Do Environmental Right-to-Know Laws Affect Markets? Capitalization of Information in the Toxic Release Inventory

Ralph Mastromonaco

April 2, 2013

Department of Economics
University of Oregon
533 Prince Lucien Campbell Hall
1285 University of Oregon
Eugene, OR 97403
ralphm@uoregon.edu
Phone: 541-346-4671
Fax: 541-346-1243
Abstract

This paper investigates how information contained in the U.S. Environmental Protection Agency’s Toxic Release Inventory (TRI) program, one of the largest environmental right-to-know programs, affects prices in the housing market. I use a reduction in the reporting requirements for several chemicals in 2000 and 2001 as a quasi-experiment to test for housing price changes near existing firms who must now report. Using a difference-in-differences specification, I find that listing a firm in the Toxic Release Inventory as a lead user lowers nearby housing prices up to 2.5%. The results suggest that housing market participants do capitalize into prices the information contained in the TRI.

Keywords: Right-to-know laws, Toxic Release Inventory, quasi-experimental, difference-in-differences, environmental quality, hedonics, risk perceptions
1 Introduction

Public disclosure laws are designed to provide the public with information not normally included in the free exchange of goods and services. These “right-to-know” laws have been called for in many areas of the economy. As of 2008, New York City required all chain restaurants to disclose caloric content on restaurant menus with the intent to combat obesity by simply requiring that information be provided to consumers. This requirement has been expanded nationally by the Patient Protection and Affordable Care Act, however, several studies have found no evidence that caloric content on menus affects total purchased calories (Dumanovsky et al. (2011); Swartz et al. (2011)). In California, Proposition 37, on the November 2012 ballot, would have required all genetically modified foods to be labeled as “GMO”: genetically modified organisms. Proponents argued that consumers have a right to know what they are eating. Opponents countered that the GMO label would frighten and mislead consumers, citing an American Medical Association report affirming the lack of scientific evidence differentiating GMO from non-GMO (Morisy (2012)). Understanding how consumers process information that they have a “right-to-know”, both on a detailed, continuous scale, like caloric content, and on a discrete scale, like a GMO label, is an important issue for policy makers in many areas.

Public disclosure laws have been an important component of environmental policy for several decades (Konar and Cohen (1997)). While right-to-know environmental regulations have roots in ethical, legal, and medical arguments, economists dating back to Coase (1960) have argued that public disclosure of the type and quantity of pollution can reduce deadweight losses. The “Coase Theorem” states, in part, that with full information and no transactions costs, bargaining between the generator of an externality and those that bear the burden will result in an efficient outcome. Proponents of disclosure laws cite the spirit of Coase theorem, arguing that by using disclosure, pollution can be reduced using market-
based incentives (as opposed to comparatively expensive command-and-control regulation) as polluting firms face pressure to abate from an informed public. Critics argue that disclosed information is not easily understood by the public, is either ignored completely or misunderstood, and comes at great cost to firms; arguments similar to those made by opponents of Proposition 37. The effectiveness of any disclosure policy hinges on how consumers and households use the disclosed information. Accordingly, understanding how information about environmental and neighborhood amenities influences household behavior remains an active area of research.¹

Perhaps the most prominent right-to-know law is the Emergency Planning and Community Right to Know Act (EPCRA), passed by Congress on the heels of the Union Carbide disaster in Bhopal, India.² EPCRA created the Toxic Release Inventory (TRI), which requires certain firms to report annual emissions of toxic chemicals, to provide transparency about the presence, type, and quantity of hazardous chemicals to the communities that were most likely to be impacted by their release. Policy makers designed the TRI partly based on the “market-based regulation” idea: if the EPA provides detailed, facility-level information on toxic chemicals and emissions to the public, firms will be incentivized to reduce the amount of pollution they produce via public pressure. The goal of pollution reduction would be hard to attain if households do not make use of the data provided in the TRI. The goal of this paper is to further the understanding of how information is utilized by the households most likely to be impacted by toxic releases.

Since the inception of the TRI, toxic emissions have fallen in the United States. For example, from 1989 to 1999, emissions in the U.S. have fallen 40 percent.³ The EPA reports that disposal and releases of covered chemicals have fallen approximately 30 percent

²EPCRA is also known as the “Superfund Amendments and Reauthorization Act”.
³See Bui and Mayer (2003) for discussion.
between 2001 and 2010 (U.S. Environmental Protection Agency (2010)). However, evidence that the public internalizes information on toxic emissions, for example in the housing and stock markets, is mixed.⁴ Accordingly, it is difficult to claim that emissions are falling as a result of public pressure if its unclear that households and investors respond to emissions data.

In light of the mixed evidence on household reaction to emissions data, the goal of this research is to determine whether households react to information content in the TRI not directly related to emissions. Part of the original intent of EPCRA was to inform households about the presence of toxic chemicals in their communities and to prepare them for the possibility of an accidental release of hazardous materials. Site reporting requirements to the TRI are based upon onsite quantities of reportable chemicals. Having a nearby firm listed in the TRI informs households of the quantity and types of emissions as well as the threat of a potential accidental spill of chemicals. If households are more sensitive to living near TRI facilities for fear of catastrophe rather than chronic exposure to toxic air emissions, evaluating the impact of the TRI program solely on emissions data might be insufficient. Furthermore, if there is significant stigma or public pressure associated with exceeding the reporting thresholds for the TRI, the measured reduction in aggregate emissions seen in the data may be a result of numerous firms’ incentives to reduce their chemical usage to just below the reporting requirements.

