns conducted by the US Conference of Mayors (USCM)[24] beginning in 1993, 36% of surveyed respondents in 1998 were communities with populations of less than 50,000.[25] Thus it is important to keep in mind that our results are appli-

22. Still, this leaves a possibility for differential impacts between treatment and control if, for example, the unobserved brownfield site is across the street from the one that we do observe. This would bias our results. If cleaning a site attracts outside investment, then our estimates would be biased upward. On the other hand, site cleanup may crowd out investment in nearby areas and thereby reduce the estimated impact from cleanup.

23. The table of comparisons is included in the appendix. To make the comparison of samples across studies as accurate as possible, we apply appropriate time period and geographic limits to our data. We recover mean attributes of housing samples from three recent papers using data from Cook County, IL, New York County, NY, and Hamilton County, OH. It should be noted that our sample explicitly excludes any site that belongs on the National Priorities List (NPL), as NPL sites are not considered by the Brownfields Program, whereas the sample of sites from these other papers may include those that belong to state Superfund programs.

24. The US Conference of Mayors surveyed cities regarding the status of brownfield sites in the United States (US Conference of Mayors 2010). The first survey was conducted in 1993, and the most recent was conducted in 2010. For the 2010 report, the cities that responded (approximately 99) accounted for about 29,624 sites across the nation, spanning approximately 45,437 acres of land.

25. The US Census Bureau defines "rural" as all areas not included within an urban area, which include urban areas of 50,000 people or more and urban clusters of between 2,500 and 50,000 people.

cable to urban brownfields, which accounted for more than two-thirds of the USCM survey of brownfields around the nation. We assess the systematic differences in brownfield site attributes by housing data availability in the summary statistics.

With respect to the data cuts given the available sample of brownfields, removing houses over 5 kilometers away ensures that site impact, a localized disamenity, is picked up in housing prices. Next, removing houses near multiple brownfields within 5 kilometers, while restrictive, ensures a clean exposure definition given that sites have different cleanup time lines. Our summary statistics (to be described in sec. 4.3) reveal that houses near multiple brownfields are less desirable than those near at most one site within 5 kilometers. This suggests that we are potentially removing less desirable neighborhoods from our sample. If one abstracts from issues of cleanup timing (i.e., supposing all sites in a neighborhood are cleaned at once), then conditional on proximate neighborhoods being a valid control group for the treated areas, the impacts from our quasi-experimental setup should still control for time-varying neighborhood differences. The potential bias, however, in omitting these clustered sites could stem from additional feedback effects from cleaning all sites at once, or, alternatively, stigma from cleaning some but not all sites. The direction of the bias would thus additionally require consideration of these feedback and stigma effects. While important for informing policy, we should note that the estimates recovered from our model can only describe the impacts from cleanup that preclude any spatial feedback effects between sites.

Finally, an additional concern is that sites that are denied funding can still be cleaned and developed. Our results using DID will not be affected since we only use the sample of brownfields that were awarded with cleanup under the Brownfields Program. For DD-NNM, which includes property sales around nonawarded sites, this would cause us to incorrectly assign some of the treatment observations as controls and would thus tend to attenuate our estimated impacts.

4.2. Defining the Treatment Buffer

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.[26] That is, houses located within a certain distance of a brownfield are considered to be in the treatment group, while houses located outside of that distance (where the site has no effect regardless of cleanup) are designated as controls. As the effects of hazardous waste sites such as those on the National Priorities List decrease very quickly with distance from the site (Adler et al. 1984; Kohlhase 1991; Kiel 1995), the price shocks that would affect the trend of the treated group would arguably affect that of the control group as well. Ultimately,

26. Linden and Rockoff (2008) estimate the impact of sex offender arrival in Mecklenberg County, NC.

this common trend assumption is untestable. However, this paper provides graphical evidence next and specification tests in the results section that allow us to better assess its validity.

We begin by estimating a pair of price functions over distance from the nearest brownfield site—one for pre-cleanup transactions and one for post-cleanup transactions. The distance at which the pre-cleanup and post-cleanup price functions converge is where brownfield cleanup no longer affects 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 functions, we use a local linear polynomial estimator (Fan and Gijbels 1996), which is described in detail in the appendix.[27] 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 2 describes the average square footage of houses sold at each distance from a brownfield site before and after cleanup. It is clear from this figure that houses sold before cleanup of brownfield sites within approximately 2 kilometers tend to be larger than those sold in that same buffer after cleanup. We therefore control parametrically[28] for house attributes before recovering the nonparametric relationship between house prices and distance in figure 3. Figure 3 also controls parametrically for year effects to allow for general inflationary trends and differences in brownfield characteristics, including the proposal scores, proposal type, and the number of times the sites are assessed.[29]

Figure 3 provides evidence in support of the assumption that houses that are far enough from brownfields represent a valid control group. While we find that houses at all distances have higher prices on average after cleanup, we find that this difference is no longer statistically significant outside of 2,070 meters. Taking the treatment group to be defined by a 2,070-meter buffer, the DID estimator will compare the average change in prices before and after cleanup inside the buffer with the similarly defined change outside the buffer.

Given the definition of the treatment and control groups, a natural way to check whether the common trend assumption is reasonable is to compare the price trends

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

28. In practice, we recovered residuals from a linear regression of housing price on house attributes, brownfield attributes, and year fixed effects, and then used the residuals in a nonparametric regression on distance.

29. Under certain circumstances, additional testing may be advised by a licensed site professional, and a supplemental site assessment is conducted in addition to the required assessment. Recognizing those sites that demand additional testing may control for differences in the severity of contamination at sites.

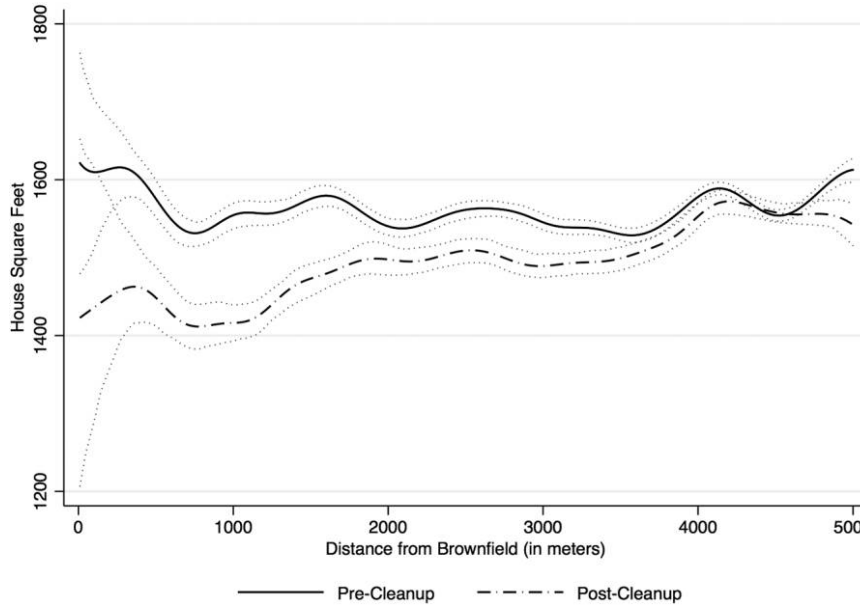

Figure 2. Average square footage of houses transacted by distance from brownfield before versus after remediation with 99% confidence intervals.

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

Figure 4 plots the prices of treatment (i.e., inside 2,070 meters) and control houses against time relative to the cleanup date.[30] The trends pre- and post-treatment are sim-

30. As was the case when generating figure 3, we parametrically control for housing attributes, year effects, and brownfield characteristics before nonparametrically estimating price as a function of time relative to the cleanup period.

