# Continuous Spatial Difference-in-Differences for Welfare Evaluation*

Sean McCulloch
*Brown University* †

21 January 2026

[Click here for most recent version]

*Abstract:* Many events that occur at a location in space affect neighborhood amenities. Researchers frequently employ difference-in-differences to identify the impact of these events on nearby house prices. I analyze this setting in a potential outcomes model where the treatment is continuous in distance to the treatment site. This reveals an implicit assumption of homogeneous treatment effects needed to identify the price effect's derivative with respect to distance using difference-in-differences. In the hedonic model, this derivative is what recovers homeowner marginal willingness to pay and so difference-in-differences hedonic estimates no longer have a clear welfare interpretation without this assumption. However, I introduce an alternative assumption and estimator to bound the price derivative. Instead of requiring treatment effect homogeneity, I assume price effects are concave in distance from the treatment site, reflecting an underlying intuition for many applications in how spillovers decay with distance. I then propose an estimating procedure and apply it to three key applications from the literature: toxic industrial plant openings (TRI), Low Income Housing Tax Credit (LIHTC) developments, and sex offender move ins. For each application, using the new procedure changes the interpretation of the spatial treatment's effect. Finally, I compare values across amenities. Local homeowners would be willing to pay at least 13.1 million dollars for a toxic industrial plant not to exist, while the analogous figure for a low income housing development is 6.7 million dollars. For sex offender move ins, I find parallel trends violations.

*I thank Fabian Bald (discussant), Kyle Butts, Joseph Gyourko, Jon Roth, and Matthew Turner as well as seminar participants at Brown University, the University of Pennsylvania, the European Meeting of the UEA, the AERE summer meeting, the UEA Summer School, and the North American Meeting of the UEA for helpful comments. I thank the Population Studies and Training Center at Brown University, which receives funding from the NIH, for training support (T32 HD007338) and for general support (P2C HD041020). Previous versions of this paper were circulated as "A Unified Approach to Evaluating Neighborhood Amenities."

†Department of Economics, Brown University, Blistein House, Providence, RI 02906. email: sean_mcculloch@brown.edu.

## 1 Introduction

A large literature in urban and environmental economics uses difference-in-differences (DiD) to evaluate neighborhood amenities occurring at a location in space (Linden and Rockoff, 2008, Currie et al., 2015, Diamond and McQuade, 2019, among others). For example, an industrial plant opens at a particular address and spews pollutants into the air that affect nearby homeowners. These studies evaluate spillovers onto the nearby house prices surrounding a treatment site. Because these treatment sites significantly impact house prices, this suggests they also impact homeowners' welfare. However, there is no consensus in the literature on how to estimate these price effects and how such estimates translate into statements about welfare.

While most DiD applications are interested in estimating level treatment effects, in hedonics the price effect's *derivative* is what provides information about individuals' preferences and welfare. In particular, the classic result of Rosen (1974)'s model is that the first order condition from a homeowner's maximization problem relates their marginal willingness to pay (MWTP) for a housing characteristic to the hedonic price function's derivative. Thus, to make statements about welfare using the hedonic model, the estimand of interest is a continuous price function's derivative. However, the recent applied econometrics literature has raised the issue that DiD estimates do not identify the derivative under treatment effect heterogeneity (D'Haultfœuille et al., 2023, Callaway et al., 2024a, de Chaisemartin and D'Haultfœuille, 2025). Applied to this particular setting, a DiD estimate's derivative with respect to distance to a treatment site would not capture the true hedonic price derivative because comparisons across DiD estimates also include an additional selection bias term.

In this paper, I provide a framework for the spatial setting that can explain how to interpret DiD estimates under treatment effect heterogeneity and can connect these estimates to welfare statements using the hedonic model. I develop a bounding argument that does not impose treatment effect homogeneity instead exploiting context-specific restrictions. Finally, I outline an estimating procedure and illustrate the method by revisiting the three applications cited above.

To build an econometric foundation, I first formalize this setting in a potential outcomes model where the treatment is continuous in distance to the treatment site and describe the implicit homogeneity assumption needed to identify the treatment effect derivative. To do this, I adapt Callaway et al. (2024a)'s framework for continuous DiD to this spatial setting, which defines the causal parameters of interest and identifying assumptions explicitly. Callaway et al. (2024a) defines the *average causal response* as a continuous treatment effect's derivative and raises the issue that identifying the average causal response by DiD requires a strong assumption of treatment effect homogeneity across different levels of treatment intensity, or in the spatial context across distances to the treatment site. Callaway et al. (2024a) propose the *strong parallel trends* assumption, which assumes there is no heterogeneity in price effects across distances to the treatment site. A violation to this assumption is likely in many spatial applications because treatment sites are often selected. For example, urban planners might only site polluting industrial plants in neighborhoods where pollution would least affect house prices. Unfortunately, in this case the DiD price estimate's derivative no longer has a clear interpretation as a structural object because it is contaminated by an additional selection bias term.

However, I introduce an alternative assumption and estimator to bound the structural derivative. To build intuition, I first show the price derivative is identified under treatment effect heterogeneity if potential outcomes are linear in distance from the treatment site. In this special case, a simple "rise over run" estimator, which divides the level price effect by its distance from the location where spillovers end, identifies the derivative. I then relax the linearity assumption to a concavity assumption. A concavity assumption may be plausible in many spatial applications. For example, it may be reasonable to assume that the impact of an industrial plant's pollution on price are concave in distance from the plant because pollution levels decay exponentially moving away from the plant. In this case, the simple "rise over run" estimator is a conservative lower bound on the price derivative.

To provide a welfare interpretation to the econometric estimates, I then make explicit the connection between the average causal response parameter defined in the econometric framework and the hedonic model. I formalize what difference-in-differences can identify from Rosen (1974)'s first order condition when we do not assume treatment effect homogeneity. The average MWTP at each value of a housing characteristic is equal to the average causal response. As discussed above, this estimand is not identified without additional assumptions. However, under the concavity assumption, difference-in-differences bound the average marginal willingness to pay (MWTP) for proximity to a treatment site.

I then construct bounds on an aggregate welfare measure by adopting the modeling framework of Diamond and McQuade (2019)—except for their assumption of strong parallel trends. I conduct nonmarginal analysis by placing parametric assumptions on homeowners' utility functions (Bajari and Benkard, 2005) and estimate the aggregate willingness to pay (WTP) for the treatment site following a similar calculation. While this parameter is only point identified under homogeneous treatment effects, using the alternative price derivative estimator described above provides a conservative bound on homeowner aggregate willingness to pay. Finally, I use a revealed preference argument as in the prior hedonic literature (Ekeland et al., 2004, Banzhaf, 2021) to complete the bounds by adding an upper bound on the aggregate willingness to pay while allowing heterogeneous treatment effects. In this paper's empirical applications these bounds are fairly tight. In my leading empirical application, for example, I obtain bounds of [$13.1M, $19.5M] on the WTP to avoid a polluting industrial plant, compared with a point estimate of $16.4M under the strong assumption of homogeneous price effects.

For the empirical applications, I outline a nonparametric estimation procedure that yields a continuous price effect estimate. I implement the procedure in two stages. In the first stage, I group individual house sales into bins by distance from the treated site, giving an average sale price at each bin similar to Butts (2023). Binning the data forms repeated cross sections at various distances from the treatment site. In the second stage, I construct difference-in-differences comparisons from the bin level averages and then smooth the estimates using nonparametric sieve regression (Chen et al., 2024) as recommended in Callaway et al. (2024a). In both stages, the nonparametric methods employ data-driven tuning parameter choices following Cattaneo et al. (2024) and Chen et al. (2024), respectively.

I use this procedure to revisit three key spatial treatments studied in the literature: industrial plants that emit toxic pollutants, Low Income Housing Tax Credit (LIHTC) developments, and sex offender move ins. For each of these, revisiting the application with the new econometric procedure, and often far more observations, changes the interpretation of the spatial treatment's effect. I find negative effects on nearby housing prices for all three types of treatment event, though the estimate for sex offenders may reflect parallel trends violations rather than an actual effect. I estimate a negative price effect of 3% directly next to an industrial plant in the EPA's Toxic Release Inventory (TRI) but with price effects observable over a much larger geographic area than studied in Currie et al. (2015). I find a moderate, but precisely estimated, decline in house values around LIHTC developments with houses right next to a development losing 2% of house value. In contrast to the previous literature, I do not find much evidence of heterogeneous price effects by neighborhood income level. Finally, for sex offender move ins, I measure small negative price effects close to the sex offender's new address. However, these estimates are imprecise and event study plots show evidence of parallel trends violations.

I connect the price effects for TRI plants and LIHTC developments to a structural hedonic model to estimate the aggregate welfare impacts on local homeowners and compare amenity values across these two distinct types of treatment event. It is important to note that without stronger assumptions these estimates are not a complete accounting of all possible channels that could affect welfare. However, these impacts do provide a useful quantification of the localized spillovers onto nearby homeowners, which are economically meaningful. Quantifying aggregate impacts facilitates comparison of values across amenities: Local homeowners would be willing to pay at least 13.1 million dollars for a toxic industrial plant not to exist, while the analogous figure for a LIHTC development is 6.7 million dollars. Given the large number of TRI plants and LIHTC developments across the U.S., these externalities appear to be important. A back of the envelope calculation suggests that the negative externalities totaled over the whole nation are at least on par with the direct fiscal costs of major government infrastructure projects like the U.S. highway system. This highlights the importance of regular accounting of the value of non-marketed amenities for guiding policy.

### A *Relation to the literature*

There is currently no consensus in the literature on how researchers should estimate spatial difference-in-differences and how their estimates relate to welfare. Most applied studies run a standard DiD specification after specifying a binary threshold for being treated. Butts (2023) formalizes a multi-ring approach with a data-driven choice of these rings.[1] In contrast to most studies, Diamond and McQuade (2019) estimate continuous price effects. As in this study, they also use nonparametric methods but instead calculate empirical derivatives and then integrate up to a level price effect. They then show how a continuous price effect connects to a structural hedonic model to estimate welfare impacts. However, none of these papers consider treatment effect heterogeneity by distance to the treated site, which is key to interpreting these estimates in a hedonic model. I provide an econometric framework that accommodates treatment effect heterogeneity and explicitly connects the estimands defined in that framework to the hedonic

[1] See Butts (2023) for more description of earlier studies on multi-ring approaches.

model. In addition, I propose an estimation procedure that employs data-driven tuning parameter choices following Cattaneo et al. (2024) and Chen et al. (2024).

This paper builds off the recent applied econometrics literature on DiD with a continuous treatment in the presence of heterogeneous treatment effects. I start by showing how Callaway et al. (2024a)'s framework fits this spatial setting. As in recent work by D'Haultfœuille et al. (2023) and de Chaisemartin et al. (2025), I then exploit a connection between DiD and an older literature on panel data correlated random coefficient models (Chamberlain, 1982, 1992, Graham and Powell, 2012).[2] This older literature shows the structural derivative can be identified under heterogeneous treatment effects if the researcher imposes functional form restrictions. While Callaway et al. (2024a) do not consider functional form restrictions on potential outcomes, I show that if price effects are linear in distance to a treatment site, then average marginal effects are identified. I then further develop a bounding argument if linearity is violated but price effects are concave in distance from the plant, a natural assumption in many spatial settings.

A few other recent papers consider similar bounding arguments based on shape restrictions (D'Haultfœuille et al., 2023, de Chaisemartin et al., 2025). Most closely related, de Chaisemartin et al. (2025) consider an "average of slopes" estimand similar to the rise over run estimator described above and note that under concavity the average of slopes is a bound on counterfactual treatment effects. I employ a similar insight to lower bound average marginal effects, justifying concavity by using the context-specific assumption of spatial decay with distance. In contrast to de Chaisemartin et al. (2025), I also construct an upper bound by exploiting additional restrictions from the economic model's equilibrium conditions.

This paper also contributes to a large literature on hedonic property valuation, originating with Rosen (1974) who provides an economic interpretation for how house prices relate to individual preferences. First, I formalize how DiD estimates relate to homeowner MWTP in Rosen (1974)'s first order condition under treatment effect heterogeneity. Then, I employ a structural hedonic model closely following Diamond and McQuade (2019)'s model who calculate the total willingness to pay of nearby homeowners for a LIHTC development. Relative to Diamond and McQuade (2019), I connect these welfare estimands explicitly to the causal parameters defined in an econometric framework that allows treatment effect heterogeneity. Another important study of hedonics in difference-in-differences settings is Banzhaf (2021) who shows that the estimated price effects can measure an informative lower bound on the general equilibrium welfare change. However, it is worth noting that adding up reduced form price effects to form a lower bound is anti-conservative when price changes are negative (because both quantities are negative, the inequality reverses for absolute values). In this paper, I show how the hedonic modeling of both Diamond and McQuade (2019) and Banzhaf (2021) can relate a continuous DiD estimate to form complementary statistics to bound both sides of a welfare measure in the presence of treatment effect heterogeneity. Finally, by quantifying aggregate price and welfare impacts, this paper's method allows for comparison of hedonic values across completely different kinds of amenities.

The rest of this paper proceeds as follows. Section 2 presents an intuition for estimating

[2]See Arkhangelsky and Imbens (2024) for more details on the equivalence between panel data models and the recent applied econometrics literature on DiD.

continuous price effects from a spatial DiD and motivates how these price effects connect to the hedonic model under homogeneous treatment effects. Section 3 develops the econometric framework and discusses identification of the price derivative. Section 4 presents the structural hedonic model. Section 5 outlines the estimating procedure. As an example, Section 6 applies the procedure to estimate the local welfare impact from a polluting industrial plant (Currie et al., 2015). Section 7 compares price and preference estimates across applications. Section 8 concludes.

## 2 Intuition with homogeneous treatment effects

To develop intuition for the key estimand, in this section I first motivate the DiD estimator in the simpler case with homogeneous treatment effects and a fully separable price function. In this case, the derivative of the DiD estimator is the hedonic price function derivative and therefore identifies homeowners' MWTPs. In the latter Sections 3 and 4, I formalize these assumptions and consider a bounding argument that does not assume homogeneous treatment effects.

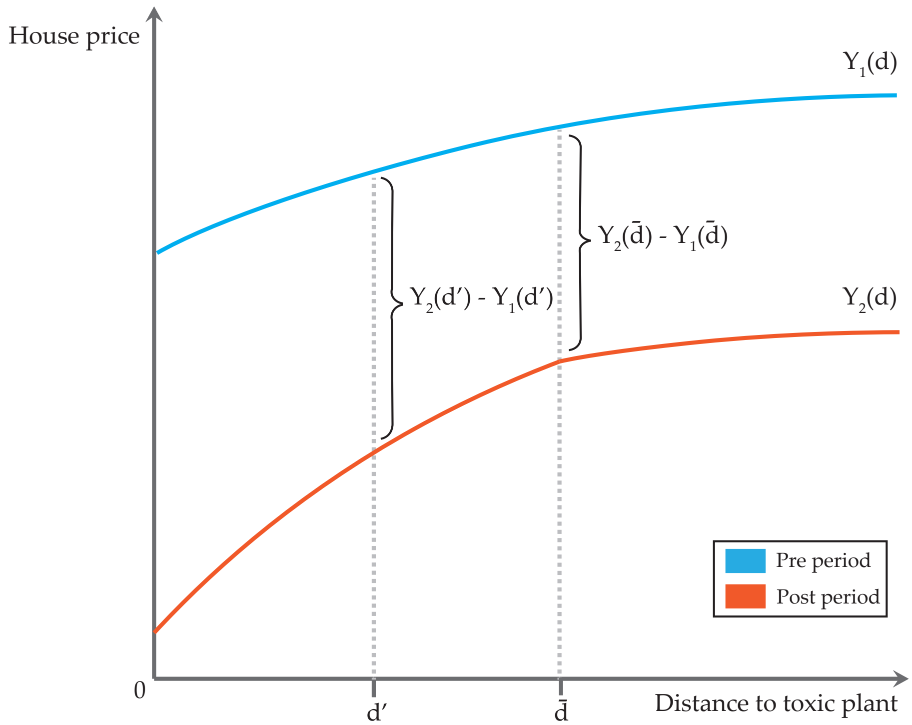
To fix ideas, in the first few sections I use the opening of an industrial plant that emits toxic pollutants as a running example. Figure 1 presents a stylized diagram of housing prices before and after the opening of this toxic industrial plant as a function of distance to the plant's site. There are two time periods $t \in \{1, 2\}$. In the second period, an industrial plant opens that spews toxic chemicals into the air. The blue line plots housing prices in the pre period and the orange line shows prices in the post period. Housing prices are lower close to the site even before the plant opens reflecting that the plant is sited in an area with systematically lower residential land quality. In the post period houses depreciate in value everywhere, but those closer to the plant depreciate more, with the additional contribution of pollution from the plant to housing values.

The black braces illustrate the construction of a price effect estimate for houses located $d'$ from the plant using a continuous DiD strategy. Let $\overline{d}$ be the distance at which pollution from the plant is no longer relevant. The price depreciation $(Y_2(d) - Y_1(d))$ for houses at $d'$ is larger in magnitude than the price depreciation for houses at $\overline{d}$ because of the closer proximity of $d'$ to the new plant. Let $\Delta t$ be a time trend common to all locations. Taking the difference of the two price changes yields a price effect estimate for being located $d'$ away from a toxic plant,

$$(Y_2(d') - Y_1(d')) - (Y_2(\overline{d}) - Y_1(\overline{d})) = (\Delta t + DiD(d')) - (\Delta t) = DiD(d'). \tag{1}$$

That is, $DiD(d')$ is the price effect from being $d'$ from a toxic plant on houses that are located $d'$ from a plant. Evaluating $DiD(d)$ at each location $d \in [0, \overline{d}]$ then gives a continuous price effect estimate.

Figure 1: Continuous spatial difference-in-differences

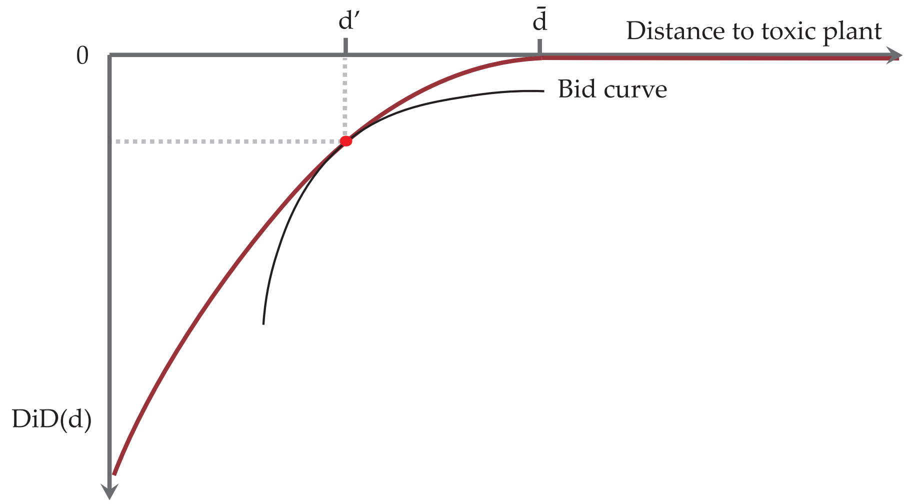



*Note: This figure illustrates the construction of a price effect from a continuous spatial difference-in-differences. The figure plots house price on distance to the toxic plant before and after the plant's opening. The blue line shows the gradient in the pre period and the orange line shows the gradient in the post period. The vertical distance between the gradients, marked by curly braces, shows the price change over time at two different locations $d'$ and $\bar{d}$, the distance at which pollution from the plant is no longer relevant. A DiD estimate for the effect at $d'$ is* $DiD(d') = (Y_2(d') - Y_1(d')) - (Y_2(\bar{d}) - Y_1(\bar{d}))$.

The red line in Figure 2 shows the $DiD(d)$ function from Equation (1), based on the differences between the pre and post period pricing gradients in Figure 1. Assuming treatment effect homogeneity, estimation of a price effect by differences-in-differences can identify the hedonic pricing function after treatment in the *ex post* equilibrium as explained in Banzhaf (2021). To develop intuition, assume for now that $DiD(d)$ is the post period hedonic price function with respect to $d$. Sections 3 and 4 will define the assumptions needed for identification formally.

The key insight of Rosen (1974)'s model is that homeowners' bid (indifference) curves lie tangent to the pricing function at each value of a housing characteristic. Figure 2 illustrates this tangency condition for one evaluation point $d'$. The black line shows the bid curve of homeowners who live $d'$ from the plant in the post period lying tangent to the pricing function at that location. As a result, the derivative of $DiD(d)$ evaluated at each $d$ provides information about individual homeowners' marginal willingnesses to pay for proximity to the treated site. Hence, in this paper, the derivative of the price function is the key estimand. A more detailed formulation of Rosen (1974)'s tangency condition is presented later in the model in Section 4.

Figure 2: DiD identifies the ex post hedonic price function under treatment effect homogeneity




*Note: This diagram shows the hypothetical DiD function constructed from the pre and post period pricing gradients in Figure 1. The red line shows price effects from the DiD plotted on distance to the toxic plant. With homogeneous treatment effects, DiD identifies the hedonic price function in the ex post equilibrium. The black line shows the bid (indifference) curve of homeowners who live $d'$ from the plant in the post period lying tangent to the pricing function at that location.*

In the next section, I provide an econometric framework adapting Callaway et al. (2024a) that formalizes the intuition from this section and generalizes to multiple treatment sites/time periods. In this framework, I define the identifying assumptions. This makes transparent the need for additional assumptions on treatment effect heterogeneity needed to identify the price derivative with respect to proximity to the plant.

## 3 Econometric framework

A researcher wants to study the effect of toxic plant openings on local house prices. The researcher observes $M$ distinct toxic plant openings indexed by $1, 2, \ldots, M$ and individual house sales occurring in time periods t = 1,2,...,$T$. Plants open at different times but all house sales close to one of the plants have the same treatment time, and so each location characterizes a group with the same treatment timing in a staggered adoption design. Let $G_j \in \{1, 2, \ldots M, \infty\} = \mathcal{G}$ be the plant that house sale $j$ is associated with, which is that housing unit's *treatment site-timing group*.[3] Following convention, I set $G = \infty$ for units that are untreated across all time periods. As formalized below, these are houses far enough away from the plant to not be affected by its spillovers. Finally, define $g^t : \mathcal{G} \mapsto \mathcal{T}$ as a function that maps a plant to the time period when it first opened.

For each plant there is a pre period and a post period. In the post period, units receive a dose of pollution $D = d$ based on their proximity to the toxic plant measured by distance from the treated site. Let $\mathcal{D}$ denote the support of $D$. Proximity is a one to one mapping to pollution

[3]This is a finer partition than technically required to form comparisons within the same timing group more similar to the grouping in a stacked DiD (Cengiz et al., 2019, Wing et al., 2024), but as discussed later using only comparisons within the same location is desirable to defend the assumptions needed for identification.

exposure and therefore fully characterizes the continuous treatment in this setting. Finally, let $X$ denote a house's pretreatment covariates.

Following the notation and framework of Callaway et al. (2024a), define $Y_{j,t}(g, d)$ as the potential outcome for house $j$ near plant $g$ in time period $t$. This is the house price (or log transformed house price) if house $j$ was $d$ distance away from a treatment site in time period $t$. The realized outcome is $Y_{j,t} = Y_{j,t}(G_j, D_j)$. I now define treated and untreated potential outcomes. In contrast to most DiD applications, in the spatial context we often do not know ex ante which units are untreated. For untreated potential outcomes to be well defined, we must place an assumption on how far away the plant affects house prices.

**Assumption 1** (Spatial extent of treatment effects). *As in the previous section, denote $\overline{d}$ where $0 < \overline{d} < \infty$ as the spatial extent of treatment effects, so units farther than $\overline{d}$ from a plant are untreated in every time period. For any $d \geq \overline{d}$, $Y_{j,t}(g, d) = Y_{j,t}(\infty, \overline{d})$.*

Researchers can corroborate this assumption either using application-specific knowledge or visually inspecting the estimated price effects. As discussed later in this section, correct specification of $\overline{d}$ is important, especially under treatment effect heterogeneity. Now, I write a unit's untreated potential outcome as $Y_{j,t}(\overline{d}) \equiv Y_{j,t}(\infty, \overline{d})$. Additionally, define $W_{j,t} = D_j 1\{t \geq g^t(G_j)\}$, which is the dose that unit $j$ experiences in time period $t$ and is equal to 0 for all units not yet treated in time period $t$.

Callaway et al. (2024a) define two treatment effect estimands of interest: the *level treatment effect* $Y_{j,t}(g, d) - Y_{j,t}(\overline{d})$ and the *causal response* $\frac{\partial Y_{j,t}}{\partial d}(g, d)$, the derivative of the potential outcomes function. In the DiD framework, we consider averages of these types of effects. For treated units,

$$ATT(g, t, d|g', d') = \mathrm{E}[Y_t(g,d) - Y_t(\overline{d})|G = g', D = d']$$

is the average effect of being $d$ distance from plant $g$ in time period $t$ for units that are located $d'$ from the plant $g'$. When $d = d'$ and $g = g'$, this is the observed $ATT$. Similarly, define the average causal response parameter as

$$ACRT(g, t, d|g',d') = \frac{\partial ATT(g, t, l|g', d')}{\partial l}|_{l=d} = \frac{\partial \mathrm{E}[Y_t(g, l)|G = g', D = d']}{\partial l}|_{l=d}$$

For identifying individuals' preferences according to Rosen (1974)'s tangency condition, the object of interest is the hedonic price function's derivative, so $ACRT(g,t,d|g,d)$ and its aggregates are the key estimands.

Next, I define sufficient assumptions for identification of the $ATT$, which I repeat from the appendices of Callaway and Sant'Anna (2021) and Callaway et al. (2024a). First, while ideally the unit of observation would be individual houses, researchers rarely observe multiple sales of the same house over a short period of time within a small geographic area. Hence, I set up the econometric framework based on repeated cross sections of a geographic area in given time periods, i.e. a distance ring surrounding a treatment site in the pre or post period. The following assumption from Callaway and Sant'Anna (2021) ensures identification with repeated cross sections:

**Assumption 2** (Random sampling for repeated cross sections). *Conditional on* $T = t$*, the data are independent and identically distributed from the distribution of* $(Y_t, G_1, G_2, \ldots, G_M, D, T, X)$*, for all* $t = 1, \ldots, T$*, with* $(G_1, G_2, \ldots, G_M, D, X)$ *being invariant to* $T$*.*

This assumption rules out compositional changes, which is often cited as an important concern in the context of housing sales. Depending on the particular application, changes to the composition of housing characteristics during the study time period may be a significant threat to identification.

**Assumption 3** (Support and continuously differentiable treatment). *(a) The support of* $D$ *is* $\mathcal{D} = [0, d_U]$ *with* $0 < \overline{d} < d_U$*, where* $d_U < \infty$ *is an upper bound on the spatial extent considered. Furthermore,* $\mathrm{P}(D > \overline{d}) > 0$ *and* $dF_{D|G}(d|g) > 0$ *for all* $(g, d) \in \mathcal{G} \times [0, \overline{d}]$

*(b) Assume for all* $g \in \mathcal{G}$ *and* $t \in 2, \ldots, T$*,* $\mathrm{E}[\Delta Y_t | G = g, D = d]$ *is continuously differentiable in* $d$ *on* $[0, \overline{d}]$*.*

**Assumption 4** (No anticipation / staggered adoption). *(a) For all* $g \in \mathcal{G}$ *and* $t \in 1, \ldots, T$ *with* $t < g^t(g)$*,* $Y_{j,t}(g, d) = Y_{j,t}(\overline{d})$*.*

*(b)* $W_{j,1} = 0$ *almost surely. For* $t = 2, \ldots, T, W_{j,t-1} = d > 0$ *implies that* $W_{j,t} = d$*.*

**Assumption 5** (Parallel trends). *For all* $g \in \mathcal{G}, t = 2, 3, \ldots, T$ *and* $d \in \mathcal{D}$*,* $\mathrm{E}[\Delta Y_t(\overline{d}) | G = g, D = d] = \mathrm{E}[\Delta Y_t(\overline{d}) | G = \infty, D = \overline{d}]$

Assumption 3 defines the support of treatment. Assumption 4 says that units do not anticipate treatment in ways that affect their outcomes, so before their treatment date we observe units' untreated potential outcomes. Assumption 5 is a parallel trends assumption. Note that in contrast to a binary DiD, this parallel trends assumption must hold across the continuum of distances $\mathcal{D}$, requiring that the average paths of untreated potential outcomes are the same for all groups and for all distances across all time periods.

Under these assumptions, the Average Treatment Effect on the Treated $ATT(g,t,d|g,d)$ is identified. As in Callaway et al. (2024a),

**Theorem 3.1** *Under assumptions 1 to 5, the* $ATT(g,t,d|g,d)$ *is identified.*

An explicit expression is given in Appendix B. Proofs for this section are in Appendix B.

### A *Nonidentification of the* $ACRT(g,t,d|g,d)$

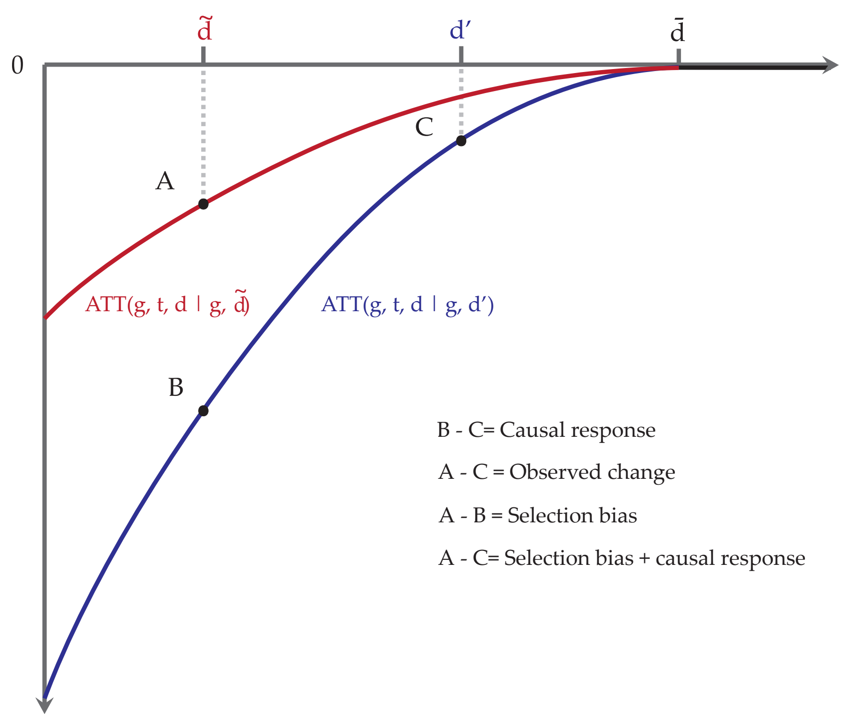
Thus far, I have recapped assumptions that are either standard to the binary DiD setup or place restrictions on the treatment's support. In a continuous DiD framework, Callaway et al. (2024a) discuss the need for additional restrictions on treatment effect heterogeneity along the dosage level in order to identify the $ACRT(g,t,d|g,d)$. In this section, I first review the key threat to identification discussed by Callaway et al. (2024a). I then discuss the *strong parallel trends* assumption they propose and why it may not hold in many spatial contexts.[4] Unfortunately,

[4]Callaway et al. (2024a) note explicitly that this assumption is unlikely to hold in many applications stating, "Before doing that, it is worth mentioning that we are not proposing [strong parallel trends] as an assumption that empirical researchers should readily adopt; in fact, in many applications, [Strong parallel trends] may be a strong or implausible assumption. Rather, our aim is to clarify that many natural target parameters in DiD applications with a continuous treatment require stronger assumptions than parallel trends..."

without this assumption, the $ATT$ function's derivative no longer has a clear interpretation as a structural object. However, in the next subsection I suggest methods to bound the welfare measure under alternative assumptions.

To illustrate the identification issue, Figure 3 shows hypothetical $ATT$ curves for units located at two distinct distances from a treatment site. The red curve represents the treatment effects for units located closer to the treatment site at $\widetilde{d}$ and the blue curve represents the treatment effects for units farther away at $d'$. In this example, units located farther away from the treatment site at $d'$ have a larger negative treatment effect at all counterfactual levels of $d$ compared to units in fact located at $\widetilde{d}$. The issue with this heterogeneity in treatment effects is that comparisons across two different dosage levels might be biased for the $ACRT$ because the comparison of $ATT$s at the two different levels will include the "jump" between $ATT$ response curves. For this reason, the $ACRT(g,t,d'|g,d')$, or the slope of the $ATT$ at $d'$, may not be identified.

Figure 3: Non-identification of the *ACRT* with treatment effect heterogeneity




*Note: This figure shows the average treatment effect on the treated $ATT(g,t,d|g,d)$ under counterfactual distances to the treatment site for the type of units located at two distinct distances (similar to Figure 3 in Callaway et al. (2024a)). The blue curve represents the treatment effects for units at $d'$ and the red curve represents the treatment effects for units located closer to the treatment site at $\widetilde{d}$. The type of housing unit at $d'$ is more affected by pollution at all pollution levels than the type of housing unit at $\widetilde{d}$. The three black points, A,B, and C, illustrate how using the difference between the observed $ATT(g,t,d|g,d)$ values may not estimate the causal response for moving a unit at $d'$ from $d'$ to $\widetilde{d}$ levels of pollution because $B-C < A-C$.*

To give a concrete example for a polluting industrial plant, say that houses closer to the plant's site have fewer trees in their neighborhood to start off with compared to houses farther away. Further imagine that toxic pollution kills trees and homeowners' positive valuations of trees are capitalized into house values. Then houses in neighborhoods with fewer trees would be less affected by the negative effects of pollution at any given distance from an operating plant. In this scenario, units at $\widetilde{d}$ will be less affected by treatment than a counterfactual in which the units at $d'$ were exposed to $\widetilde{d}$ levels of pollution.

The three black points, A,B, and C, illustrate how using the difference between the observed $ATT(g{,}t{,}d|g{,}d)$ values may not estimate the causal response for moving a unit at $d'$ from $d'$ to $\widetilde{d}$ levels of pollution. Imagine a researcher wants to know the counterfactual from changing units with dosage $d'$ to $\widetilde{d}$. In this case, they would like to know $ATT(g{,}t{,}\widetilde{d}|g{,}d') - ATT(g{,}t{,}d'|g{,}d')$, or $B - C$. They propose estimating this causal response using the observed $ATT$ values by $ATT(g{,}t{,}\widetilde{d}|g{,}\widetilde{d}) - ATT(g{,}t{,}d'|g{,}d')$, or $A - C$. But, the true causal response is more negative than suggested by the difference between the observed $ATT$ values, $B - C < A - C$.

This issue also applies to marginal comparisons. Figure 4 presents a pathological example where treatment effect heterogeneity causes the price derivative to have the wrong sign at most distances. The red line shows an $ATT(g{,}t{,}d|g{,}d)$ function at its observed values. The gray dotted lines show several counterfactual $ATT$ response functions for different types of housing units. In particular, each gray line is the counterfactual for the type of housing unit at the black dot's distance where the counterfactual response function intersects the red line. The slopes of the gray lines represent the true $ACRT$. In this example the observed $ATT$ function's derivative is negative at most distances, despite the true $ACRT$ being positive at every distance. A researcher interpreting the observed $ATT$ function's derivative as homeowners' MWTPs would find that most homeowners prefer to live closer to an industrial plant, despite all homeowners in this example actually preferring to live farther away. So, it is at least possible to imagine scenarios in which treatment effect heterogeneity causes the observed $ATT$ derivative to estimate preferences that differ vastly from homeowners' true preferences.[5]

More formally, I present one of the primary theorems in Callaway et al. (2024a) on non-identification of the $ACRT$ from the $ATT$ derivative.