Generally speaking, the existing literature on TRI emissions valuation can suffer from two empirical problems.⁵ First, basic cross-sectional hedonic analysis that measures the implicit price for emissions or TRI site proximity could be subject to omitted variables bias if unobserved housing or neighborhood quality is spatially correlated with TRI firm locations. Second, panel data models that try to difference away this unobserved heterogeneity might

---

⁴ See Hamilton (1995); Khanna et al. (1998); Bui (2005); Bui and Mayer (2003); Banzhaf and Walsh (2008); Konar and Cohen (1997)

⁵ One notable exception is Sanders (2011), who also employs a difference-in-differences estimator
produce a corresponding decrease in variation in the data. Cross-sectional variation in emissions might be especially salient to households whereas year-to-year changes in emissions might not be detectable or important to residents. These issues might produce insignificant estimates of the coefficient on TRI site proximity or emissions exposure in a typical hedonic property value model, leading to conclusions that housing markets do not capitalize information in the TRI.

To help clarify the effect of information in the TRI on households, I use discontinuous changes in reporting thresholds for several of the most toxic chemicals to design a quasi-experimental empirical model. In 2000, the EPA lowered the reporting requirement for manufacturing or possessing “Persistent, Bioaccumulative and Toxic” chemicals (PBTs). In general, these thresholds were reduced from thousands of pounds per year to between ten and one hundred pounds per year. In 2001, lead and lead compounds were designated as PBTs and the threshold for reporting was lowered accordingly, from ten thousand pounds per year to one hundred pounds per year. As a result, firms that were using more than the new threshold but less than the old threshold were no longer exempt from TRI reporting. If housing market participants utilize the information provided in the TRI program, it is likely that these shocks to information sets will have corresponding effects on housing prices.

I use a difference-in-differences (DID) estimator to exploit this regulatory change to test for whether the listing of existing firms using the affected chemicals that previously did not report to the TRI affects the prices of houses in proximity to those firms. In this setup, houses in proximity to TRI sites that were not affected by the changes in reporting thresholds serve as a control group. The treatment group are houses in similar proximity to a firm that had to report to the federal program for the first time as a result of the change in regulations. By constructing a panel dataset of housing transactions, I am able to improve the performance of the DID estimator by controlling for house-specific unobserved quality.

The TRI database is provided by EPA and is supplemented by data on point-source
emissions from the California Air Resources Board (CARB). This important supplemental database allows me to identify facilities that were previously in existence but did not have to report to the TRI program. Identification in the quasi-experiment requires knowledge of whether facilities that appear for the first time in the TRI data after the regulatory change were in fact already in existence. The TRI data reports whether a facility possesses the affected chemicals and I match these facilities to the CARB data to determine the earliest reported date of existence for that firm. This provides a set of “treated” sites for the quasi-experimental approach.

Using micro-data on individual housing transactions in the San Francisco-Oakland-San Jose Metropolitan Statistical Area, I find that the provision of new information in the form of TRI reporting does have a significant and negative impact on housing prices. The release of data identifying a plant as a TRI plant that uses lead lowers prices within three kilometers by up to 2.5%. The regulatory change for PBTs in 2000 is less robust, but there is some evidence that housing prices were negatively affected by this change as well. This difference could have to do with the salience of the toxicity of lead and the relative obscurity of the other chemicals initially designated as PBTs.

These results suggest that information about the local presence of toxic chemicals is indeed capitalized into housing prices. Interestingly, since most of the “treated” sites actually have very little to no annual emissions, the results concur with a study by Bui and Mayer (2003) that finds changes in housing prices are not related to changes in TRI emissions. However, my results imply a different conclusion about the relationship between the TRI program and the housing market. While levels and changes in emissions in the TRI do not appear to be strongly related to housing prices, information in the TRI about the types of chemicals being housed in large quantities on-site does have an impact on housing prices.

The implication of this and previous research is that firms may have perverse incentives not necessarily to reduce emissions, but rather to avoid the reporting requirements alto-
together. As a result, large declines in the amount of emissions nationwide reported in the TRI could be a result of firms striving to get underneath reporting thresholds. Toxic air emissions fall as a consequence of the quantities of reportable on-site chemicals being reduced.

## 2 Toxic Release Inventory Background

Section 313 of the Emergency Planning and Community Right-to-know Act required the EPA to create the Toxic Release Inventory by collecting and publicizing information relating to the possession and release of certain toxic chemicals by facilities in certain industries. The legislation was passed on the heels of the Union Carbide disaster in Bhopal, India, where a leak of methyl isocyanate gas was responsible for the deaths of thousands of people. The TRI serves the dual purpose of informing local emergency planning officials of the specific potential risks at covered facilities as well as informing the public at large, who were given the “right-to-know” by EPCRA.