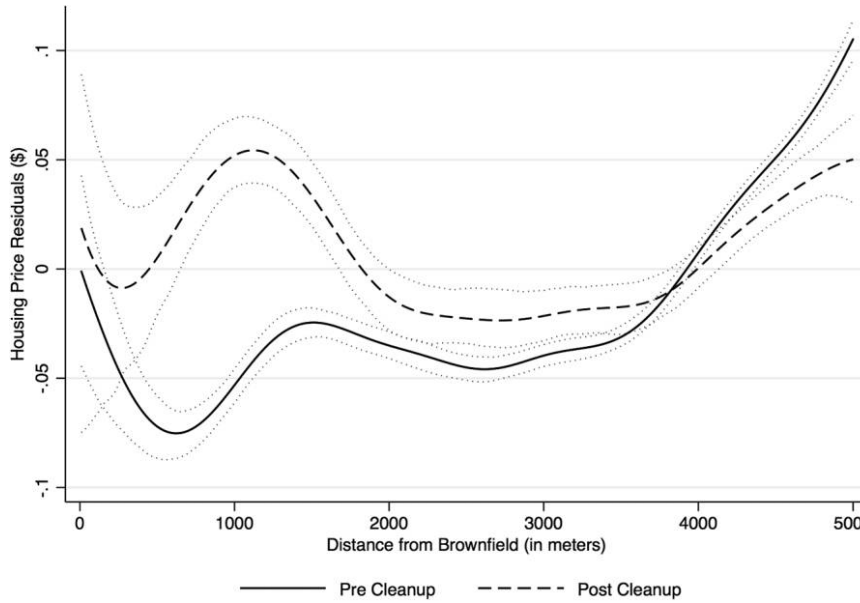

Figure 3. Nonparametric price function estimates before and after remediation with 99% confidence intervals.

ilar for the two groups. While both groups exhibit a jump at the point of treatment, suggesting that some of the treatment may spill out into the control group, the discontinuity for the control group going from pre- to post-cleanup (−0.58%) is smaller than that in the treatment group (6.67%). The differences-in-differences approach measures the jump in the treatment group relative to that in the control group.

4.3. Summary Statistics

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

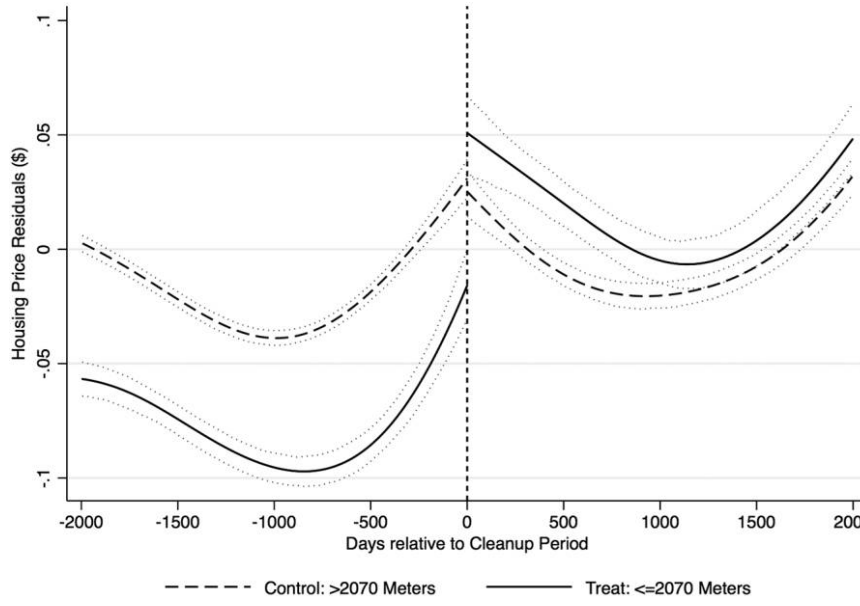

Figure 4. Nonparametric price function estimates relative to the cleanup period for treatment and control houses with 95% confidence intervals.

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 2 also compares attributes of houses near multiple versus one site only. It finds that houses near multiple brownfields are (mean difference in parentheses) less expensive ($3,838), older (12 years), smaller (23 square feet), and are more likely to be condominiums (4%) and multifamily homes (6%) than those near at most one site within 5 kilometers. While these differences are statistically significant, we make this trade-off for the sake of cleanly identifying site exposure.

Table 3 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 2,070 meters of a brownfield) and the control group (between 2,070 meters and 5 kilometers of a brownfield). Columns 5 and 6 test for equality of group means. Although we reject the equality of means for many attributes, we do take comfort in the fact that the differences are far smaller than comparisons between houses within 5 kilometers to houses in the rest of the county (table A3 in the appendix). We take table 3 as evidence that there are important differences between treatment and control groups that should be accounted for parametrically in the DID specification.

Table 1. Brownfield Attributes by Availability of Housing Data

	With Dataquick			Without Dataquick				
Variable	Mean (1)	SD (2)	Obs. (3)	Mean (4)	SD (5)	Obs. (6)	*t*-Statistic (7)	Reject? (8)
Funded and unfunded:								
Petroleum	.27	.45	401	.17	.38	777	–4.02	Y
Hazardous substances	.75	.43	401	.85	.35	777	4.26	Y
Proposal score (std.)	78.38	11.88	401	76.03	11.38	512	–3.04	Y
Funded only:								
Petroleum	.30	.46	239	.18	.39	201	–2.76	Y
Hazardous substances	.73	.45	239	.85	.36	201	3.00	Y
Proposal score (std.)	84.54	4.71	239	84.85	4.68	201	.69	N
Property size (acres)	10.83	31.29	239	12.91	44.53	197	.57	N
Ready for reuse	.43	.50	239	.44	.50	201	.25	N

Note. The table uses the sample of brownfield sites that applied to the Brownfields Program and compares average brownfield attributes by availability of housing data for both funded and nonfunded sites, as well as for funded sites only. "Petroleum" and "Hazardous substances" are dummy variables that refer to the cleanup application type (sites may apply for both types of grants). std. = standardized.

Table 4 provides a yearly breakdown of cleanup starts and completions for the brownfields that were awarded cleanup grant funding.[31] Since the Brownfields Law was only recently enacted in 2002, many cleanup completions occur toward the end of the window of observations, which limits the number of post-cleanup transactions we have to work with. The average cleanup duration for all brownfields for which we can calculate durations is approximately 15 months (or 444 days) with a standard deviation of 451 days.[32] 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.

31. There are two sites that began cleanup before 2002—one for areas with Dataquick coverage and one for areas without. These are likely from pilot programs that received funding before the formal program began.

32. Table A4 reports cleanup duration by toxin found and media of contamination.

Table 2. Housing Attributes by Proximity to Number of Brownfields

	Near One Site		Near Multiple Sites			
Attributes	Mean (1)	SD (2)	Mean (3)	SD (4)	*t*-Statistic (5)	Reject? (6)
Price	218,479	153,032	214,641	150,706	−17.97	Y
Real estate rate (county)	10.18	4.64	10.20	4.31	2.48	Y
Age	46.78	31.49	58.85	34.22	259.07	Y
Square footage	1,565.92	684.63	1,542.84	757.77	−22.53	Y
Bathrooms	1.95	.84	1.83	.93	−92.87	Y
Bedrooms	3.04	1.05	3.02	1.34	−14.00	Y
Sold in year built	.04	.19	.03	.16	−43.59	Y
Condominium	.15	.35	.19	.39	83.12	Y
Multifamily	.05	.21	.10	.31	153.83	Y
Single family	.80	.40	.70	.46	−167.85	Y
Mobile	.00	.05	.00	.03	−27.75	Y
Miscellaneous	.00	.05	.00	.06	24.55	Y
Observations	876,693		1,186,858			

Note. The table compares mean attributes of houses located within 5 kilometers of at most one brownfield site to those located near multiple brownfield sites within 5 kilometers.