**Theorem 3.2** *Under assumptions 1 to 5, the $ATT$ derivative recovers a mix of causal effect parameters and selection bias terms. In particular, for $d \in [0{,}\overline{d}]$,*

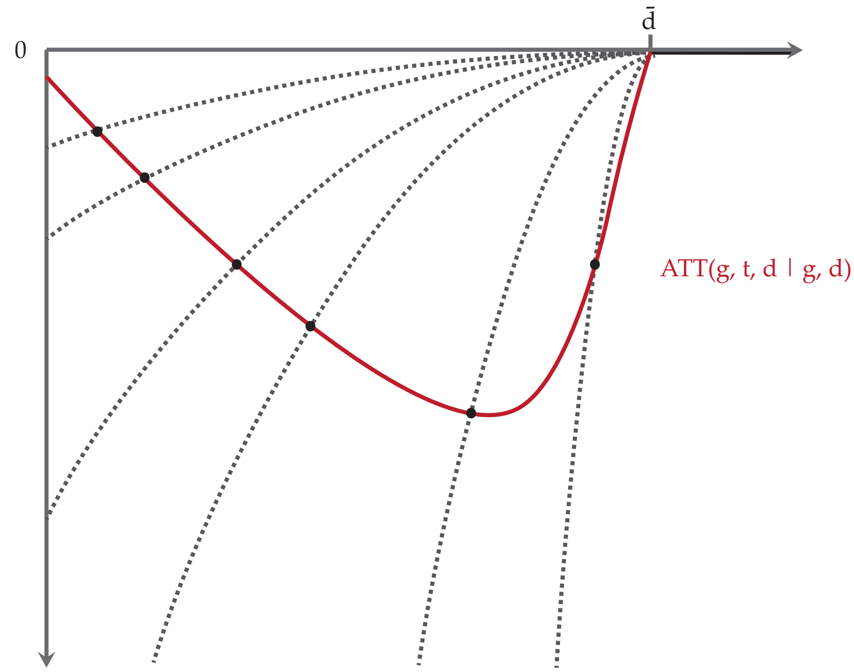
$$\frac{\partial \mathrm{E}[\Delta Y_t|G=g, D=d]}{\partial d} = \frac{\partial ATT(g{,}t{,}d|g{,}d)}{\partial d} = ACRT(g{,}t{,}d|g{,}d) + \underbrace{\frac{\partial ATT(g{,}t{,}d|g{,}l)}{\partial l}|_{l=d}}_{\textit{selection bias}}.$$

So estimating the $ACRT$ by the derivative of the $ATT$ curve results in the $ACRT$ plus an additional selection bias term.

What should a researcher do if they want to use DiD price effect estimates to make statements about welfare? As suggested in Callaway et al. (2024a), one option to recover a structural interpretation is to invoke a stronger assumption that limits treatment effect heterogeneity along the distance margin, such as the "strong parallel trends" assumption.

[5]Figure 4 also highlights a significant issue if the researcher incorrectly measures the spatial extent where the treatment effects end. Say, going from left to right, a researcher stops measuring treatment effects at the second to last black dot. By only comparing $ATT$ estimates at treated distance levels, the researcher would incorrectly estimate a positive effect from the treatment because of selection bias.

Figure 4: A pathological $ATT$ function




*Note: This figure shows an example where treatment effect heterogeneity causes the price derivative to have the wrong sign at most distances. The red line shows the $ATT(g,t,d|g,d)$ at its observed values. The gray dotted lines show several counterfactual $ATT$ response functions for different types of housing units. In particular, each gray line is the counterfactual for the type of housing unit at the black dot's distance where the counterfactual response function intersects the red line. At most distances the derivative of the $ATT$ function is negative, despite the true $ACRT$ being positive at every distance.*

**Assumption 6** (Strong parallel trends). *For all $g \in \mathcal{G}$, $t = 2,3,\ldots,T$ and $d \in \mathcal{D}$,*

$$\mathrm{E}[Y_t(g,d) - Y_t(\bar{d})|G = g] = \mathrm{E}[Y_t(g,d) - Y_t(\bar{d})|G = g, D = d].$$

Assumption 6 says that units at all distance levels will have the same $ATT$ response function or $ATT(g,t,d|g,d) = ATT(g,t,d|g)$. Effectively, this rules out treatment effect heterogeneity, e.g. selection on gains, which might not be plausible in many settings. However, with this assumption, identification of the $ACRT$ and therefore its interpretation as a structural object is recovered.

**Theorem 3.3** *Under assumptions 1-6, the $ACRT(g,t,d|d)$ is identified. That is, for $d \in [0,\bar{d}]$,*

$$\begin{aligned}\frac{\partial \mathrm{E}[\Delta Y_t|G = g, D = d]}{\partial d} = \frac{\partial ATT(g,t,d|g,d)}{\partial d} &= ACRT(g,t,d|g,d) + \underbrace{\frac{\partial ATT(g,t,d|g,l)}{\partial l}\Big|_{l=d}}_{=0} \\ &= ACRT(g,t,d|g,d).\end{aligned}$$

The discussion above suggests it is not hard to imagine situations where treatment site selection will imply heterogeneity by distance. For example, town planners may only allow industrial plants in areas where a plant's negative externalities would least affect housing values. In this case, a researcher may not be comfortable imposing the strong parallel trends assumption. In the next subsection I discuss how researchers can still bound the $ACRT$ under an alternative assumption.

## B *Lower bound on the* $ACRT$

In this subsection, I first motivate the bounding argument graphically. I then show that the $ACRT$ is point identified under restricted forms of treatment effect heterogeneity. In particular, if potential outcomes are linear in distance, then a simple estimator identifies the $ACRT$. I then weaken the linearity assumption to concavity. In this case, the simple estimator is a bound on the $ACRT$.

I will consider a "rise over run" estimator that takes the observed $ATT$ estimate and divides by its distance from $\overline{d}$. Figure 5 illustrates the intuition. As in Figure 4 the red line is the observed $ATT$ function. The gray dotted line is the counterfactual $ATT$ response function for units in the subpopulation $d'$ away from the treatment site. The blue dotted lines illustrate the rise over the run at $d'$ relative to $\overline{d}$.

To start, Panel (a) imagines the true counterfactual response function $ATT(g,t,d|g,d')$ as linear. In this special case, the observed $ATT(g,t,d'|g,d')$ identifies the causal response because we can "connect the dots" between 0 treatment effect at $\overline{d}$ and the $ATT(g,t,d'|g,d')$ at $d'$. The $ACRT$ or the price derivative is then simply the average rise over the common run: $\frac{ATT(g,t,d'|g,d')}{d'-\overline{d}}$. Unfortunately, we may expect a linearity assumption to be restrictive. As an equilibrium object, theory suggests the hedonic price function is generically nonlinear (Ekeland et al., 2004).

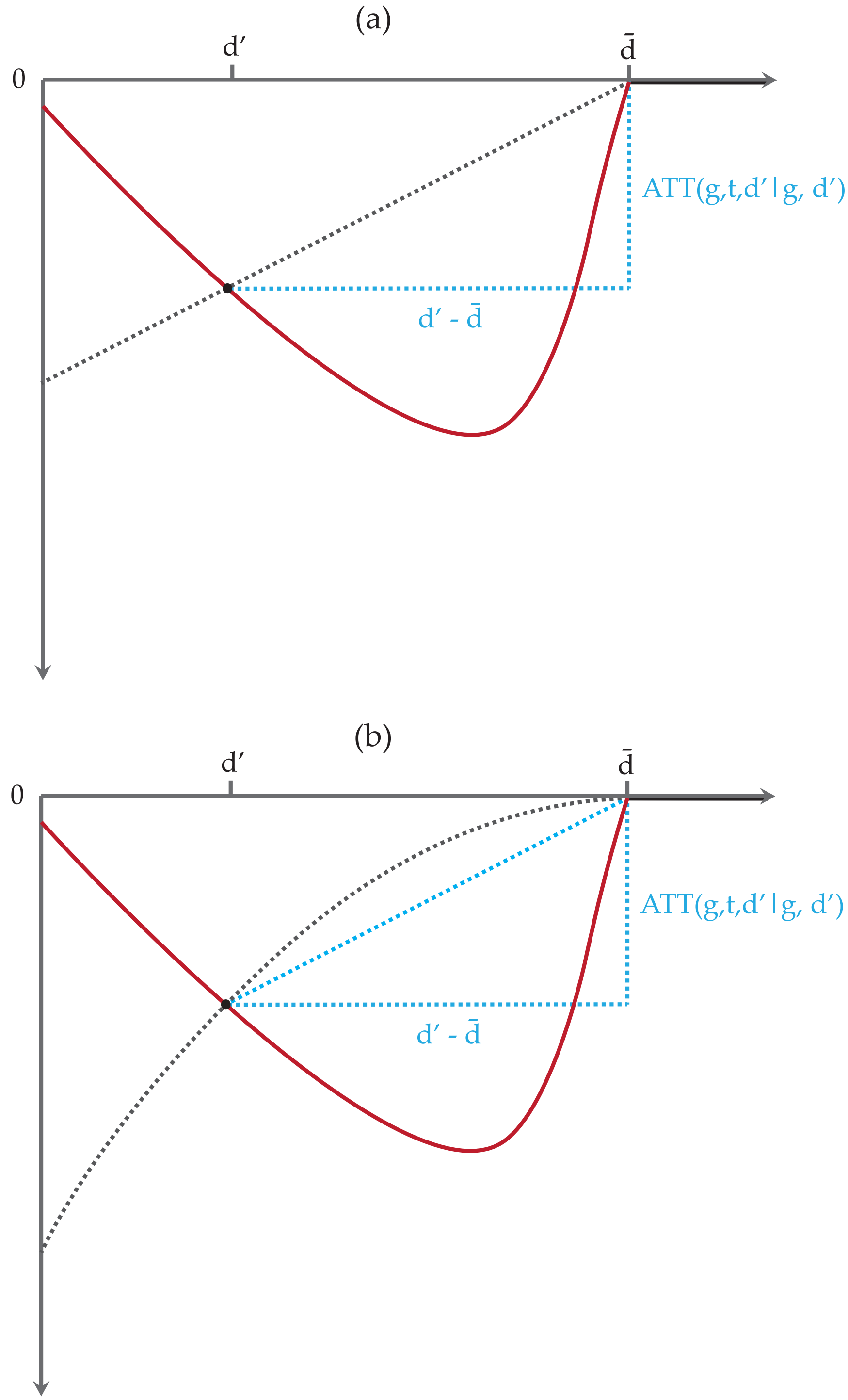
However, this rise over run estimator for the $ACRT$ can still be useful. Panel (b) shows the counterfactual $ATT$ response function again as a nonlinear function. The price derivative is no longer identified, but in this example with extreme selection bias the "connect the dots" approach yields a price derivative estimate that is much closer to the true price derivative than the observed $ATT$ function's derivative. Furthermore, a weaker assumption of concavity implies the rise over run estimator will always underestimate the $ACRT$. In Figure 5 this can be seen by the diagonal blue dotted line's slope being flatter than the gray dotted line's slope.

I argue a concavity assumption is plausible in many spatial DiD applications, where the ring DiD design was ex ante motivated by the expectation that treatment effects dissipate with distance from the treatment site. For example, imagine a house is several kilometers away from an industrial plant. A counterfactual price loss if the plant was sited a few blocks closer to the house may not be very large. On the other hand, imagine a house starts a few blocks away from an industrial plant. Moving the plant directly next to the house could have a significant negative effect on its value. In other words, we may expect pollution to impose convex costs getting closer and closer to the treatment site. It should be caveated that concavity is also an assumption and it might not hold depending on the application. For example, a subway station opening increasing prices for houses within walking distance of the station but decreasing prices for houses a little farther away would violate concavity. Still, for many spatial applications, the researchers' initial motivation for using the ring DiD research design was because they expected treatment effects to dissipate moving away from the treatment site.

The next few results formalize the intuition from Figure 5.

**Assumption 7** (Linear potential outcomes). *Assume a housing unit $j$'s potential outcomes is linear in distance to the treated site and $\overline{d}$ is now the minimum distance where spillovers from the plant are no longer*

Figure 5: Linear approximation to the $ACRT$




*Note: This figure illustrates that the rise over the run identifies the $ACRT$ when potential outcomes are linear in Panel (a) and is a conservative estimate when the counterfactual $ATT$ response functions are concave in Panel (b). As in Figure 4 the red line is the observed $ATT$ function. The gray dotted line is the counterfactual $ATT$ response function for units of the type at $d'$. The blue dotted lines illustrate the rise over the run at $d'$ relative to $\overline{d}$.*

*relevant. That is, for $D_j \in \mathcal{D}$ we have*

$$Y_{j,t} = \begin{cases} Y_{j,t}(\overline{d}) + \gamma_j(D_j - \overline{d}) & \text{if } D_j < \overline{d} \\ Y_{j,t}(\overline{d}) & \text{if } D_j \geq \overline{d} \end{cases}$$

*where $\gamma_j$ is a unit-specific slope coefficient.*

This assumption says that housing units' potential outcomes follow a linear correlated random coefficients model. Under this model, the average of the unit-specific slope coefficients $\gamma_j$ is the average causal response $ACRT$. Note that Assumption 7 now requires that the researcher exactly specify the minimum distance threshold $\overline{d}$ where treatment effects no longer matter, whereas for Assumption 1 any $\overline{d}$ above the minimum distance would suffice.

The following theorem shows that under Assumption 7 an alternative "rise over run" estimator identifies the $ACRT$.

**Theorem 3.4** *Under assumptions 1 to 5 and Assumption 7, the $ACRT$ is identified. In particular, for $d \in [0, \overline{d}]$,*

$$\frac{ATT(g,t,d|g,d)}{d - \overline{d}} = ACRT(g,t,d|g,d).$$

While similar estimators for the causal response are proposed in D'Haultfœuille et al. (2023), de Chaisemartin et al. (2025) and in an earlier literature on panel data correlated random coefficients models (Chamberlain, 1982, 1992, Graham and Powell, 2012), in this paper I show that this argument also holds in the framework of Callaway et al. (2024a). Hence, identifying or bounding the causal response in this way may be of independent interest in this literature or to applied researchers. Finally, it is worth noting that in this spatial setting point identification also required strengthening Assumption 1 to knowledge of the exact minimum distance threshold $\overline{d}$. However, in other contexts where zero dosage is clearly defined, this assumption is not needed, and the bound under concavity described below also does not require exact knowledge of the distance threshold.

As noted above, linearity may be a restrictive assumption. A weaker assumption is concavity.

**Assumption 8** (Concave ATT response functions). *Assume the ATT functions for each treatment group subpopulation $d' \in [0,\overline{d}]$ are concave on the interval $[d', \overline{d}]$. That is, for any $\lambda \in [0,1]$ we have*

$$ATT(g,t,(1-\lambda)d' + \lambda\overline{d}|g,d') \geq (1-\lambda)ATT(g,t,d'|g,d') + \lambda ATT(g,t,\overline{d}|g,d').$$

Recall in this spatial setting the treatment intensity defined relative to $\overline{d}$ is decreasing in $d$, so the x-axis is effectively inverted. Another way to view Assumption 8 is that it says the costs of pollution are convex in the amount of pollution exposure.

The next theorem shows the rise over the run is less than the $ACRT$ if the $ATT$ response functions are concave. This follows from a standard theorem for 1-dimensional concave functions.

**Theorem 3.5** *Under assumptions 1 to 4 and Assumption 8 for $d \in [0, \overline{d}]$ we have*

$$\frac{ATT(g,t,d|g,d)}{d - \overline{d}} \leq ACRT(g,t,d|g,d).$$

Hence, under concavity the rise over run estimator is a lower bound for the true $ACRT$. Also, in the case that the researcher incorrectly specifies $\overline{d}$ above the minimum, the rise over the run is still a lower bound on the $ACRT$.

In Section 4, I will use Assumption 8 and Theorem 3.5 to bound homeowners' willingnesses to pay. Interpreted in the hedonic model, a lower bound on the price derivative underestimates homeowners' MWTP to avoid proximity to an industrial plant. I then pair this result with a revealed preference argument as in the prior hedonic literature (Ekeland et al., 2004, Banzhaf, 2021) to bound both sides of the welfare measure. I postpone formal discussion of these bounds to Section 4 after I introduce the hedonic model's assumptions. Before proceeding to the hedonic model, I briefly consider the issue of aggregating causal parameters across treatment sites.

### C *Aggregating across groups and time periods*

In this subsection, I consider aggregated causal effect parameters that average over treatment sites and time periods. In practice, researchers will typically need to aggregate subgroup estimates in order to have statistical power. Following Callaway et al. (2024a), among units that ever participate in treatment define

$$ATT^{agg}(d|d) = \sum_{g\in\mathcal{G}} \sum_{t=2}^{T} w(g,t,d) ATT(g,t,d|g,d)$$

where the weights are nonnegative and sum to one.

A key takeaway from Callaway and Sant'Anna (2021) and Callaway et al. (2024a) is that by constructing the subgroup $ATT$ estimates manually, the researcher can now choose the weighting of effects across subgroups rather than having the weights be determined by OLS, which at best does not target a sensible summary parameter and at worst might lead to well known problems with negative weighting.[6] There are several possible reasonable weighting choices. For example, the simplest choice is to weight the $ATT$ estimates equally. In Section 5, I discuss alternative weighting choices in more detail.

Finally, define the aggregated causal response parameter as the derivative of the aggregated $ATT$ function,

$$ACRT^{agg}(d|d) = \frac{\partial ATT^{agg}}{\partial l}(l|d)|_{l=d}.$$

Depending on the assumption imposed, the $ACRT^{agg}$ is identified by the derivative of the observed $ATT^{agg}$ function or the rise over the run.

$$\textbf{Under strong parallel trends}\text{: } ACRT^{agg}(d|d) = \frac{\partial ATT^{agg}}{\partial d}(d|d);$$
$$\textbf{Under linear potential outcomes}\text{: } ACRT^{agg}(d|d) = \frac{ATT^{agg}(d|d)}{d - \overline{d}}.$$

As discussed directly below, in the hedonic model the price function derivative identifies individuals' MWTPs and so the $ACRT^{agg}(d|d)$ is the key estimand.

[6]See for example Borusyak and Jaravel (2017), de Chaisemartin and D'Haultfoeuille (2020), and Goodman-Bacon (2021), among others.

## 4 Structural hedonic model of housing choice

In this section, I present a structural hedonic model of housing choice for proximity to a spatially targeted treatment following the setup in Diamond and McQuade (2019) closely and maintaining most of their modeling assumptions. However, relative to Diamond and McQuade (2019), I impose weaker assumptions on the hedonic price function that allow for heterogeneous treatment effects and parallel shifts in the price function over time.

First, I introduce the canonical setup from Rosen (1974) and show formally how Rosen (1974)'s tangency condition relates average homeowner MWTP to the $ACRT$ estimand in Section 3. I then characterize nonmarginal WTPs by relying on a parametric assumption on homeowner's utility functions following Bajari and Benkard (2005). Finally, I construct bounds on local homeowners' aggregate willingness to pay for the plant not to exist using results from Section 3. As in Section 3, I discuss how weaker assumptions affect the interpretation of the welfare estimand.

In addition, a common empirical practice is to estimate the price change in logs and then multiply the $ATT$ estimates by a baseline price level to convert from logs back to levels, as in this paper's empirical applications. However, when allowing for heterogeneous price effects, this does not identify the level price change without an additional assumption. In the final part of this section, I formalize the required assumption to identify preferences in dollars using the $ATT$ for the log-transformed house price.

**Setup:** In Rosen (1974)'s classic model, a differentiated good is described by a vector of product characteristics, $\boldsymbol{z} = (z_1, z_2, \ldots, z_K)$. For a house, these characteristics may consist of structural attributes such as the lot size or the number of bedrooms, or neighborhood qualities such as proximity to a toxic industrial plant. Define $z_1 = d$ as radial distance to a toxic plant and $\boldsymbol{z_{-d}}$ as a vector collecting all other housing characteristics. In the previous section, house prices $Y_{j,t}$ were considered in a potential outcomes framework. Now, for a house $j$ with characteristics $\boldsymbol{z}$ in time period $t$, the potential outcomes $Y_{j,t}(d)$ are defined by a mapping from all housing characteristics

$$Y_{j,t}(d) = Y_t(\boldsymbol{z}) = Y_t(d, \boldsymbol{z_{-d}})$$

where as before $t \in \{1,2\}$ represents the pre versus the post period. In each period, the equilibrium matchings of consumers and producers leads to the observed function between house prices and characteristics, termed the hedonic price function.

**MWTP:** For each homeowner $i \in \mathcal{I}$, her utility maximization depends on her house's characteristics and consumption of a numeraire good, $c$.

$$\max_{\mathbf{z},c} u_{i,t}(\boldsymbol{z}) + c \;\; s.t. \;\; w_{i,t} - Y_t(\boldsymbol{z}) - c = 0$$

where $w_{i,t}$ is homeowner wealth. Maximization of the utility function implies that in equilibrium consumers choose the vector of characteristics $\boldsymbol{z_i^*}$ such that each characteristic $z_k$ satisfies

$$\frac{\partial u_{i,t}(\boldsymbol{z_i^*})}{\partial z_k} = \frac{\partial Y_t(\boldsymbol{z_i^*})}{\partial z_k}. \tag{2}$$

For distance to an industrial plant in particular

$$\frac{\partial u_{i,t}(d_i^*, \boldsymbol{z_{i,-d}^*})}{\partial d} = \frac{\partial Y_{j,t}(d_i^*, \boldsymbol{z_{i,-d}^*})}{\partial d}. \tag{3}$$

Thus, a consumer's marginal willingness to pay (MWTP) for distance to an industrial plant is the hedonic price function's partial derivative evaluated at her equilibrium product choice. Equation (2) is the key insight of Rosen (1974) and is the formal statement of the tangency condition suggested in Figure 2. It is worth noting that an implicit assumption of this first order condition is that homeowners choose from a continuum of housing choices and therefore can always solve for an interior solution.[7]

However, before we can relate DiD estimates to homeowners' MWTPs, we still require an additional assumption on the price function.

**Assumption 9** (Price function separability). *Assume the price function takes the following form:*

$$Y_t(\boldsymbol{z}) = m(d, \boldsymbol{z_{-d}}) + h_t(\boldsymbol{z_{-d}})$$

*where $m(.)$ is the contribution to price from proximity to an industrial plant and $h_t(\boldsymbol{z_{-d}})$ is the contribution from the other housing characteristics.*

This assumption says the change in the price function over time follows an additive fixed effects model. Similar assumptions on the data generating process are common in the panel data literature[8] and a stronger separability assumption is assumed in Diamond and McQuade (2019). $Y_{j,t}$ admits limited forms of nonseparable treatment effect heterogeneity through $m(d, \boldsymbol{z_{-d}})$, whereas a completely additively separable price function, i.e. $m(d,\boldsymbol{z_{-d}}) = m(d)$, would imply strong parallel trends.[9]

With this assumption, homeowners' average marginal willingnesses to pay in the post period relates to the $ACRT$. In this section, for expositional clarity I consider only two time periods and suppress the notation $g$ and $t$. The following proposition formalizes the connection between Rosen (1974)'s tangency condition and DiD estimates. Proofs for this section are in Appendix D.

**Proposition 1** (Rosen (1974)'s first order condition and the $ACRT$). *Suppose assumptions 1 to 5, Assumption 9, and the standard hedonic model hold. Then*

$$\mathrm{E}[\frac{\partial u_{t=2}}{\partial d}|D = d] = \mathrm{E}[\frac{\partial Y_{t=2}}{\partial d}|D = d] = ACRT(d|d).$$

*Furthermore, Section 3 suggested three alternative assumptions for identifying or bounding the $ACRT$:*
*(a) If strong parallel trends holds, then*

$$\mathrm{E}[\frac{\partial u_2}{\partial d}|D = d] = \frac{\partial ATT}{\partial d}(d|d).$$

*(b) If instead linear potential outcomes holds, then*

$$\mathrm{E}[\frac{\partial u_2}{\partial d}|D = d] = \frac{ATT(d|d)}{d - \overline{d}}.$$

[7] Bajari and Benkard (2005) discuss weakening the assumption of complete, continuous product choice, and show that it also leads to bounds instead of point identification on preferences.

[8] See for example Graham and Powell (2012).

[9] To see this, consider the definition of the $ATT$ under Assumption 9: $ATT(d|d') = \mathrm{E}[Y_{t-2}(d) - Y_{t-2}(\overline{d})|D = d'] = \mathrm{E}[m(d, -z)|D = d'] = \mathrm{E}[m(d)]$ for any $d'$.

*(c) If instead concave $ATT$ functions holds, then*

$$\mathrm{E}[\frac{\partial u_2}{\partial d}|D=d] \geq \frac{ATT(d|d)}{d-\bar{d}}.$$

Proposition 1 says that, in the post period equilibrium, the average marginal willingness to pay at each distance $d$ is equal to the $ACRT$ at $d$. As discussed in Section 3, the $ATT$ function's derivative only identifies the $ACRT$ under strong parallel trends, but the alternative rise over run estimator lower bounds the $ACRT$ under concavity. Note that once the price function allows for heterogeneous treatment effects we can no longer identify individual-level marginal willingnesses to pay, instead only identifying averages within each distance level.

**Nonmarginal WTP:** Although Rosen's first order condition in (3) relates each individual's MWTP with the hedonic price function's derivative at the chosen equilibrium bundle, in order to conduct nonmarginal welfare analysis, we need to characterize individuals' entire bid curves. Bajari and Benkard (2005) suggest making a parametric assumption on individuals' utility functions. Then, for each individual, her entire bid curve is recoverable from the one product choice made in equilibrium. In particular, they suggest a utility function that is log linear in the product characteristics.

For individual $i$, assume their utility maximization problem in time period $t$ is:

$$\begin{aligned} \max_{\mathbf{z},c} u_{i,t} = \beta_{i,d}\log(1+d) + \beta_{i,2}\log(z_2) + ... + \beta_{i,k}\log(z_k) + c \\ \text{s.t. } Y_t(\boldsymbol{z}) + c \leq w_{it} \end{aligned} \tag{4}$$

where $1+d$ is used so that utility is defined at zero.[10] Then individual $i$'s taste parameter for distance to the toxic plant $\beta_{i,d}$ can be solved by

$$\beta_{i,d} = (1+d_i^*)\frac{\partial Y_t(d_i^*, \boldsymbol{z^*_{i,-d}})}{\partial d}. \tag{5}$$

We can now consider homeowners' utility change under counterfactual pollution exposures. Consider the counterfactual where a homeowner is no longer exposed to the industrial plant but their house is otherwise the same. For a homeowner in the affected geographic area, her willingness to pay for the plant not to exist is

$$WTP_i = \beta_{i,d}\log(\frac{1+\bar{d}}{1+d_i^*}).$$

For homeowners farther than $\bar{d}$ away, their willingness to pay for the plant not to exist is $WTP_i = 0$. By assumption, an immediate consequence of Proposition 1 and Equation (5) is that the $ACRT$ identifies the average nonmarginal willingnesses to pay at each distance $d$.

**Lemma 4.1** *Suppose assumptions 1 to 5, Assumption 9, the standard hedonic model hold, and homeowners' utility functions are log linear according to Equation (4). Then*

$$\mathrm{E}[WTP|D=d] = (1+d)ACRT(d|d)\log(\frac{1+\bar{d}}{1+d}).$$

[10] For all empirical results, I assess sensitivity to this choice of constant in Appendix Table A3. Differences in WTP estimates across different choices of constant are marginal.

*As in Proposition 1(a),(b), and (c), depending on the assumption imposed, one of the two $ACRT$ estimators from Section 3 used in place of the $ACRT$ in the right hand side expression will identify or lower bound* $\mathrm{E}[WTP|D=d]$.

## A *Totaling to aggregate welfare impacts*

To create an aggregate welfare measure, we can add up individual homeowners' willingnesses to pay for the plant not to exist:

$$\text{Agg WTP} = \sum_i WTP_i = \sum_i \beta_{i,d} \log(\frac{1+\overline{d}}{1+d_i^*}).$$

This sum can be rewritten as the number of affected homeowners times the average willingness to pay by integrating over $[0,\overline{d}]$. Define $N_H$ as the number of affected housing units and $f(d)$ as the PDF of housing units. Then,

$$\text{Agg WTP} = N_H \int_0^{\overline{d}} \mathrm{E}[WTP|D=d] f(d)\,\mathrm{d}d.$$

Plugging in Lemma 4.1 yields an expression for the aggregate willingness to pay in terms of the $ACRT$.

**Proposition 2** (Aggregate willingness to pay). *Suppose assumptions 1 to 5, Assumption 9, the standard hedonic model hold, and homeowners' utility functions are log linear according to Equation (4). Then the aggregate willingness to pay for the plant not to exist is*

$$\text{Agg WTP} = N_H \int_0^{\overline{d}} (1+d) ACRT(d|d) \log(\frac{1+\overline{d}}{1+d}) f(d)\,\mathrm{d}d. \tag{6}$$

*As in Proposition 1 and Lemma 4.1, depending on the assumption imposed, one of the two $ACRT$ estimators from Section 3 used in place of the $ACRT$ in the right hand side expression will identify or lower bound* Agg WTP.

In Section 3, I argued that a concavity assumption was more plausible in many spatial applications than either a homogeneous treatment effects or a linear potential outcomes assumption. Under concavity, the right hand side expression evaluated using the rise over run as the $ACRT$ estimator is a lower bound on the aggregate willingness to pay.

We can also bound the aggregate willingness to pay on the other side using a similar argument as in the prior hedonic literature. Banzhaf (2021) shows the total of the reduced form price effects from a policy change is a lower bound on the aggregate welfare change. Note the sign reversal as I defined $WTP_i$ as the willingness to pay for the plant *not* to exist. The aggregate price change is

$$\Delta\text{Price} = N_H \int_0^{\overline{d}} ATT(d|d) f(d)\,\mathrm{d}d. \tag{7}$$

Reversing signs, $-\Delta$Price is an upper bound on the aggregate willingness to pay for the plant not to exist. It is worth noting that the aggregate price change is not necessarily conservative. While for amenities the aggregate price change will underestimate the magnitude of the welfare change, for disamenities like pollution exposure it will overestimate the magnitude.

The following result formalizes the bounding argument for homeowners' aggregate WTP.

**Theorem 4.2** (Bounds on Agg WTP). *Suppose assumptions 1 to 5, Assumption 9, the standard hedonic model hold, and homeowners' utility functions are log linear according to Equation (4). Under concave $ATT$ functions (Assumption 8), homeowners' aggregate willingness to pay is bounded:*

$$
\begin{aligned}
&N_H \int_0^{\overline{d}} (1+d) \frac{ATT(d|d)}{d-\overline{d}} \log\left(\frac{1+\overline{d}}{1+d}\right) f(d)\, \mathrm{d}d \quad \leq \text{Agg WTP} \\
&\leq -N_H \int_0^{\overline{d}} ATT(d|d) f(d)\, \mathrm{d}d.
\end{aligned}
$$

Theorem 4.2 says homeowners' aggregate willingness to pay for the plant not to exist is bounded between evaluating the willingness to pay using the rise over run estimator and the negative aggregate price change.

Here, I provide an intuition for the proof. First, the lower bound follows from the previous results and the properties of integrals because $\frac{ATT(d|d)}{d-\overline{d}} \leq ACRT(d|d)$ and all terms are positive. For the upper bound, the proof relies on a revealed preference argument exploiting the hedonic model's equilibrium conditions. In the canonical hedonic model, homeowners maximize over a complete, continuous set of housing choices. In this setting, homeowners in the affected geographic area could have chosen a house with characteristics $(\overline{d}, \boldsymbol{z^*_{i,-d}})$ but instead chose $(d^*_i, \boldsymbol{z^*_{i,-d}})$ close to the plant. If the price difference was less than their willingness to pay for the plant not to exist, they would have a profitable deviation by choosing a house with $(\overline{d}, \boldsymbol{z^*_{i,-d}})$. Hence, $Y_t(\overline{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}}) \geq WTP_i$ for every homeowner $i$ to be in equilibrium. Observe that the $ATT$ is the negative of the average of the left hand side expression conditioning on treatment group subpopulation $D = d$. As a result, $-ATT(d|d) \geq E[WTP|D = d]$. The upper bound then follows by the definition of the Agg WTP and the properties of integrals. An explicit proof is given in Appendix D.

The aggregate willingness to pay estimand is a useful welfare measure with a clear interpretation. It is the amount homeowners' living close to the plant in the post period would pay for the plant not to exist. This result holds under weaker assumptions than those in the prior literature: Diamond and McQuade (2019) assume treatment effect homogeneity and a constant hedonic price function over time.

However, without stronger assumptions that the price function is fully separable and time constant, the aggregate willingness to pay is not a complete measure of all possible effects on homeowner welfare. Homeowners may reoptimize their optimal bundle of other housing characteristics after the plant opens, and this reoptimization is not accounted for in the aggregate willingness to pay for the plant. However, while for interpreting the estimand researchers face a tradeoff between stronger conclusions and more credible assumptions, in practice the expression for the estimand does not change. After reversing signs, Equation (6) above coincides with the solution of Diamond and McQuade (2019)'s model.

**Log vs. level house prices:** Before proceeding to the empirical applications, I discuss using DiD estimates for log-transformed house prices to estimate monetary measures of homeowners' preferences. There are several good reasons researchers often use log-transformed house prices instead of level prices as the outcome. For example, the parallel trends and separability assumptions may be more plausible using proportional changes, and in estimation it may limit the influence of outliers. A common way to get a price effect in dollars from a log-transformed outcome is to multiply the estimates by average house prices. However, this transformation requires an additional assumption. Below, I formalize the conditions under which this works.

Throughout the remaining analysis, let $\tilde{ATT}(d|d)$ and $\tilde{ACRT}(d|d)$ refer to the analogous DiD estimands for the log-transformed outcome. That is,

$$\tilde{ATT}(d|d) = \mathrm{E}[\log Y_2(l) - \log Y_2(\bar{d})|D = d], \quad \tilde{ACRT}(d|d) = \frac{\partial \tilde{ATT}(l|d)}{\partial l}|_{l=d}.$$

Suppose separability (Assumption 9) instead holds for the log house price:

$$\log Y_t(d, \mathbf{z_{-d}}) = m(d,\mathbf{z_{-d}}) + h_t(\mathbf{z_{-d}}).$$

The house price derivative with respect to $d$ is

$$\frac{\partial Y_t(d, \mathbf{z_{-d}})}{\partial d} = \frac{\partial m(d,\mathbf{z_{-d}})}{\partial d} Y_t(d, \mathbf{z_{-d}})$$

and the $\tilde{ACRT}(d|d)$ is

$$\begin{aligned} \tilde{ACRT}(d|d) &= \frac{\partial}{\partial l}\mathrm{E}[\log Y_2(l) - \log Y_2(\bar{d})|D = d]|_{l=d} \\ &= \mathrm{E}[\frac{\partial \log Y_2(d)}{\partial d}|D = d] = \mathrm{E}[\frac{\partial m(d, \mathbf{z_{-d}})}{\partial d}|D = d]. \end{aligned}$$

Rosen's first order condition relates to the log house price. From utility maximization the first order condition for a homeowner $i$ is

$$MWTP_i = \frac{\partial Y(d_i^*, \mathbf{z}^*_{\mathbf{i},-\mathbf{d}})}{\partial d} = \frac{\partial m(d_i^*,\mathbf{z}^*_{\mathbf{i},-\mathbf{d}})}{\partial d} Y(d_i^*, \mathbf{z}^*_{\mathbf{i},-\mathbf{d}}).$$

First, notice that the average causal response using the log-transformed outcome identifies homeowner average MWTP as a *percentage* of house price. Dividing both sides by the house price,

$$\frac{MWTP_i}{Y(d_i^*, \mathbf{z}^*_{\mathbf{i},-\mathbf{d}})} = \frac{\partial m(d_i^*,\mathbf{z}^*_{\mathbf{i},-\mathbf{d}})}{\partial d}.$$

Then, as before take conditional expectations on both sides:

$$\mathrm{E}[\frac{MWTP}{Y(d, \mathbf{z_{-d}})}|D = d] = \mathrm{E}[\frac{\partial m(d,\mathbf{z_{-d}})}{\partial d}|D = d] = \tilde{ACRT}(d|d)$$

where the left hand side is the average MWTP as a percentage of house value.