The TRI program, however, is simply an accounting and reporting program; there are no limits or controls on the releases of chemical compounds.\(^6\) For each reporting year, eligible firms are required to file, by July 1st of the following year, a “Form R,” detailing their chemical use profile. The TRI was seen as a market-based regulation as opposed to a command-and-control regulation. Instead of government regulators dictating the types and quantities of chemicals that could be released, the goal was to reduce emissions by shedding light on the activities of firms, who would in turn react to public pressure by reducing emissions.

Data under the TRI program first became available in 1987. Since then, there have been

---

\(^6\) However, several chemicals that firms are required to report under the TRI program are regulated under other statutes, such as the Clean Air Act Amendments (CAAA) or the Resource Conservation and Recovery Act (RCRA).
four major changes to the program. First, the “Phase 1” expansion added 286 chemicals to the inventory list, bringing the total to 602 chemicals. Second, “Phase 2” expanded the range of facilities to include non-manufacturing facilities within the same industry classification codes. At the time, EPA estimated that this would require an additional 6,000 new facilities to begin participation in the program. Phase 1 became effective in the 1995 reporting year data and Phase 2 became effective in the 1998 reporting year data. Third, effective in the 2000 reporting year, EPA reduced the reporting threshold for PBTs from several thousands of pounds per year to between ten and one hundred pounds per year, depending on the chemical. Lastly, effective in the 2001 reporting year, the minimum threshold for usage of lead and lead compounds was reduced from 10,000 pounds per year to 100 pounds per year. These last two regulatory changes are the focus of this empirical analysis.

3 Literature Review

Several papers have turned their attention to the Toxic Release Inventory and its relationship to the housing market. Bui and Mayer (2003) study the relationship between changes in emissions and changes in housing prices in Massachusetts. Using a first-differenced hedonic approach, they find no significant statistical relationship between changes in emissions and prices from 1987 to 1992. Their results call into question the notion that public pressure

---

7In addition, there was a voluntary emissions reduction program, the “33/50” initiative, coordinated under the TRI program from 1991-1995. Gamper-Rabindran (2006) found little evidence that this program effectively reduced emissions.

8Also effective in 2000, the courts ruled phosphoric acid was not subject to reporting under EPCRA, and was dropped from TRI.

9A separate group of papers have focused on how stock market returns and firms were affected by information on releases. Hamilton (1995) finds that stock market returns on the day TRI data were initially released were negatively correlated with emissions quantity. Khanna et al. (1998) finds that negative stock market returns were correlated with decreases in reported emissions over time but an increase in off-site waste disposal. Bui (2005) discusses an econometric issue that could affect the robustness of the findings of the previous studies while finding no significant relationship between stock market returns and emissions for petroleum firms.
has led to the observed decline in reported TRI emissions since the program’s inception. Oberholzer-Gee and Mitsunari (2006) examine the effect of the initial release of the TRI data on local residents’ risk perceptions. Using micro-data instead of aggregate price data, they find that prices in very close proximity to TRI facilities were significantly affected by the revelation of the quantity of emissions released.

In a paper closely related to this one, Sanders (2011) examines the impact of the addition of new industries to the TRI program that took place in 1998. The author also employs a difference-in-differences estimator, testing for changes in zip-code level median prices that are attributable not to a change in environmental quality, but a change in information. However, there are important differences between his study and this one that should be enumerated. First, by defining treatment as a zip-code with an industry that must report for the first time in 1998 that also sees a substantial increase in reported annual emissions, the estimated capitalization effects confirm that large increases in reported emissions are noticed by the housing market. The test in this study ask a more general question: does simply being listed in the TRI, regardless of the amount of emissions, have an impact on local housing prices? Second, I employ housing level micro-data that allows me to explicitly consider the relationship between capitalization effects and distance to relevant treatment sites. Gamper-Rabindran and Timmins (2012) demonstrate in the context of Superfund sites that, within a census tract, houses with prices at the median are farther from nearby waste sites than houses with prices in lower percentiles. Third, I explicitly consider the lag period between policy announcement and policy effect date, demonstrating that the market can adjust prior to the release of data. In spite of the differences between the two papers, our results are complementary, and reinforce our separate conclusions.

Currie et al. (2013) examine the assumption of full information in hedonic pricing models by utilizing the TRI program. Their research is guided by similar intuition: If households are unaware of the presence of toxic facilities, then associated price changes cannot be related
to WTP for environmental quality. Rather than looking at how information contained in the TRI affects housing prices, they examine whether opening and closing of plants that were ever included in the TRI have an effect on prices and determine that the operation of toxic facilities negatively impact housing prices. While they do not explicitly control for the release of TRI data, their results and conclusions support my findings of negative price effects.

Brooks and Sethi (1997) examine the characteristics of communities that have higher exposure to TRI emissions. Their results indicate that communities with higher levels of emissions tend to have higher proportions of black residents, less active voters, more renters, and lower educated households. Banzhaf and Walsh (2008) examine the “Tiebout Hypothesis” which states that households will “vote with their feet” and move to the location that provides the best combination of prices and public goods. Looking at the changes in TRI emissions and changes in population between 1990 and 2000, they find robust evidence that population flows are positively correlated with reductions in pollution. In the context of their model, if households place a negative value on toxic air emissions, they will migrate to areas that have lower amounts of toxic air. Their results suggest that TRI emissions do enter household decision making.