Table 3. Housing Attributes by Treatment Status

	Treat (≤ 2,070 m)		Control (>2,070 m)			
Attributes	Mean (1)	SD (2)	Mean (3)	SD (4)	*t*-Statistic (5)	Reject? (6)
Price	192,306	130,325	198,788	142,330	8.97	Y
Real estate rate (county)	9.96	3.62	10.64	4.09	32.89	Y
Age	52.94	36.06	45.01	31.51	−47.50	Y
Square footage	1,559.49	689.47	1,560.25	690.73	.21	N
Bathrooms	1.83	.78	1.88	.81	12.05	Y
Bedrooms	3.09	1.19	3.00	1.01	−18.23	Y
Sold in year built	.04	.19	.05	.21	8.41	Y
Condominium	.13	.34	.15	.36	9.59	Y
Multifamily	.09	.29	.04	.20	−44.27	Y
Single family	.77	.42	.81	.40	15.31	Y
Mobile	.00	.03	.00	.04	3.47	Y
Miscellaneous	.00	.04	.00	.03	−1.72	N
Observations	46,090		204,719			

Note. The table compares mean attributes of houses by location inside versus outside of the treatment buffer (2,070 meters). The sample includes all houses within 5 kilometers of an awarded brownfield. Attributes are taken from houses selling before cleanup.

Table 4. Time Line of Brownfield Start and Completion Frequencies

	With Dataquick		Without Dataquick	
	Starts (1)	Completions (2)	Starts (3)	Completions (4)
2000	1			
2001			1	
2002	1		1	
2003	4		2	
2004	23	6	17	5
2005	37	12	23	13
2006	35	35	36	18
2007	23	24	28	26
2008	30	17	34	27
2009	30	23	17	33
2010	8	22	8	14
2011	1		1	2

Note. The table gives the cleanup start and completion frequencies by year for brownfields with and without housing data.

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. Not surprisingly, cleanups begun later in the sample are less likely to be completed. There is, however, a significant fraction of cleanups with petroleum contamination begun early in the sample that have not been completed by 2012.[33]

4.4. Neighborhood Turnover

As suggestive evidence of neighborhood turnover that could alter the equilibrium hedonic price function over time, we use a subsample of the data from Massachusetts for which we can obtain the income and race of the primary home buyer to examine changes in race and income distributions before and after cleanup.[34] Tabulations of the data (table 5) show that there are more minorities and low income owners in areas close to brownfields (i.e., within the treatment buffer), an outcome that is consistent

33. Table A5 of the appendix describes the fraction of cleanups initiated in each year that were not completed by 2012. There were no cleanups initiated in 2012 from the pool of awarded sites between 2002 and 2008.

34. Housing transactions in Dataquick are merged to data from the Home Mortgage Disclosure Act (HMDA) based on the Census tract, loan amount, and lender name of each property transaction. For details, see Bayer et al. (2016).

Table 5. Spatial Distribution of Income before Cleanup

	Inside Buffer (1)	Outside Buffer (2)	Diff (In) – Diff (Out) (3)
A. Race:			
White	73.1	87.6	−14.5
Asian/Pacific Island	3.38	4.46	−1.08
Black	14.95	3.53	11.42
Hispanic	7.59	3.58	4.01
Other	.99	.84	.15
B. Income quintiles:			
1	22.97	19.26	3.71
2	21.08	19.73	1.35
3	20.37	19.91	.46
4	19.58	20.11	−.53
5	16	21	−5

Note. The table reports the average differences in race and income (%) before cleanup for different race categories and segments of the income distribution. Income quintiles are based on the distribution of household incomes of all houses sold before cleanup.

with many studies in the environmental justice literature.[35] We find that neighborhoods that are cleaned experience a relative increase of 16.1% in the share of white homeowners compared to nearby neighborhoods unaffected by the presence of brownfield sites (table 6, panel A).[36] With regard to income, table 6, panel B, shows increases in the middle and lower class (first, second, and third quintiles) following cleanup.

We can assess whether cross-sectional comparisons are better able to control for resorting in populations in response to cleanup by comparing the pre- and post-cleanup race distributions from table 6, panel A, to that by award status and year. In table 7, we present relative changes in race distributions by award status (panel A) and by award status and year (panel B). Demographic "changes" using cross-sectional variation in exposure to a cleaned site are much smaller than changes using temporal variation, suggesting that the cross-sectional comparisons are less likely to confound changes in preferences from sorting.

Although we cannot explain these shifts in race and income,[37] we can be more certain that dynamic forces that alter the socioeconomic makeup of communities are at

35. See Been and Gupta (1997), Pastor, Sadd, and Hipp (2001), Baden and Coursey (2002), and Wolverton (2009).

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

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

Table 6. Change in Race and Income Distribution after Cleanup

	Inside Buffer			Outside Buffer			
	Post = 0 (1)	Post = 1 (2)	Diff (In) (3)	Post = 0 (4)	Post = 1 (5)	Diff (Out) (6)	Diff (In) – Diff (Out) (7)
A. Race:							
White	73.1	89.03	15.93	87.6	87.43	–.17	16.1
Asian/ Pacific Island	3.38	1.66	–1.72	4.46	2.88	–1.58	–.14
Black	14.95	5.23	–9.72	3.53	4.66	1.13	–10.85
Hispanic	7.59	3.95	–3.64	3.58	4.86	1.28	–4.92
Other	.99	.13	–.86	.84	.17	–.67	–.19
B. Income:							
1	20	32.65	12.65	20	32.41	12.41	.24
2	20	23.98	3.98	20	23.16	3.16	.82
3	20	20.92	.92	20	18.5	–1.5	2.42
4	20.03	12.37	–7.66	20	15.35	–4.65	–3.01
5	19.97	10.08	–9.89	20	10.59	–9.41	–.48

Note. The table gives average changes in race and income (%) after cleanup in neighborhoods near brownfields relative to neighborhoods far from brownfields over the same period. Relative differences are reported for different race categories and segments of the income distribution.

play, which can signify preference shifts at the communities of interest, further motivating the use of our DD-NNM estimator.

5. EMPIRICAL RESULTS

5.1. Cross-Sectional Estimates

Table 8 reports the results of a cross-sectional specification, where exposure is based on award status. We restrict the comparison to be between houses that are in the vicinity of brownfields—some of which have been cleaned, others of which have not. We find that the value of cleanup is negative at –11%. The counterintuitive sign of this effect may be a result of omitted variables bias if cleanup grants are targeted toward struggling neighborhoods. Table 8 suggests that unobservable neighborhood attributes may be correlated with their cleanup status, necessitating a different empirical approach.

5.2. Fixed Effects Estimates

Next, we use fixed effects at either the brownfield site or house level to control for time-invariant unobservables, which may be the source of bias that leads to the counterintuitive results found in the cross-sectional specifications. The fixed effects specification

Table 7. Change in Race Distribution by Award Status

	A. Overall						
	Inside Buffer			Outside Buffer			
	Award = 0 (1)	Award = 1 (2)	Diff (In) (3)	Award = 0 (4)	Award = 1 (5)	Diff (Out) (6)	Diff (In) – Diff (Out) (7)
White	93.44	89.03	–4.41	90.17	87.43	–2.74	–1.67
Asian/Pacific Island	2.05	1.66	–.39	3.72	2.88	–.84	.45
Black	1.9	5.23	3.33	2.51	4.66	2.15	1.18
Hispanic	1.87	3.95	2.08	2.89	4.86	1.97	.11
Other	.74	.13	–.61	.71	.17	–.54	–.07

	B. By Year			
Diff(in) – Diff(out)	2008	2009	2011	2012
White	1.8	1.74	.59	.66
Asian/Pacific Island	.88	.21	.02	2.54
Black	–3.8	2.09	–2.7	1.11
Hispanic	1.3	–3.79	2.11	–4.3
Other	–.19	–.22	–.02	0

Note. The table gives average changes in race (%) by award status for neighborhoods near brownfields relative to neighborhoods far from brownfields. Relative differences are reported for different race categories. Panel A gives the overall relative changes and Panel B gives the changes for each year.