However, without an additional assumption the average MWTP in dollars is not identified from multiplying $\tilde{ACRT}(d|d)$ by the mean house price.[11] The average MWTP in dollars is

$$\mathrm{E}[MWTP|D = d] = ACRT(d|d) = \mathrm{E}[\frac{\partial m(d,\mathbf{z_{-d}})}{\partial d} Y(d, \mathbf{z_{-d}})|D = d]$$

where the right hand side is not identified without the following assumption.

[11] Issues with estimating an average treatment effect by first estimating the effect in logs and multiplying by a baseline mean are discussed in more detail in Chen and Roth (2024), Remark 8.

**Assumption 10** *(Uncorrelated proportional response) For all $d \in \mathcal{D}$,*

$$\mathrm{E}[\frac{\partial m(d,\mathbf{z_{-d}})}{\partial d} Y(d,\mathbf{z_{-d}})|D=d] = \mathrm{E}[\frac{\partial m(d,\mathbf{z_{-d}})}{\partial d}|D=d]\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d].$$

This assumption says that, conditional on a treatment group subpopulation $d$, percent changes in house price from the plant are uncorrelated with level house prices. With this additional assumption,

$$\mathrm{E}[MWTP|D=d] = ACRT(d|d) = \tilde{ACRT}(d|d)\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d]$$

where $\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d]$ is the average post period house price $d$ distance away from a plant.

If Assumption 10 holds, then the average causal response in logs multiplied by the average house price is the average causal response in levels. As a result, $\tilde{ACRT}(d|d)\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d]$ can be substituted for $ACRT(d|d)$ throughout the propositions and theorems developed in this section, with the other assumptions (1-9) instead applied to the log-transformed house price.

I now write equations for the aggregate willingness to pay and price change measures in terms of the DiD estimands for the log-transformed house prices. Using the log-transformed $\tilde{ACRT}(d|d)$, the estimating equation for the aggregate willingness to pay is

$$\text{Agg WTP} = N_H \int_0^{\overline{d}} (1+d)\tilde{ACRT}(d|d)\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d]\log(\frac{1+\overline{d}}{1+d})f(d)\,\mathrm{d}d. \tag{8}$$

Using the log-transformed $\tilde{ATT}(d|d)$, the aggregate price change is

$$\Delta\text{Price} = N_H \int_0^{\overline{d}} \tilde{ATT}(d|d)\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d]f(d)\,\mathrm{d}d \tag{9}$$

where $ATT(d|d) = \tilde{ATT}(d|d)\mathrm{E}[Y(d,\mathbf{z_{-d}})|D=d]$ under Assumption 10 and log changes approximating percent changes.

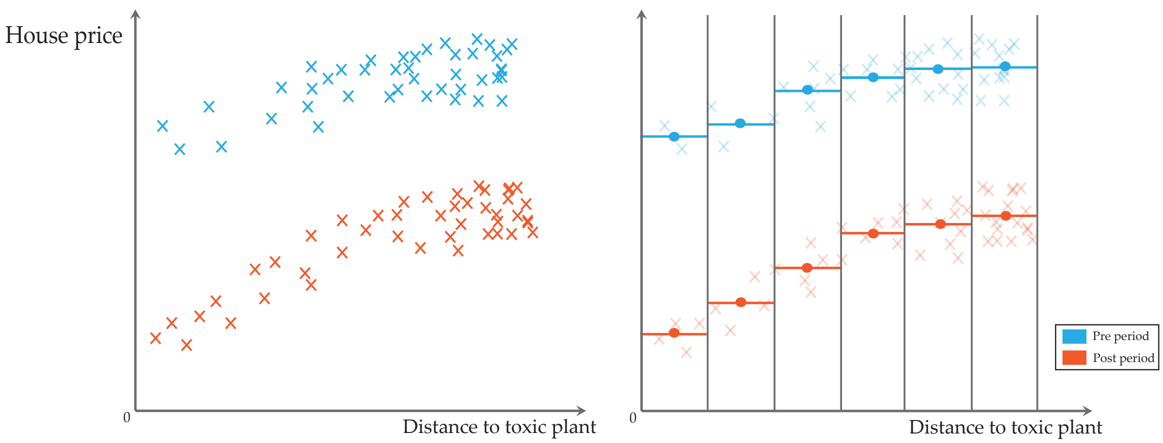
Equations (8) and (9) are the estimating equations I take to the data. In the primary specifications, I use log-transformed house prices as the outcome, which imposes Assumption 10. For each application in this paper, I report three estimates: (i) a lower bound using the rise over run estimator as the $ACRT$ in Equation (8), (ii) an upper bound using the aggregate price change, and (iii) a benchmark estimate assuming homogeneous treatment effects by using the $ATT$ derivative as the $ACRT$. In addition, I report homeowner's average willingness to pay as a percentage of house price, which does not impose Assumption 10. All terms in these equations can be estimated given data on treatment site locations, house sales, and the density of housing units where the key inputs to these equations are simple transformations of $\tilde{ATT}(d|d)$.

## 5 Estimation strategy

I now outline an estimating procedure a researcher could apply to housing sale microdata that yields a continuous $ATT$ function estimate. This procedure has two stages. In the first stage, house sales are grouped into different rings or bins by distance to a polluting plant, giving an average sale price at each bin. Binning the data forms repeated cross sections at different distances from the treatment site. I then construct $2 \times 2$ DiD comparisons from the bin level averages yielding price effect estimates at various $d$ from a treatment site. In the second stage, I trace out a smooth relationship using a nonparametric sieve regression of the DiD comparisons on $d$. In both stages, this procedure employs data-driven tuning parameter choices.

**First stage:** Figure 6 returns to the intuition for the continuous DiD estimator in the simple example in Figure 1 but now imagines estimating pre and post period pricing gradients using housing sale microdata. The left panel shows individual house sales distributed around the pre and post period pricing gradients with some unit-level noise. In the right panel, the sales have been grouped into different bins by distance to the plant's site and then aggregated to house price averages indicated by the constant fit lines.

Figure 6: Binning diagram




*Note: This diagram conceptualizes how binning data by its distance to the plant site forms a nonparametric approximation of the pre and post period pricing gradients. The left plot shows hypothetical realizations of housing sale microdata in the pre and post periods. On the right, the microdata has been aggregated to distance bin averages indicated by the constant fit lines.*

An advantage of aggregating to distance bins is that we can compare averages of the same bin over time forming repeated cross sections that fit into a standard DiD framework. Let $\overline{Y}_{g,t,d}$ be the (log) average house price in the bin at distance $d$ away from the toxic plant $g$ in time period $t$. In this example, $t \in \{1,2\}$. Consider the first bin closest to the toxic plant as the bin at distance $d'$ and the bin farthest away from the plant at distance $\overline{d}$. The price depreciation for the $d'$ bin is $\Delta Y_{g,d'} = (\overline{Y}_{g,2,d'} - \overline{Y}_{g,1,d'})$ and the price depreciation at $\overline{d}$ is $\Delta Y_{g,\overline{d}} = (\overline{Y}_{g,2,\overline{d}} - \overline{Y}_{g,1,\overline{d}})$. Then, a measure of the treatment effect at $d'$ is given by

$$ATT(g,2,d'|d') = (\overline{Y}_{g,2,d'} - \overline{Y}_{g,1,d'}) - (\overline{Y}_{g,2,\overline{d}} - \overline{Y}_{g,1,\overline{d}}).$$

For this hypothetical plant, we have six DiD comparisons, one for each of the six distance bins where the control bin $\overline{d}$ is 0 by construction. In the actual empirical applications, I choose the optimal number of bins following Cattaneo et al. (2024).

In the application in Section 6, I apply this binning procedure across a large number of polluting plants. I estimate pre and post period pricing gradients around over 11,000 industrial plants. Binning the data around these plants forms hundreds of thousands of $ATT$ estimates at various distances from the plants' sites, which I use to trace out a smooth relationship between price and distance to a polluting plant.

**Second stage:** After constructing $ATT$ estimates, estimating a smooth price effect is now a 1-dimensional regression of the price changes on distance to the treatment site. For this regression, I follow Callaway et al. (2024a)'s recommendation to use the nonparametric sieve estimator from Chen et al. (2024). This estimator avoids functional form assumptions and has desirable statistical properties.[12] Chen et al. (2024) compute the optimal sieve dimension $\hat{K}$ for regressions on flexible $\hat{K}$-dimensional transformations of $d$, which I denote by $\psi^{\hat{K}}(d)$. Then, the second stage regression of the $ATT$ estimates on $d$ is

$$A\hat{T}T(g,2,d|,d) = \psi^{\hat{K}}(d_{g,d})'\gamma_{\hat{K}} + \epsilon_{g,d}, \tag{10}$$

and the estimator for $ATT^{agg}(d|d)$ is

$$A\hat{T}T^{agg}(d|d) = \psi^{\hat{K}}(d_{g,d})'\hat{\gamma}_{\hat{K}}. \tag{11}$$

One could also use a simpler parametric regression or other nonparametric estimators to smooth the price effect estimates.

### A *Aggregation weights*

Equation (10) weights the $ATT$ estimates equally, but there are many reasonable weighting choices for aggregating the treatment effect parameters. In this particular context, I argue that it is feasible to gain significant improvements in efficiency by weighting by the $ATT$ estimates' inverse variances. In the empirical applications, the number of observations per bin can vary dramatically both across and within treatment sites, especially for bins that are equally spaced and for bins close to the treatment site.[13] For this regression, we have knowledge of the data generating process by calculating the bin means from the house sale microdata. Consider the case above with two time periods. An estimator of $ATT$ in site $g$ using the bin means is

$$A\hat{T}T(g,2,d) = (\hat{\overline{Y}}_{g,2,d} - \hat{\overline{Y}}_{g,1,d}) - (\hat{\overline{Y}}_{g,2,\overline{d}} - \hat{\overline{Y}}_{g,1,\overline{d}})$$

where $\hat{\overline{Y}}_{g,t,d}$ is a sample bin mean with an estimated variance $\hat{\sigma}^2_{g,t,d}$. An estimator of the $ATT$ variance is $var(A\hat{T}T(g,t,d)) = \hat{\sigma}^2_{g,2,d} + \hat{\sigma}^2_{g,1,d} + \hat{\sigma}^2_{g,2,\overline{d}} + \hat{\sigma}^2_{g,1,\overline{d}}$ where all terms can be estimated using the standard errors of the means in the data. Then, the normalized weights are

$$\hat{w}(g,2,d) = \frac{1/var(A\hat{T}T(g,2,d))}{\sum_{g\in\mathcal{G}} 1/var(A\hat{T}T(g,2,d))},$$

and we can estimate $ATT^{agg}(d|d)$ as in Equation (10) using $\hat{w}(.)$ as regression weights.

[12] Chen et al. (2024) chooses the sieve dimension such that $ATT(d|d)$ converges at the fastest possible rate in sup-norm and yields uniform confidence bands that are asymptotically more precise than undersmoothing. More details are in Callaway et al. (2024a) and Chen et al. (2024).

[13] Appendix C discusses the rationale for this weighting choice in more detail. In particular, it shows empirically the variation in how densely populated bins are with housing sales and formally tests for heteroskedasticity around polluting industrial plants.

Because there are multiple reasonable weighting choices, in the appendix I show all price effect estimates under several different weighting choices. In general, estimates based on equal weighting are similar but less precise, and this discrepancy looks to be driven by estimates derived from small numbers of house sales.

## 6 Application: Impact of a polluting plant (Currie et al., 2015)

In this section, I illustrate the construction of a local welfare change estimate with an empirical example by returning to Currie et al. (2015) who study the effect of toxic industrial plant openings on house values. Overall, I find similar price effect estimates within the same spatial range as the original paper but the continuous DiD reveals a much wider spatial extent of these price effects. Finally, I extend their analysis by incorporating these price effects into the structural hedonic model to estimate local homeowners' aggregate willingness to pay for a polluting industrial plant not to exist.

### A *Data*

Estimating the aggregate welfare measures requires only three data sources. The main data sources are a list of treatment events with geographic coordinates/dates and housing transaction microdata. Additionally, getting an estimate of the number of affected houses requires a complete count of housing units in these areas.

**Toxic plants (EPA):** As in Currie et al. (2015), I identify plants that emit toxic pollutants using the EPA's Toxic Release Inventory (TRI) database. The EPA requires industrial plants to self report their emissions if they handle more than a threshold amount of toxic chemicals. While the self reported emission quantities themselves are unreliable, the TRI is useful as a registry to identify plants that emit pollutants with their precise geographic locations.

Currie et al. (2015) are able to measure plants' exact operating dates by linking the TRI data to the U.S. Census Longitudinal Business Database. Unfortunately, this database is confidential and not available to me. Instead, I rely on imputation using when a plant first appears in the TRI database as the event date and restricting to plants that appear in the TRI database throughout the duration of a plant's post period. On the other hand relative to Currie et al. (2015), I have housing transaction data from more states over a longer time frame, implying my estimates are based on many more plant openings and housing transactions. Nevertheless, in comparable specifications from Currie et al. (2015), I find similar price effect estimates.

**Housing sales (CoreLogic):** I use data on single family home sales from the CoreLogic Deed database as the primary outcome. CoreLogic is a private data aggregator that provides detailed records on the near universe of single family home transactions in most metropolitan areas. This housing transaction microdata includes a rich collection of parcel-level housing characteristics and the exact coordinate location and date of each sale. I use transactions from 1990 to 2020 restricting to arms length transactions and dropping extreme outliers.[14] All house sales are adjusted to 2023 dollars.

[14]Specifically, I drop sales recorded as less than $10,000 or greater than $10 million dollars as in Currie et al. (2015). I also log transform prices to limit the influence of outliers.

**Housing unit counts (U.S. Census):** Not all houses in an area will transact during the study time period, so to get an estimate of the total affected housing units I add up the housing unit counts in 2010 U.S. census blocks surrounding the toxic plants. I retrieve counts separately by whether housing units are owner or renter occupied. In this paper, because I do not use rental listing data, I focus on the density of owner occupied housing units.

**Summary statistics:** Panel A of Table 1 presents summary statistics on the plants used in the analysis sample. In total, the price effect estimates are based on 11,849 distinct plants that first appeared in the TRI data from 1995-2015 with the typical plant first appearing in 2005. The final row shows the average annual self reported emissions for these plants was around 18,457 pounds per year. Although the self reported quantities are unreliable, this average suggests that these plants do emit significant amounts of pollutants over the study time period.

Table 1: Plant characteristics and housing sales / units by distance to plant site

| | | | | | | |
|---|---|---|---|---|---|---|
| **Panel A. Plant characteristics** | | | | | | |
| Number of plants | 11,849 | | | | | |
| Average year of opening | 2005 | | | | | |
| Avg. annual emissions (lbs) | 18,457 | | | | | |
| | Distance to toxic plant site (km) | | | | | |
| **Panel B. Housing sales** | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 |
| Avg. house price (2023 $) | $321,537 | $350,266 | $360,961 | $375,839 | $382,229 | $393,012 |
| Num. plant × house sales | 3,029,638 | 12,556,605 | 21,862,672 | 30,144,086 | 37,186,807 | 43,141,011 |
| Num. plant × distance bins | 27,894 | 51,819 | 54,881 | 56,036 | 56,315 | 56,315 |
| **Panel C. Housing units** | | | | | | |
| Num. owner occ. units | 3,936,235 | 15,902,543 | 27,660,928 | 38,031,320 | 47,101,033 | 55,373,611 |

*Note: This table presents summary statistics on treatment sites used in the sample and the average sale prices and owner occupied housing unit counts of nearby houses by distance to the treatment site. Panel A describes 11,849 distinct plants that first appeared in the TRI data from 1995-2015. Panel B shows the average house price in real 2023 dollars by distance to the toxic plant site. Panel C tabulates totals of owner occupied housing unit counts in nearby census blocks by distance to the toxic plant site.*

Panel B shows the average house prices by distance from the toxic plants within 6 kilometers of the plants' sites.[15] On average, house values are substantially higher farther away from the plant site, suggesting the locations selected for industrial plants are of systematically lower residential land quality. The final two rows of panel B show how aggregating to distance bin averages reduces the large number of individual housing transactions to a more manageable analysis dataset.

[15]It is not obvious how far away from a treated site a researcher should consider potential price effects ex ante. For TRI plant sites, I began by considering houses within 2 miles or roughly 3.2 kilometers of the treated sites following the choice of Currie et al. (2015). However, in the price effect estimates I observed a strong gradient at least out to 3.5 kilometers leading me to extend the range considered to 6 kilometers where the gradient flattens. An advantage of approaches like those outlined in Diamond and McQuade (2019), Butts (2023) or in this paper, is that they allow visual inspection of the treatment effect curve to assess where the treatment effects decay to 0.

Panel C tabulates totals of owner occupied housing units in nearby census blocks by distance to the toxic plant site. A key consideration for evaluating a spatial treatment's impact is that there are typically many more housing units farther away than close in, as the area of a circle grows geometrically with its radius. For example, around all treatment site locations in my sample there are only 3.9 million owner occupied housing units within 1 kilometer, while there are 55.4 million owner occupied housing units 5-6 kilometers away from the treatment sites. As a result, even small price effects far away from a treatment site can still have a big economic impact, as they affect large numbers of housing units.

### B *Results*

In this section, I first show the construction of DiD bin comparisons following the procedure outlined in Figure 6 for a single plant. I then present estimates using all of the 11,849 plants, which results in 303,260 DiD bin comparisons at various distances from a plant site. Finally, I smooth these binned comparisons with a nonparametric sieve regression following Chen et al. (2024).

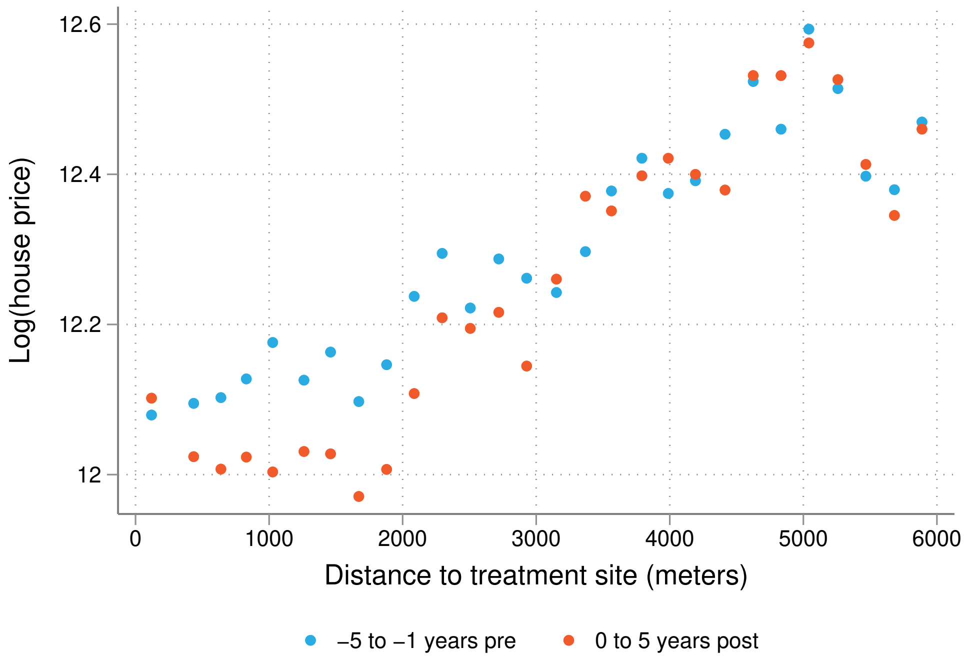
Figure 7 illustrates the binning procedure for just one of the treatment sites. The top graph plots log house price averages by distance to an industrial plant located in the Providence, RI metro area which first appeared in the TRI database in 1996. The blue dots show the averages by bin in the pre period from 1991 to 1995 and the orange dots show the averages in the post period from 1996 to 2001. The bins were chosen following Cattaneo et al. (2024) as the IMSE optimal number of equally spaced bins in the pre period.[16,17] In the bottom graph, the blue triangles show comparisons at each bin from taking the difference between the orange and the blue dots and then subtracting the analogous difference in the bin farthest away from the plant (rightmost bin). This procedure yields a distinct difference-in-differences estimate for each bin. Estimates using only comparisons around a single plant are noisy, but suggest a gradient with larger price declines closer to this plant.

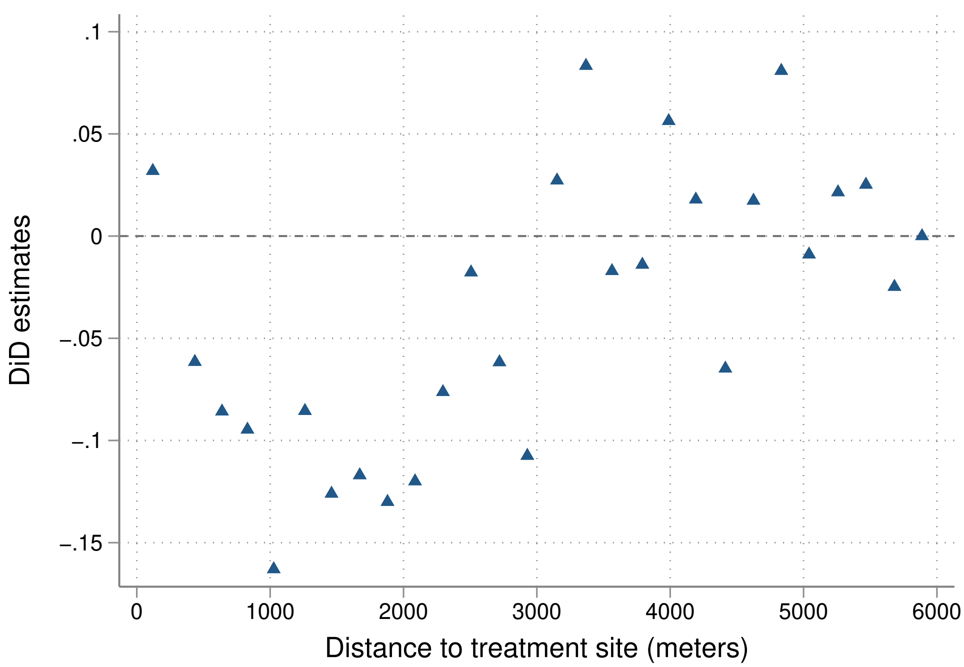
I repeat this procedure separately for each of the 11,849 plants in the sample allowing for a different choice of bins at every treatment site. Performing this procedure separately by treated site ensures that the comparisons are between housing units with the same treatment timing in the same neighborhood. Comparing within treatment timing group avoids issues with negative weighting in staggered adoption designs well documented in the literature. Comparing within the same neighborhood lends more credibility to the parallel trends assumption. The output of this binning process across all sites is a large number of price effect estimates at various distances $d$ from the plants. In a final step, I regress these price effect estimates on $d$ to trace out a continuous $ATT$ function.

[16]Using quantile bins, which is standard for bin scatter plots and suggested in Butts (2023), leads to similar results. However, because the density of housing sales increases moving away from the treatment site the number of bins close to $d = 0$ is often sparse using quantile bins. For each type of spatial treatment, I also present estimates using quantile bins in Appendix A1.

[17]Software for this method is available at `https://nppackages.github.io/binsreg/`.

Figure 7: Binning example




*Note: The top graph plots log house price averages by distance to a single industrial plant located in the Providence, RI metro area that first appeared in the TRI data in 1996. The blue dots show the averages by bin from 1991 to 1995 while the orange dots show the averages from 1996 to 2001. In the bottom graph, the blue triangles show the comparisons at each bin generated by taking the difference between the orange and the blue dots and then subtracting the analogous difference in the bin farthest away from the plant (rightmost bin). This procedure yields a distinct difference-in-differences estimate for each distance bin.*

*Price effects for all plants*

Figure 8 shows price effect estimates as a smooth function of distance to the toxic plant site using comparisons from all 11,849 plants. The top graph compares distance bin average house prices in the 0 to 5 years after a plant opens relative to -5 to -1 years before. I call these estimates the "short run" price effects. The bottom graph shows the comparisons 6 to 10 years after a plant opens again relative to -5 to -1 years before, or the "long run" price effects.[18] In both plots, the red line is the fit line from a nonparametric sieve regression of the 303,260 DiD bin comparisons on distance to the plant site. This regression uses the data-driven choice of sieve dimension of Chen et al. (2024) and weights by the DiD estimates' inverse variances.[19,20] The shaded red areas are 95% uniform confidence bands. As a second check, blue dots are bin scatter plots of the 303,260 comparisons following Cattaneo et al. (2024), in which the data is now binned for a second time for visualization.

In both the short run and long run estimates, an industrial plant opening causes house prices to decline. Furthermore, plants' negative externalities affect houses closer to the plant more. In the "short run", which considers house prices 0 to 5 years after the opening of a plant, houses within a kilometer of the plant site lose 1% of house value with negative estimated effects smaller in magnitude extending out to 3–3.5 kilometers. In the "long run", which considers house prices 6 to 10 years after the opening of a plant, the relationship becomes more pronounced. Houses within a kilometer of the plant site lose 2-3% of house value with estimated effects extending out to around 4 kilometers.

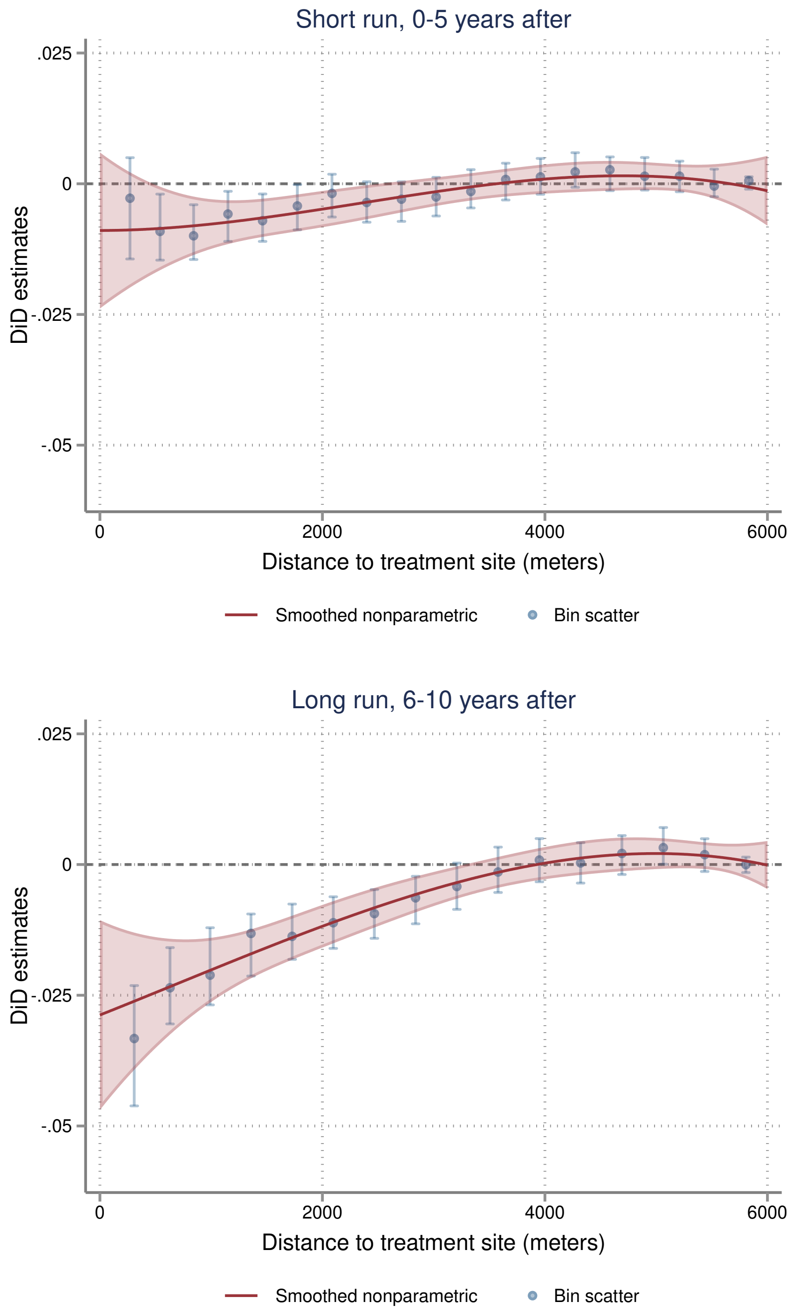
In this paper, I interpret the long run price effect estimate or the bottom plot in Figure 8 as the main price effect estimate. Hence, I use the red line as the key input to characterize aggregate local welfare and price impacts according to Equations (8) and (9).

[18]Pooling the observations into three time periods creates some disconnect from a standard DiD regression which includes year fixed effects. However, the procedure can be adapted to include treatment site×sale year fixed effects by absorbing year fixed effects in the estimation of the bin level averages. This is similar to the aggregation correction suggested in Section IV.C of Bertrand et al. (2004). Appendix A8 shows price effects residualizing treatment site×sale year fixed effects. These estimates are very similar to the estimates without year fixed effects.

[19]Appendix A1 shows binscatter plots of the long run price effects using alternative binning and weighting choices. In particular, A1 compares bin scatter plots with equal weighting of the ATT estimates, equal weighting but dropping ATT estimates based on small numbers of observations, weighting by the minimum number of observations in a bin, and weighting using inverse variance weights under both equal and quantile spaced binning strategies. In all cases, I find similar price effects for TRI industrial plants though equal spaced binning without weighting is less precise. While quantile binning is standard for binscatter plots and yields precise estimates far away from the site, estimation close to the site is poor because the number of bins becomes sparse as there are far fewer housing sales close in than far away.

[20]Software for this method is available at `https://rdrr.io/github/JeffreyRacine/npiv/man/npiv-package.html`.

Figure 8: Price effects as a function of distance to the toxic plant

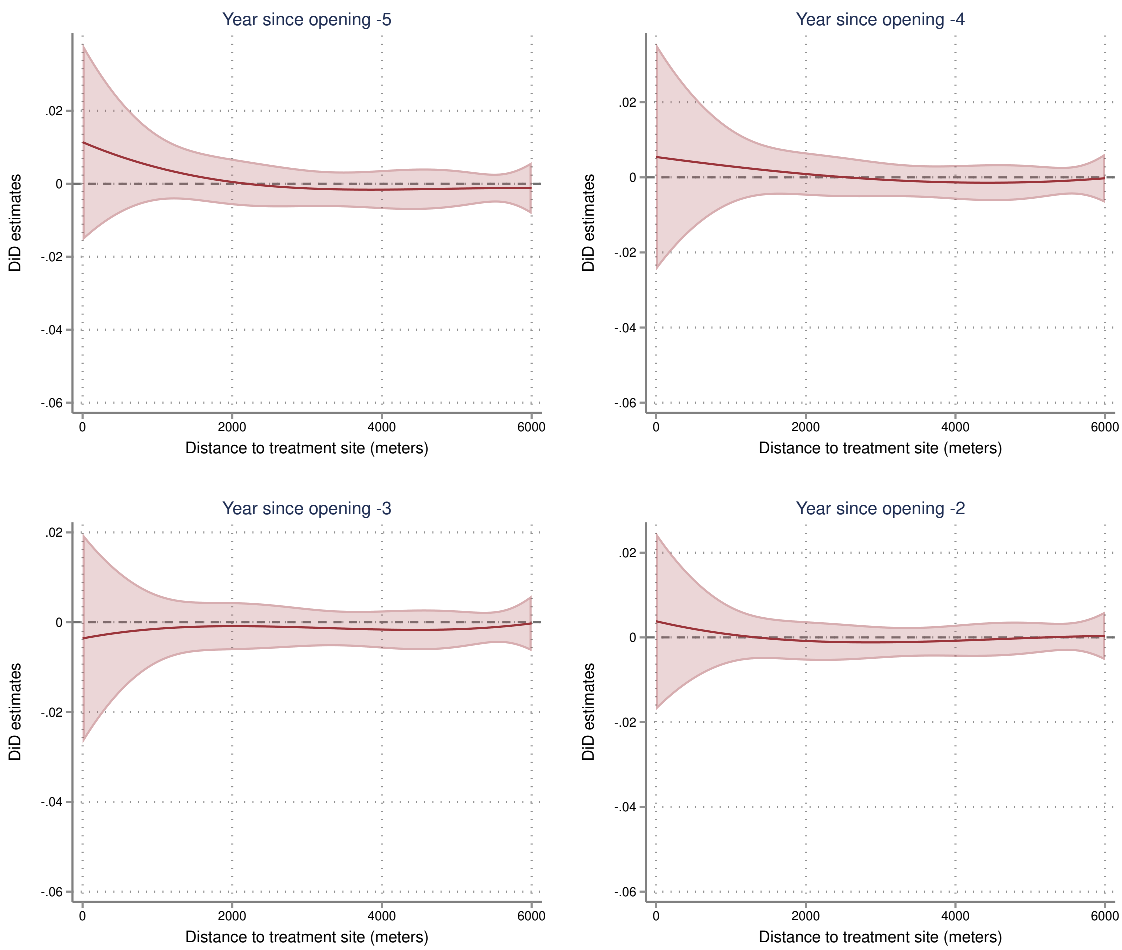



*Note: This figure shows price effect estimates as a smooth function of distance to the toxic plant site. The top graph compares averages prices in the distance bins 0 to 5 years after a plant opens relative to the prices in the same distance bin -5 to -1 years before. The bottom graph shows the analogous comparisons 6 to 10 years after a plant opening relative to -5 to -1 years before. In both plots, the red line is the fit line from a nonparametric sieve regression of 303,260 DiD bin comparisons surrounding 11,849 plants on distance to the plant site following Chen et al. (2024) and weighting by the DiD estimates' inverse variances. The shaded red areas are 95% uniform confidence bands. Blue dots are bin scatter plots of the 303,260 comparisons (binning the data a second time) with confidence intervals clustered by plant site following Cattaneo et al. (2024). For the bin estimates, the intervals are not centered exactly on each bin because they are robust bias-corrected intervals.*

*Event studies*

Before proceeding to the welfare calculation, I first present event study plots. The key assumption of the difference-in-differences strategy is the parallel trends assumption. That is, in absence of treatment the price trend for houses close to the site would have been the same as the price trend for the houses slightly farther away but unaffected by the toxic plant. However, with a continuous treatment, the price effect is now estimated along a continuum of different levels, thereby implying a continuum of possible pre trend checks. Below, I present a generalization of the event study to the continuous setting.

Figure 9: Continuous DiD event study




*Note: This figure is a pre trends check by performing a similar analysis as in Figure 8 but using data prior to the treatment date. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in Figure 8. In Appendix Figure A9 I show the analogous plots for all years in the post period.*

To check for pre trends, Figure 9 performs a similar analysis as in Figure 8 but using data prior to the treatment date. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in Figure 8. This pretrend check does not show evidence of significant parallel trends violations. If the treated and control units were on parallel trends in the pre period, we would see no relationship between price and distance in the pre period years. In years -4, -3, and -2, the estimated relationship is essentially a flat line. Year -5 shows a moderate decreasing gradient, but given the confidence band a null effect can not be ruled out. Additionally, in each pre period year, the estimate is small in magnitude relative to the price effect observed in the post period. Appendix Figure A9 shows the analogous plots for all years -5 to +10, which shows negative price effects in the post periods similar in magnitude to the price effects in Figure 8.