Lastly, hedonic models using quasi-experimental designs and panel data are increasingly being used in the valuation of environmental quality. Chay and Greenstone (2005) use a mix of these approaches to estimate marginal willingness to pay for air quality. Greenstone and Gallagher (2008) use a regression discontinuity design based on the Hazardous Ranking Score assigned to Superfund sites at the inception of the program and finds that listing sites on the National Priorities List had no significant impact on prices. However, Kuminoff and Pope (2012) have shown that hedonic valuation studies that use non-marginal changes in amenities over time can be subject to “conflation bias” if the assumption of a time-invariant hedonic price function is violated. As a result, the capitalization effects documented in
quasi-experimental econometric models that use temporal variation in amenities might not translate into estimates of marginal willingness to pay. Since this paper employs such econometric techniques, I am careful not to make statements about marginal willingness to pay. Rather, I maintain that capitalization effects, whether attributable to changing hedonic price functions or a negative willingness to pay for proximity to TRI sites, are evidence of the use of information provided in the Toxic Release Inventory.

4 Data

The data used in this paper come from several sources. First, TRI data is taken directly from EPA.\textsuperscript{10} Second, facility location and existence data is supplemented by data from the California Air Resource Board, pursuant to the Air Toxics “Hot Spots” Emissions Inventory program.\textsuperscript{11} Lastly, housing transactions data were purchased from Dataquick Information Systems. In this section, I describe each of these datasets in turn.

4.1 TRI Data

The TRI database contains the quantities, types, and release pathways (air, water, off-site, etc.) of toxic chemicals released by reporting facilities, as well as the latitude and longitude of those facilities. While this information is self-reported to the EPA, the agency has the ability to levy civil penalties for violations of ECPRA, and can force the rectification of those violations. TRI facility release and location data is available for each year from 1987 to 2009. However, not every facility reports in every year. This could be a result of production activities ceasing, production activities being reduced such that the total quantity of toxic chemicals falls below the threshold, or a plant switching production processes or outputs so that it no longer uses these toxic chemicals. As a practical matter, I will treat each facility

\textsuperscript{10}Data available at http://www.epa.gov/tri/tridata/data/basicplus/index.html
\textsuperscript{11}Searchable database at http://www.arb.ca.gov/ei/disclaim.htm
as remaining in existence in years between the first observed year and the last observed year. Using the longitude and latitude of each plant, I can match the facilities panel data set to the housing data set.

The accuracy of reported emissions in the TRI have been met with skepticism. For example, de Marchi and Hamilton (2006) conclude that reductions in reported emissions in the TRI cannot be reconciled against data reported by EPA ambient pollution monitors. Moreover, they specifically single out lead as an unreliable chemical in terms of accurate reporting. These problems have contributed to the difficulty researchers have had in identifying statistical relationships between emissions and prices. The distrust in reported emissions levels amongst practitioners enhances the appeal of my empirical design. The econometric model does not rely on reported emissions. Instead, my research design depends only upon the presence of a plant in the TRI data.

Two important features of the TRI data need to be highlighted. First, the cycle of data collection and data release by the EPA results in TRI data being made public several months after the end of the reporting year. Reports are generally due to the EPA by July of the following year. According to the EPA, the reporting year 2000 data were released in May, 2002, and the reporting year 2001 data were released in June, 2003. The policy change that affected PBT reporting for the year 2000 was announced on October 29, 1999, and the policy change that lowered the lead threshold was announced January 17, 2001. These dates correspond to the days that the final rules were each entered into the Federal Register. I separately control for the possible effects that the announcement and the data release can have on housing prices.

4.2 CARB Data

The California Air Resource Board (CARB) maintains a database of facilities and point source emissions similar to the TRI. This database was created as part of California’s Air
Toxics “Hot Spots” Information and Assessment Act. This legislation has many of the same stated goals as the TRI program, namely to identify toxic chemicals in communities and to inform the public. Generally speaking, any facility that manufactures, uses or releases toxic chemicals and also releases more than 10 tons of criteria air pollutants must file with this state level inventory program. The dataset, which provides comprehensive data starting in 1996, allows me to identify facilities that were in existence prior to their first reporting date in the TRI. This is crucial for identifying the set of firms which were in existence prior to the threshold reductions for PBT and lead. Treatment sites are those that are identified as reporting to CARB prior to the threshold reductions and reporting to TRI, for the first time, after the threshold reductions.

Merging the CARB data with the TRI data based on facility name and address, I was able to match 546 of the 840 TRI sites in my sample. Of those 546 sites, 21 sites existed in the CARB data prior to 2001 and reported for the first time in 2001 in the TRI as a lead using site. Two sites existed in the CARB data prior to 2000 and reported for the first time in 2000 in the TRI as using one of the PBT chemicals affected by the policy change. These sites will be referred to as the “lead treatment sample” and the “PBT treatment sample”, respectively, throughout the remainder of this paper.