Table 8. Cross-Sectional Specification

Variables	(1) All within 5 km
CLEANUP	−.117***
	(.00274)
Constant	11.50***
	(.0103)
Observations	469,928
R^2	.488
Controls:	
Year fixed effects	Yes
Brownfield characteristics	Yes
House controls	Yes

Note. The cross-sectional specification compares houses near cleaned versus uncleaned brownfields (within 5 kilometers). Sample used includes only houses inside a 5-kilometer buffer around a funded brownfield that has been cleaned or an unfunded brownfield. *CLEANUP* = 1 if the house is near a funded site that has been cleaned. House attributes include county tax rate, age, number of bathrooms, bedrooms, square footage, and indicator variables for selling in the year built, condo, multifamily, mobile unknown/miscellaneous house type, and year of build for years 1800s, 1900–1950, 1950–2000, 2000–2005, 2005–10, and after 2010. Brownfield attributes include whether a site is assessed twice, a petroleum site, a hazardous substances site, and the standardized proposal score. Standard errors in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

uses all houses in a buffer; we consider buffers of 1,000, 2,000, 3,000, and 5,000 meters to demonstrate robustness. We also include controls for year fixed effects, house attributes, and the real estate tax rate. The results of the fixed effects specification, described in table 9, differ strikingly from the cross-sectional results, with statistically significant increases in house prices from cleanup that range between 8% and 11%, depending on the size of the buffer. Standard errors are clustered at the brownfield level. Limiting the comparison to within-house differences with house fixed effects, we find that cleanup yields a statistically significant increase in house values of 9% at the 5% level using standard errors clustered at the brownfield site.

5.3. Difference-in-Differences Estimates (DID)

While they are able to deal with time-invariant unobservable neighborhood attributes, fixed effects do nothing to control for time-varying unobservables that may be correlated with brownfield cleanup. Estimates would still be biased if, for example, cleanup

Table 9. Fixed Effects

Variables	$b = 1{,}000$ (1)	$b = 2{,}000$ (2)	$b = 3{,}000$ (3)	$b = 5{,}000$ (4)	$b = 5{,}000$ (5)
INTERIM × *TREAT*	.0718*	.101**	.0939**	.101***	.113**
	(.0401)	(.0404)	(.0432)	(.0381)	(.045)
POST × *TREAT*	.0984*	.115**	.0968**	.0842**	.0926**
	(.0551)	(.0451)	(.0418)	(.0371)	(.042)
Observations	18,686	64,652	136,480	370,910	193,421
R^2	.674	.671	.686	.687	.254
Controls:					
Year fixed effects	Yes	Yes	Yes	Yes	Yes
House controls	Yes	Yes	Yes	Yes	
Fixed effects	Brownfield	Brownfield	Brownfield	Brownfield	House

Note. The fixed effects specification compares houses sold before and after cleanup. Sample used includes only houses (i) around awarded brownfields and (ii) inside buffer "b" meters. *POST* = 1 if transaction occurs after nearby brownfield is cleaned. *INTERIM* = 1 if transaction occurs during cleanup. House characteristics (for specifications with brownfield fixed effects) are the same as those used in table 8. Standard errors clustered by brownfield in parentheses. *TREAT* = 1 if house is located within *b* meters of a site.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

were systematically directed toward locations that were considered disadvantaged but were improving in unobservable ways. The DID approach overcomes this problem with the "common trends" assumption—namely, that the change over time in unobservables in the control group is the same as it would have been in the treatment group in the absence of treatment. By assigning the control group to be houses in the same neighborhood as those in the treatment group, but far enough away from the site to not be affected by cleanup, we address this assumption and obtain estimates that account for any time-varying unobservables that are common to both the treatment and control groups. Moreover, by differencing over time, the DID approach also controls for time-invariant unobservables, just as the fixed effects specification did.

As described in section 3, the average treatment effect on the treated is measured by the coefficient on the interaction of the indicators for a house being in the treatment group (*TREAT*) and its transaction occurring after the cleanup has been completed (*POST*). These estimates can be found in the fifth row of table 10. It is important to note that our estimate is an average effect across all houses at various distances within the exposure buffer, as previous literature has shown that the price-distance gradient can vary depending on the nature of the undesirable land use (Braden, Feng, and Won 2011) or the type of neighborhood (Kiel and Williams 2007).

With only year fixed effects and brownfield-level controls, we find a treatment effect of 5.7% using the preferred buffer size of 2,070 meters. In a specification that includes

Table 10. Differences in Differences

	b = 2,070 Meters		
Variables	(1)	(2)	(3)
TREAT	–.0496***	–.0475***	–.0456**
	(.004)	(.003)	(.018)
INTERIM	.237***	.112***	.104***
	(.004)	(.003)	(.039)
POST	.149***	–.000276	.0747**
	(.004)	(.003)	(.037)
INTERIM × *TREAT*	–.0747***	–.0365***	–.0161
	(.009)	(.007)	(.023)
POST × *TREAT*	.0567***	.0701***	.0500*
	(.007)	(.006)	(.029)
Constant	13.97***	13.21***	11.08***
	(.022)	(.018)	(.166)
Observations	370,910	370,910	370,910
R^2	.087	.471	.380
Controls:			
Year fixed effects	Yes	Yes	Yes
Brownfield characteristics	Yes	Yes	
House controls		Yes	Yes
Brownfield fixed effects			Yes

Note. The DID specification compares price differences before and after site cleanup for houses located within a treatment buffer of 2,070 meters to differences in prices for those houses located outside of the treatment buffer. Sample used includes only houses (i) around awarded brownfields and (ii) within 5 kilometers of a site. *TREAT* = 1 if house is located within 2,070 meters of a site. *POST* = 1 if transaction occurs after nearby brownfield is cleaned. *INTERIM* = 1 if transaction occurs during cleanup. Brownfield and house controls (when included) are the same as those in table 8. There are a total of 197 sites in the brownfield fixed effects specification. Standard errors, clustered by brownfield, in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

year fixed effects, house-level and brownfield-level controls, and controls for the real estate tax, this effect increases to 7.0%. Further introducing brownfield fixed effects decreases this effect to 5.0%, which is significant at 10% after clustering standard errors at the brownfield level.[38] The coefficients on the time dummy variables, *INTERIM* and *POST* are both positive and significant in the specification with brownfield fixed effects,

38. See Bertrand, Duflo, and Mullainathan (2004) for the importance of clustering standard errors.

which suggest that at the community level, broad neighborhood improvements are simultaneously being made over time. This reinforces the importance of controlling for time-varying unobservables through the use of a control group in order to avoid overstating the impact of cleanup. The estimate for the cleanup interim interaction in our main specification, $INTERIM \times TREAT$ is now statistically indistinguishable from 0, compared to the –4.6% estimated price impact on treated houses before cleanup started ($TREAT$). We take this as evidence of expectations for neighborhood improvements to come before completion of site cleanup. Our DID estimates additionally provide some support for our interpretation of the treatment effect as the value of remediating a site back to being potentially productive (under the Brownfields Program). The average price increase for houses within the exposure buffer after cleanup (5.0%) fully recovers the depressed values of these houses compared to their outer-buffer counterparts before cleanup (–4.6%).

Before proceeding to our DD-NNM estimator, we include some robustness checks that assess the validity of our control group, the presence of housing supply impacts, and any evidence of reverse causality. We first reestimate the DID model by moving the cleanup completion date 180, 365, 720, and 1,095 days prior to the date cleanup activities actually began (table 11). If the control houses served as a valid comparison group for the treated houses, then one would not expect any impact based on these falsified cleanup dates. If there were, then it would suggest that the treatment and control groups had differential trends before cleanup commenced. Table 11 shows that the coefficients on the treatment effect ($POST \times TREAT$) are insignificant, providing some evidence that the current specification successfully controls for temporal and spatial confounders.