I report two other event study plots in the appendix as additional checks. In Appendix A13, I show an event study plot that pools together the $ATT$ estimates into two distance groups to reduce the number of partitions following Callaway et al. (2024b). In Appendix A14, I show the results from a standard binary stacked DiD event study regression. These figures also do not show evidence of parallel trends violations with similar negative price effects to Figure 8.

*Aggregate price and welfare impact estimates*

Recalling Equation (8) from Section 4,

$$\text{Agg WTP} = N_H \int_0^{\overline{d}} (1+d) A\tilde{CRT}(d|d) \mathrm{E}[Y(d, \mathbf{z_{-d}})|D=d] \log(\frac{1+\overline{d}}{1+d}) f(d)\, \mathrm{d}d.$$

Each term is as defined before:

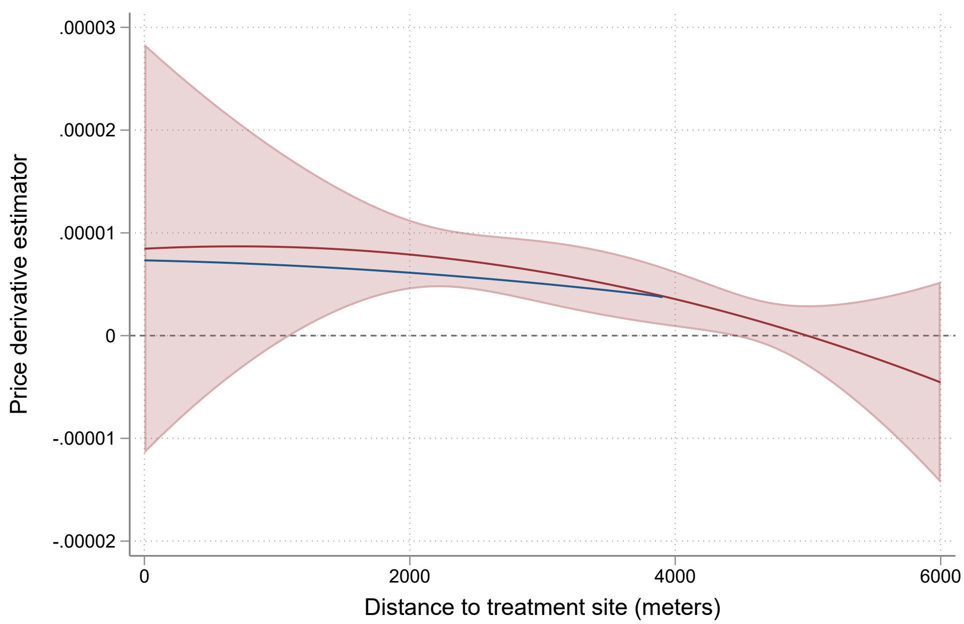
| | | |
|---|---|---|
| $N_H$ | – | Total number of owner occupied housing units affected |
| $\overline{d}$ | – | Spatial extent of price effects |
| $A\tilde{CRT}(d\|d)$ | – | Log price function derivative |
| $\mathrm{E}[Y(d, \mathbf{z_{-d}})\|D=d]$ | – | Average house price at distance $d$ |
| $f(d)$ | – | Density of owner occupied housing units at distance $d$. |

All terms can be estimated from the data. $A\tilde{CRT}(d|d)$ is the primary estimand, which I estimate based on the red line in bottom graph of Figure 8. Depending on the assumption imposed, I use either the red line's derivative $\frac{\partial A\tilde{T}T(d|d)}{\partial d}$ or its rise over run $\frac{A\tilde{T}T(d|d)}{d-\overline{d}}$. Because I measure price changes in logs rather than levels, I multiply by the average house price at each distance $d$, denoted $\mathrm{E}[Y(d, \mathbf{z_{-d}})|D=d]$. To estimate $\mathrm{E}[Y(d, \mathbf{z_{-d}})|D=d]$, I regress post period housing prices on distance to the treated site using the same nonparametric sieve regression procedure (Chen et al., 2024). For the other terms, I set $\hat{\overline{d}} = 3{,}928$ meters, the point where the long run price effect first crosses 0 in Figure 8. Within the $\hat{\overline{d}}$ meters of a plant, on average there are $\hat{N}_H = 6{,}939$ owner occupied housing units in nearby U.S. Census blocks affected per site. Finally, I estimate $\hat{f}(d)$ by standard kernel density estimation using owner occupied housing unit counts in nearby census blocks.

Figure 10 shows estimates for each of the terms used in the aggregate welfare calculation. Panel (a) shows estimates for the price derivative. The red line is the $A\tilde{T}T(d|d)$'s derivative with uniform confidence bands constructed following Chen et al. (2024). A nice feature of this estimation method is that the data driven choice of sieve dimension for the $ATT$ is the same for its derivative inheriting the desirable statistical properties outlined in Chen et al. (2024). The derivative estimate is always positive over the affected area suggesting that homeowners are willing to pay to be farther away from a toxic industrial plant. Assuming homogeneous treatment effects, this estimate would imply a homeowner near the treatment site's MWTP is $\sim$1% of their house's value to be 1 kilometer farther away from a plant (0.00001 $\times$ 1,000). The blue line shows the price derivative estimate using the rise over the run, which is a lower bound on the MWTP even in the presence of treatment effect heterogeneity if a plant's externalities decay exponentially with distance.[21] In this application, the rise over the run is similar to the $ATT$ derivative though modestly smaller in magnitude.

Figure 10: TRI, components of aggregate welfare calculation

(a) Price derivative estimates $\frac{\partial \tilde{ATT}}{\partial d}(d|d)$ and $\frac{\tilde{ATT}(d|d)}{d-\hat{\bar{d}}}$



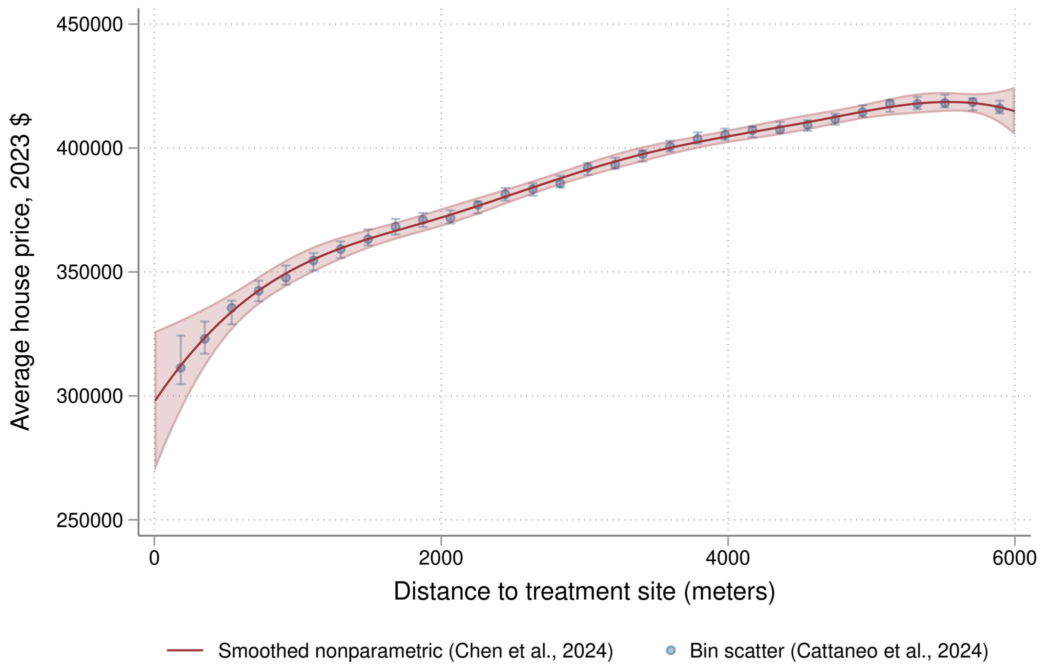

Panel (b) estimates the relationship between the average house price and distance to a plant site in the post period. I regress average house prices in the post period on distance to the treatment site and plant-specific fixed effects.[22] As expected, house prices are significantly lower near the site of a plant, likely reflecting in large part the impact of other factors on property values other

[21]This estimate is trimmed close to $\bar{d}$. As described in Graham and Powell (2012), an average partial effect using the rise over the run is only irregularly identified. This is because for $d$ close to $\bar{d}$ the denominator, or the run, is close to 0. To deal with small denominator effects, I follow Graham and Powell (2012) and trim the distance range close to $\bar{d}$ by following the rule of thumb bandwidth in Graham and Powell (2012), which in this application is only 24 meters.

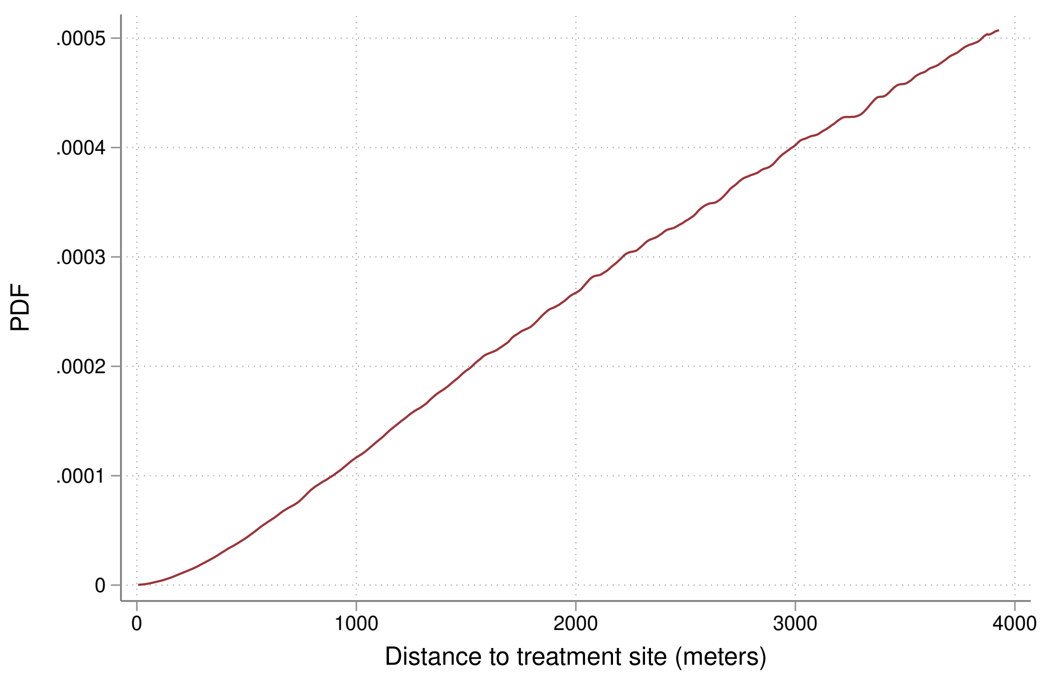
[22]To control for plant fixed effects, I follow the procedure in Chen et al. (2024)'s application. Specifically, I regress housing prices on distance to the treatment site and plant-specific fixed effects under an initial basis choice. I then apply Chen et al. (2024) as before but using house prices minus the estimated plant fixed effects from the initial regression as the outcome variable. As an additional check, the blue scatter plot also shows the estimates from `binsreg` with covariate adjustment following Cattaneo et al. (2024).

Figure 10: continued

(b) Average house price level $\widehat{\mathrm{E}[Y|D=d]}$




(c) Owner occupied housing unit density $\hat{f}(d)$




*Note: This figure plots estimates of each term used in the Aggregate Willingness to Pay Equation (8) for TRI industrial plants as a function of distance to the treatment site. Panel (a) shows the price derivative estimates. Panel (b) shows the average house price in the post period. Panel (c) plots the PDF of owner occupied housing units.*

than the plant itself. House prices are $\sim$ \$300,000 close to the plant and increase to $\sim$ \$420,000 at 6 kilometers away. Finally, Panel (c) shows the empirical density of owner occupied housing units on distance to the plant site. This plot confirms the intuition from the summary statistics that there are many more housing units far away from the plant's site than located close in.

After plugging in each term, Equation (8) is a 1-dimensional numerical integration problem that I solve using standard methods. To start, as a baseline I evaluate Equation (8) using $\frac{\partial A\tilde{T}T(d|d)}{\partial d}$ as the estimator for homeowners' MWTPs. As discussed in Sections 3 and 4, this assumes homogeneous treatment effects. I calculate homeowners would be willing to pay \$16,411,924 for the plant not to exist. Strengthening assumptions further to a time constant hedonic price function as in Diamond and McQuade (2019), this estimate says a plant opening causes an aggregate welfare loss of \$16,411,924 dollars to nearby homeowners.

However, $\frac{\partial A\tilde{T}T(d|d)}{\partial d}$ does not identify the price derivative under unobserved treatment effect heterogeneity. Sections 3 and 4 discussed an alternative restriction. In particular, if treatment effects are concave in distance from the plant, the average rise over the common run is a conservative lower bound on the price derivative. Evaluating Equation (8) instead using the rise over the run estimator, I calculate homeowners would be willing to pay \$13,086,846 dollars for the plant not to exist. As anticipated, the rise over the run estimate appears to be conservative.

In addition, I calculate the aggregate price loss to homeowners to bound the other side of the welfare measure. By Equation (9),

$$\Delta\text{Price} = N_H \int_0^{\bar{d}} A\tilde{T}T(d|d)\mathrm{E}[Y(d, \mathbf{z}_{-\mathbf{d}})|D = d]f(d)\,\mathrm{d}d.$$

As shown in Section 4, the negative price change is an upper bound on homeowners' willingness to pay for the plant not to exist. In other words, we expect the aggregate price loss to overestimate the magnitude of a toxic plant's welfare impact. Using the long run price effect estimate from Figure 8 as $A\tilde{T}T(d|d)$ and estimating $\hat{N}_h$, $\mathrm{E}[\widehat{Y|D = d}]$, and $\hat{f}(d)$ as above, I find a toxic plant causes an aggregate price loss of -\$19,459,895 dollars to local homeowners. Here, consistent with theory, the negative price loss provides an informative but non-conservative bound on the welfare measure.

Taken together, the welfare statistics bound homeowners' willingness to pay between \$13.1 and \$19.5 million dollars for a polluting industrial plant not to exist. The point estimate assuming homogeneous treatment effects falls between these two bounds at \$16.4 million dollars.

## 7 Additional applications: LIHTC developments and sex offenders

In this section, I use the econometric procedure to evaluate two more key applications in the hedonic literature: low income housing developments and sex offender move ins. First, I briefly review the data sources I use to characterize treatment events for each application. Both applications have significantly more treatment sites than the original studies. Then, I present price effect estimates for low income housing developments. For sex offender move ins, I find evidence of parallel trends violations. Finally, I compare price and welfare impacts between TRI industrial plants and LIHTC developments on a consistent basis.

## A *Background*

*LIHTC developments*

The Low Income Housing Tax Credit is a large government program that incentivizes construction of affordable housing for low income renters. To date, it has funded approximately 1/5 of the multifamily developments in the U.S. and houses 2% of the nation's population. See, e.g. Diamond and McQuade (2019) and Soltas (2024), for a more detailed background on the LIHTC program and its funding allocation process. Briefly, developers receive tax credits for new construction of multifamily units if those units satisfy certain criteria for low income housing. In particular, a certain percentage of the building's tenants must have below the area median gross income (AMGI). Developers usually sell these tax credits to investors to fund the development and those investors receive the credits over a ten year period.

The U.S. Department of Housing maintains a database on LIHTC developments.[23] The LIHTC data provide information on 53,032 developments that have been funded since 1986. I restrict to developments with nonmissing geographic information and with funding allocated from 1995 to 2015. In addition, I can only study sites where I observe nearby housing sales before and after funding allocation. After these restrictions, my final estimation sample has 17,895 LIHTC sites.

*Sex offender move ins*

Linden and Rockoff (2008) and Pope (2008)'s studies on the effect of sex offender move ins on house prices were some of the ring DiD design's earliest applications. The former study uses data from Mecklenburg County in North Carolina and the latter uses data from Hillsborough County in Florida. Both of these U.S. states are well suited to this analysis in that state law requires sex offenders to register their address to the local sheriff within 10 day of their move in.

I collected the current sex offender registries for the wholes of North Carolina and Florida,[24] which contain all sex offenders' addresses and their move in dates in these states. It is important to note that the current registries reflect only the most recent address change of each sex offender and not a complete history of moves. As a result, my sample is skewed toward sex offenders that have lived in the same location for longer periods of time, whereas the typical sex offender moves frequently. After dropping sex offenders that are in prison, deceased, or no longer in state and restricting to move ins prior to 2015 with nearby housing sales, I am left with 9,193 sex offender move ins in my estimation sample with 1,701 coming from North Carolina and 7,492 from Florida.

Appendix Tables A1 and A2 presents summary statistics on house sales and housing units close to LIHTC developments and sex offender move ins analogous to Table 1.

## B *Price effects by application*

*LIHTC developments*

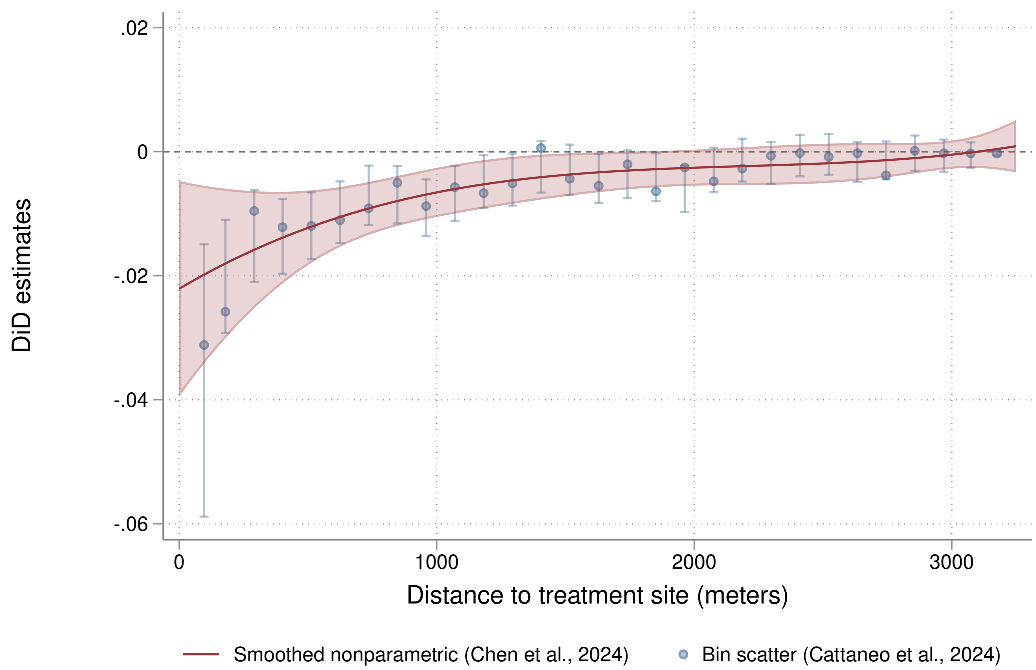
I now present price effect estimates for a LIHTC development following the same procedure as in Section 6. The red line in Figure 11 is the main price effect estimate from comparing average house prices in the 6 to 10 years after a LIHTC development's funding allocation relative to -5

[23] LIHTC data are publicly available at `https://www.huduser.gov/portal/datasets/lihtc/property.html`.

[24] North Carolina's registry is available at `https://sexoffender.ncsbi.gov/stats.aspx`. Florida's registry is available at `https://offender.fdle.state.fl.us/offender/sops/registryDownload.jsf`.

to -1 years before allocation. The red line is the fit line from a nonparametric sieve regression of 324,007 DiD bin comparisons on distance to a LIHTC development around 17,895 LIHTC sites. As before, this regression uses a data-driven choice of sieve dimension following Chen et al. (2024) and weights by the DiD estimates' inverse variances.[25] Blue dots are bin scatter plots following Cattaneo et al. (2024), which provides a second check of the smoothed estimate. I find a small but precise decline in house values of 2% directly next to a development with effects decaying to 0 around 2.5-3 kilometers.

Figure 11: LIHTC developments, long run price effect, all LIHTC sites




*Note: This figure shows the long run price effect estimate as a smooth function of distance to a LIHTC development. The red line is the fit line from a nonparametric sieve regression of 324,007 DiD bin comparisons from 17,895 LIHTC developments on distance to the LIHTC site following Chen et al. (2024) and weighting by the DiD estimates' inverse variances. The shaded red areas are the 95% uniform confidence bands. Blue dots are bin scatter plots of the 324,007 comparisons with confidence intervals clustered by treatment site following Cattaneo et al. (2024).*

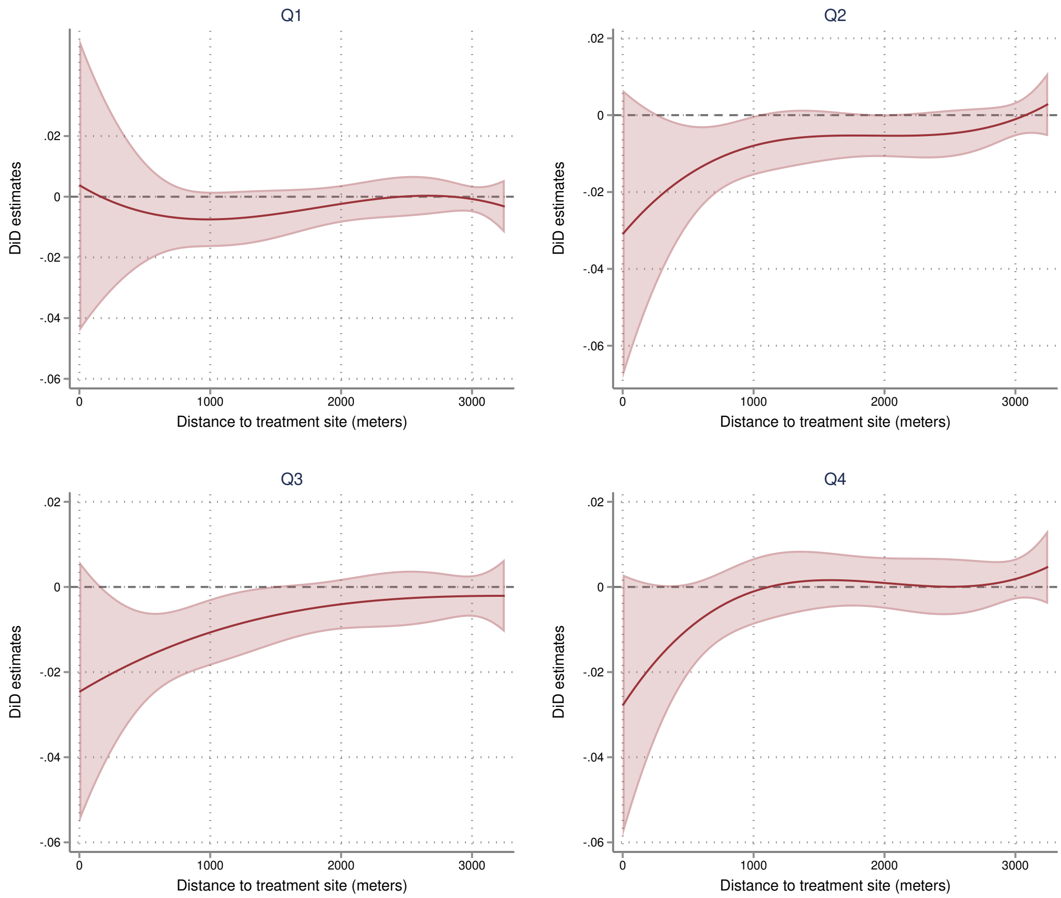
Appendix Figure A10 presents a continuous DiD event study for LIHTC developments analogous to the event study for TRI industrial plants (Figure 9). As in the first application, I do not find strong visual evidence that would suggest parallel trends violations. More discussion is provided in the appendix.

Figure 12 shows price effect estimates splitting the sample of treatment sites by the household income quartile of a LIHTC site's census block group. I find relatively limited evidence of heterogeneity by neighborhood household income. In the first income quartile, the estimate

[25] As with TRI industrial plants, I also present long run price effect estimates under an array of alternative weighting and binning schemes, and residualizing treatment site $\times$ sale year fixed effects. These plots are presented in Appendix A2 and Appendix A8.

looks like a null effect. In the second, third, and fourth income quartiles, I find a similar price effect estimate showing a moderate price decline of 2-3% directly next to a LIHTC development.

Figure 12: LIHTC developments, price effects by neighborhood household income quartile



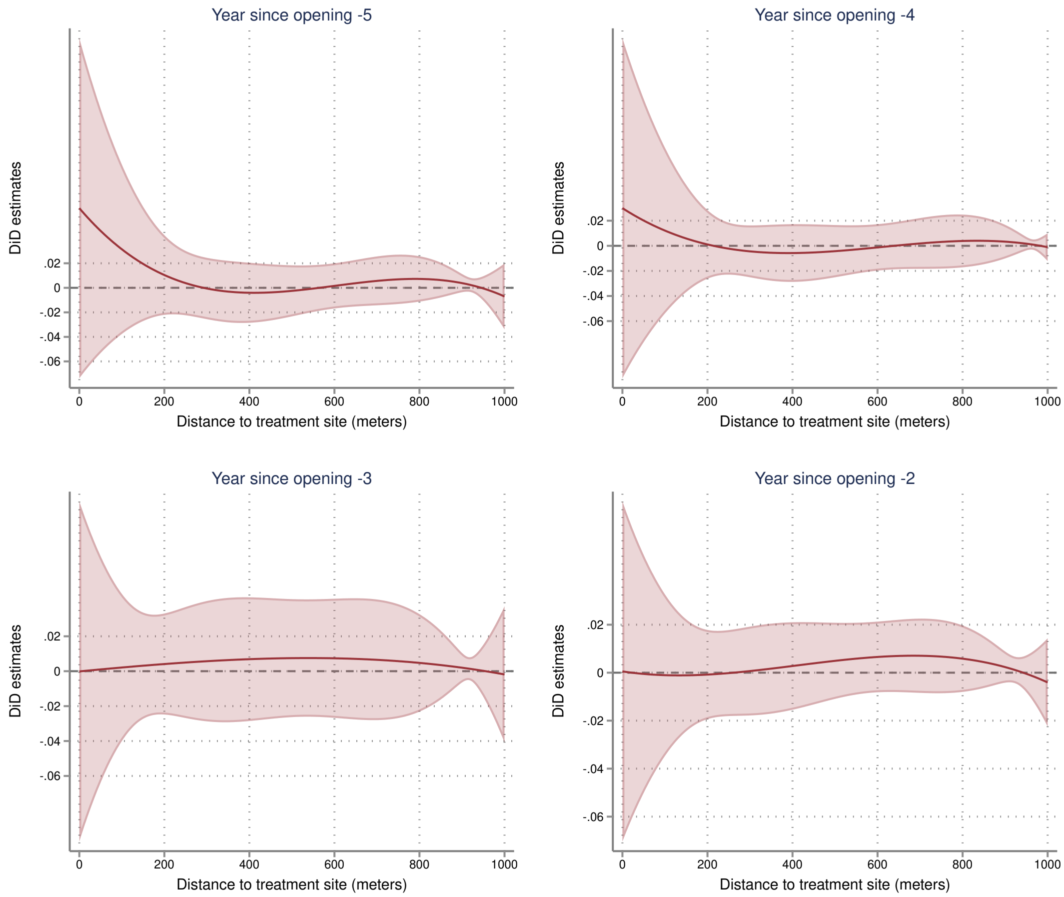

*Note: This figure performs a similar analysis as in Figure 11 but splitting the sample of treatment sites into quartiles by the median household income in a site's census block group in the 2008-2012 5-year American Community Survey (ACS). The income quartile cutoffs are $31,657, $45,627, and $64,509 in 2023 dollars.*

*Sex offender move ins have parallel trends violations*

In contrast to TRI plants and LIHTC developments, for sex offender move ins I find evidence of parallel trends violations. Figure 13 is the continuous event study plot for sex offender move ins. While the confidence bands imply there is a lot of uncertainty,[26] prices for housing units close to the sex offender appear to be trending downward even prior to the event date, suggesting sex offenders may be moving into parts of the neighborhood where houses were already depreciating in value.

[26]Precision can be improved by aggregating more comparisons together, e.g. by going back to a binary estimation of the treatment effect and calculating the weighted mean $ATT$ for all bins within 200 meters of the sex offender's address (roughly comparable to the 0.1 mile treated group in Linden and Rockoff (2008)). Evidence of parallel trends violations are also visible in these binary estimates.

Figure 13: Sex offender move ins, event study




*Note: This figure performs an event study similar to Figure 9 for sex offender move ins. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using nonparametric sieve regression.*

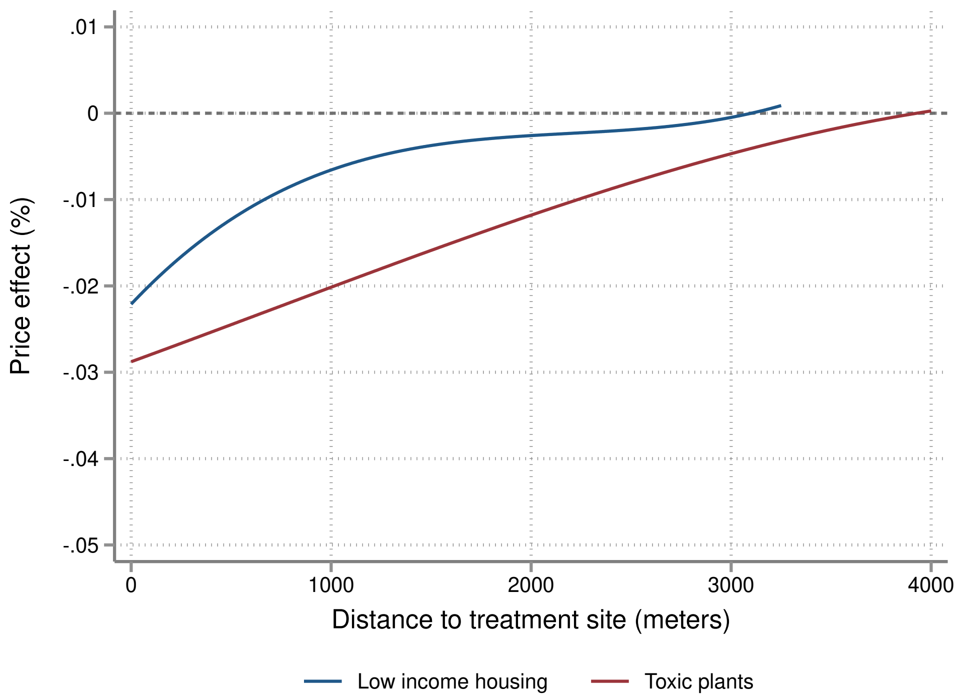
## C *Comparing price and welfare impacts across treatment types*

I now compare local price and welfare impacts for TRI industrial plants and LIHTC developments on the same scale. I do not report results for sex offenders in the main paper because the price effect estimates may be due to parallel trends violations.[27]

Figure 14 plots the long run price effect estimates on the same y and x axes. The blue line shows the price effect estimate for LIHTC developments and the red line shows the effect for TRI industrial plants. The price effect from a TRI industrial plant exceeds that of a LIHTC development, in magnitude and the affected geographic area's spatial scale.

[27]For completeness, Appendix Figure A18 and Appendix Table A4 also present a similar welfare analysis for sex offender move ins. However, caution should be taken when interpreting those results.

Figure 14: Long run price effects on the same scale




*Note: This figure shows the price effect estimates from Figures 8 and 11 on the same y and x axes. The blue line shows the price effect estimate for LIHTC developments and the red line shows the effect for TRI industrial plants.*

Table 2 presents estimates of the price and welfare impacts on local homeowners from TRI industrial plants and LIHTC developments. The first column reports the estimates for TRI plants and the second column for LIHTC developments. In each panel, the first two rows show heterogeneity-robust bounds on homeowners' willingnesses to pay to avoid the disamenity from the treatment site following the bounding argument in Section 4. Assuming treatment effect concavity, the first row is a lower bound on the willingness to pay from using the rise over the run estimator as the average MWTP. The second row shows the upper bound, which is the negative of the average price change. The third row in each panel compares the bounds to a benchmark that assumes homogeneous treatment effects (strong parallel trends) and uses the $ATT$ derivative as the average MWTP.

The first panel shows the average willingness to pay to avoid the disamenity from the treatment site as a percentage of house value. Following the argument of Theorem 4.2, the average amount a homeowner within the area affected by a toxic industrial plant would be willing to pay to not be exposed to the plant is bounded between 0.5% and 0.76% of their house's value. The analogous figures for LIHTC developments is 0.23% to 0.36%. Assuming homogeneous treatment effects, the estimated average willingness to pay to avoid the treatment site is 0.63% of house value for TRI plants and 0.32% for LIHTC developments.

The second panel is similar but now estimated in dollars by multiplying by average house prices. As noted in Section 4, converting from logs to levels in this way requires Assumption 10. As in the first panel, the first two rows in this panel are the bounds where the first row uses the rise over the run estimator for the price derivative and the second row is the average price change. These correspond to the integrals evaluated in the Aggregate WTP Equation (8) and the Aggregate Price Change Equation (9). The third row in this panel compares to using the $ATT$ derivative as the estimator for homeowner MWTP. The average amount a homeowner within an area affected by a toxic industrial plant would be willing to pay to not be exposed to the plant is bounded between \$1,886 and \$2,804. The analogous figures for LIHTC developments is \$815 to \$1,277. Again, the estimates assuming homogeneous treatment effects lie within these bounds at \$2,365 and \$1,113 for TRI and LIHTC sites, respectively. The final row in this panel shows the estimated number of affected homeowners or $\hat{N}_H$ in equations (8) and (9).

The third panel of the table shows the aggregate impacts per site attained by multiplying the average WTP estimates and price changes in dollars by the number of homeowners, $\hat{N}_H$. Scaling up the figures by the number of homeowners, implies that the aggregate homeowner WTP per site is 13.1-19.5 million dollars for TRI industrial plants and 6.7-10.5 million dollars for LIHTC developments. As suggested by the price effect estimates in Figure 14, the welfare loss per site from a TRI industrial plant's externality is significantly larger than that of a LIHTC development. The last row of the third panel has an estimate of the number of treatment sites currently operating in the United States. There are 15,899 TRI sites and 43,466 LIHTC sites in the United States.[28]

The final panel shows back of the envelope estimates of the total externalities nationally based on multiplying the aggregate impacts per site by the estimates for the number of treatment sites currently operating in the United States. Because both types of treatment affect many homeowners and there are numerous sites across the U.S., at the national level these effects appear to be important. The total national welfare externalities are at least \$208 billion dollars for TRI industrial plants and \$291 billion dollars for LIHTC developments. While a LIHTC development has about half the aggregate price and welfare impact per site compared to a TRI industrial plant, there are more than double the number of LIHTC sites across the nation leading to a larger total externality among local homeowners.

I conclude discussion of this table with two remarks about the estimates assuming homogeneous treatment effects. First, for these two applications, the homogeneous treatment effect estimates are within the heterogeneity-robust bounds, suggesting there may not be a huge bias from assuming homogeneous effects. However, if the degree of selection bias is large, there is no guarantee that in other applications an estimate using the $ATT$ derivative will lie within the bounds. Second, imposing stronger assumptions so that the price function is time constant with homogeneous treatment effects as in Diamond and McQuade (2019) leads to stronger welfare statements. In this case, the third row in each panel is a point identified welfare loss among homeowners. For example, the third row of the third panel implies a polluting industrial plant causes a welfare loss of 16.4 million dollars under these stronger assumptions.