Table 1 enumerates the different site types and the average facility-level toxic releases over the period 2001 to 2009 of sites in each category. 353 sites are control TRI sites. From

---

12 Criteria air pollutants are those regulated under the U.S. Clean Air Act. They include ozone, particulate matter, carbon monoxide, sulfur dioxide, nitrogen oxides, and lead.

13 Of the 294 facilities that cannot be matched, 12 reported lead use and reported initially in 2001 and 1 reported PBT use and reported initially in 2000. I cannot verify if these sites existed prior to the reporting threshold changes and they are treated as new facilities. The results could be biased towards zero if in fact these facilities are treatment facilities, and the control group contains many houses that are actually treatment houses. I run the analysis separately treating these 13 facilities as treatment sites and find that the DID estimates actually are slightly less negative. However, they are not significantly or economically different, and the conclusions of this research are not affected.
the table it's clear that the treatment sites of both types have significantly fewer emissions than the average control site. Additionally, a higher fraction of lead treatment sites report no emissions at all. This has important implications for the interpretation of any estimated treatment effects. With these differences in emissions profiles, it is argued that any price effects attributed to the new information represent the housing market’s sensitivity to either the presence of toxic chemicals or a general stigma that accompanies meeting reporting requirements for the TRI. Unlike previous studies, the link between emissions and prices is not explicitly under examination.

[Figure 1 about here.]

Figure 1 shows the geographic distribution of TRI sites in the Bay Area. The map displays Census 2000 census tracts, colored according to median home values, with darker areas having higher prices. The small circular dots represent TRI firms in the control group. Lead treatment sites are represented with triangles and PBT treatment sites are represented with stars. The sites are largely distributed around the bay, with a large cluster of firms at the southeast end, near San Jose. There does appear to be a spatial relationship between tracts with lower prices and density of TRI sites.

4.3 Housing Data

The housing data provides a record of each single family housing transaction, attached and detached, that took place between 1996 and 2008 for the San Francisco - Oakland - San Jose Metropolitan Statistical Area. This area effectively comprises the counties of Alameda, Contra Costa, Marin, San Francisco, San Mateo and Santa Clara. The dataset contains many observable characteristics for each house (e.g. number of bedrooms, square footage, etc.) as well as the transaction price, loan amount, transaction date, latitude and longitude
coordinates and the year 2000 census tract. Each property is uniquely identified in the data which allows the creation of a panel data set. In an effort to remove outliers, houses observed in the top and bottom 1% of the price and square footage distribution are dropped,\textsuperscript{14} as well as the top 1% of the number of bedrooms and the number of bathrooms distribution. Houses with missing attribute or location data are also dropped from the dataset.

An unfortunate feature of the transactions data is that Dataquick will overwrite the characteristics recorded for a given property in previous transactions if a newer transaction is recorded with different and presumably updated information. However, in certain circumstances, if the renovation is on the scale of a large addition or major construction, the transaction will be flagged as having such an improvement. The implication for panel analysis is not being able to reliably observe changes to properties, since any moderate change made to the property is retroactively applied to all records in the data. As a result, all observable characteristics will drop out of any repeat sales analysis. To combat the presence of homes that likely have changed in substantial ways, homes that are observed to appreciate (depreciate) more than 50\% on an annualized basis, have the major construction data flag, transact with a loan amount greater than the transaction price by $5,000, or are observed to transact twice or more in any twelve month span are dropped from the sample.

\textsuperscript{14}In the dataset, the top percentile of the price distribution is approximately $1.6 million and the bottom percentile was approximately $40,000.
facilities, whether or not they are undergoing a shock to information.

Table 2 provides summary statistics for the various housing samples used in estimation. Column (1) describes the sample of houses that are not near any TRI sites and are consequently excluded from this study. Column (2) details the control sample which is the set of all houses with at least one TRI site fewer than three kilometers away. Columns (3) and (4) summarize the treatment groups of houses that are within three kilometers of PBT or Lead treatment sites, respectively.

Houses not near any TRI facilities appear to have very different observables than the control and treatment groups. These houses tend to be newer houses on much larger lots and be larger in terms of square footage, the number of bathrooms, and the number of bedrooms. Accordingly, these houses also have much higher prices. If unobservables are correlated to some extent with observables, then these houses could bias a difference-in-differences estimator.

Column (2) describes the control group of houses which look more like the PBT and Lead treatment groups. It should be noted that the PBT and Lead treatment groups are different from each other, and share a different relationship to the control group. PBT treatment houses tend to be on much smaller lots, have fewer bedrooms and bathrooms. However, lead treatment houses are newer, larger, and more expensive relative to the control group. If PBT treatment houses are in neighborhoods that have a priori worse unobservables than the control group, and those unobservables translate into worse pre-policy trends, the difference-in-differences estimator could be biased downwards, overestimating any negative price effects. Conversely, if the lead treatment houses are in neighborhoods that have a priori better unobservables, my difference-in-differences estimates would be biased upwards, rendering any negative estimates conservative.