Another potential concern, even though time frames are relatively short, is that our estimated impacts incorporate changes in housing supply as a result of cleanup. One scenario is if housing supply increases as a result of cleanup. In this case, one should expect to see the estimated cleanup impact decrease as we allow for a longer post-treatment period. To test for this, we limited the post-cleanup transactions to being at most 1, 2, and 3 years after cleanup in the main DID specification. The estimated treatments from these regressions are presented in columns 2–3 of table 12. ATT estimates are not statistically different and stable across specifications. We admit that this test of supply is weak, as it does not exclude there being, for example, concurrent changes in demand and supply over time as a result of cleanup or any long-run impacts of cleanup.

Finally, if brownfield remediation was part of a larger effort to revitalize the specific area in which the site is located (i.e., the houses located within the buffer), then we should see anticipatory effects before the site is cleaned. To test for this, we created additional time dummy variables that correspond to 1, 2, and 3 years before cleanup starts. If there is reverse causality, then the coefficients on leads of "treatment" impacts should be positive and significant. The regression in column 5 of table 12 shows that none of them are, lending support that it is brownfield cleanup that yields house price

Table 11. Falsification Test—*D* Days before Actual Cleanup Date

Variables	$D = 180$ (1)	$D = 365$ (2)	$D = 730$ (3)	$D = 1{,}095$ (4)
TREAT	–.0469***	–.0462**	–.0448**	–.0468**
	(.018)	(.018)	(.018)	(.018)
POST	.0845***	.0732**	.0354	–.00747
	(.031)	(.031)	(.030)	(.029)
POST × *TREAT*	.0263	.0221	.0166	.0187
	(.022)	(.022)	(.022)	(.021)
Constant	11.08***	11.07***	11.07***	11.08***
	(.166)	(.166)	(.164)	(.162)
Observations	370,910	370,910	370,910	370,910
R^2	.253	.251	.248	.248
Number of sites	197	197	197	197

Note. The above falsification tests move the cleanup date *D* days before the actual date of cleanup, and reestimates the DID specification with brownfield fixed effects. As it is unclear what an "Interim" period should be with the falsified dates, we do not separately identify cleanup interim effects. Controls used for falsification tests are the same as those used in column 3 of table 10. Standard errors, clustered by brownfield, in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

increases and not house prices increases as a result of a general revitalization that led to brownfield cleanup.

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

Both the fixed effects and DID approaches rely on the strong assumption that the hedonic price function remains stable over time. If cleanup activities initiate neighborhood turnover, the identities of those living in close proximity to the site may change, and with them, marginal willingness to pay may change as well. In fact, Kuminoff and Pope (2014) demonstrate that estimates of the hedonic price function may provide no information about MWTP. As such, one needs a method that both controls for unobservables that may be correlated with cleanup activities while not relying on variation in cleanup status over time. The difference-in-differences nearest neighbor matching estimator described in section 3 is designed to do this by controlling for brownfield, house, and neighborhood characteristics.

Estimates of the average treatment effect on cleanup (π) are recovered in two stages. Stage 1 uses a bias-corrected nearest neighbor matching estimator to recover the effect of locating near and far from awarded sites using the set of comparable houses around nonawarded sites. In doing so for both houses inside and outside of the buffer, we can

Table 12. Robustness

		At Most Y Years Post Cleanup			
Variables	Main Estimate (1)	$Y = 1$ (2)	$Y = 2$ (3)	$Y = 3$ (4)	Anticipatory Effect (5)
Lead 3 years					.0171
					(.020)
Lead 2 years					–.0119
					(.021)
Lead 1 year					.00255
					(.024)
POST × TREAT	.0500*	.0484**	.0486*	.0459*	.0511*
	(.029)	(.024)	(.026)	(.027)	(.030)
Constant	11.08***	11.04***	11.04***	11.05***	11.08***
	(.166)	(.171)	(.170)	(.170)	(.164)
Observations	370,910	314,047	328,844	340,489	370,910
R^2	.380	.391	.390	.388	.381
Number of sites	197	193	194	195	197

Note. Column 1 gives the main estimate from the DID specification (from col. 3 of table 10). Columns 2–4, respectively, estimate the ATT by limiting the post-treatment period to at most 1, 2, and 3 years after cleanup. Column 5 includes leads in treatment impact of 1, 2, and 3 years to test for any anticipatory effects. Controls used are the same as those used in column 3 of table 10. Standard errors, clustered by brownfield, in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

correct for any unobserved differences between awarded and nonaward sites. Panel A of table 13 gives average estimates from stage 1 for our preferred buffer size of 2,070 meters using $J = 5$ matches. The estimates under Inside Treatment Buffer of panel A are derived by averaging over the treatment effect from comparing houses inside the treatment buffer of cleaned sites to houses inside the treatment buffers of uncleaned sites ($\beta_2 + \pi$). Estimates under Outside Treatment Buffer of panel A are the averaged treatment effects derived from comparing houses in the control groups of cleaned sites to houses in the control groups of uncleaned sites (β_2). We cannot consider results for the years 2004–7, since yearly data on tract-level attributes, which are used to control for neighborhood composition, are only available starting in the year 2008.

The matching estimates for houses outside the treatment buffer are negative and significant across all years, ranging from –8.8% to –21.9% for $J = 5$ matches, which is consistent with the idea that cleanup targets neighborhoods that are worse off. Several of the within-buffer estimates are insignificant, as the sample sizes have been reduced by a fair amount after limiting the data to post-cleanup sales by year. The per-

Table 13. Stage 1 Difference-in-Differences Nearest Neighbor Matching

	A. 5 Matches								
	Inside Treatment Buffer					Outside Treatment Buffer			
Year	Est. (1)	SE (2)	Obs. (3)	% Exact (4)	Year	Est. (5)	SE (6)	Obs. (7)	% Exact (8)
Y2008	.120***	(.0443)	4,404	60.92	Y2008	–.163***	(.0254)	17,737	63.45
Y2009	.150***	(.0396)	5,131	59.88	Y2009	–.0882***	(.0226)	20,363	68.39
Y2010	–.0883**	(.0395)	4,341	50.11	Y2010	–.212***	(.0240)	17,273	63.28
Y2011	–.0623	(.0409)	3,967	55.73	Y2011	–.219***	(.0283)	16,065	69.84
Y2012	–.0153	(.0454)	3,317	60.66	Y2012	–.163***	(.0274)	13,400	67.1
	B. 10 Matches								
Y2008	.170***	(.0391)	4,404	57.82	Y2008	–.131***	(.0197)	17,737	59.63
Y2009	.146***	(.0351)	5,131	58.81	Y2009	–.0988***	(.0178)	20,363	64.14
Y2010	–.0514	(.0345)	4,341	49.29	Y2010	–.210***	(.0195)	17,273	58.76
Y2011	–.0743**	(.0333)	3,967	51.53	Y2011	–.230***	(.0237)	16,065	65.63
Y2012	–.0516	(.0387)	3,317	56.81	Y2012	–.183***	(.0213)	13,400	61.29

Note. The table gives bias-adjusted nearest neighbor matching estimates using 5 and 10 nearest neighbors based on the Mahalanobis metric at a buffer of 2,070 meters. Matches are restricted to be in the same state. Column "%Exact" gives the percentage of the observations that match exactly on State. Match variables: House attributes include the county tax rate, age, number of bedrooms, bathrooms, square footage, and indicator variables for selling in the year built, condo, multifamily, mobile, unknown/miscellaneous house type, and year of build for years 1800s, 1900–1950, 1951–2000, 2001–2005, 2006–2010, and after 2010. Brownfield attributes include the standardized proposal score, site longitude and latitude, and indicator variables for assessed twice, petroleum site, and hazardous substances site. Neighborhood attributes at the tract level include median income, %white, %black, %Asian, %Hispanic, %blue collar, %high school graduate, %college graduate, %with child under 18, %speaks English, %speaks Spanish, %speaks Indo-European language, and %US citizen. Standard errors in parentheses and calculated according to Abadie and Imbens (2006).