[28]The estimate for the number of TRI sites is the number of plants in the TRI database reporting positive pollution emissions in 2022. The estimate for the number of LIHTC sites is the number of developments in the HUD database still being monitored as of April 2024.

Table 2: Total local price and welfare impacts on homeowners

| | Toxic industrial plants | Low income housing |
|---|---|---|
| **LB:** Avg. WTP per homeowner (%), $\frac{A\tilde{T}T}{d-\bar{d}}$ | 0.50% | 0.23% |
| | (0.091) | (0.074) |
| **UB:** Avg. $-\Delta$Price per homeowner (%) | 0.76% | 0.36% |
| | (0.124) | (0.103) |
| **SPT:** Avg. WTP per homeowner (%), $\frac{\partial A\tilde{T}T}{\partial d}$ | 0.63% | 0.32% |
| | (0.095) | (0.066) |
| **LB:** Avg. WTP per homeowner ($), $\frac{ATT}{d-\bar{d}}$ | $1,886 | $815 |
| | (296) | (273) |
| **UB:** Avg. $-\Delta$Price per homeowner ($) | $2,804 | $1,277 |
| | (398) | (373) |
| **SPT:** Avg. WTP per homeowner ($), $\frac{\partial ATT}{\partial d}$ | $2,365 | $1,113 |
| | (313) | (229) |
| # of affected homeowners per site | 6,939 | 8,219 |
| **LB:** Agg. WTP per site, $\frac{ATT}{d-\bar{d}}$ | $13.1 mil. | $6.7 mil. |
| | (3.84) | (2.92) |
| **UB:** Agg. $-\Delta$Price per site | $19.5 mil. | $10.5 mil. |
| | (6.18) | (4.11) |
| **SPT:** Agg. WTP per site, $\frac{\partial ATT}{\partial d}$ | $16.4 mil. | $9.1 mil. |
| | (7.53) | (2.46) |
| # of sites in the U.S. | 15,899 | 43,466 |
| **LB:** Total WTP, $\frac{ATT}{d-\bar{d}}$ | $208 bil. | $291 bil. |
| **UB:** Total -$\Delta$Price | $309 bil. | $456 bil. |
| **SPT:** Total WTP, $\frac{\partial ATT}{\partial d}$ | $261 bil. | $398 bil. |

*Note: This table presents estimates of the welfare and price impacts on local homeowners from TRI industrial plants and LIHTC developments. The first column has estimates for TRI industrial plants and the second column has estimates for LIHTC developments. The first two rows in each panel show heterogeneity-robust bounds. The first row shows the lower bound on the avg. willingness to pay among homeowners as a percentage of house value for a treatment site to not exist from using the rise over the run as the avg. MWTP. The second row shows the upper bound, which is the negative of the avg. price change as a percentage of house value. The third row compares to an estimate that assumes homogeneous treatment effects (strong parallel trends) and uses the ATT derivative as the avg. MWTP. Rows 4-6 convert to monetary estimates by multiplying by the avg. house price at each distance. In the next panel, aggregate WTP and price changes are calculated by multiplying the avg. WTP and price change by the # of affected homeowners. The final three rows are back of the envelope estimates of the total externalities nationally based on multiplying the aggregate impacts by an estimate for the number of treatment sites currently in the U.S. (in the row directly above). Standard errors were calculated by bootstrap resampling treatment sites for 1,000 replicates.*

### D *Discussion*

It is important to discuss the limitations on what is not captured by these local welfare impact estimates.

In particular, while these estimates are well suited as an estimate of the spillovers onto nearby homeowners, for both types of treatment event there is reason to believe the estimates mainly capture costs without consideration of the potential benefits. In the case of industrial plants, these plants may provide jobs to the larger geographic area, which might not show up in a localized pricing gradient, but could increase welfare overall.[29] For LIHTC developments, the obvious beneficiaries are low income tenants who now have greater access to affordable housing, but whose preferences will not be captured by the valuation of nearby owner occupied houses.

In addition, the structural model in Section 4 involves several assumptions. In this paper, I focused on relaxing one assumption—the strong parallel trends assumption to identify the hedonic price derivative. I then provided statistics to bound an aggregate welfare measure without this assumption. While these statistics appear to be informative, their derivation required other assumptions on separability of the hedonic pricing function, homeowners' utility functions, price effect concavity, and other assumptions in the standard hedonic model.

With these limitations in mind, these estimates may be valuable for local municipal policymakers who frequently have to weigh tradeoffs in decisions such as whether to site an industrial plant in their town. In this case, an industrial plant has both positive and negative externalities—while benefits like the projected number of jobs provided by a plant are often available to policymakers, the value of non-marketed neighborhood characteristics like pollution are harder to quantify. This paper fills this gap by connecting the pricing gradients around TRI industrial plants to a structural hedonic model.

For LIHTC developments, in contrast to the previous literature I did not find much evidence that low income housing development raises property values in poor neighborhoods (Baum-Snow and Marion, 2009, Diamond and McQuade, 2019). While this result warrants further study, it potentially has important implications to the debate over LIHTC's efficacy as a policy program. Soltas (2024) models developer behavior and estimates that LIHTC subsidies add few net units to the housing stock. On the other hand, low income households do benefit from these subsidies. However, if LIHTC development can neither be justified as place-based policy nor as increasing the housing stock, that provides greater evidence that comparable programs such as housing vouchers might provide similar benefits to low income tenants at less fiscal cost.

Finally, quantifying welfare impacts using this procedure facilitates comparison of values across different amenities: Local homeowners would be willing to pay at least 13.1 million dollars for a polluting industrial plant to not exist, while the analogous figure for LIHTC developments is 6.7 million dollars. Given the large number of TRI industrial plants and LIHTC developments in the U.S., these externalities are important. A back of the envelope calculation estimates national externalities of at least 201 billion dollars for TRI industrial plants and 291 billion dollars for LIHTC developments.[30] For context, in the most recent Interstate Highway Cost Estimate in

[29] As noted in Currie et al. (2015), a fact suggestive of the potentially large size of these benefits is that the typical investment in a TRI industrial plant can be in excess of 500 million dollars, dwarfing the 13.1-19.5 million dollar negative externality reported in Table 2.

[30] While due to evidence of parallel trends violations I do not report estimates for sex offender move ins in the main paper, the external effects from sex offenders could still be economically important. Even assuming a price impact of approximately -1.3% over just 200 meters, sex offenders still cause 190 billion dollars in externalities nationally. Despite small price impacts per treatment site, there are many more sex offenders in the U.S. than there are either TRI industrial plants or LIHTC developments.

1991, the total cost to build the entire U.S. interstate highway system was estimated at 129 billion dollars or 285 billion in 2023 dollars.[31] Hence, systematic evaluation of non-marketed housing characteristics, which encompasses many environmental and neighborhood amenities, may be useful for local, state, or national accounting.

## 8 Conclusion

A large literature in economics employs ring DiD designs to characterize the effect of neighborhood amenities on house prices. However, there is no consensus in the literature on how to estimate these effects and how they relate to welfare. In hedonics, the price derivative characterizes welfare. By adapting Callaway et al. (2024a)'s econometric framework for continuous DiD to this setting, I highlight an implicit assumption of treatment effect homogeneity needed for a DiD derivative to identify the price derivative. I then develop an approach that does not assume treatment effect homogeneity. Instead, I combine a shape restriction consistent with spatial decay and the hedonic model's equilibrium conditions to bound the impact. I illustrate this approach revisiting three key applications with more treatment sites and observations. Finally, I compare price and homeowner preference estimates across applications on a consistent basis.

I conclude by noting that this framework and estimating procedure can apply to many more types of spatial treatments, allowing for comparison of values across different amenities. For example, to evaluate policies such as EPA cleanup sites including Superfund and brownfield remediation sites or other EPA locations of interest[32] where concrete cost estimates might allow a complete cost-benefit analysis. Other applicable examples include power plants, transit stations, windmills, and so on.[33] In general, there is now a large literature of spatial DiD estimates with reduced form price effects, but these estimates often do not make a clear connection to an underlying economic model. In this paper, I propose a framework for spatial DiD that explicitly accommodates the continuous setting and relates DiD estimates to welfare.

[31] `https://www.fhwa.dot.gov/infrastructure/50estimate.cfm`

[32] For studies of EPA sites, see Greenstone and Gallagher (2008), Gamper-Rabindran and Timmins (2013), Haninger et al. (2017) and Marcus (2021), among others.

[33] See, e.g., Davis (2011), Gibbons and Machin (2005), and Dröes and Koster (2016).

## References


Arkhangelsky, D. and Imbens, G. (2024). Causal models for longitudinal and panel data: a survey. *The Econometrics Journal*, 27(3):C1–C61.

Bajari, P. and Benkard, C. L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6):1239–1276.

Banzhaf, S. (2021). Difference in differences hedonics. *Journal of Political Economy*, 129(8):2385–2414.

Baum-Snow, N. and Marion, J. (2009). The effects of low income housing tax credit developments on neighborhoods. *Journal of Public Economics*, 93(5–6):654–666.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust difference-in-difference estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs. *Working Paper*.

Butts, K. (2023). Difference-in-differences with geocoded microdata. *Journal of Urban Economics: Insights*, 133.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. (2024a). Difference-in-differences with a continuous treatment. *American Economic Review*, forthcoming.

Callaway, B., Goodman-Bacon, A., and Sant'Anna, P. H. C. (2024b). Event studies with a continuous treatment. *AEA Papers and Proceedings*, 114:601–605.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Cattaneo, M., Crump, R., Farrell, M., and Feng, Y. (2024). On binscatter. *American Economic Review*, 114(5):1488–1514.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics*, 18(1):5–46.

Chamberlain, G. (1992). Efficiency bounds for semiparametric regression. *Econometrica*, 60(3):567–596.

Chen, J. and Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Chen, X., Christensen, T., and Kankanala, S. (2024). Adaptive estimation and uniform confidence bands for nonparametric structural functions and elasticities. *The Review of Economic Studies*, forthcoming.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2):678–709.

Davis, L. W. (2011). The effect of power plants on local housing values and rents. *The Review of Economics and Statistics*, 93(4):1391–1402.

de Chaisemartin, C., D'Haultfoeuille, X., Pasquier, F., Sow, D., and Vazquez-Bare, G. (2025). Difference-in-differences for continuous treatments and instruments with stayers. *Manuscript*.

de Chaisemartin, C. and D'Haultfœuille, X. (2025). Treatment-effect estimation in complex designs under a parallel-trends assumption. *Manuscript*.

de Chaisemartin, C. and D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Diamond, R. and McQuade, T. (2019). Who wants affordable housing in their backyard? An equilibrium analysis of low-income property development. *Journal of Political Economy*, 127(3):1063–1117.

Dröes, M. I. and Koster, H. R. (2016). Renewable energy and negative externalities: The effect of wind turbines on house prices. *Journal of Urban Economics*, 96:121–141.

D'Haultfœuille, X., Hoderlein, S., and Sasaki, Y. (2023). Nonparametric difference-in-differences in repeated cross-sections with continuous treatments. *Journal of Econometrics*, 234(2):664–690.

Ekeland, I., Heckman, J., and Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of Political Economy*, 112(1):S60–S109.

Gamper-Rabindran, S. and Timmins, C. (2013). Does cleanup of hazardous waste sites raise housing values? evidence of spatially localized benefits. *Journal of Environmental Economics and Management*, 65(3):345–360.

Gibbons, S. and Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics*, 57:148–169.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Graham, B. S. and Powell, J. L. (2012). Identification and estimation of average partial effects in "irregular" correlated random coefficient panel data models. *Econometrica*, 80(5):2105–2152.

Greenstone, M. and Gallagher, J. (2008). Does hazardous waste matter? evidence from the housing market and the superfund program. *Quarterly Journal of Economics*, 123(3):951–1003.

Haninger, K., Ma, L., and Timmins, C. (2017). Estimating the impacts of brownfield remediation on housing property values. *Journal of the Association of Environmental and Resource Economists*, 4(1):197–241.

Linden, L. and Rockoff, J. (2008). Estimates of the impact of crime risk on property values from megan's laws. *American Economic Review*, 98(3):1103–1127.

Marcus, M. (2021). Going beneath the surface: Petroleum pollution, regulation, and health. *American Economic Journal: Applied Economics*, 13(1):1–37.

Pope, J. (2008). Fear of crime and housing prices: Household reactions to sex offender registries. *Journal of Urban Economics*, 64(3):601–614.

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

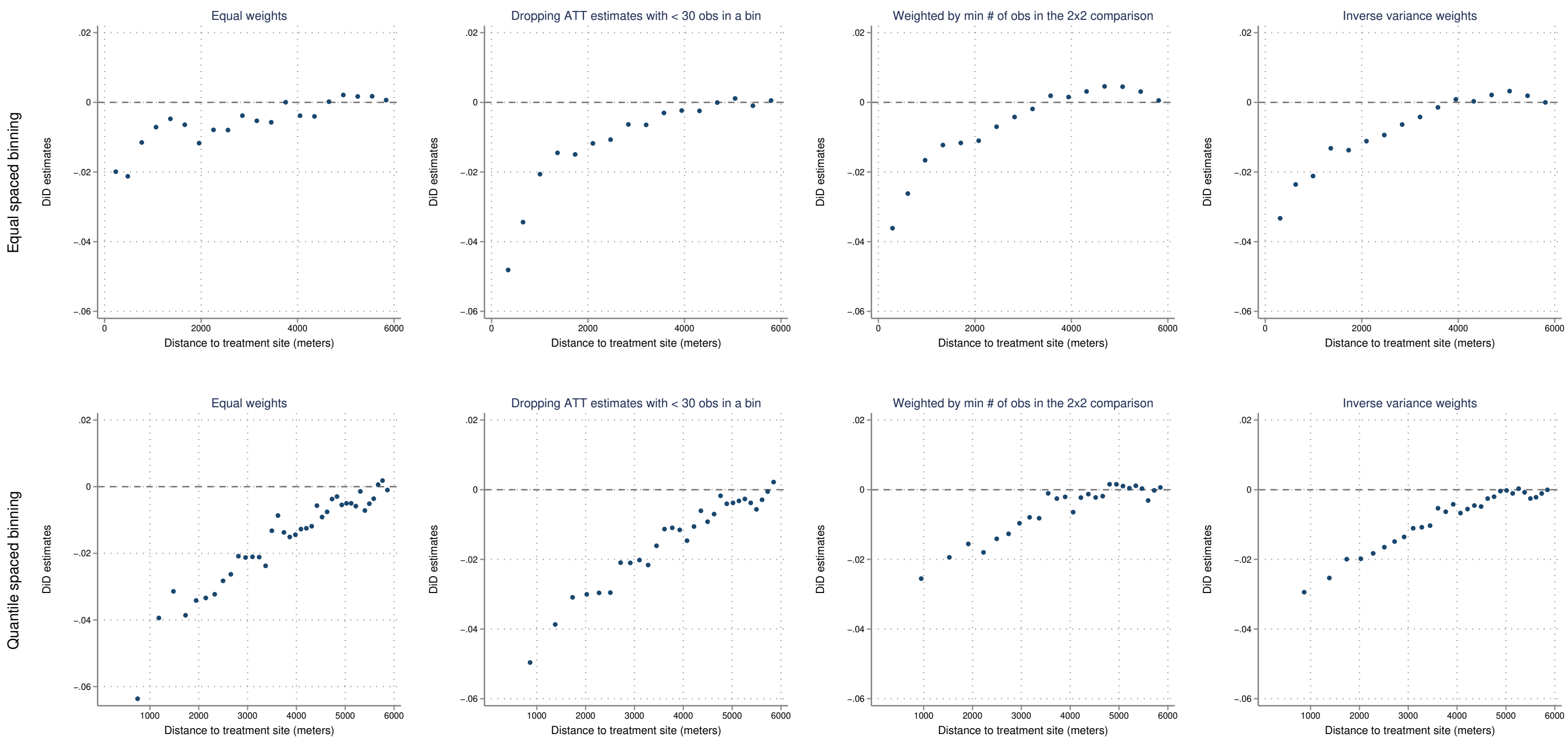
Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.

Soltas, E. (2024). Tax incentives and the supply of low-income housing. *manuscript*.

Wing, C., Freedman, S. M., and Hollingsworth, A. (2024). Stacked difference-in-differences. *NBER Working Paper*, (32054).

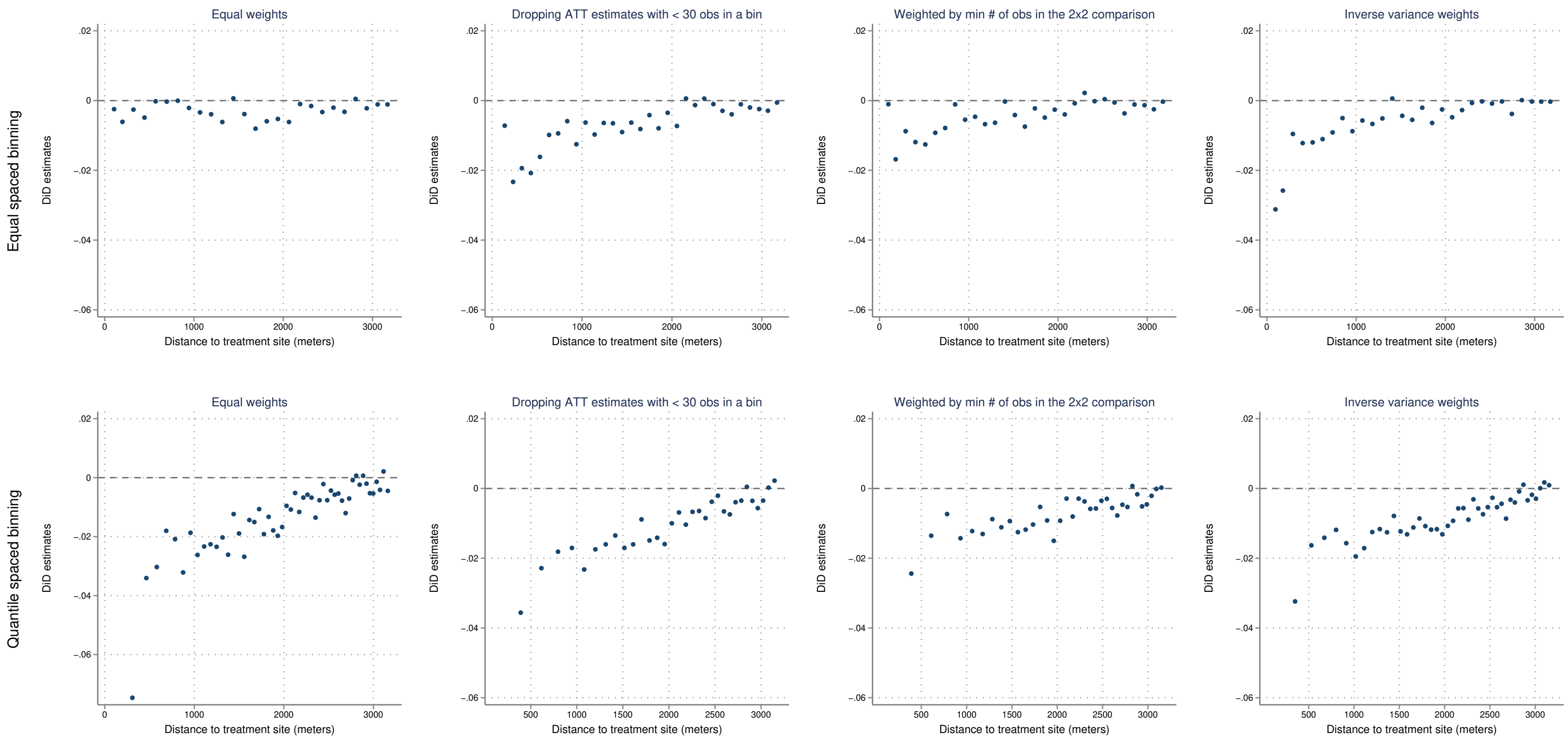
# Appendix A Appendix figures

Figure A1: TRI long run price effects under alternative binning and weighting choices



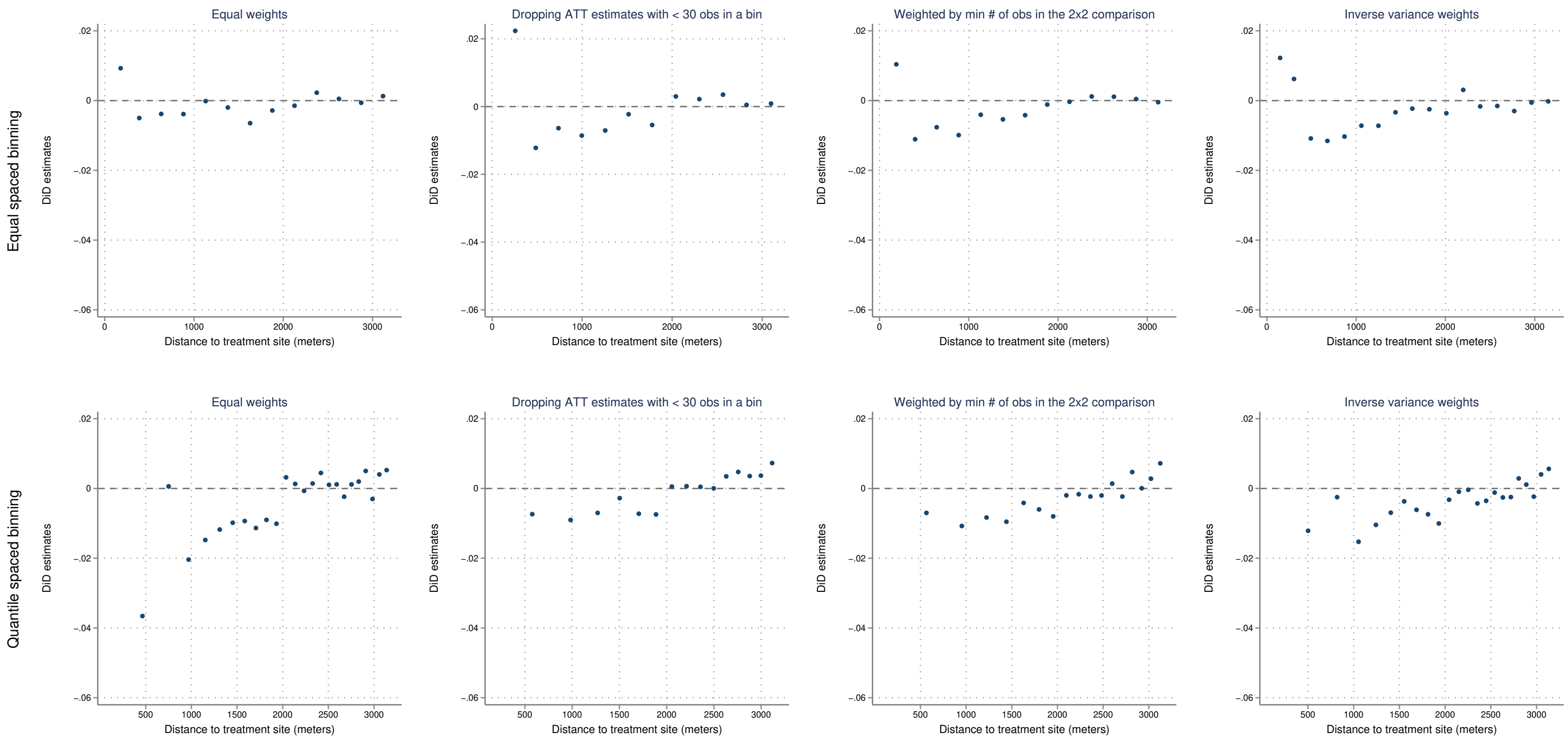

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A2: LIHTC long run price effects under alternative binning and weighting choices



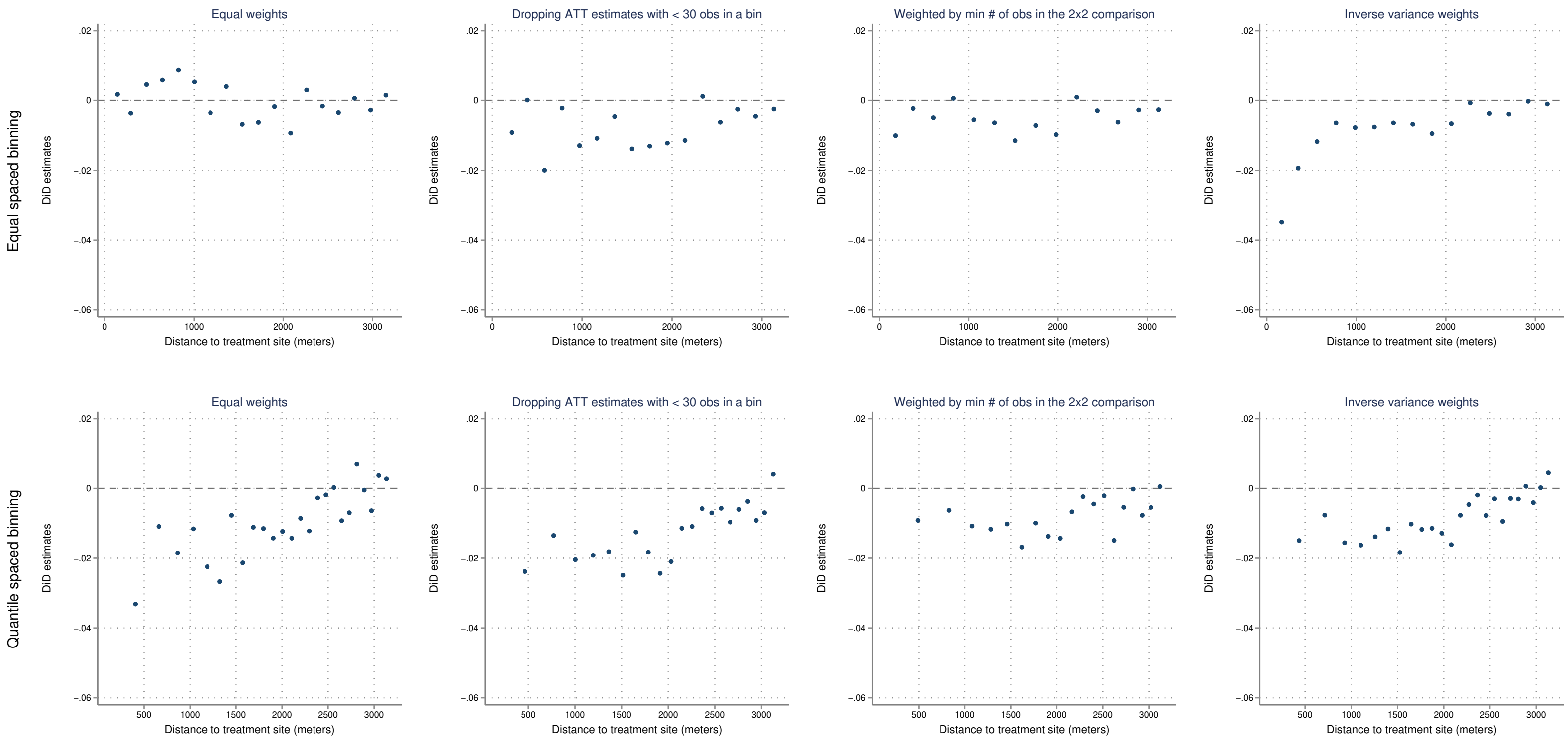

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A3: LIHTC Q1 long run price effects under alternative binning and weighting choices



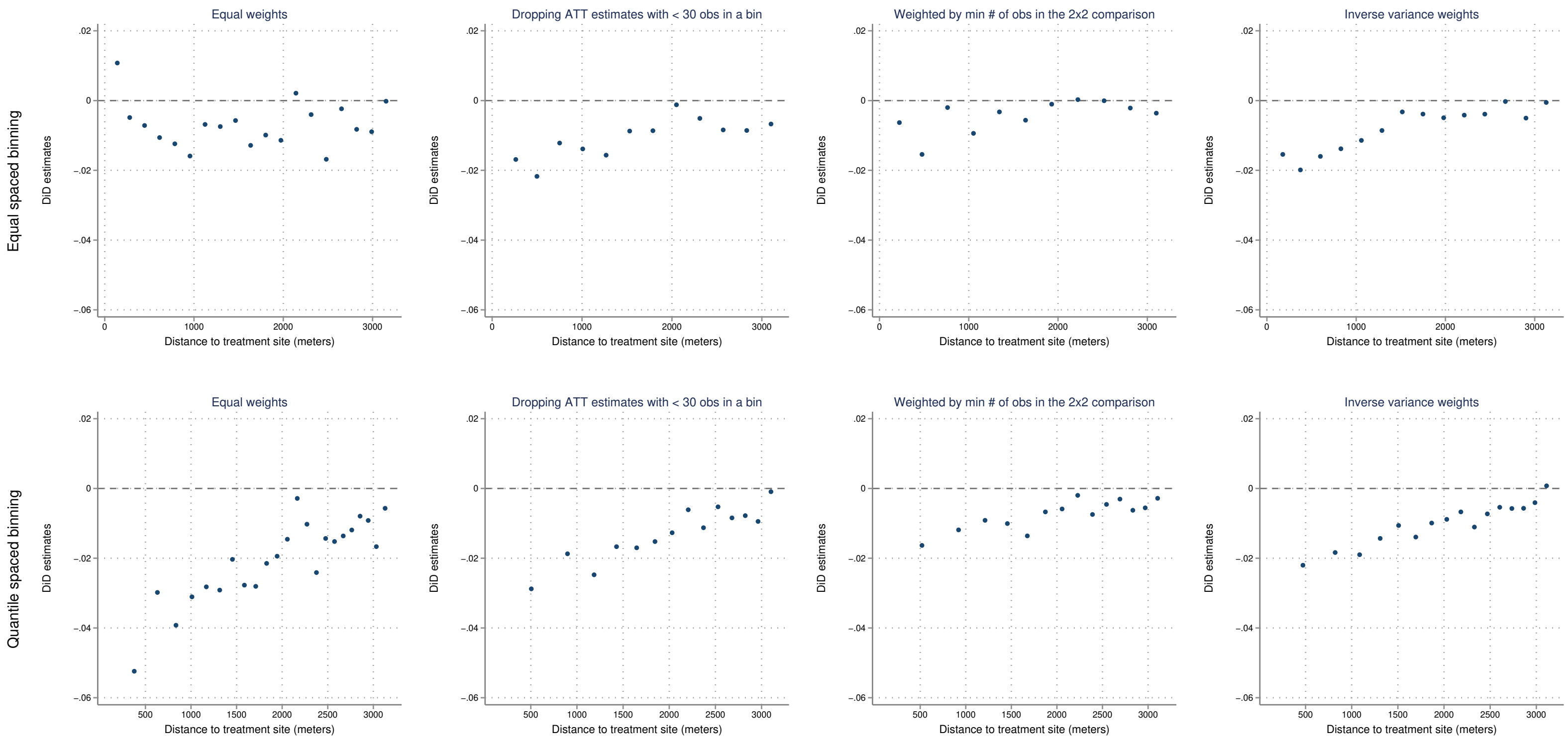

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A4: LIHTC Q2 long run price effects under alternative binning and weighting choices



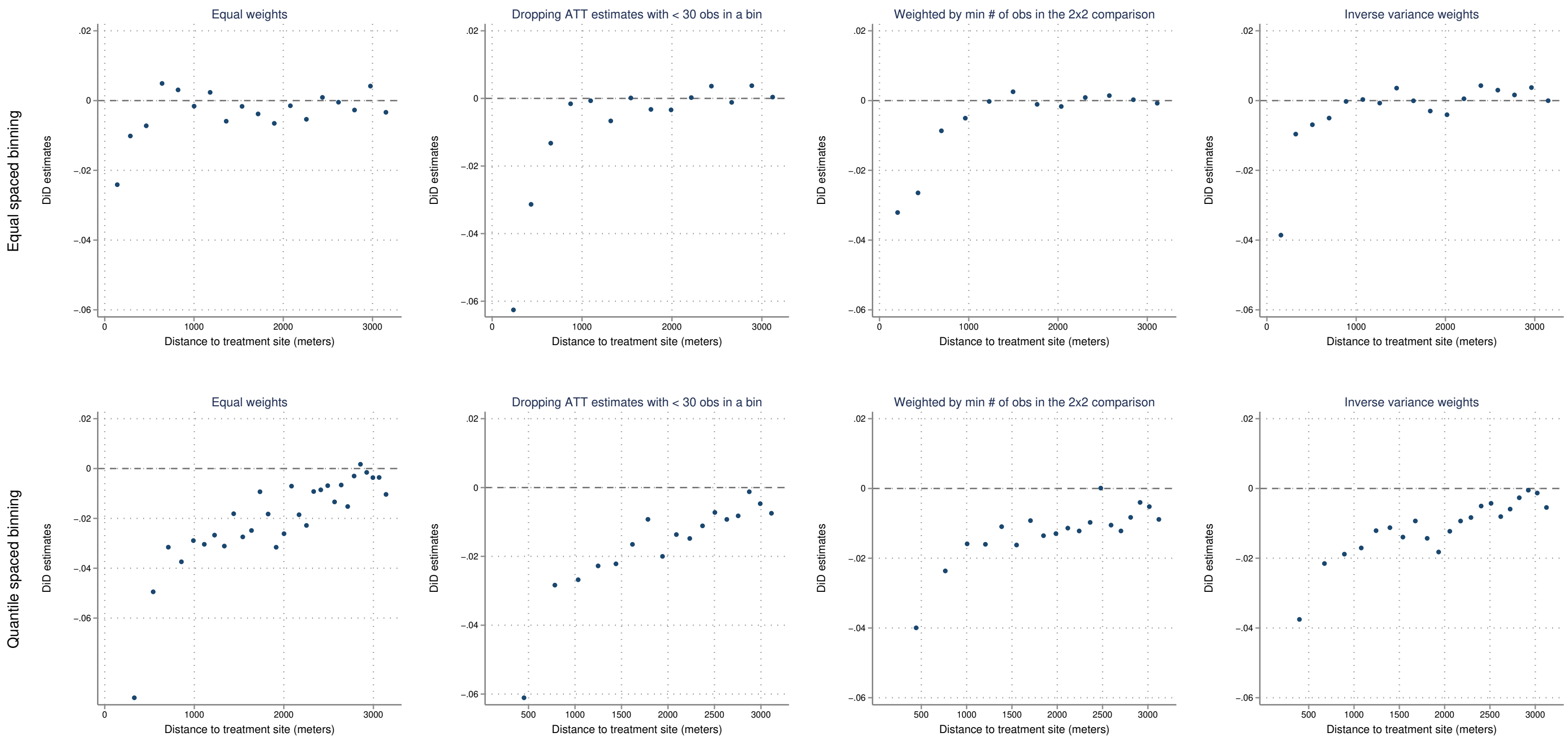

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A5: LIHTC Q3 long run price effects under alternative binning and weighting choices