A closer examination of Columns (3) and (4) reveals some interesting details about the distribution of treatment and control sites throughout the region. First, there are no houses
that are within three kilometers of an affected PBT site and not also within three kilometers of a control site. This may call into question the ability of the model to identify the treatment effect of the PBT policy change since houses near PBT sites are also always in proximity to regular control sites. The variation in the data needed to identify the treatment effect for the lead policy change is greater, as approximately 5% of houses near lead treatment sites are not near control sites. The apparent clustering of TRI facilities suggests that the neighborhoods that contain the facilities, and consequently the houses within those neighborhoods, might be unobservedly different than other neighborhoods.

4.4 **Price Trends**

![Figure 2 about here.]

![Figure 3 about here.]

The differences-in-differences estimator requires the fulfillment of the “common trends” assumption, which states that the control and treatment groups must have similar trends before the policy intervention. Otherwise, the difference in outcomes post-policy could simply be a result of the continuation of pre-existing differential trends. In the context of TRI threshold reductions, house prices in the control group need to be on the same trend as houses that are in the treatment group.

Figure 2 depicts the price trends and sales volume for groups of houses affected by the PBT policy change. Panel 2a displays the price trends for the PBT treatment and control houses and Panel 2b considers the sales volume for each group. The price time series is constructed by a local-polynomial smoothing of the average monthly residuals, by treatment status, from a regression of the natural log of price on house age, house fixed effects, and monthly dummies. In effect, each line depicts how the price variation that is not explained
by observable characteristics changes over time for both the treatment and control group of houses. Figure 3 presents the same data for the houses affected by the lead threshold reduction. To be clear, in each picture, the control group consists of houses within three kilometers of only a non-affected TRI site, while the treatment group consists of houses within three kilometers of a respective PBT or Lead treatment site. Sales volume is calculated as the percent change in monthly transactions, by group, relative to December 1997.

Each graph has a vertical blue, dashed line and a vertical red, solid line. The dashed lines represent the respective announcement dates of the policy changes for PBT and lead, while the solid lines represent the respective dates that the affected datasets were released to the public. As detailed above, the PBT policy changes took place before the lead policy changes.

Prior to the PBT policy announcement, prices in both groups were following relatively flat trajectories. After announcement, there is a discontinuous drop in prices followed by a brief boom that quickly dissipates in the treatment group. There isn’t a second discontinuity in prices after the release of the data detailing which firms are new to TRI reporting. After the release in data, prices fall for many months, eventually recovering. The trajectory of sales volume does not appear to reveal any meaningful patterns. Post-policy change, housing prices fell as quantity transacted fell, suggesting a decrease in demand.

It is clear from Panel 2a that there is a meaningful drop in prices when comparing prices before the announcement of the policy and after the effective date when data is released. Prior to the policy intervention, the treatment and control groups seem to be following a similar trend. This suggests that the “global” treatment effect would be identified and likely to be negative. However, the transition in prices after the announcement is puzzling. I will discuss the transition period in more detail below.

Panel 3a depicts more compelling evidence in support of the common trends assumption. In the months prior to the lead threshold reduction, both groups of houses have an upward
trending price path. Upon announcement of the new policy, there is a discontinuous drop in prices in the treatment group. Interestingly, after the initial drop, there is a steady reduction in prices until the data release date. When the data is released, there appears to be a smooth reversal of trajectory, during which housing prices in the treatment group and control group appreciate for a few years afterwards. For these comparison groups, Panel 3a suggests that the treatment and control groups had similar trends prior to the policy announcement and returned to similar trends after an adjustment period to the policy change. This pattern suggests that the differences-in-differences estimator would provide an accurate estimate of the overall effect on prices of the policy change, and that the effect is likely to be negative.

The nature of the policy intervention in both cases resulted in a prolonged adjustment period (approximately 30 months) where it was known to market participants that new firms will be listed in the TRI, but it was not known which firms would be affected. The patterns observed in the PBT treatment group are hard to explain, and can be a function of the small sample size of houses that fell into this group in the study area.

The pattern in Panel 3a does have an intuitive explanation. At the time the policy is announced, it is likely that the only individuals that knew which firms would be affected were the employees and owners of said firms. If future listing in the TRI was perceived to negatively influence housing values in the future, affected firm owners and employees would want to exit the housing market in advance of the public release of data to preserve their equity if they lived near their place of employment. Additionally, as time progressed, it is possible that the spread of information and the increase in houses for sale contributed to a downward trend in prices near the sites. By the time the data is actually released, prices had adjusted almost completely, and the local market continued on the trend from the pre-announcement era.

Panel 3b supports this notion. Between announcement and data release, prices were falling at the same time there was a large increase in the amount of transactions of treat-
ment houses. Falling prices accompanied by increases in quantity suggest that there was an outward expansion in supply. This is precisely what one would expect to see if “insiders,” who also supply housing in these areas, start to sell their properties to “outsiders,” whose demand for housing is unaffected since they are likely unaware of the happenings at local plants.

Admittedly, it is difficult to verify the details of this diffusion process empirically beyond documenting the trajectory of prices and quantities and their implications during the policy adjustment period. However, for the identification strategy to be invalidated because of correlated unobservables or omitted variables, it must be the case that there is an omitted variable that is differentially lowering prices in close proximity to facilities that will soon appear in the Toxic Release Inventory and not affecting prices in close proximity to facilities that are already reporting. It is difficult to imagine what such an omitted variable would be given the geographic dispersion of lead treatment sites, let alone to control for it. Scorse (2007) documents that firms who cease to be identified as a “Top 10 Worst Polluter” in terms of TRI emissions in their state respond with a decrease in reported emissions reductions. However, it is unlikely that treatment plants are responding strategically to new reporting requirements in such a way that would lower housing prices around them, before the data detailing their toxicity is released.