* $p < .10$.

** $p < .05$.

*** $p < .01$.

centage with exact matches assesses the quality of the matches used for each year. These range from 50.1% to 60.9% for the within-buffer estimates, and 63.3% to 69.8% for the outside-buffer estimates. This highlights the importance of using a bias-corrected estimator and matching on measures of neighborhood quality. Panel B of table 13 gives average results from stage 1 using $J = 10$ matches and suggests that results are fairly robust to using matches of different sizes.

Stage 2 then recovers the effect of cleanup on the treated from a regression of individual treatment effects (stage 1) on a buffer indicator and controls. We include brownfield fixed effects to control for any unobserved factors that may be correlated

with cleanup assuming that these unobservables affect the entire neighborhood but do not do so differentially by buffer status. Stage 2 cleanup estimates find statistically significant impacts that range from 10.2% to 13.3% (see table 14).[39] Estimates with $J = 10$ matches are similar and range from 11.8% to 15.2%.[40]

These results suggest that we can indeed interpret our results as implying a positive and significant willingness to pay for brownfield remediation (i.e., a welfare interpretation). The largest of our fixed effects estimators, which estimates an 11.5% increase in housing values, is similar to the smallest of the statistically significant DD-NNM estimates of 10.2% in 2011 and 10.8% in 2008.[41] Compared to the estimates in other years, the larger DD-NNM estimates, which range from 13.2% to 15.2%, are between 16% and 32% higher than our largest fixed effects estimate.

In comparison 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. Although capitalization does not separate equilibrium changes to the hedonic price function from MWTP, there is a close relationship between the two if capitalization is recovered using a DID framework. In particular, Banzhaf (2015) shows the capitalization estimate from DID to be a lower bound on the willingness to accept (WTA) the absence of some policy of interest, which in this case refers to cleanup. Intuitively, he describes the DID estimator as identifying the direct effect of treating a particular house with an increase in some amenity g when the rest of the market has already been treated with the policy of interest. In other words, the DID estimator is the integral under the post-treatment hedonic price gradient for the policy change, represented by

$$\int_{g^0}^{g^1} \frac{dP^1}{dg}\, dg$$

in figure 5 taken from Banzhaf (2015). The DD-NNM estimator, in recovering the entire area under the post-treatment Hicksian demand curve, provides the Equivalent Surplus (ES) for a change in treatment status, or, from figure 5,

$$\int_{g^0}^{g^1} h\left(g, u^1\right)\, dg,$$

where u^1 refers to the post-treatment utility of some house treated with cleanup. The difference between our estimators (5% vs. 10%–15%) would thus be represented by the

39. The year 2010 estimate of 6.0% is statistically insignificant with standard errors clustered at the brownfield level.

40. The smallest estimates of 9.4% and 9.7%, respectively, in years 2010 and 2011 are statistically insignificant.

41. The estimate in year 2010 is lower, but comparable, at 6.3% is insignificant (SE 0.07).

Table 14. Stage 2 Difference-in-Differences Nearest Neighbor Matching

	A. 5 Matches				B. 10 Matches		
Year	Est. (1)	SE (2)	*N* (3)	Year	Est. (4)	SE (5)	*N* (6)
Y2008	.108**	(.0437)	6,672	Y2008	.152***	(.0464)	6,672
Y2009	.133*	(.0685)	7,751	Y2009	.144**	(.0699)	7,751
Y2010	.0633	(.0733)	7,925	Y2010	.0938	(.0670)	7,925
Y2011	.102*	(.0586)	8,033	Y2011	.0965	(.0581)	8,033
Y2012	.132**	(.0626)	6,662	Y2012	.118*	(.0654)	6,662

Note. Stage 2 of DD-NNM first recovers individual treatment effects on all houses located around awarded sites from stage 1, computed based on the nearest matches found from nonawarded sites. Treatment effects are then regressed on a dummy variable for locating inside the treatment buffer, house/neighborhood controls, and site-specific dummy variables. Standard errors, clustered by brownfield, in parentheses.

* $p < .10$.

** $p < .05$.

*** $p < .01$.

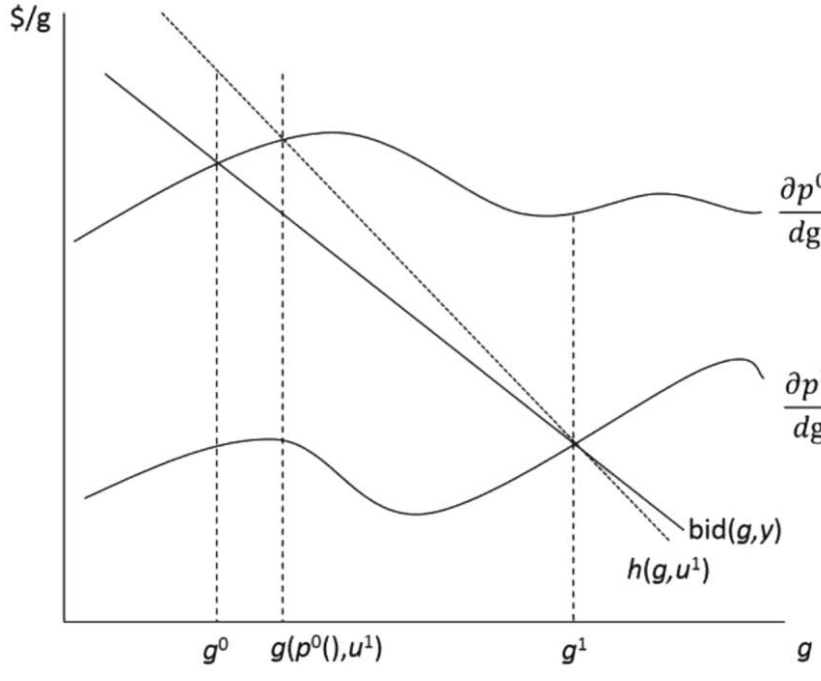

Figure 5. Lower bound for WTA changes in public goods (Banzhaf 2015)

area beneath the post-treatment Hicksian demand curve and above the post-treatment hedonic price gradient for the range of the treatment (g^0 to g^1). This difference, termed the "indirect effect" in Banzhaf (2015), can be ascribed to the adjustment in the hedonic price function as a result of the policy that holds the disamenity (a brownfield) at its original, unremediated state for some house. Viewed from the lens of the program evaluation literature, this represents spillovers from other houses under treatment even when that house has not been treated. While this direct effect is a lower bound on ES, the wedge between ES and this direct effect may still be large. Our various sample limitations aim to reduce the size of this wedge, including removing exposure to multiple sites and removing renovated houses that may mediate additional impacts on the equilibrium hedonic price function.

Capitalization of amenity changes into local housing prices is clearly of importance to city planners for the purposes of estimating expected tax revenue. That this measure using DID also provides a lower bound on ES implies that it may still be useful in welfare analysis. In particular, while studies in the past have emphasized estimating ex ante price functions to predict and bound (from above) the impacts from some amenity change, one can use the ex post hedonic price function and the DID estimate it yields to conduct a retroactive benefit-cost analysis that recovers a conservative estimate of benefits. Still, the lower bound from DID may underestimate benefits by a large margin, and thus caution must be exercised for assuming time-constant hedonic price functions in policy evaluation if the policy under consideration induces large enough changes such that the population considered before treatment is inherently different from the one after.