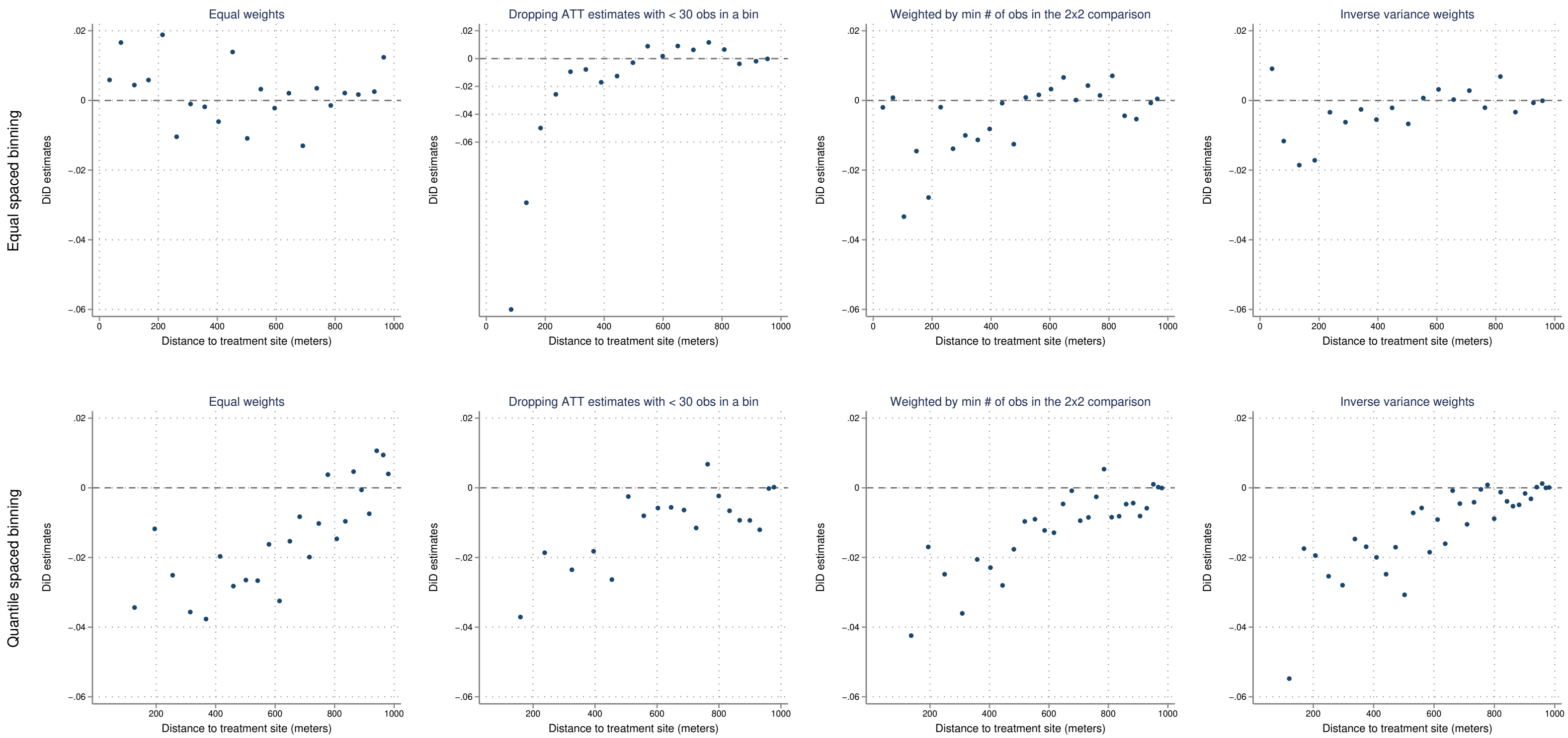

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A6: LIHTC Q4 long run price effects under alternative binning and weighting choices



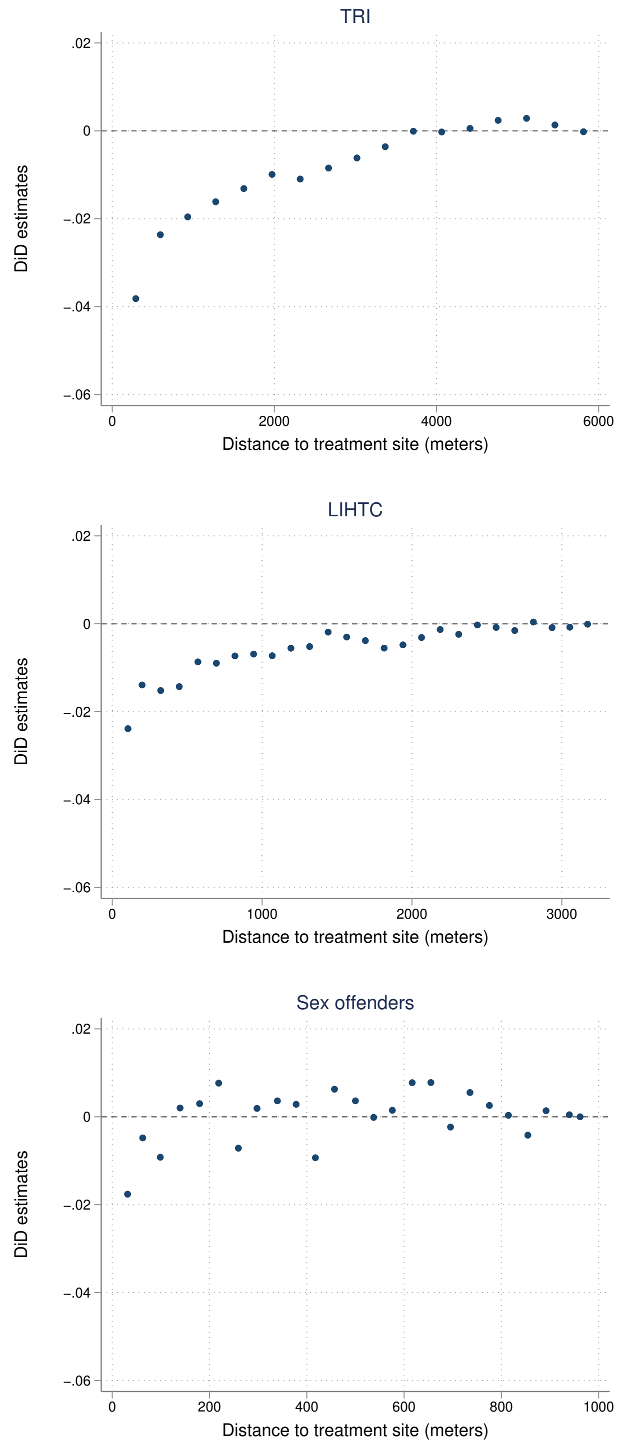

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A7: Sex offender move ins long run price effects under alternative binning and weighting choices



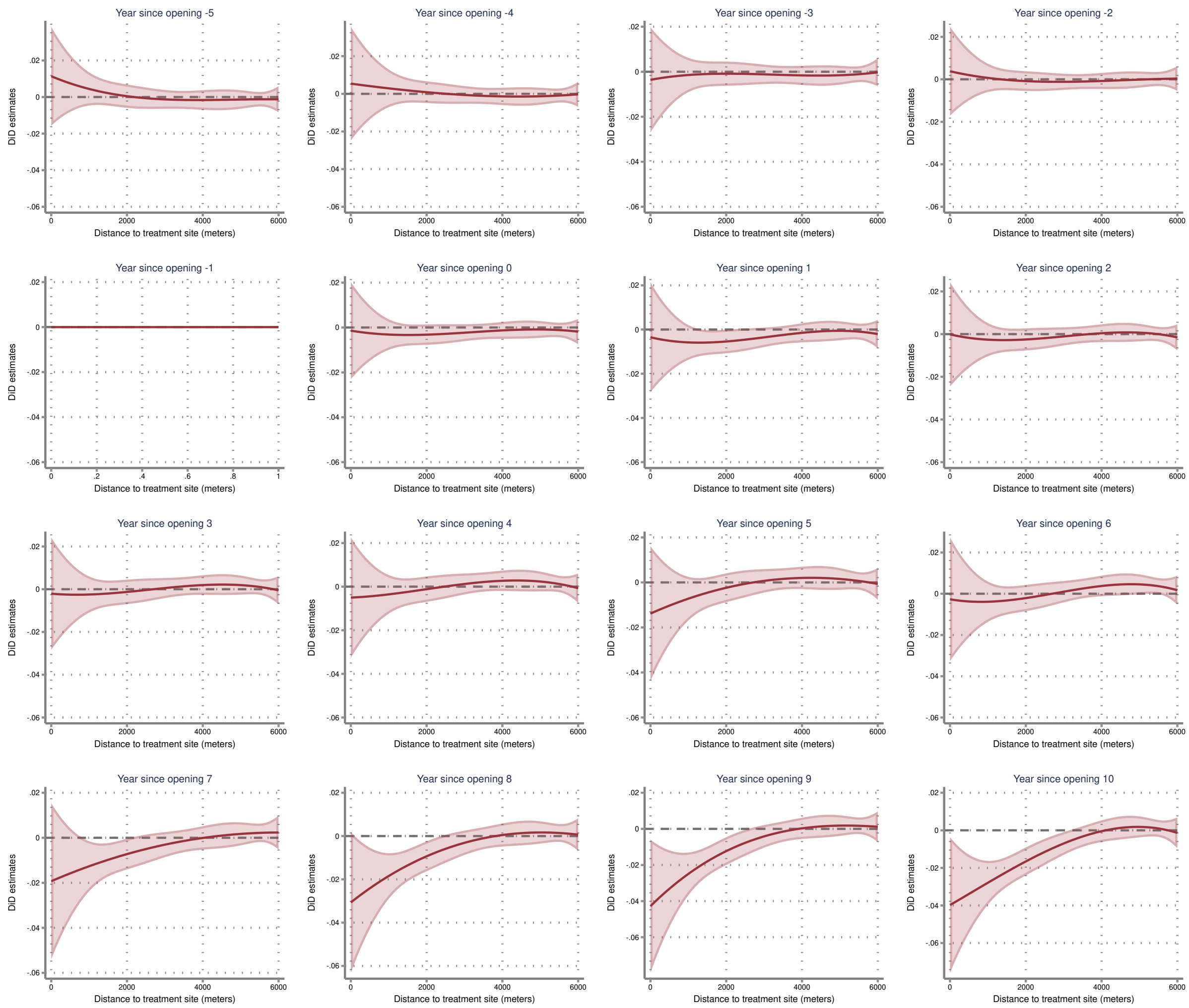

*Note: This figure shows the ATT estimates as in Figure 8 under alternative binning and weighting choices. The top row show bin scatter plots using equally spaced bins and the bottom row shows plots using quantile spaced bins. The first column weights each ATT estimate equally. The second column also weights ATT estimates equally but drops ATT estimates with less than 30 observations in any of the 2×2 sample averages. The third column weights by the minimum number of housing sales across the 2×2 sample averages. The fourth column weights ATT estimates by their inverse variances.*

Figure A8: Long run price effects residualizing treatment site× sale year fixed effects



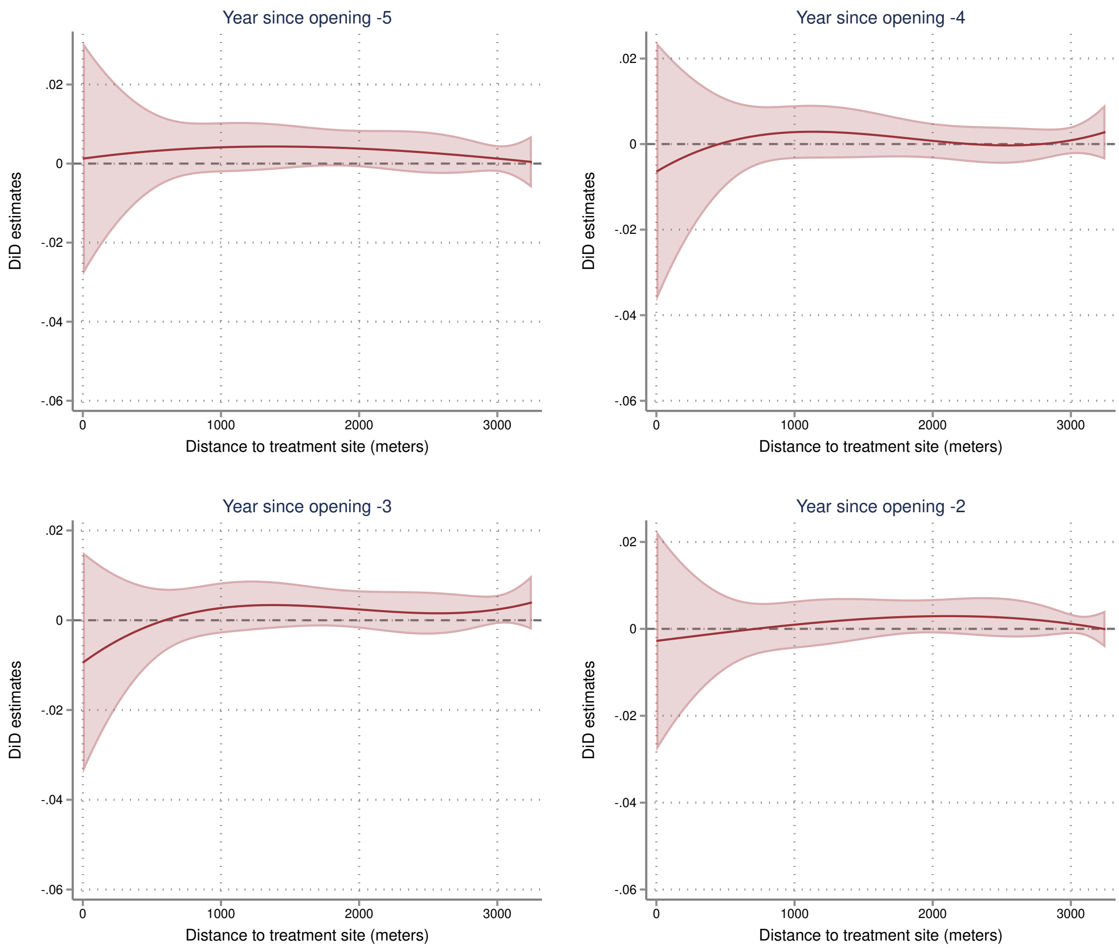

*Note: This figure shows the ATT estimates as in Figure 8 estimating bin level average sale prices in the first stage with treatment site× sale year fixed effects following Cattaneo et al. (2024). As in Figure 8, the data is binned a second time for visualization and averages are weighted by the ATT estimates' inverse variances. The top row shows a binscatter plot of the estimates for TRI industrial plants. The middle row plots the estimates for LIHTC sites. The bottom plot shows estimates for sex offender move ins.*

Figure A9: TRI, full continuous event study




*Note: This figure shows the smoothed price effect estimates as in Figure 9 for all years. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in the long run price effect estimate.*

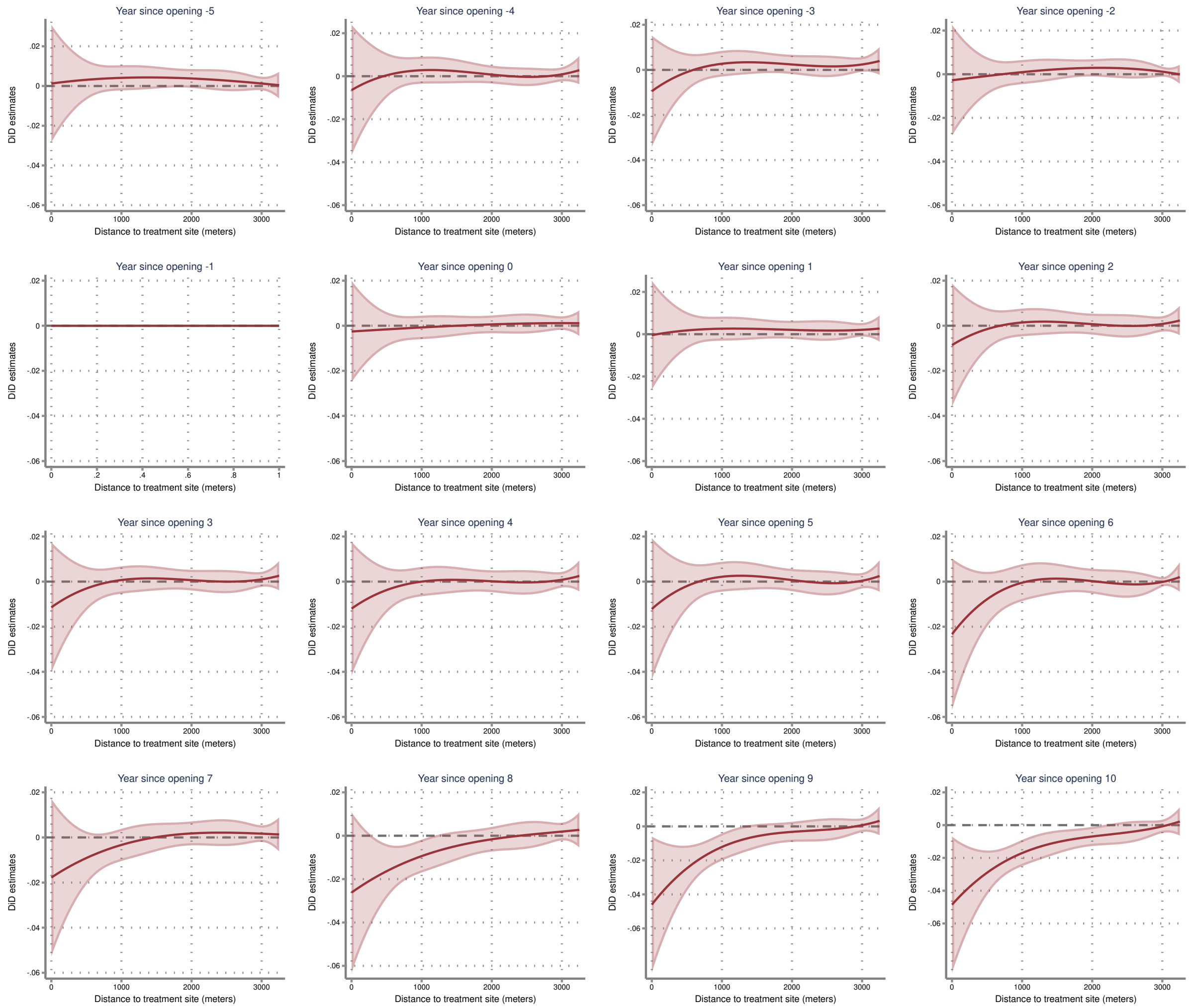
Figure A10: LIHTC developments, event study, all LIHTC sites




*Note: This figure performs a similar analysis as in Figure 11 but using data prior to the treatment date. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in Figure 11. Appendix Figure A11 presents event study plots for all years in the post period.*

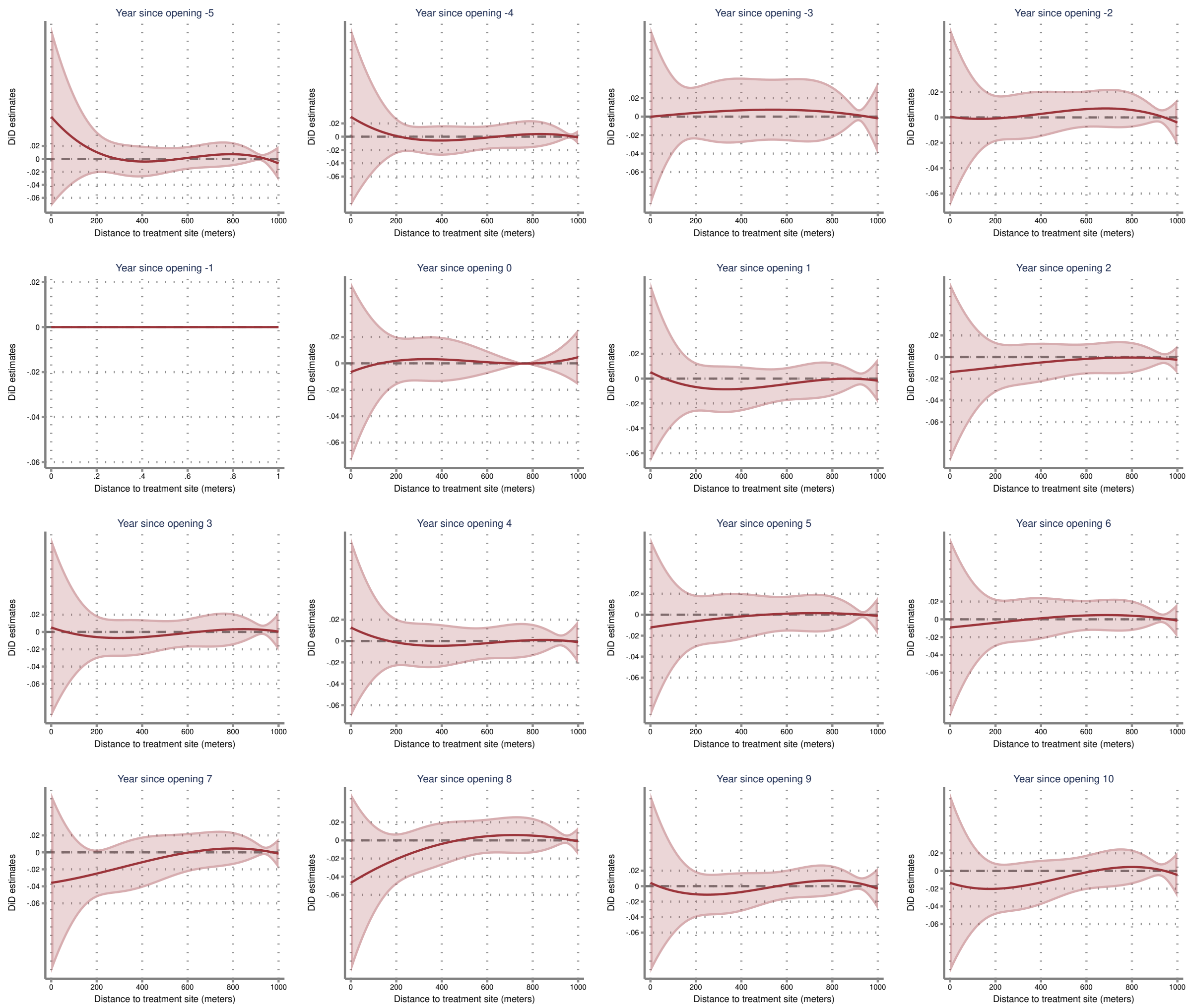
Figure A10 presents the continuous DiD event study for LIHTC developments analogous to the event study for TRI industrial plants (Figure 9). For each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in Figure 11. I do not find strong visual evidence that would suggest parallel trends violations. In the top left graph, the relationship five years before the allocation of funding is essentially a flat line. Years -4 and -3 might suggest a modest increasing gradient close to the site. However, in year -2 the estimated relationship again looks like a flat line and, given the confidence bands, I can not rule out a null effect in all four pre period years. Appendix Figure A11 presents event study plots for all years in the post period.

Figure A11: LIHTC, full continuous event study



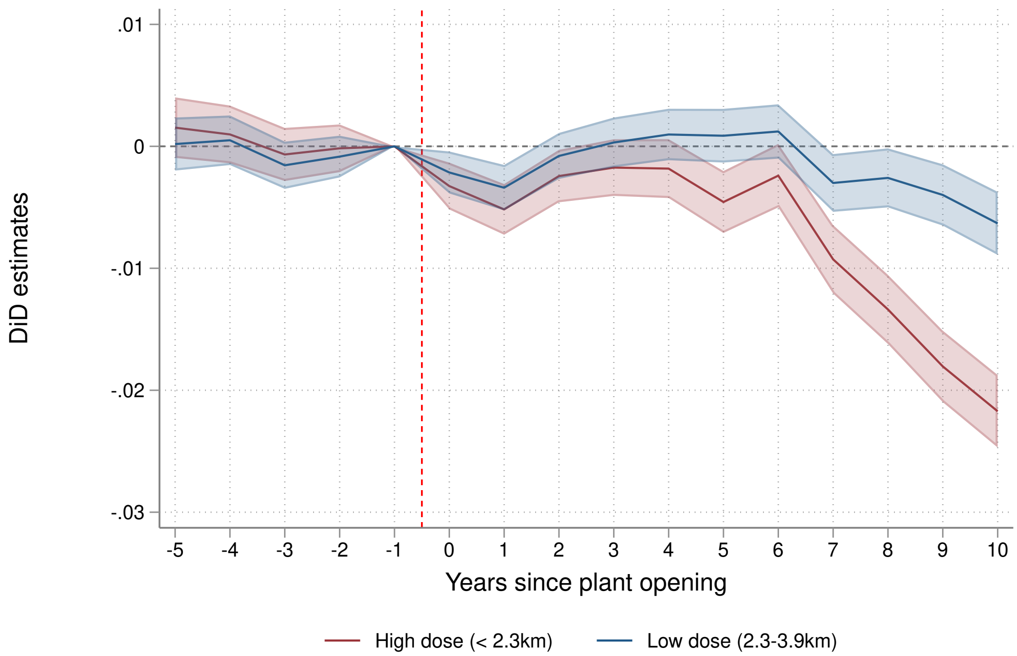

*Note: This figure shows the smoothed price effect estimates as in Figure A10 for all years. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in the long run price effect estimate.*

Figure A12: Sex offender move ins, full continuous event study



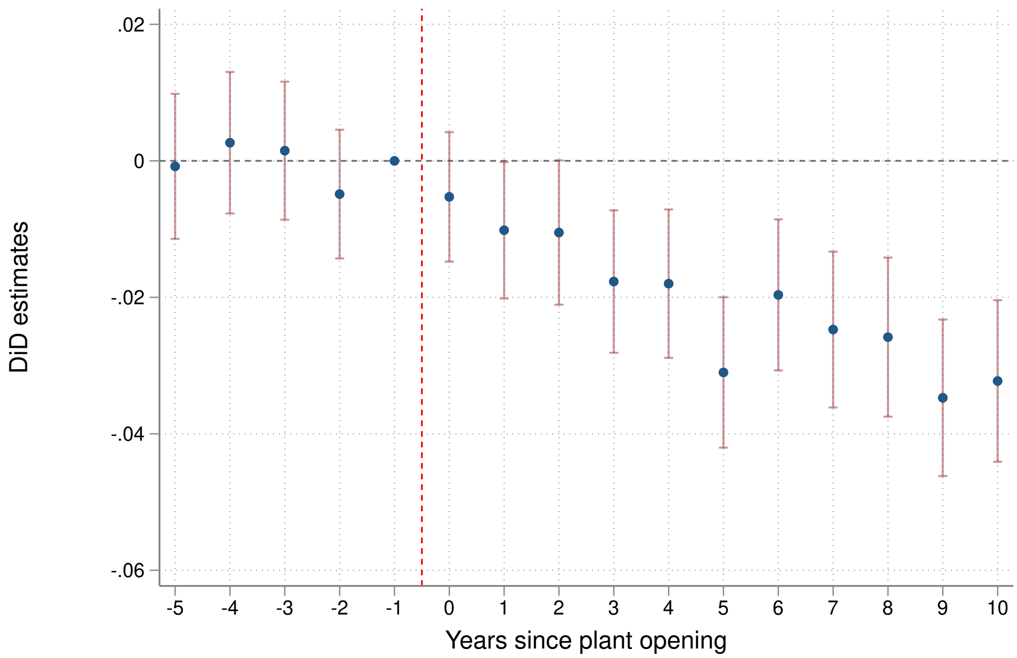

*Note: This figure shows the smoothed price effect estimates as in Figure 13 for all years. Specifically, for each year of the pre period, I form DiD comparisons at every distance bin relative to average house prices in the year before treatment. I then regress these comparisons on distance to the treatment site using the same nonparametric sieve regression procedure as in the long run price effect estimate.*

Figure A13: Alternative event study (Callaway et al., 2024b), TRI




*Note: This figure shows an event study for TRI industrial plants pooling together ATT estimates into two distance groups following Callaway et al. (2024b). I split the estimates into high and low treatment dose groups at 2,307 meters, the median bin distance from a plant within the treated area. The solid red line is the estimate for "high dose" areas (<2.3km), while the blue line is the estimate among "low dose" areas (2.3-3.9km). Each point is a weighted average of the ATT estimates within the dose group. Shaded areas are 95% pointwise confidence intervals.*

Figure A14: Traditional event study, TRI




*Note: This figure shows the results of a traditional event study. Each point is the coefficient from a stacked DiD regression that compares price trends in each year for distance bins within 1 kilometer of a plant vs. 4–5 kilometers away from a plant. Standard errors are clustered by plant site. The traditional event study is estimated using the 4,715 plant sites with a balanced panel in the inner and outer rings in each year.*

Figure A14 is based on the following regression model:

$$\log(saleprice)_{rgt} = \alpha_{gt} + \rho_{gr} + \sum_{k=-5,\, k\neq -1}^{10} \beta_k \times 1(t - t^* = k, r = 1) + \epsilon_{rgt}$$

where $\log(saleprice)_{rgt}$ is the average log sale price of houses in ring $r$ near plant $g$ in year $t$, $\alpha_{gt}, \rho_{gr}$ are treatment site $\times$ sale year and treatment site $\times$ ring fixed effects, and $r = 1$ if the average price is in the inner ring. $\beta_k$ are the parameters of interest plotted in Figure A14.

Figures A15, A16, and A17 plot every component that goes into the aggregate local price and welfare change calculations for each type of spatial treatment. In each figure:

- Panel (a) shows the main price effect estimate comparing 6-10 years after treatment to -5 to -1 years before. The red line is the fit line from a nonparametric sieve regression of DiD bin comparisons on distance to the treatment site following Chen et al. (2024) weighting by the $ATT$ estimates' inverse variances. The shaded red areas are the 95% uniform confidence bands. Blue dots are bin scatter plots of the comparisons following Cattaneo et al. (2024) (binning the data a second time). Confidence intervals on the binned scatter plot are calculated with standard errors clustered by treatment site.

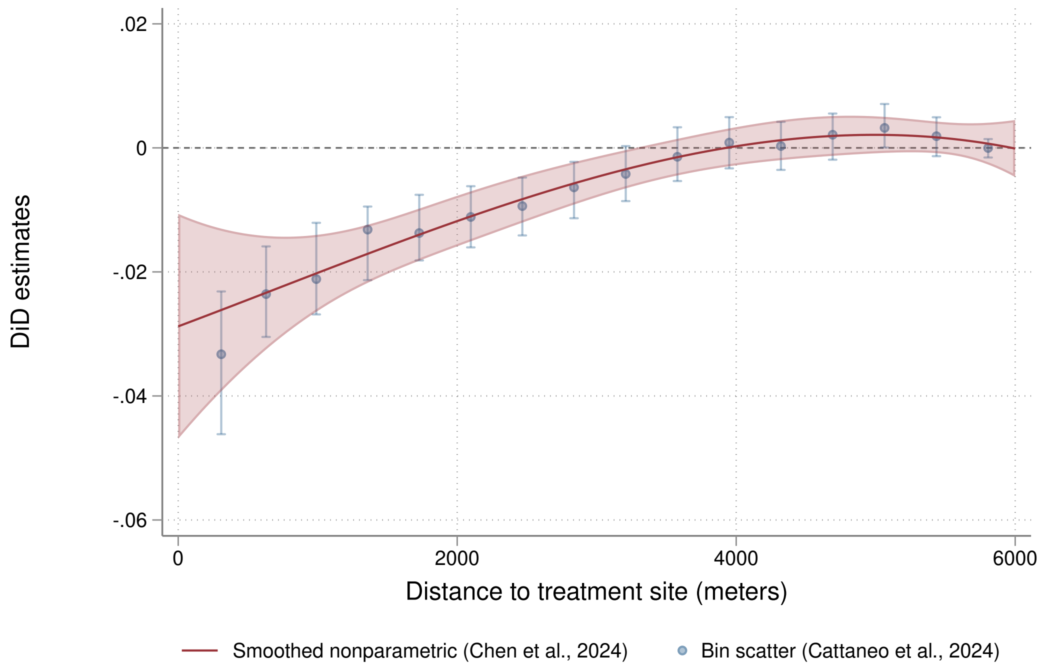
- Panel (b) shows the derivative estimate with 95% uniform confidence bands following Chen et al. (2024), or the derivative of the red line in panel (a).

- Panel (c) estimates the relationship between the average house price and distance to the treatment site in the post period by regressing average house prices in the post period on distance to the treatment site and treatment site location specific fixed effects. To control for the fixed effects, I follow the procedure in Chen et al. (2024)'s application. Specifically, I regress housing prices on distance to the treatment site and treatment site location specific fixed effects under an initial basis choice. I then apply Chen et al. (2024) as before but using house prices minus the estimated treatment site location fixed effects from the initial regression as the outcome variable. As an additional check, the blue scatter plot also shows the estimates from `binsreg` with covariate adjustment done following Cattaneo et al. (2024).

- Panel (d) is the PDF of owner occupied housing units on distance to the treatment site within the spatial extent of treatment effects using standard kernel density estimation.