5 Estimation

The empirical model stems from the hedonic price theory of Rosen (1974). Under hedonic price theory, the price of a product, or in this case a house, is a function of the attributes of the product in the information set of market participants, both observable and unobservable. With many observations of housing transactions, including sufficient variation in attributes
and transaction prices, it is possible to estimate a marginal price for an additional unit of each attribute.

The empirical question in this paper is whether the emissions and toxic chemical data provided in the TRI program enters the information set of housing market participants. I use the hedonic price model to test for a change in the implicit price for proximity to sites forced to report to the TRI program as a result of the exogenous rule change. The existence and visual impact of the site is within the information set of buyers and sellers both before and after the regulatory shift. Once the rule changed, the requirement that several existing plants report to the TRI program for the first time introduces new information about the emissions and toxicity of those firms to buyers and sellers. If this information was already known, or if the information seemed inconsequential to households, the implicit price for proximity to such a site should not be significantly different after the regulation changes. However, since the visual disamenity and knowledge of the existence of the site does not change, finding a significant change in the implicit price for proximity to these sites implies a change in the information set and thus a capitalization into home prices of the information provided by the TRI program.

5.1 Difference in Differences Model

To prevent mixing of treatments, I ignore any houses that are in proximity to both PBT treatment sites and lead treatment sites. I run separate regressions to estimate the respective treatment effects. I specify the log price of house $h$ in time $t$ as:

$$\ln P_{ht} = \beta X_{ht} + \gamma TRI_{ht} + \alpha TREAT_{ht} + \lambda_1 ANN_{ht} + \lambda_2 POLICY_{ht} +$$

$$\beta_1 TREAT_{ht} \times ANN_{ht} + \beta_2 TREAT_{ht} \times POLICY_{ht} + \eta_h + \epsilon_{ht} \quad (1)$$
In Equation (1), $X_{ht}$ are the observable attributes of house $h$, which can include observable housing attributes and time dummies. $TRI_{ht}$ is a variable denoting the proximity of house $h$ to a TRI site whose reporting status is unaffected by the regulatory changes. $TREAT_{ht}$ is a variable measuring the proximity of house $h$ at time $t$ to a either a lead treatment site or a PBT treatment site.

$ANN_{ht}$ and $POLICY_{ht}$ are indicator variables that equal one if house $h$ sold after the announcement date and policy effective date (data release) of the respectively policies. Note that these period variables are mutually exclusive: the announcement indicator is “turned off” for transactions after the data release. $(\beta_1, \beta_2)$ identify the respective price effects, relative to the pre-announcement period, of the policy announcement and data releases which reveal that nearby sites are now listed in the TRI. The difference $(\beta_2 - \beta_1)$ details the incremental price effect of the data release, conditional on announcement occurring. $\eta_h$ is a house-specific unobservable attribute that is controlled for with house fixed effects.

If changes in housing prices in each group are not significantly different from each other, the null hypothesis that information is not used by households cannot be rejected. This is the intuition behind the DID estimator. Since the “treatment” sites were in existence before they began reporting to the TRI program, all of the visual attributes of the firm and neighborhood can be assumed to be constant before and after listing in the TRI. Subtracting the change in prices seen in the control group from the change in prices of the treatment group amounts to a differencing away of both site-specific heterogeneity and any price trends affecting housing markets around facilities that handle and release toxic chemicals. The resulting difference can be interpreted as the effect that TRI site listing alone has on housing prices.

### 5.2 Radius of Impact

[Figure 4 about here.]
Intuitively, there is likely a point where houses are simply too far away from the sites in question to be affected by the policy change in a meaningful way. The definition of the proximity variable can influence the results. If defined to broadly, houses outside of the latent radius of impact, which are expected to be unaffected by the policy, will “dilute” the estimate of the treatment effect, since houses that aren’t actually being treated are classified as treated. Conversely, if proximity is defined too narrowly, houses inside the latent radius of impact will be included in the control group, attenuating the difference between control and treatment groups.

I use the empirical relationship between distance and price effects to select an appropriate radius of impact for use in the proximity variables. To do so, I use a procedure used by Davis (2011). First, I regress the natural logarithm of housing prices on a vector of observable attributes (excluding TRI site proximity) and year and zip code fixed effects. Then, I run a local linear regression of the price residuals on distance from the nearest TRI site. Figure 4 scatter plots these coefficients, as well as a fitted 8th degree polynomial.

The price residuals contain the unexplained variation in prices after controlling for all observables and year and zip code fixed effects. The local linear regression non-parametrically estimates the slope of the relationship between prices and distance from TRI sites. As Figure 4 shows, close proximity to TRI sites has a negative effect on prices. This negative effect decreases in magnitude as distance increases. Following the fitted polynomial, the effect reaches zero between two and three kilometers, and appears to stay very close to zero beyond three kilometers.