6. CONCLUSION

We find that property value increases accompanying cleanup averaging from 5.0% to 11.5%; for a welfare interpretation that does not rely on the intertemporal stability of the hedonic price function, a double-difference matching estimator finds even larger effects of up to 15.2%. In comparison to the literature, our results generally lie within the range of estimated impacts (3%–36%) from previous work that examines the effect of brownfield sites on nearby property values (Ihlanfeldt and Taylor 2004; Mihaescu and vom Hofe 2012; Linn 2013; Savchenko and Braden 2014).[42] Although all of these

42. Linn (2013) finds property price increases of 3% from listing and certification of brownfield sites in Cook County, IL. Mihaescu and vom Hofe (2012) find assessed value reductions of 20%–22% for houses within 1,000 feet of sites in Hamilton County, OH. A working paper by Savchenko and Braden (2014) finds that brownfield cleanups in New York, NY, increases housing prices by 8% or 1% depending on whether sustainable forms of redevelopment take place. While most previous studies focus on the impacts on residential properties, Ihlanfeldt and Taylor (2004) estimate the effect that brownfield listing has on the price-distance gradient for commercial and industrial properties in Fulton County, GA. They find the largest impacts

studies examine the impacts on prices of a nearby brownfield or a change in its status, there are important differences in the definition of exposure and the samples used for estimation that prevent direct comparison of the results. First, previous papers often defined exposure to be proximity to the nearest brownfield site, regardless of the number of brownfield sites nearby. Our analysis limits exposure to houses near at most one site within a 5-kilometer buffer for the sake of a cleaner exposure definition. Second, our sample includes a set of sites taken from multiple states, whereas prior studies focus on a handful of counties. Third, while our sample of sites has larger geographic coverage, brownfields in our sample are defined more narrowly, namely, they must have previously been assessed for contamination, are state owned (or partnered with a nonprofit organization), and cannot belong on the NPL or fall under the Toxic Substances and Control Act. The types of sites included from the previous papers can include those with much higher contamination risk; on the other hand, our discussions with program officers from various states also suggest that voluntary programs also contain sites for which contamination is only perceived. In this sense, compared to the sample of sites from voluntary state programs used in previous papers, our sample of sites can be considered a subset.

6.1. Cost-Benefit Analysis

Our results allow us to address the simple question, "Is brownfield remediation worth it?" In answering this question, we take a conservative approach. First, we take our most conservative estimate of the cleanup effect—the difference-in-differences estimate based on a 2,070 meter treatment buffer (5.0%), rather than the larger estimates generated by the fixed effects and DD-NNM specifications. Next, we take a conservative estimate of the value of housing that sold inside the treatment buffer prior to cleanup. Ideally, we would like to measure the total value of all housing units inside each buffer prior to the start of cleanup, but we do not observe every house sell during that pre-cleanup period. Rather than try to impute values for houses that we do not see transact during that period, we take the conservative approach of aggregating the value of only the houses that do sell in the 5 years prior to the start of cleanup inside the treatment buffer. We are able to construct this aggregate value for 51 of the brownfields—$4,158,640,384. Multiplying by a cleanup impact of 5.0% yields an estimate of the aggregate increase in housing value owing to cleanup of $207,932,016. This represents an average benefit value of $4,077,098 per site, with a median of $2,291,315. Figure 6 plots the distribution of benefits across sites. The Northeast Midwest Institute (NEMW) estimates an average cost of brownfield cleanup to be $602,000 based on

---

for office and apartment buildings, where they find price increases of, respectively, 36% and 26% by moving from 0.5 to 2 miles away from site, and the smallest impact on the price-distance gradient for industrial sites (3%).

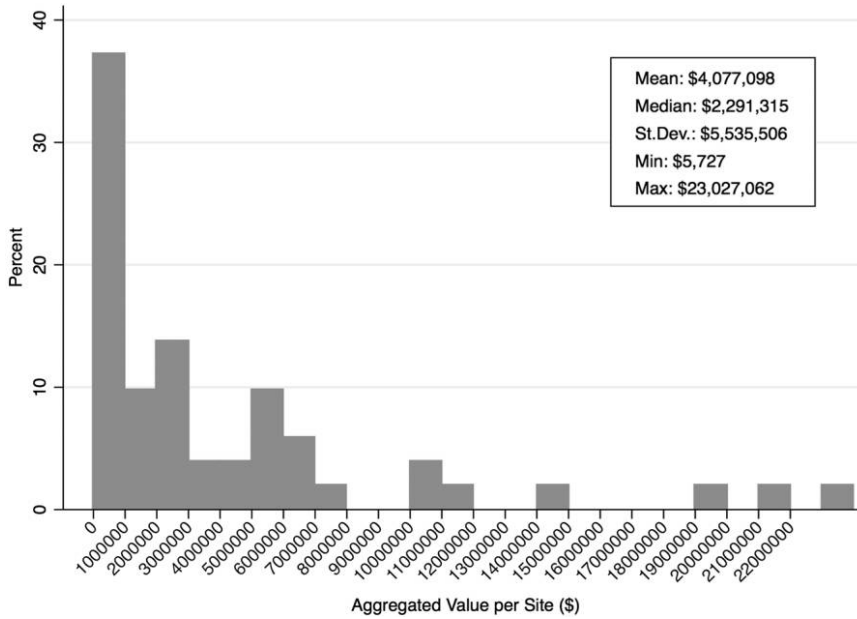

Figure 6. Distribution of benefits across brownfields: aggregate value of houses sold 5 years prior to cleanup period.

cleanup data provided by the EPA (Paull 2008). Although the smallest of our benefits estimate is below the estimated cleanup costs (17 of 51 brownfields have estimated benefits less than $600,000), the benefits for the majority of the cleaned sites still far exceed the cost.[43] Furthermore, brownfield remediation should easily pass a cost-benefit test if we considered all the properties located inside the treatment buffer, a larger treatment buffer, or one of our larger treatment effect estimates.

6.2. Limitations

The EPA Brownfields Program provides grants to assess and clean up properties the "expansion, redevelopment, or reuse of which may be complicated by the presence or potential presence of a hazardous substance, pollutant, or contaminant." In this paper, we quantify the benefits associated with these remediation activities using property value hedonic techniques. As is typically the case in property value hedonic applications, omitted neighborhood attrtibutes have the potential to bias these estimates. Indeed, our evidence suggests that neighborhoods that successfully clean brownfields

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

under the program may be worse in other unobserved dimensions. As such, we offer a slate of quasi-experimental approaches to overcome this problem, including simple neighborhood fixed effects, a difference-in-differences approach that relies on a treatment and control group defined by geographic proximity, and a difference-in-differences nearest neighbor matching estimator that exploits the advantages of our treatment and control group definitions while not requiring that the hedonic price function remain stable over time. Furthermore, our paper offers the added benefit of external validity given our unique data, as it is the first to use a nationally representative sample of brownfield sites considered by the EPA federal Brownfields Program.

Before concluding, we acknowledge a few limitations of our analysis. First, looking at the price of housing in close proximity to brownfield sites will not capture equilibrium effects that are realized elsewhere in the urban area—that is, cleanup of brownfields may have impacts on local labor markets and on particular housing markets far from the brownfield in question. We will fail to capture these effects to the extent that they appear in other parts of the city. Given the size of a typical brownfield (relative to the size of an urban area), this may not be much of a practical issue. We also ignore interactions between the owner-occupied residential and other real estate markets for the sake of simplicity, but this would be an interesting dimension for further study. 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 2013).

Second, our approach will also not capture the value accrued to other types of properties (e.g., rental and nonresidential) or the health benefits from remediation that people are not aware of (and, hence, are not reflected in house purchase decisions and transactions prices). While the benefits from reduced impacts to human health may generally be large, in contrast to other nuisances (Superfund sites, TSDFs, or other toxic waste exposure), we do not expect this to be as much of an issue for brownfield sites, making property value hedonics a good approach in this context.

Third, the net benefits of remediation are clearly more complicated than what we recover. This is because (i) we assume that the effect is constant at every point to the edge of the buffer, (ii) we do not look at variation in the type of reuse, and (iii) we only examine what happens in the short run in the few years after cleanup (after which there may be large feedback effects that will result in additional benefits).[44] As time passes, feedback effects have the potential to affect the welfare discussion. Endogenous neighborhood amenities (e.g., the Starbucks Effect [O'Sullivan 2005]) can be influenced by the cleanup and have subsequent effects on home values themselves. The cleanup can have interesting equity impacts as well, forcing up rents and driving low-

44. Things are further complicated because we cannot see if unawarded sites are actually cleaned up, although that would bias our estimate of the value of cleanup downward, making our estimates lower bounds.

income residents out (i.e., environmental gentrification), with negative welfare consequences (Banzhaf and McCormick 2007).