Figure A15: TRI, components of aggregate welfare calculation

(a) Price effect estimate $A\hat{T}T^{agg}(d|d)$




(b) Derivative estimate $\frac{\partial A\hat{T}T^{agg}}{\partial d}(d|d)$

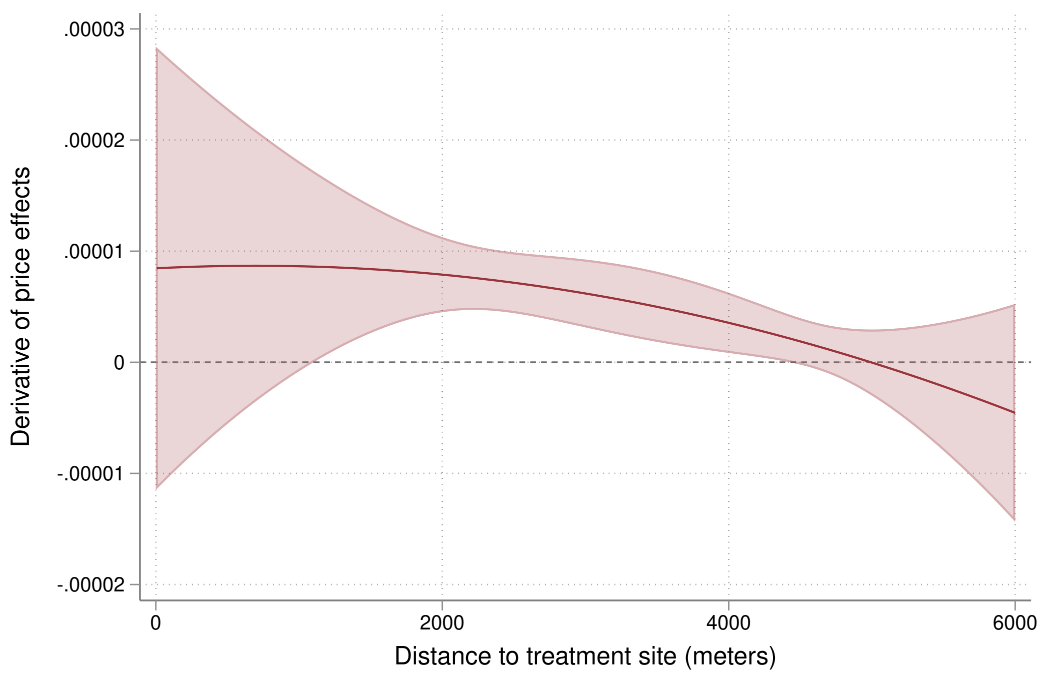
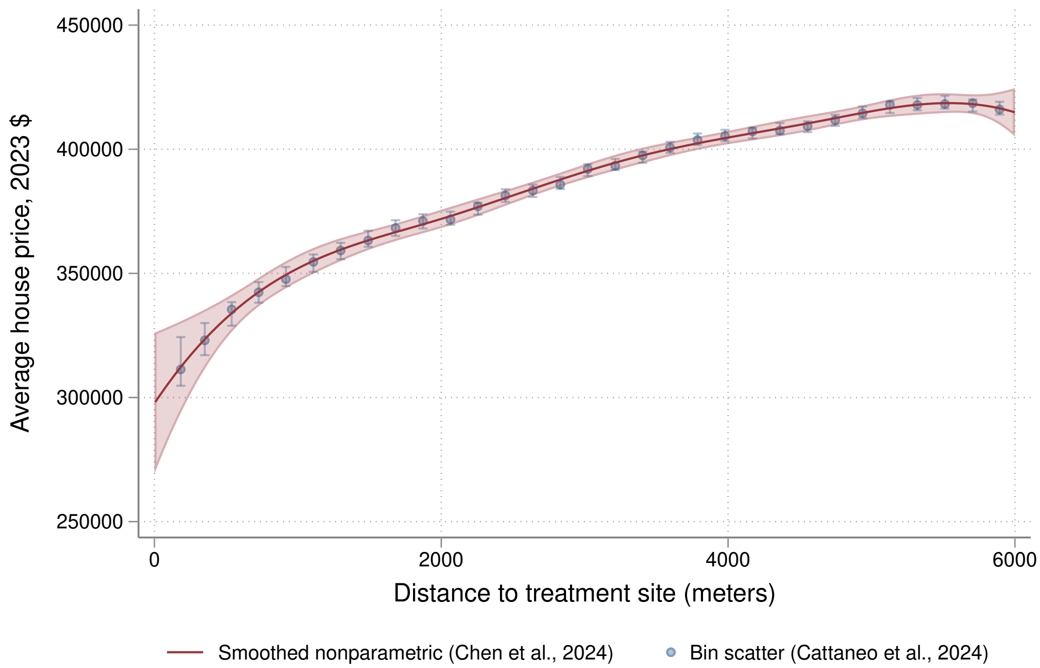


Figure A15: continued

(c) Average house price level $\widehat{E[Y|D=d]}$




(d) Owner occupied housing unit density $\hat{f}(d)$

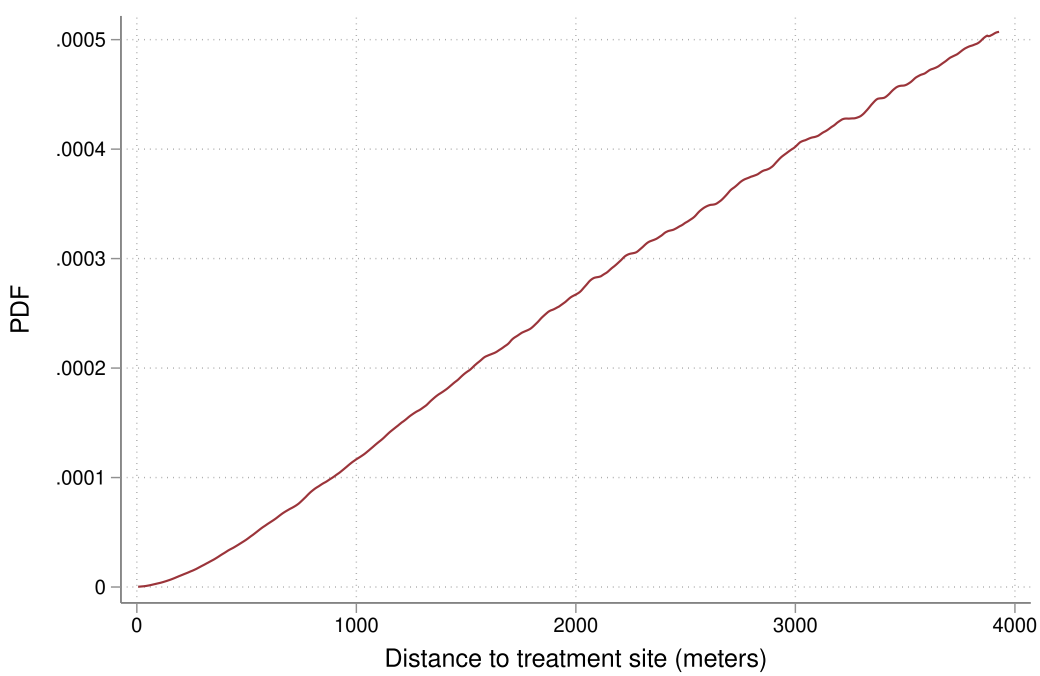
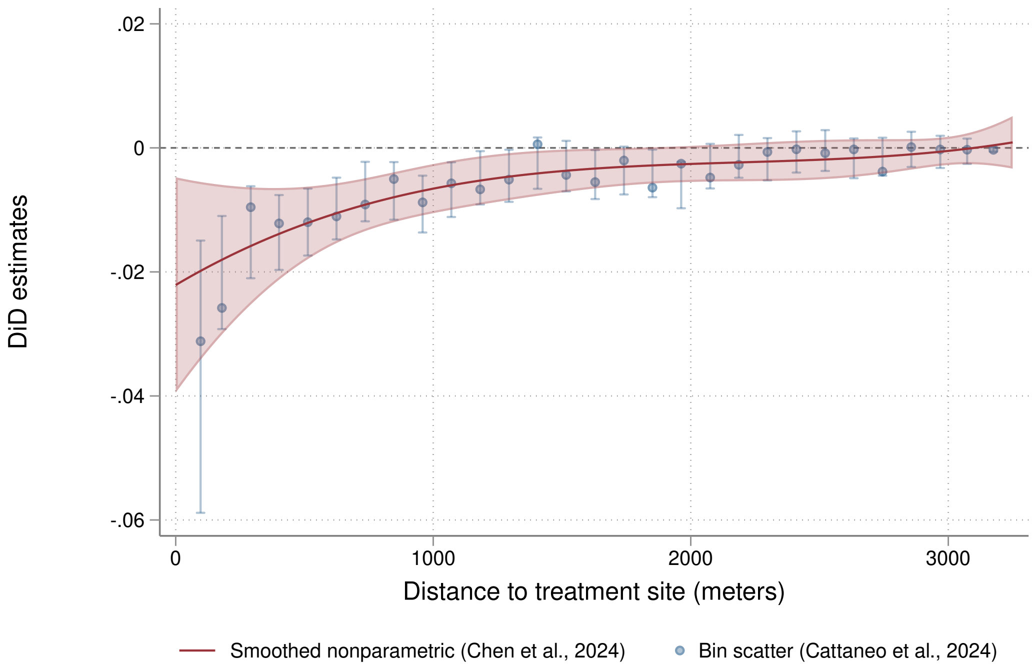


Figure A16: LIHTC, components of aggregate welfare calculation

(a) Price effect estimate $\hat{ATT}^{agg}(d|d)$




(b) Derivative estimate $\frac{\partial \hat{ATT}^{agg}}{\partial d}(d|d)$

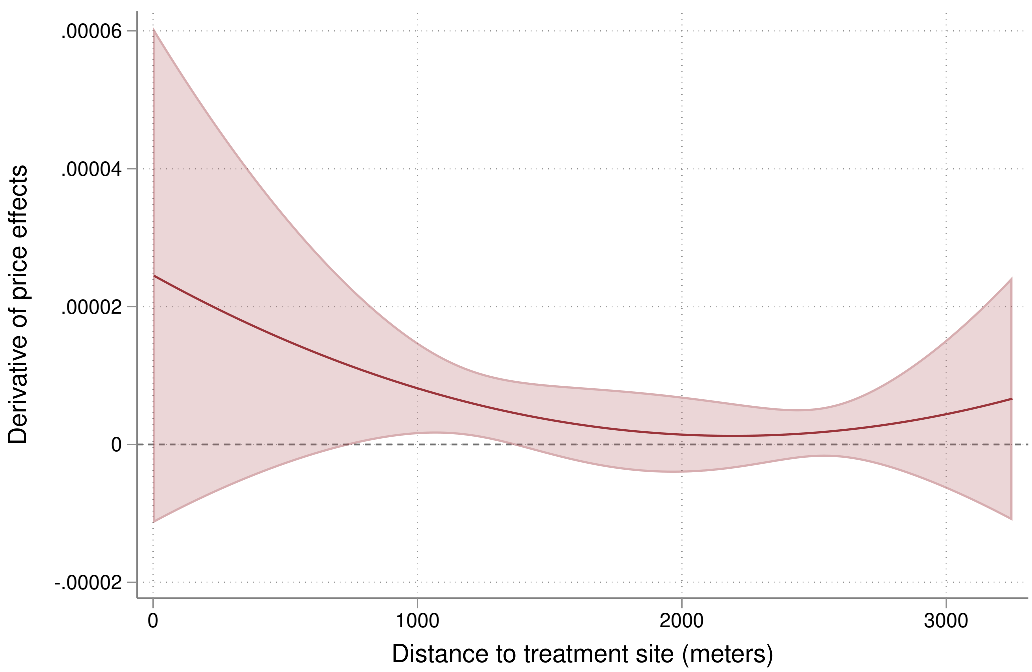
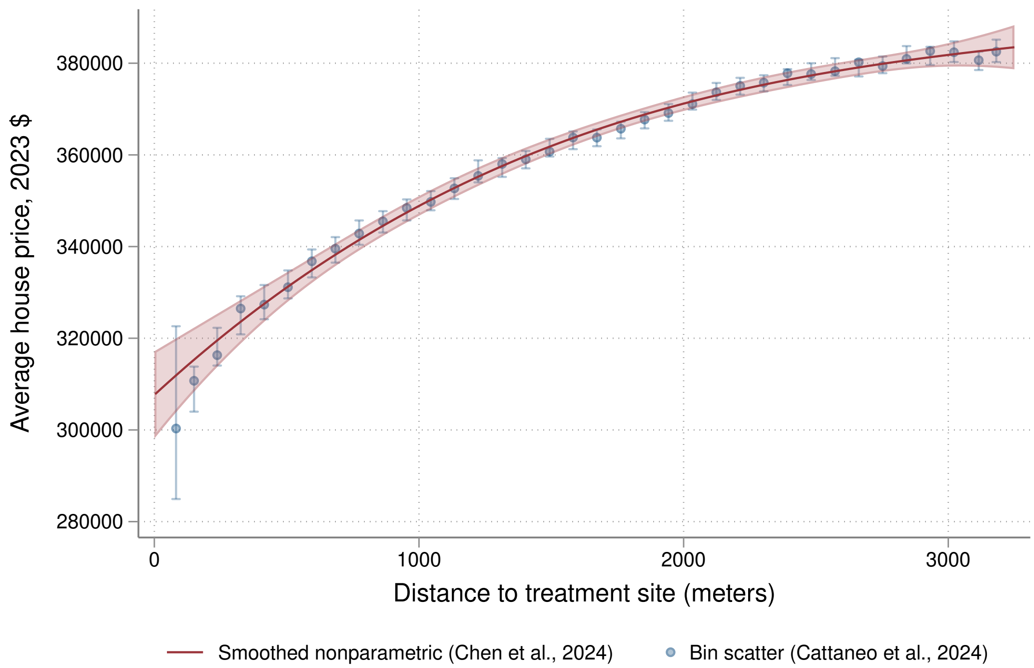


Figure A16: continued

(c) Average house price level $\widehat{\mathrm{E}[Y|D=d]}$




(d) Owner occupied housing unit density $\hat{f}(d)$

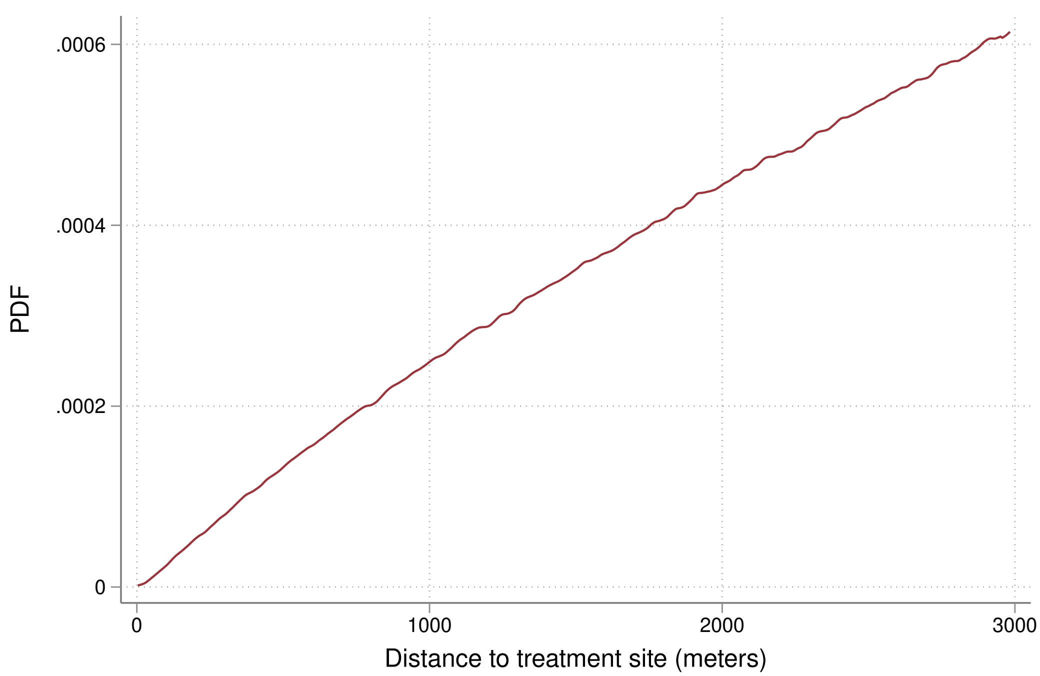
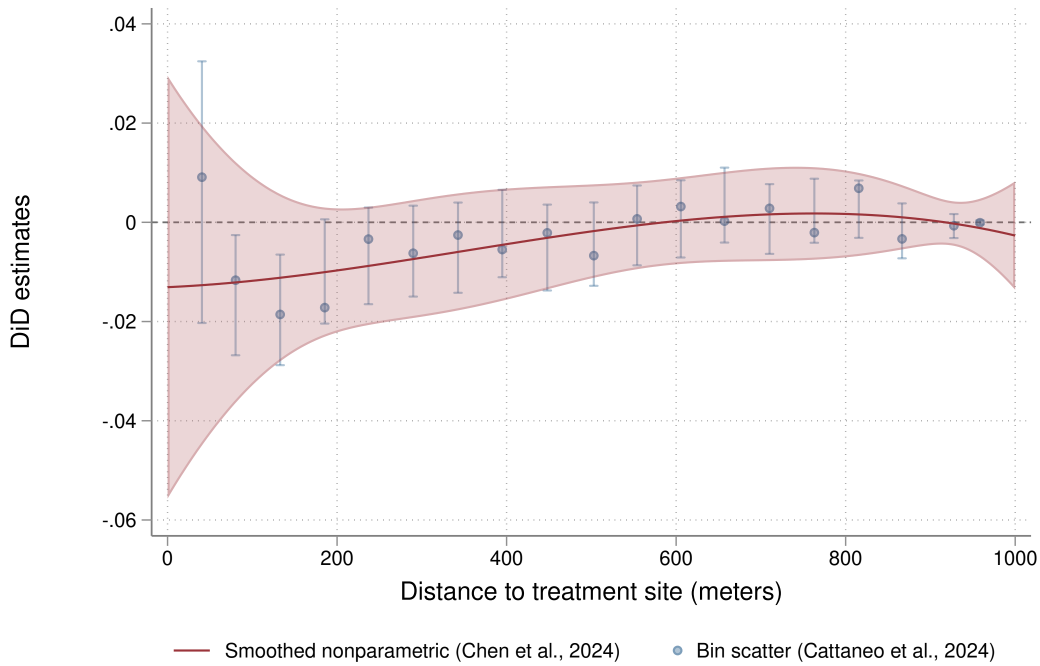


Figure A17: Sex Offenders, components of aggregate welfare calculation

(a) Price effect estimate $\hat{ATT}^{agg}(d|d)$




(b) Derivative estimate $\frac{\partial \hat{ATT}^{agg}}{\partial d}(d|d)$

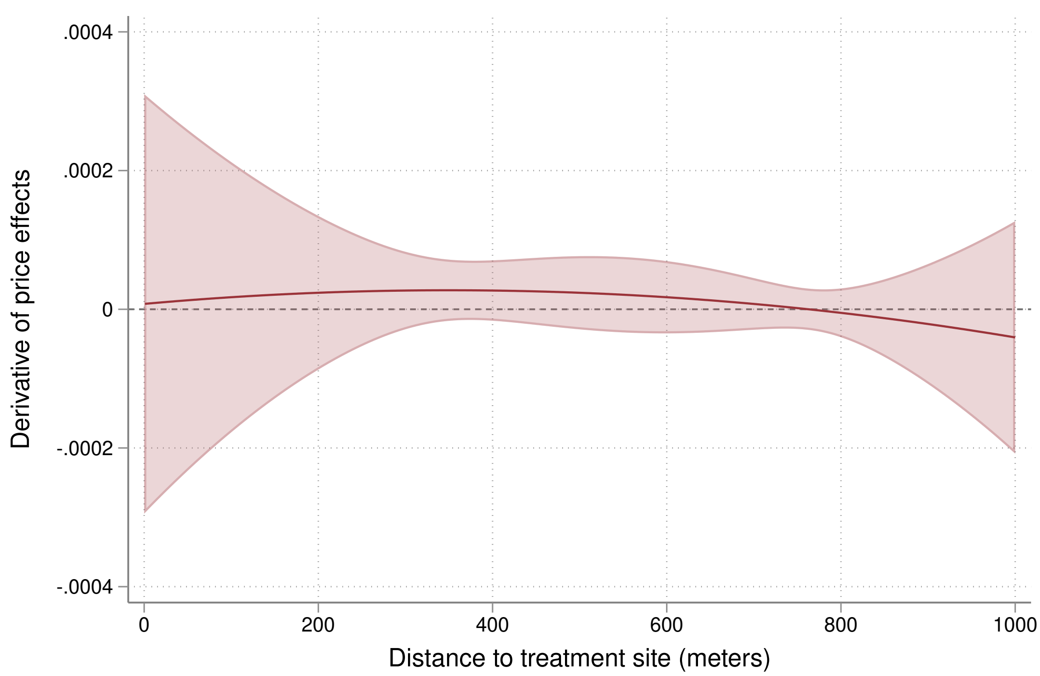
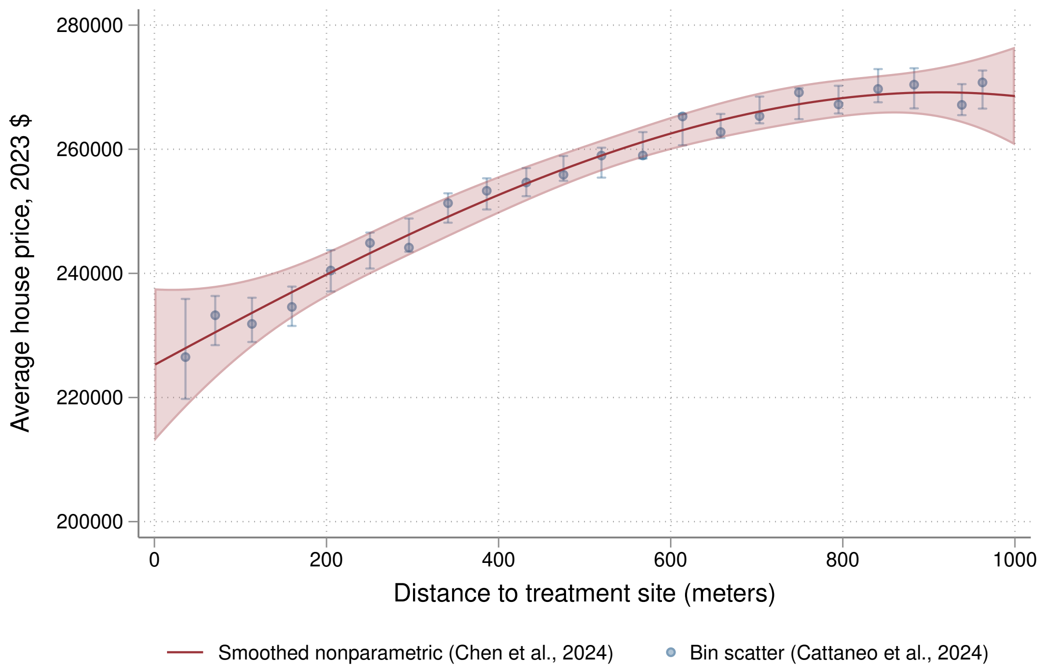


Figure A17: continued

(c) Average house price level $\widehat{E[Y|D=d]}$




(d) Owner occupied housing unit density $\hat{f}(d)$

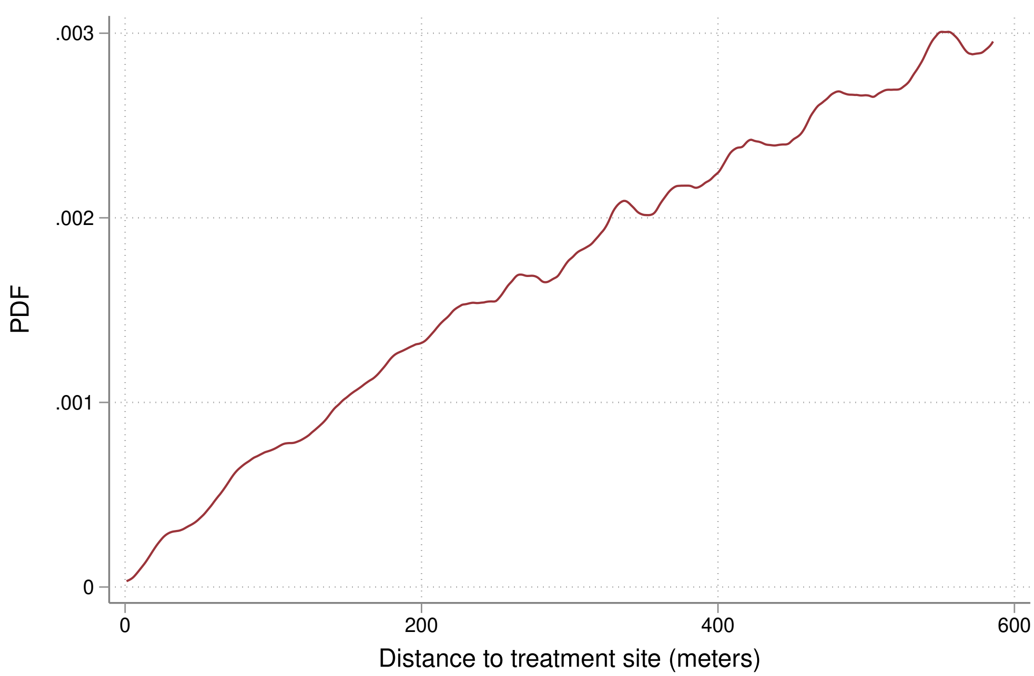
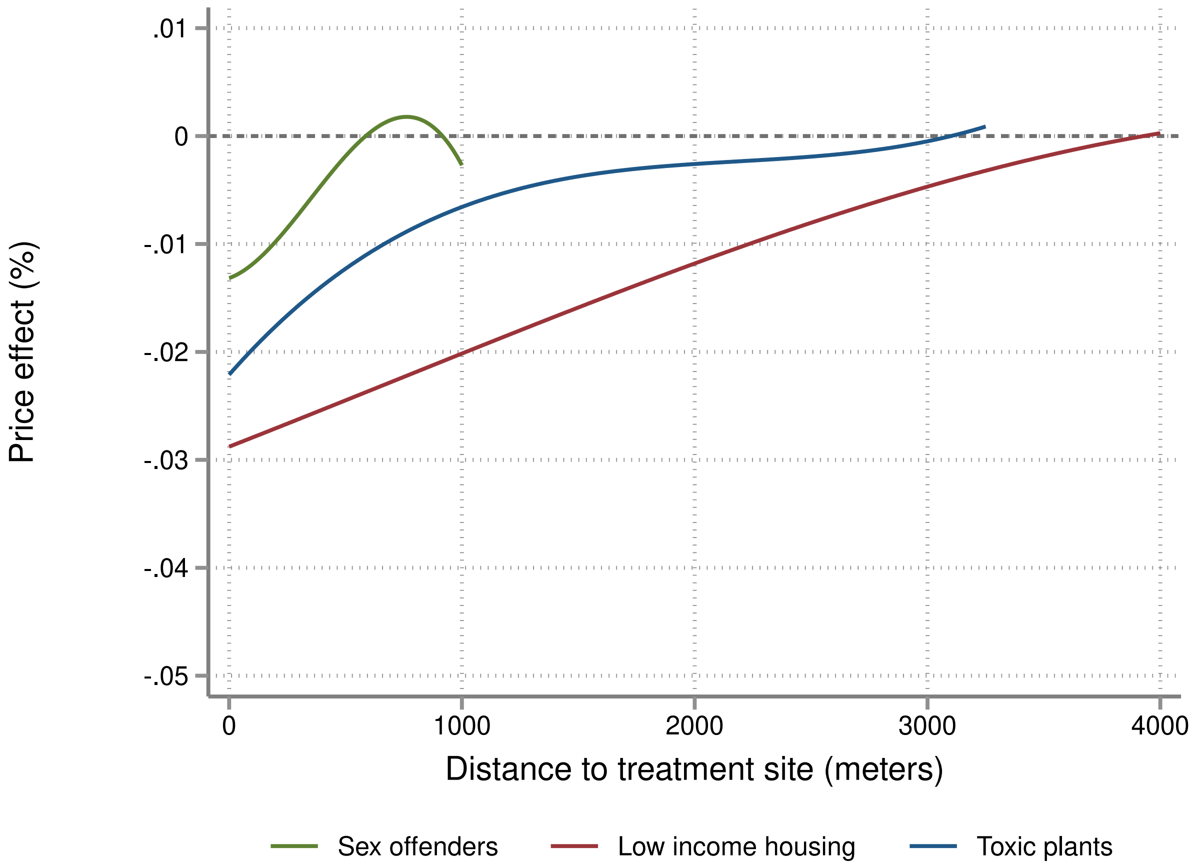


Figure A18: Long run price effects on the same scale including sex offender estimates




*Note: This figure shows the price effect estimates from Figures 8, 11, A17 on the same y and x axises analogous to Figure 14. The green line shows the price effect estimate for sex offender move ins, the blue line shows the price effect estimate for LIHTC developments, and the red line shows the effect for TRI industrial plants.*

## Appendix tables

### *A. Summary statistics*

Tables A1 and A2 presents summary statistics on site locations, housing sales, and housing unit counts for LIHTC developments and sex offenders respectively.

Panel A of Table 2 reports that there are 17,895 LIHTC sites included in my estimation sample in which the typical LIHTC development was allocated funding in 2005. Panel B shows average housing prices by distance to the treated site in five equally spaced distance ranges extending out to 3,250 meters. As with TRI industrial plant locations, LIHTC developments tend to be sited in neighborhoods with lower house values compared to areas farther away from the development. The average house price is $328,051 within 650 meters of a development and increases to $381,402 for houses 2,600 to 3,250 meters away. Panel C shows the number of owner occupied housing units where, as before, there are far more housing units farther away than close in to the treated sites. There are approximately 8 million owner occupied housing units within 650 meters, while there are 57 million units 2,600 to 3,250 meters away. Compared to TRI industrial plant sites, LIHTC sites have more housing units nearby, reflecting that LIHTC developments tend to be located in more urban areas.

Table A1: Summary statistics on LIHTC developments

| | | | | | |
|---|---|---|---|---|---|
| **Panel A. Treatment sites** | | | | | |
| Number of sites | 17,895 | | | | |
| Average year of treatment event | 2005 | | | | |
| | Distance to treatment site | | | | |
| **Panel B. Housing sales** | 0 to 650m | 650 to 1,300m | 1,300 to 1,950m | 1,950 to 2,600m | 2,600 to 3,250m |
| Average house price (2023 $) | $328,051 | $342,294 | $352,419 | $365,661 | $381,402 |
| Number of site × house sales | 4,980,723 | 14,428,167 | 22,028,002 | 28,499,351 | 34,414,231 |
| Number of site × distance bins | 49,426 | 67,803 | 68,782 | 68,978 | 69,018 |
| **Panel C. Housing unit counts** | | | | | |
| Number of owner occupied housing units | 8,077,607 | 22,855,868 | 35,401,470 | 46,512,737 | 57,218,527 |

*Note: This table presents summary statistics on treatment sites used in the sample and the average sale prices and owner occupied housing unit counts of nearby houses by distance to the treatment site. Panel A describes 17,895 LIHTC developments from 1995-2015. Panel B shows the average house price in real 2023 dollars by distance to the treatment site. Panel C tabulates totals of owner occupied housing unit counts in nearby census blocks by distance to treatment site.*

Table 3 presents the same information for sex offender move ins for North Carolina and Florida. In total, I observe 9,193 sex offender moves from 1995 to 2015 in these two states with an average move year of 2009. This average is closer to the end of the time range because I only observe the current registry of sex offenders and not a complete history of moves. Panel B shows average house prices in five equally spaced distance ranges within 1,000 meters of a sex offender's address, consistent with a smaller spatial extent as considered in Linden and Rockoff (2008) and Pope (2008). As with the other two types of treatment events, sex offenders live in areas with systematically lower house values. The average house price is $233,907 within 200 meters compared to $305,425 for houses 800 to 1,000 meters away. Panel C tabulates housing unit counts by distance to the sex offenders' addresses. In my sample areas, there are approximately 300,000 owner occupied housing units within 200 meters, while there is roughly 1.8 million units 800 to 1,000 meters away. Given the smaller spatial extent considered, for each sex offender who moves into a neighborhood, any estimated price effects will affect far fewer homes than a TRI industrial plant or a LIHTC development.

Table A2: Summary statistics on sex offender move ins

| | | | | | |
|---|---|---|---|---|---|
| **Panel A. Treatment sites** | | | | | |
| Number of sites | 9,193 | | | | |
| Average year of treatment event | 2009 | | | | |
| | Distance to treatment site | | | | |
| **Panel B. Housing sales** | 0 to 200m | 200 to 400m | 400 to 600m | 600 to 800m | 800 to 1,000m |
| Average house price (2023 $) | $233,907 | $258,563 | $285,061 | $296,376 | $305,425 |
| Number of site × house sales | 377,047 | 951,218 | 1,377,936 | 1,764,400 | 2,103,548 |
| Number of site × distance bins | 13,059 | 18,666 | 19,830 | 21,354 | 23,334 |
| **Panel C. Housing unit counts** | | | | | |
| Number of owner occupied housing units | 300,567 | 762,615 | 1,134,528 | 1,478,115 | 1,793,093 |

*Note: This table presents summary statistics on treatment sites used in the sample and the average sale prices and owner occupied housing unit counts of nearby houses by distance to the treatment site. Panel A describes 9,193 sex offender move ins from 1995-2015. Panel B shows the average house price in real 2023 dollars by distance to the treatment site. Panel C tabulates totals of owner occupied housing unit counts in nearby census blocks by distance to treatment site.*

Table A3: Robustness check for sensitivity to choice of constant in utility function

| | Value of c | $\frac{ATT}{d-\bar{d}}$ | $\frac{\partial ATT}{\partial d}$ |
|---|---|---|---|
| **Panel A. TRI** | 0 | \$13.085 mil. | \$16.409 mil. |
| | .1 | \$13.085 mil. | \$16.409 mil. |
| | 1 | \$13.087 mil. | \$16.412 mil. |
| | 10 | \$13.108 mil. | \$16.438 mil. |
| | 100 | \$13.309 mil. | \$16.692 mil. |
| **Panel B. LIHTC** | 0 | \$6.694 mil. | \$9.142 mil. |
| | .1 | \$6.694 mil. | \$9.143 mil. |
| | 1 | \$6.696 mil. | \$9.147 mil. |
| | 10 | \$6.714 mil. | \$9.193 mil. |
| | 100 | \$6.884 mil. | \$9.619 mil. |

*Note: This table presents estimates of the aggregate welfare impacts varying the constant used within the log transformation in the utility function, i.e.* $c + d$*. The first column shows the constant used. The second column shows the welfare estimate using the rise over run estimator as the estimate for homeowner MWTP. The third column shows the welfare estimate using the ATT derivative as homeowner MWTP. For specifications with c=0, the equation integrates to the next closest evaluation point to 0, which is 6 and 3 meters for TRI and LIHTC sites, respectively.*

Table A4: Total local price and welfare impacts on homeowners

| Treatment event | Toxic industrial plants | Low income housing | Sex offenders |
|---|---|---|---|
| Mean WTP per homeowner, $\frac{ATT}{d-\bar{d}}$ | \$1,886 | \$815 | \$807 |
| | (296) | (273) | (277) |
| Mean WTP per homeowner, $\frac{\partial ATT}{\partial d}$ | \$2,365 | \$1,113 | \$834 |
| | (313) | (229) | (341) |
| Mean $\mid\Delta$Price$\mid$ per homeowner | \$2,804 | \$1,277 | \$1,246 |
| | (398) | (373) | (418) |
| # of affected homeowners per site | 6,939 | 8,219 | 250 |
| Agg. WTP per site, $\frac{ATT}{d-\bar{d}}$ | \$13.1 mil. | \$6.7 mil. | \$0.2 mil. |
| | (3.84) | (2.92) | (0.16) |
| Agg. WTP per site, $\frac{\partial ATT}{\partial d}$ | \$16.4 mil. | \$9.1 mil. | \$0.2 mil. |
| | (7.53) | (2.46) | (0.22) |
| Agg. $\mid\Delta$Price$\mid$ per site | \$19.5 mil. | \$10.5 mil. | \$0.3 mil. |
| | (6.18) | (4.11) | (0.26) |
| # of sites in the U.S. | 15,899 | 43,466 | 917,771 |
| Total WTP, $\frac{ATT}{d-\bar{d}}$ | \$208 bil. | \$291 bil. | \$185 bil. |
| Total WTP, $\frac{\partial ATT}{\partial d}$ | \$261 bil. | \$398 bil. | \$191 bil. |
| Total $\mid\Delta$Price$\mid$ | \$309 bil. | \$456 bil. | \$285 bil. |

*Note: This table presents estimates of the price and welfare impacts on local homeowners from three distinct types of treatment events that could affect neighborhood amenities. The first column has estimates for TRI industrial plants, the second column has estimates for LIHTC developments, and the third column has estimates for sex offender move ins. Caveat: sex offender move ins are included for completeness but are based on imprecise estimates with possible parallel trends violations. The first two rows show the mean willingness to pay among homeowners in the affected geographic area for a treatment site to not exist using either the rise over the run or the ATT derivative as the average MWTP. The third row shows the absolute value of the average price change. Aggregate WTP and price changes are calculated by multiplying the mean WTP and price change by the # of affected homeowners. The final three rows are back of the envelope estimates of the total externalities nationally based on multiplying the aggregate impacts by an estimate for the number of treatment sites currently in the U.S. (in the row directly above). The estimate for the # of TRI sites is the number of plants in the TRI database reporting positive pollution emissions in 2022. The estimate for the # of LIHTC sites is the number of developments in the HUD database still being monitored as of April 2024. The estimate for the # of sex offenders is from a 2019 report by the National Center for Missing and Exploited Children (NCMEC). Standard errors were calculated by bootstrap resampling treatment sites for 1,000 replicates.*

## Appendix B Econometric proofs

This appendix provides proofs for the theorems on identification in Section 3 of the main text. These proofs are an adaptation of the proofs from the appendixes of Callaway et al. (2024a) to fit the spatial application. I then include additional proofs for identification under a linear correlated random coefficients model and a lower bound under $ATT$ concavity.

I begin in the same setup as Section 3. There are $M$ distinct treatment sites indexed by $1, 2, \ldots, M$ and individual house sales $j \in \mathcal{J}$ occurring in time periods t = $1,2,\ldots,T$. Let $G_j \in \{1, 2, \ldots M, \infty\} = \mathcal{G}$ be the site that the sale $j$ is associated with, which is that housing unit's *treatment site-timing group*. Following convention, I set $G = \infty$ for units that are untreated across all time periods. Finally, define $g^t(G) : \mathcal{G} \mapsto \mathcal{T}$ as a function that maps a site to the time period when it first opened.

For each site there is a pre period and a post period. In the post period, units receive a dose $D = d$ based on their distance from the treated site. Let $\mathcal{D}$ denote the support of $D$. Finally, let $X$ denote a house's pretreatment covariates.

Define $Y_{j,t}(g, d)$ as the potential outcome for house $j$ near plant $g$ in time period $t$. This is the house price if house $j$ was $d$ distance away from a treatment site in time period $t$. The realized outcome is $Y_{j,t} = Y_{j,t}(G_j, D_j)$. I now define treated and untreated potential outcomes. For untreated potential outcomes to be well defined, we must place an assumption on how far away the plant affects housing units.

**Assumption 1** (Spatial extent of treatment effects). *Denote $\overline{d}$ where $0 < \overline{d} < \infty$ as the spatial extent of treatment effects, so units farther than $\overline{d}$ from a plant are untreated in every time period. For any $d \geq \overline{d}$, $Y_{j,t}(g, d) = Y_{j,t}(\infty, \overline{d})$.*

Now I write a unit's untreated potential outcome as $Y_{j,t}(\overline{d}) \equiv Y_{j,t}(\infty, \overline{d})$. Additionally, define $W_{j,t} = D_j 1\{t \geq g^t(G_j)\}$, which is the dose that unit $j$ experiences in time period $t$ and is equal to 0 for all units not yet treated in time period $t$.

In the DiD framework, we consider averages of individual effects. In particular,

$$ATT(g, t, d|g', d') = \mathrm{E}[Y_t(g,d) - Y_t(\overline{d})|G = g', D = d']$$

is the average effect of being $d$ distance from plant $g$ in time period $t$ for units that are located $d'$ from the plant $g'$. Similarly, define the average causal response parameter as

$$ACRT(g, t, d|g',d') = \frac{\partial ATT(g, t, l|g', d')}{\partial l}|_{l=d} = \frac{\partial Y_t(g, l|G = g', D = d')}{\partial l}|_{l=d}$$

Next, I repeat the assumptions from Section 3.

**Assumption 2** (Random sampling for repeated cross sections). *Conditional on $T = t$, the data are independent and identically distributed from the distribution of $(Y_t, G_1, G_2, \ldots, G_M, D, T, X)$, for all $t = 1, \ldots, T$, with $(G_1, G_2, \ldots, G_M, D, X)$ being invariant to $T$.*

**Assumption 3** (Support and continuously differentiable treatment). *(a) The support of $D$ is $\mathcal{D} = [0,d_U]$ with $0 < \overline{d} < d_U$, where $d_U < \infty$ is an upper bound on the spatial extent considered. Furthermore, $\mathrm{P}(D > \overline{d}) > 0$ and $dF_{D|G}(d|g) > 0$ for all $(g,d) \in \mathcal{G} \times [0,\overline{d}]$*

*(b) Assume for all $g \in \mathcal{G}$ and $t \in 2,\dots,T$, $\mathrm{E}[\Delta Y_t|G = g, D = d]$ is continuously differentiable in $d$ on $[0,\overline{d}]$.*

**Assumption 4** (No anticipation / staggered adoption). *(a) For all $g \in \mathcal{G}$ and $t \in 1,\dots,T$ with $t < g^t(g)$, $Y_{j,t}(g, d) = Y_{j,t}(\overline{d})$.*

*(b) $W_{j,1} = 0$ almost surely. For $t = 2,\dots,T, W_{j,t-1} = d > 0$ implies that $W_{j,t} = d$.*

**Assumption 5** (Parallel trends). *For all $g \in \mathcal{G}, t = 2,3,\dots,T$ and $d \in \mathcal{D}$, $\mathrm{E}[\Delta Y_t(\overline{d})|G = g, D = d] = \mathrm{E}[\Delta Y_t(\overline{d})|G = \infty, D = \overline{d}]$*

Under this setup and Assumptions 1–5, the $ATT$ is identified.

**Theorem Appendix B.1** *Under assumptions 1 to 5, the $ATT$ is identified. In particular,*

$$ATT(g,t,d|g,d) = \mathrm{E}[Y_t - Y_{g-1}|G = g, D = d] - \mathrm{E}[Y_t - Y_{g-1}|G = g, D = \overline{d}]$$

*Proof:* This result and its proof are essentially the same as in the standard binary DiD setting and is similar to the proof in section SB of Callaway et al. (2024a) but using the parallel trends Assumption 4 above.

$$\begin{aligned} ATT(g,t,d|g,d) &= \mathrm{E}[Y_t(g,d) - Y_t(\overline{d})|G = g, D = d] \\ &= \mathrm{E}[Y_t(g,d) - Y_{g-1}(\overline{d})|G = g, D = d] - \mathrm{E}[Y_t(\overline{d}) - Y_{g-1}(\overline{d})|G = g, D = d] \\ &= \mathrm{E}[Y_t(g,d) - Y_{g-1}(\overline{d})|G = g, D = d] - \mathrm{E}[Y_t(\overline{d}) - Y_{g-1}(\overline{d})|G = g, D = \overline{d}] \\ &= \mathrm{E}[Y_t - Y_{g-1}|G = g, D = d] - \mathrm{E}[Y_t - Y_{g-1}|G = g, D = \overline{d}] \end{aligned}$$

The first equality is the definition of the $ATT$, the second equality follows by adding and subtracting $\mathrm{E}[Y_{g-1}(\overline{d})|G = g, D = d]$, and the third equality follows by Assumption 5. Finally, note that by linearity of the expectation operator and by assuming random sampling from repeated cross sections, each term in the final equation can be estimated by the observed sample averages.

*Nonidentification of the ACRT*

A key result from Callaway et al. (2024a) noted in the main text is that under standard assumptions the derivative is not identified by the $ATT$ derivative.

**Theorem Appendix B.2** *Under assumptions 1 to 5, the $ATT$ derivative recovers a mix of causal treatment effect parameters and selection bias terms. In particular,*

$$\frac{\partial \mathrm{E}[\Delta Y_t|G = g, D = d]}{\partial d} = \frac{\partial ATT(g,t,d|g,d)}{\partial d} = ACRT(g,t,d|g,d) + \underbrace{\frac{\partial ATT(g,t,d|g,l)}{\partial l}|_{l=d}}_{\text{selection bias}}.$$

*Proof:* First, notice that

$$\frac{\partial \mathrm{E}[\Delta Y_t|G = g, D = d]}{\partial d} = \frac{\partial ATT(g,t,d|g,d)}{\partial d} = \lim_{h \to 0} \frac{ATT(g,t,d + h|g,d + h) - ATT(g,t,d|g, d)}{h}$$

is not the $ACRT$, which is given by

$$ACRT(g,t,d|g,d) = \lim_{h \to 0} \frac{ATT(g,t,d + h|g,d) - ATT(g,t,d|g, d)}{h}.$$

Instead, we have

$$
\begin{aligned}
\lim_{h \to 0} \frac{ATT(g,t,d+h|g,d+h) - ATT(g,t,d|g,d)}{h} &= \lim_{h \to 0} \frac{ATT(g,t,d+h|g,d) - ATT(g,t,d|g,d)}{h} \\
&\quad + \lim_{h \to 0} \frac{ATT(g,t,d+h|g,d+h) - ATT(g,t,d+h|g,d)}{h} \\
&= ACRT(g,t,d|g,d) + \left. \frac{\partial ATT(g,t,d|g,l)}{\partial l} \right|_{l=d}
\end{aligned}
$$

However, we can recover identification of the $ACRT$ by the $ATT$ derivative with additional assumptions. In particular, consider the *strong parallel trends* assumption suggested in Callaway et al. (2024a).

**Assumption 6** (Strong parallel trends). *For all $g \in \mathcal{G}$, $t = 2,3,\dots,T$ and $d \in \mathcal{D}$,*

$$
\mathrm{E}[Y_t(d) - Y_t(\overline{d})|G = g] = \mathrm{E}[Y_t(d) - Y_t(\overline{d})|G = g, D = d].
$$

Assumption 6 says that units at all distance levels will have the same $ATT$ response function or $ATT(g,t,d|g,d) = ATT(g,t,d|g)$. This rules out any form of treatment effect heterogeneity along the distance margin. Under this assumption, the $ACRT$ is identified.

**Theorem Appendix B.3** *Under assumptions 1-6, the $ACRT(g,t,d|d)$ is identified. That is,*

$$
\begin{aligned}
\frac{\partial \mathrm{E}[\Delta Y_t|G = g, D = d]}{\partial d} = \frac{\partial ATT(g,t,d|g,d)}{\partial d} &= ACRT(g,t,d|g,d) + \underbrace{\frac{\partial ATT(g,t,d|g,l)}{\partial l}|_{l=d}}_{=0} \\
&= ACRT(g,t,d|g,d).
\end{aligned}
$$

*Proof:*

$$
\begin{aligned}
\frac{\partial ATT(g,t,d|g,d)}{\partial d} &= \lim_{h \to 0} \frac{ATT(g,t,d+h \mid g,d+h) - ATT(g,t,d \mid g,d)}{h} \\
&= \lim_{h \to 0} \frac{\mathrm{E}[Y_t(g,d+h) - Y_t(\overline{d}) \mid G = g, D = d+h]}{h} - \frac{\mathrm{E}[Y_t(g,d) - Y_t(\overline{d}) \mid G = g, D = d]}{h} \\
&= \lim_{h \to 0} \frac{\mathrm{E}[Y_t(g,d+h) - Y_t(\overline{d}) \mid G = g] - \mathrm{E}[Y_t(g,d) - Y_t(\overline{d}) \mid G = g]}{h} \\
&= \lim_{h \to 0} \frac{\mathrm{E}[Y_t(g,d+h) - Y_t(g,d) \mid G = g]}{h} \\
&= \lim_{h \to 0} \frac{\mathrm{E}[Y_t(g,d+h) - Y_t(g,d) \mid G = g, D = d]}{h} \\
&= ACRT(g,t,d|g,d)
\end{aligned}
$$

The first equality follows from the definition of the derivative. The second equality follows from the definition of the $ATT$. The third equality follows from the strong parallel trends assumption. The fourth equality uses the linearity of the expectation operator. The fifth equality uses the strong parallel trends assumption again. In the last line, the result follows by taking the limit as $h \to 0$ and the definition of the $ACRT$.

In the next few results, I consider alternative assumptions on treatment effect heterogeneity to identify or bound the $ACRT$. First, I show the $ACRT$ is identified by the rise over the run in the special case that units' potential outcomes are linear. Second, I show that an alternative assumption of concave counterfactual $ATT$ response functions imply that the estimator under linearity is less than or equal to the true $ACRT$. Third, I review a result where assuming selection on gains implies that the $ATT$ derivative is an upper bound

*Point identifying the $ACRT$ by assuming linearity*

**Assumption 7** (Linear potential outcomes). *Assume a housing unit $j$'s potential outcomes is linear in distance to the treated site and $\overline{d}$ is now the minimum distance where spillovers from the plant are no longer relevant. That is, for $D_j \in \mathcal{D}$ we have*

$$Y_{j,t} = \begin{cases} Y_{j,t}(\overline{d}) + \gamma_j (D_j - \overline{d}) & \text{if } D_j < \overline{d} \\ Y_{j,t}(\overline{d}) & \text{if } D_j \geq \overline{d} \end{cases}$$

*where $\gamma_j$ is a unit-specific slope coefficient.*

This assumption says that housing units' potential outcomes follow a linear correlated random coefficients model. Under this model, the average of the unit-specific slope coefficients $\gamma_j$ are the average causal response $ACRT$.

The following theorem shows that under Assumption 7 the alternative "rise over run" estimator identifies the $ACRT$.

**Theorem Appendix B.4** *Under assumptions 1 to 5 and Assumption 7, the $ACRT$ is identified. In particular, for $d \in [0, \overline{d}]$,*

$$\frac{ATT(g,t,d|g,d)}{d - \overline{d}} = ACRT(g,t,d|g,d).$$

*Proof:* First, recall the definition of the $ATT$ function:

$$ATT(g,t,d|g,d) = \mathrm{E}[Y_t(g,d) - Y_t(\overline{d})|G = g, D = d].$$

Now plugging in the linear potential outcomes from Assumption 7,

$$ATT(g,t,d|g,d) = \mathrm{E}[Y_t(\overline{d}) + \gamma\ (d - \overline{d}) - Y_t(\overline{d})|G = g, D = d]$$

$$\frac{ATT(g,t,d|g,d)}{d - \overline{d}} = \mathrm{E}[\gamma|G = g, D = d] = ACRT(g,t,d|g,d)$$

*Bounding the $ACRT$ by assuming concavity*

**Assumption 8** (Concave ATT response functions). *Assume the $ATT$ functions for each treatment group subpopulation $d' \in [0,\overline{d}]$ are concave on the interval $[d', \overline{d}]$. That is, for any $\lambda \in [0,1]$ we have*

$$ATT(g,t,(1-\lambda)d' + \lambda\overline{d}|g,d') \geq (1-\lambda)ATT(g,t,d'|g,d') + \lambda ATT(g,t,\overline{d}|g,d').$$

The next theorem shows the rise over the run is less than the $ACRT$ if the counterfactual $ATT$ response functions are concave. This follows from the "rooftop" theorem for 1-dimensional concave functions.

**Theorem Appendix B.5** *Under assumptions 1 to 4 and Assumption 8 we have*

$$\frac{ATT(g,t,d|g,d)}{d-\overline{d}} \leq ACRT(g,t,d|g,d).$$

*Proof:* The result follows immediately from the "rooftop" theorem for 1-dimensional concave differentiable functions. Starting from Assumption 8, we have

$$ATT(g,t,(1-\lambda)d' + \lambda\overline{d}|g,d') \geq (1-\lambda)ATT(g,t,d'|g,d') + \lambda ATT(g,t,\overline{d}|g,d').$$

Rearranging terms,

$$ATT(g,t,d' + \lambda(\overline{d}-d')|g,d') - ATT(g,t,d'|g,d') \geq \lambda(ATT(g,t,\overline{d}|g,d') - ATT(g,t,d'|g,d')).$$

Dividing both sides by $\lambda$,

$$\frac{ATT(g,t,d' + \lambda(\overline{d}-d')|g,d') - ATT(g,t,d'|g,d')}{\lambda} \geq ATT(g,t,\overline{d}|g,d') - ATT(g,t,d'|g,d').$$

Then,

$$(\overline{d}-d')\frac{ATT(g,t,d' + \lambda(\overline{d}-d')|g,d') - ATT(g,t,d'|g,d')}{\lambda(\overline{d}-d')} \geq ATT(g,t,\overline{d}|g,d') - ATT(g,t,d'|g,d').$$

Taking the limit as $\lambda \to 0$,

$$(\overline{d}-d')ACRT(g,t,d'|g,d') \geq ATT(g,t,\overline{d}|g,d') - ATT(g,t,d'|g,d').$$

Rearranging terms and noting that $ATT(g,t,\overline{d}|g,d') = 0$, we have

$$\frac{ATT(g,t,d'|g,d')}{d-\overline{d}} \leq ACRT(g,t,d'|g,d').$$

*Bounding the $ACRT$ by assuming selection on gains*

I follow Appendix SE.1 in Callaway et al. (2024a). I focus on the case assuming selection on gains approaching the treatment site. That is, for all distances $\widetilde{d}$ and $d'$ such that $\widetilde{d} < d'$ we have $ATT(g,t,d|g,\widetilde{d}) \geq ATT(g,t,d|g,d')$ for all $d$.

**Proposition 3** *Under assumptions 1-5 and suppose that for all distances $\widetilde{d}$ and $d'$ such that $\widetilde{d} < d'$ we have $ATT(g,t,d|g,\widetilde{d}) \geq ATT(g,t,d|g,d')$ for all $d$, then the following holds*

$$ACRT(g,t,d|g,d) \geq \frac{\partial ATT(g,t,d|g,d)}{\partial d} \text{ for all } d.$$

*Proof:* Recall that in Theorem 3.2 we had

$$\frac{\partial ATT(g,t,d|g,d)}{\partial d} = ACRT(g,t,d|g,d) + \underbrace{\frac{\partial ATT(g,t,d|g,l)}{\partial l}|_{l=d}}_{\text{selection bias}}.$$

The selection on gains assumption implies $\frac{\partial ATT(g,t,d|g,l)}{\partial l}|_{l=d} \leq 0$ always. Hence,

$$\frac{\partial ATT(g,t,d|g,d)}{\partial d} \leq ACRT(g,t,d|g,d).$$

Under selection on gains, the derivative of the observed $ATT$ will underestimate the $ACRT$ for being farther away from a treatment site.

## Appendix C Heteroskedasticity by distance to the treatment site

This appendix section considers the rationale for weighting $ATT$ estimates to improve precision by providing a few arguments for why we should expect heteroskedasticity to be a significant issue in this context.

In the main results of this paper, I report estimates weighting by an estimate of the subgroup $ATT$s' inverse variances. For an $ATT$ estimated around a site $g$ by bin level average house prices,

$$A\hat{T}T(g,2,d) = (\hat{\bar{Y}}_{g,2,d} - \hat{\bar{Y}}_{g,1,d}) - (\hat{\bar{Y}}_{g,2,\bar{d}} - \hat{\bar{Y}}_{g,1,\bar{d}})$$

where $\hat{\bar{Y}}_{g,t,d}$ is a sample bin mean with estimated variance $\hat{\sigma}^2_{g,t,d}$. I suggest an estimator for the $ATT$ estimate's variance:

$$var(AT\hat{T}(g,t,d)) = \hat{\sigma}^2_{g,2,d} + \hat{\sigma}^2_{g,1,d} + \hat{\sigma}^2_{g,2,\bar{d}} + \hat{\sigma}^2_{g,1,\bar{d}}$$

where all terms can be estimated using the standard errors of the means in the data. Then, the regression weights are

$$\hat{w}(g,2,d) = \frac{1/var(AT\hat{T}(g,2,d))}{\sum_{g \in \mathcal{G}} 1/var(AT\hat{T}(g,2,d))}.$$

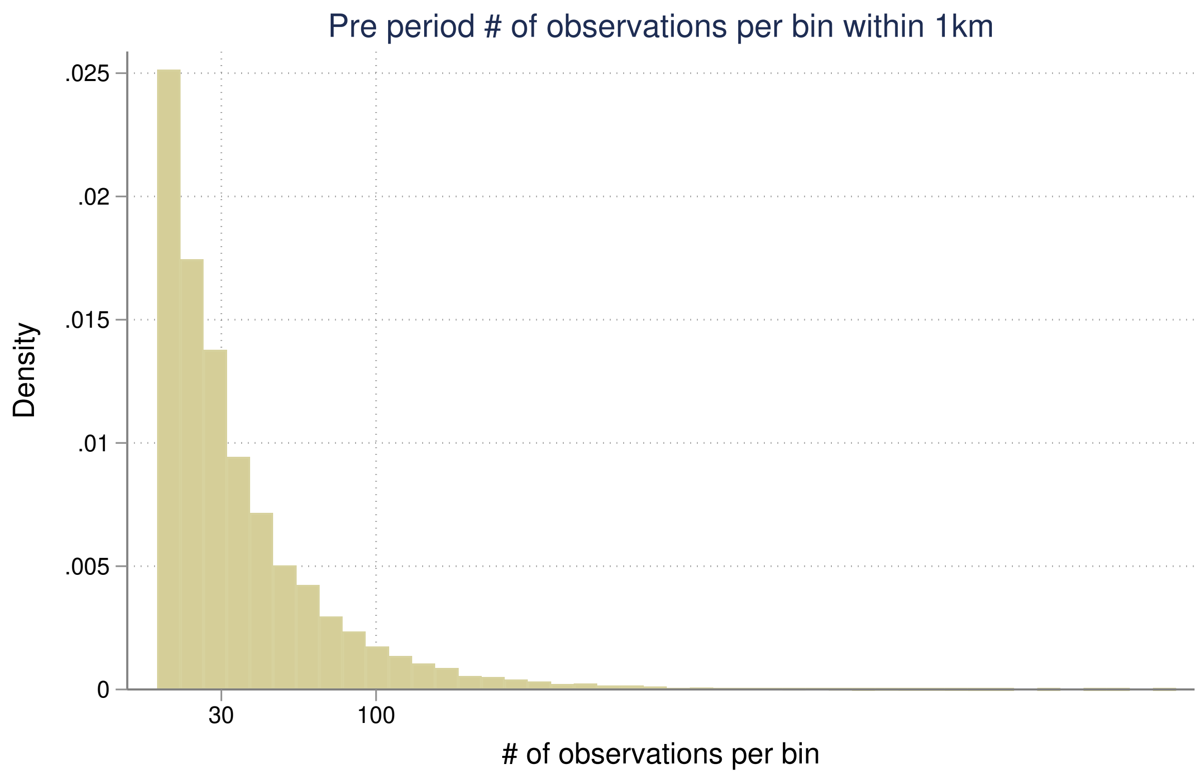
While weighting to target precision can sometimes do more harm than good, Solon et al. (2015) provides some practical guidance for when weighting is more likely to be beneficial. In particular, if the within group sample size is highly variable and small for some groups, then weighting in proportion to the within group sample size can significantly improve precision.

I provide three pieces of evidence for weighting to target precision in this context. First, I argue that we should expect the data generating process to exhibit heteroskedasticity on proximity to the site $d$. I then show evidence that housing sales are in fact sparsely populated close to the TRI industrial plant sites. Finally, I follow the suggestion in Solon et al. (2015) to test for the presence of heteroskedasticity using a modified Breusch-Pagan test. This test strongly indicates the presence of heteroskedasticity in this setting.

The area of a circle scales geometrically with its radius, suggesting we should expect the number of observations to be larger for distance bins farther away from the treatment site. Consider a sequence of radii of a circle $\{d_1, d_2, \ldots, d_L\}$ that characterize the distance bins used in the main paper. Consider these bins as equally spaced so $d_2 - d_1 = m, d_3 - d_2 = m$, and so on where $m$ is constant. The area of a ring of a circle is $\pi(d^2_{i+1} - d^2_i) = \pi((d_i + m)^2 - d^2_i) = \pi(d^2_i + 2d_i m + m^2 - d^2_i) = \pi(2d_i m + m^2) = \pi 2 d_i m + m^2$. Hence, the area of the distance bins will scale at a linear rate. If on average housing sales are distributed approximately uniform across geographic space, we would expect the number of observations per bin to also scale linearly. Panel B of Table 1 is suggestive of this linear relationship as the number of observations increases for the distance ranges farther away from the treated sites. Similarly, the PDF of housing units used in the welfare calculation shows a linear relationship (Figure 10).

Figure C1 is a PDF of the number of observations per bin within 1 kilometer of a TRI industrial plant in the pre period. A significant share of the distance bins have a small number of housing sales underlying the bin level average house price. For this application of spatial DiD based around a point, it is the case that the within group sample size is highly variable and small for many groups.

Figure C1: PDF number of observations in the pre period by distance to the treated site, TRI




*Note: This figure shows the PDF of the number of observations in the pre period for distance bins within 1 kilometer of a TRI industrial plant.*

We can also test for the presence of heteroskedasticity empirically. Solon et al. (2015) suggest a modified Breusch-Pagan test. In particular, for the $ATT$ estimates surrounding the sample of TRI industrial plant sites I regress,

$$A\hat{T}T(g,t,d) = \alpha_0 + \alpha_1 d_{g,d} + u_{g,d}$$

and then get the predicted residuals $\hat{u}_{g,d}$. In a second stage, take the pre period number of obs in the bin as the within group sample size $J_{g,d}$, and regress $\hat{u}^2_{g,d}$ on the inverse of the within group sample size

$$\hat{u}^2_{g,d} = \beta_0 + \beta_1 \frac{1}{J_{g,d}} + \epsilon_{g,d}.$$

The significance of the t-stat for the coefficient $\beta_1$ indicates whether there is significant evidence of heteroskedasticity. For TRI industrial plants, the estimated coefficient from the second regression is $\hat{\beta}_1 = 1.14$ with $t = 173.9$, which suggests a high degree of heteroskedasticity.

## Appendix D Proofs for results in hedonic model section

I restate the results from Section 4 and provide proofs.

**Proposition 1** *Suppose assumptions 1 to 5, Assumption 9, and the standard hedonic model hold. Then*

$$\mathrm{E}[\frac{\partial u_{t=2}}{\partial d}|D=d] = \mathrm{E}[\frac{\partial Y_{t=2}}{\partial d}|D=d] = ACRT(d|d).$$

*Section 3 suggested three alternative assumptions for identifying or bounding the $ACRT$:*
*(a) If strong parallel trends holds, then*

$$\mathrm{E}[\frac{\partial u_{t=2}}{\partial d}|D=d] = \frac{\partial ATT}{\partial d}(d|d).$$

*(b) If instead linear potential outcomes holds, then*

$$\mathrm{E}[\frac{\partial u_{t=2}}{\partial d}|D=d] = \frac{ATT(d|d)}{d-\bar{d}}.$$

*(c) If instead concave $ATT$ functions holds, then*

$$\mathrm{E}[\frac{\partial u_{t=2}}{\partial d}|D=d] \geq \frac{ATT(d|d)}{d-\bar{d}}.$$

*Proof:* To start, take expectations on both sides of Rosen (1974)'s FOC given by Equation (3) in the main text.

$$\mathrm{E}[\frac{\partial u_2(d,\boldsymbol{z_{-d}})}{\partial d}|D=d] = \mathrm{E}[\frac{\partial Y_2(d,\boldsymbol{z_{-d}})}{\partial d}|D=d]$$

where the condition holds in the post period $t=2$. I will show that $ACRT(d|d) = \mathrm{E}[\frac{\partial Y_2(d,\boldsymbol{z_{-d}})}{\partial d}|D=d]$. Starting with the definition of the $ACRT$,

$$\begin{aligned}
ACRT(d|d) &= \left.\frac{\partial ATT(l|d)}{\partial l}\right|_{l=d} = \left.\frac{\partial}{\partial l}[ATT(l|d)]\right|_{l=d} \\
&= \left.\frac{\partial}{\partial l}\left[\mathrm{E}[Y_2(l)-Y_2(\bar{d}) \mid D=d]\right]\right|_{l=d} \\
&= \left.\frac{\partial}{\partial l}\left[\mathrm{E}[Y_2(l) \mid D=d]\right]\right|_{l=d}
\end{aligned}$$

where the first and second lines follow from the definitions of the $ACRT$ and $ATT$, and the final line follows because $Y_2(\bar{d})$ does not depend on $l$. Now, using Assumption 9,

$$\left.\frac{\partial}{\partial l}\left[\mathrm{E}[Y_2(l) \mid D=d]\right]\right|_{l=d} = \frac{\partial}{\partial l}[\mathrm{E}[m(l,\boldsymbol{z_{-d}})+h_2(\boldsymbol{z_{-d}})|D=d]]|_{l=d}$$

It follows that

$$\frac{\partial}{\partial l}[\mathrm{E}[m(l,\boldsymbol{z_{-d}})+h_2(\boldsymbol{z_{-d}})|D=d]]|_{l=d} = \frac{\partial}{\partial l}[\mathrm{E}[m(l,\boldsymbol{z_{-d}})|D=d]]|_{l=d}$$

because $h_2(.)$ does not depend on $l$. Finally, we can interchange the partial derivative and the expectation because of the regularity assumptions in Assumption 3:

$$\begin{aligned}
\left.\frac{\partial}{\partial l}\left[\mathrm{E}[m(l,\boldsymbol{z_{-d}}) \mid D=d]\right]\right|_{l=d} &= \mathrm{E}\left[\left.\frac{\partial m(d,\boldsymbol{z_{-d}})}{\partial d}\right| D=d\right] \\
&= \mathrm{E}\left[\left.\frac{\partial Y_2(d)}{\partial d}\right| D=d\right]
\end{aligned}$$

where the last line follows from Assumption 9. Hence,

$$\mathrm{E}[\frac{\partial u_2}{\partial d}|D = d] = \mathrm{E}[\frac{\partial Y_t}{\partial d}|D = d] = ACRT(d|d).$$

The subparts (a), (b), and (c) follow immediately from the theorems in Section 3. Adding Assumption 6 to the assumptions noted at the beginning of the lemma, Part (a) follows from Theorem 3.3. Instead adding Assumption 7 to the assumptions noted at the beginning of the lemma, Part (b) follows from Theorem Appendix B.4. Instead adding Assumption 8 to the assumptions noted at the beginning of the lemma,Part (c) follows from Theorem 3.5.

**Lemma Appendix D.1** *Suppose assumptions 1 to 5, Assumption 9, the standard hedonic model hold, and homeowners' utility functions are log linear according to Equation (4). Then*

$$\mathrm{E}[WTP|D = d] = (1 + d)ACRT(d|d)\log(\frac{1 + \overline{d}}{1 + d}).$$

*As in Proposition 1(a),(b), and (c), depending on the assumption imposed, one of the two estimators identify or bound the $ACRT$, and the resulting expression on the right hand side identifies or bounds the average willingness to pay by distance.*

*Proof:* This result follows from Proposition 1 and the assumed functional form on homeowner's utility functions in Equation (4). In the main text, the assumed functional form yields

$$WTP_i = \beta_{i,d}\log(\frac{1 + \overline{d}}{1 + d_i^*}) = (1 + d_i^*)\frac{\partial Y_{j,t}}{\partial d}\log(\frac{1 + \overline{d}}{1 + d_i^*}).$$

Then, taking expectations conditional on d,

$$E[WTP|D = d] = (1 + d)\mathrm{E}[\frac{\partial Y_t}{\partial d}|D = d]\log(\frac{1 + \overline{d}}{1 + d}) = (1 + d)ACRT(d|d)\log(\frac{1 + \overline{d}}{1 + d})$$

where the last equality follows by Proposition 1.

**Proposition 2** *Suppose assumptions 1 to 5, Assumption 9, the standard hedonic model hold, and homeowners' utility functions are log linear according to Equation (4). Then the aggregate willingness to pay for the plant not to exist is*

$$\text{Agg WTP} = N_H\int_0^{\overline{d}}(1 + d)ACRT(d|d)\log(\frac{1 + \overline{d}}{1 + d})f(d)\,\mathrm{d}d.$$

*Proof:* This result follows by integrating homeowners' WTPs over $[0,\overline{d}]$ and applying Lemma 4.1. Note that this assumes the functional form on homeowner's utility functions in Equation (4). In the main text, we had

$$\text{Agg WTP} = \sum_i WTP_i = N_H\int_0^{\overline{d}}\mathrm{E}[WTP|D = d]f(d)\,\mathrm{d}d.$$

Applying Lemma 4.1,

$$\text{Agg WTP} = N_H\int_0^{\overline{d}}(1 + d)ACRT(d|d)\log(\frac{1 + \overline{d}}{1 + d})f(d)\,\mathrm{d}d.$$

**Theorem Appendix D.2** *Suppose assumptions 1 to 5, Assumption 9, the standard hedonic model, and homeowners' utility functions are log linear according to Equation (4). Under concave ATT functions, homeowners' aggregate willingness to pay for the plant to not exist is bounded:*

$$N_H \int_0^{\bar{d}} (1+d) \frac{ATT(d|d)}{d-\bar{d}} \log\left(\frac{1+\bar{d}}{1+d}\right) f(d)\,\mathrm{d}d \quad \leq \text{Agg WTP}$$
$$\leq -N_H \int_0^{\bar{d}} ATT(d|d) f(d)\,\mathrm{d}d.$$

*Proof:* I start with the lower bound. The lower bound follows from Theorem 3.5, the parametric assumption on homeowners' utility functions in Equation (4), and the properties of integrals. Under Assumption 8, by Theorem 3.5 we have $ACRT(d|d) \geq \frac{ATT(d|d)}{d-\bar{d}}$ for all $d \in [0,\bar{d}]$. Because $(1+d)\log\left(\frac{1+\bar{d}}{1+d}\right)$ is positive and using the properties of integrals,

$$\begin{aligned} & N_H \int_0^{\bar{d}} (1+d) \frac{ATT(d|d)}{d-\bar{d}} \log\left(\frac{1+\bar{d}}{1+d}\right) f(d)\,\mathrm{d}d. \\ & \leq N_H \int_0^{\bar{d}} (1+d) ACRT(d|d) \log\left(\frac{1+d}{1+\bar{d}}\right) f(d)\,\mathrm{d}d \\ & = \text{Agg WTP} \end{aligned}$$

where the last line follows from Proposition 2.

I now make the argument for the upper bound. First, note that the canonical hedonic model assumes homeowners are in equilibrium after making a decision from a complete, continuous product choice set. That is, homeowners can freely choose any house across the support of $\boldsymbol{z}$ to maximize utility. In hedonic equilibrium, I will show that the price difference is greater than each homeowner $i$'s willingness to pay for the plant to not be there. That is, $Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}}) \geq WTP_i$ for any $i$ and $d^*_i \in [0,\bar{d}]$ where other characteristics $\boldsymbol{z^*_{i,-d}}$ are fixed at each homeowner's equilibrium choice.

Assume to the contrary that $Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}}) < WTP_i$ for some $i$. A complete, continuous product choice set implies the homeowner could choose $(\bar{d}, \boldsymbol{z^*_{i,-d}})$ at the price $Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}})$. From this alternative product bundle they would receive utility $WTP_i - [Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}})]$ relative to their bundle chosen in equilibrium. But $WTP_i - [Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}})] > 0$ because $Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}}) < WTP_i$. Hence, they would have a profitable deviation from their chosen bundle by choosing a house with characteristics $(\bar{d}, \boldsymbol{z_{i,-d}}^*)$, and this can not represent a hedonic equilibrium. Therefore, in equilibrium, we must have $Y_t(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_t(d^*_i, \boldsymbol{z^*_{i,-d}}) \geq WTP_i$ for all $i$.

The $ATT$ identifies average price differences in the post period. $ATT(d|d) = \mathrm{E}[Y_{t=2}(d) - Y_{t=2}(\bar{d})|D = d]$. Notice that

$$\begin{aligned} Y_{t=2}(\bar{d}, \boldsymbol{z^*_{i,-d}}) - Y_{t=2}(d^*_i, \boldsymbol{z^*_{i,-d}}) &\geq WTP_i \ \ \forall\, i \\ \implies \mathrm{E}[Y_{t=2}(\bar{d}, \boldsymbol{z_{-d}}) - Y_{t=2}(d, \boldsymbol{z_{-d}})|D = d] &\geq \mathrm{E}[WTP|D = d] \\ \mathrm{E}[Y_{t=2}(\bar{d}) - Y_{t=2}(d)|D = d] &\geq \mathrm{E}[WTP|D = d] \\ -\mathrm{E}[Y_{t=2}(d) - Y_{t=2}(\bar{d})|D = d] &\geq \mathrm{E}[WTP|D = d] \\ -ATT(d|d) &\geq \mathrm{E}[WTP|D = d] \end{aligned}$$

where the first line follows taking expectations on both sides, the second line follows because a change in $d$ holding all other characteristics constant is the potential outcome, the third line factors out $-1$, and the final line is by the definition of the $ATT$. By the properties of integrals,

$$\begin{aligned}
&-N_H \int_0^{\overline{d}} ATT(d|d) f(d)\, \mathrm{d}d \\
&\geq N_H \int_0^{\overline{d}} \mathrm{E}[WTP|D=d] f(d)\, \mathrm{d}d \\
&= \text{Agg WTP}.
\end{aligned}$$

Finally, putting together the two bounds,

$$\begin{aligned}
&N_H \int_0^{\overline{d}} (1+d) \frac{ATT(d|d)}{d-\overline{d}} \log\left(\frac{1+\overline{d}}{1+d}\right) f(d)\, \mathrm{d}d \quad \leq \text{Agg WTP} \\
&\leq -N_H \int_0^{\overline{d}} ATT(d|d) f(d)\, \mathrm{d}d.
\end{aligned}$$