In light of these limitations, the alternative specifications explored yield a consistent conclusion—averaging over the experiences at a nationally representative sample of brownfield properties, cleanup leads to housing price increases between 5.0% and 15.2%. Taking the most conservative estimate of the value of an average site cleanup, we find that it indeed passes cost-benefit analysis by an order of magnitude based on the expenditures from the Brownfields Program. Moreover, our estimate using a difference-in-differences matching estimator without time variation is consistent with a willingness to pay (i.e., welfare) interpretation, not simply a capitalization effect. Although only one part of the larger EPA Brownfields Program, cleanup of brownfield sites alone yields large increases to nearby housing values and, given the DD-NNM results, has unambiguously positive welfare impacts on communities nearby.

REFERENCES

Abadie, Alberto, and Guido W. Imbens. 2006. Large sample properties of matching estimators for average treatment effects. *Econometrica* 74 (1): 235–67.

———. 2011. Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics* 29 (1): 1–11.

Adler, Kenneth J., Robert C. Anderson, Z. Cook, Roger C. Dower, Allen R. Ferguson, and Margo J. Vickers. 1984. The benefits of regulating hazardous waste disposal: Land values as an estimator. Public Interest Economics Center, Washington, DC.

Alberini, Anna, Alberto Longo, Stefania Tonin, Francesco Trombetta, and Margherita Turvani. 2005. The role of liability, regulation and economic incentives in brownfield remediation and redevelopment: Evidence from surveys of developers. *Regional Science and Urban Economics* 35 (4): 327–51.

Baden, Brett M., and Don L. Coursey. 2002. The locality of waste sites within the city of Chicago: A demographic, social, and economic analysis. *Resource and Energy Economics* 24 (1): 53–93.

Banerjee, Abhijit V., and Esther Duflo. 2009. The experimental approach to development economics. *Annual Review of Economics* 74 (1): 151–78.

Banzhaf, H. Spencer. 2015. Panel data hedonics: Rosen's first stage and difference-in-differences as "sufficient statistics." NBER Working paper w21485, National Bureau of Economic Research, Cambridge, MA.

Banzhaf, H. Spencer, and Eleanor McCormick. 2007. Moving beyond cleanup: Identifying the crucibles of environmental gentrification. National Center for Environmental Economics Working paper 07-02, US Environmental Protection Agency, Washington, DC.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins. 2016. A dynamic model of demand for houses and neighborhoods. *Econometrica* 84 (3): 893–942.

Bayer, Patrick, Christopher Geissler, and James Roberts. 2011. Speculators and middlemen: The role of flippers in the housing market. NBER Working paper w16784, National Bureau of Economic Research, Cambridge, MA.

Been, Vicki, and Francis Gupta. 1997. Coming to the nuisance or going to the barrios? A longitudinal analysis of environmental justice claims. *Ecology Law Quarterly* 24:1–56.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119 (1): 249–75.

Bishop, Kelly, and Christopher Timmins. 2011. Hedonic prices and implicit markets: Estimating marginal willingness to pay for differentiated products without instrumental variables. NBER Working paper w17611, National Bureau of Economic Research, Cambridge, MA.

Braden, John B., Xia Feng, and DooHwan Won. 2011. Waste sites and property values: A meta-analysis. *Environmental and Resource Economics* 50 (2): 175–201.

Fan, Jianqing, and Irene Gijbels. 1996. *Local polynomial modelling and its applications*. Monographs on statistics and applied probability, vol. 66. Boca Raton, FL: CRC Press.

Gamper-Rabindran, Shanti, Ralph Mastromonaco, and Christopher Timmins. 2011. Valuing the benefits of Superfund site remediation: Three approaches to measuring localized externalities. NBER Working paper w16655, National Bureau of Economic Research, Cambridge, MA.

Greenstone, Michael, and Ted Gayer. 2009. Quasi-experimental and experimental approaches to environmental economics. *Journal of Environmental Economics and Management* 57 (1): 21–44.

Heckman, J., H. Ichimura, J. Smith, and P. Todd. 1998. Characterizing selection bias using experimental data. *Econometrica* 66 (5): 1017–98.

Ihlanfeldt, Keith R., and Laura O. Taylor. 2004. Externality effects of small-scale hazardous waste sites: Evidence from urban commercial property markets. *Journal of Environmental Economics and Management* 47 (1): 117–39.

Kiel, Katherine A. 1995. Measuring the impact of the discovery and cleaning of identified hazardous waste sites on house values. *Land Economics* 71 (4): 428–35.

Kiel, Katherine A., and Michael Williams. 2007. The impact of Superfund sites on local property values: Are all sites the same? *Journal of Urban Economics* 61 (1): 170–92.

Kohlhase, Janet E. 1991. The impact of toxic waste sites on housing values. *Journal of Urban Economics* 30 (1): 1–26.

Kuminoff, Nicolai V., and Jaren C. Pope. 2014. Do "capitalization effects" for public goods reveal the public's willingness to pay? *International Economic Review* 55 (4): 1227–50.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. 2013. The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature* 51 (4): 1007–62.

Leigh, Nancey Green, and Sarah L, Coffin. 2005. Modeling the relationship among brownfields, property values, and community revitalization. *Housing Policy Debate* 16 (2): 257–80.

Linden, Leigh, and Jonah E. Rockoff. 2008. Estimates of the impact of crime risk on property values from Megan's laws. *American Economic Review* 98 (3): 1103–27.

Linn, Joshua. 2013. The effect of voluntary brownfields programs on nearby property values: Evidence from Illinois. *Journal of Urban Economics* 78:1–18.

Mihaescu, Oana, and Rainer vom Hofe. 2012. The impact of brownfields on residential property values in Cincinnati, Ohio: A spatial hedonic approach. *Journal of Regional Analysis and Policy* 42 (3): 223–36.

O'Sullivan, Arthur. 2005. Gentrification and crime. *Journal of Urban Economics* 57 (1): 73–85.

Palmquist, Raymond B. 2005. Property value models. *Handbook of Environmental Economics* 2:763–819.

Pastor, Manuel, Jim Sadd, and John Hipp. 2001. Which came first? Toxic facilities, minority move-in, and environmental justice. *Journal of Urban Affairs* 23 (1): 1–21.

Paull, Evans. 2008. The environmental and economic impacts of brownfields redevelopment. Northeast Midwest Institute Working paper, Washington, DC.

Rosen, Sherwin. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 82 (1): 34–55.

Savchenko, Olesya, and John B. Braden. 2014. Economic value of sustainable brownfield redevelopment. Working paper, University of Illinois at Urbana-Champaign.

Siniavskaia, Natalia. 2011. Property tax rates by county and city. http://www.nahb.org/generic.aspx?sectionID=734&genericContentID =163494&channelID=311 (assessed October 16, 2014).

Taylor, Laura O. 2003. The hedonic method. In *A primer on nonmarket valuation*, vol. 3, ed. Kevin J. Boyle, Thomas C. Brown, and Patricia A. Champ, 331–93. New York: Springer Science & Business.

US Conference of Mayors. 2010. Recycling America's land: A national report on brownfields redevelopment. http://www.usmayors.org/pressreleases/uploads/November2010BFreport.pdf.

Vigdor, Jacob L. 2002. Does gentrification harm the poor? *Brookings-Wharton Papers on Urban Affairs*, 2002, 133–82.

Wolverton, Ann. 2009. Effects of socio-economic and input-related factors on polluting plants' location decisions. *BE Journal of Economic Analysis and Policy* 9 (1): 1935–82.