5.3 Defining Proximity

I use two different definitions of proximity when estimating Equation 1. First, proximity to a TRI site can be defined as an indicator variable that equals 1 if a house is within a certain
distance, $d$. In effect, this assumes that the price effect is constant for all houses within $d$ km of a TRI site. Second, to allow for the price effect to vary continuously with distance, I follow Kiel and Zabel (2001) and define proximity for house $i$ as

$$PROX_i = \max\{0, MAXDIST - d_i\}$$

(2)

where $d_i$ is the distance that house $i$ is from the TRI site and $MAXDIST$ is the maximum radius of impact. For houses that are close to the site, $d_i$ is small, $PROX_i$ will be large. For houses far from the site, $d_i$ is large, and $PROX_i$ will be close to zero.

For robustness, I estimate Equation 1 under four different specifications of proximity, guided by the local linear regression exercise. First, I estimate the main regression using the continuous proximity variable, setting $MAXDIST$ equal to 3 km and 5 km, respectively. This assumes that the effect of proximity to a TRI site declines linearly from 0 to 3 km and 0 to 5 km, respectively. Second, I estimate Equation 1 with indicator variables equal to 1 if a house is within 2 km and 3 km, respectively. This specification assumes the effect is constant within 2 and 3 kilometers, respectively.

6 Results

6.1 Proximity Variables

Table 3 presents the results of the difference in differences analysis using continuous proximity variables. Column (1) estimates Equation 1 for the PBT policy change with $MAXDIST$ set to 3 kilometers. By examining only the signs and significance, it is clear that the PBT threshold reduction policy had no significant impact on housing prices. Column (2) increases $MAXDIST$ to five kilometers. This specification allows more houses, at farther distances,
to be considered as treated. The treatment effect for data release in this specification is negative and significant. POLICY – ANN provides the estimate for the incremental price effect of releasing the TRI data, which was -3.1%.

Columns (3) and (4) repeat this exercise for the lead threshold reduction. As one might have expected from Figure 3a, there is a significantly negative effect on the price of proximity to the affected lead TRI sites both at the policy announcement date and the policy effective date. When MAXDIST = 3, the price for proximity drops 3.4% relative to the pre-announcement period at the announcement date and drops 5.3% relative to the pre-announcement period at the effective date. The incremental drop in the price for proximity when the relevant data was released was 1.9%. These estimates correspond to the price drop for a house that is two kilometers away from a treatment site. When increasing MAXDIST to 5, the regression still returns negative and significant coefficients. Under the different specifications, a house that is two kilometers from a treatment site will have different values for proximity, making these coefficients not directly comparable. The estimates in Column (4) imply that a house that is two kilometers from a treatment site experiences a drop of 5.7% in price at the announcement date and 2.7% at the effective date.\textsuperscript{15}

### 6.2 Indicator Variables

Table 4 provides the regression results when indicator variables for proximity are used. The interpretation of the coefficients in this specification is more straightforward, as the coefficient is the price effect for any house within the specified distance of a site. Column

\textsuperscript{15}Estimates in Table 3 are the marginal prices for houses with PROX = 1. When MAXDIST = 5, a house with \(d_i = 2\) will have \(PROX = 3\). Thus, multiplying the estimates by 3 makes them comparable to the estimates in Columns (1) and (2).
(1) provides the estimates of Equation 1 when the tighter radius, two kilometers, is used. The price effects of the PBT policy announcement and effective date are not distinguishable statistically from zero.

The results in Column (2) correspond to a specification of Equation 1 that uses three kilometers as the radius of impact. Interestingly, when expanding the radius of impact to two kilometers, the effect for the PBT data release is negative and significant. This result could suggest that households that sort close to the facility have a lower preference for environmental quality than those that sort farther away. As a result prices move more drastically at greater distances, up to a point.

Columns (3) and (4) replicate the analysis in the first two columns for the lead policy announcement and data release. Consistent with Table 3 there are negative price effects at both announcement and data release. The negative effect on the data release is consistent across choice of impact radius. The estimate of the price effect at two kilometers, -4.7% at announcement and an additional 3.7% when the data is released. When using the larger radius, prices fall 5.5% at announcement and an additional 2.5% upon release of the relevant data. These estimates are quite close to the estimates of the price effect in the proximity variable specification when $MAXDIST = 5$.

6.3 Discussion

The empirical results reveal several interesting points. First, the reporting threshold reduction for lead appears to have a much more robust impact on housing prices than does the threshold reductions for PBT chemicals. What can account for the difference in price patterns between the PBT policy and the lead policy? It is likely that the relative salience of lead pollution for both firms and the public contributes to the more consistent findings for lead relative to PBT chemicals. For the threshold reductions to have an impact on prices, households would need to perceive the new information provided by the new inclusion of a
site in the Toxic Release Inventory as “bad news”. Lead has long been recognized as toxic, especially for children. It is possible that finding out that a nearby facility is using or releasing lead might invoke a stronger reaction than hearing a nearby facility is using a PBT chemical such as “Hexachlorobenzene” or “Trifluralin”, since lead and its consequences are more familiar to non-experts. The relative “popularity” of lead vis-à-vis other PBT chemicals might explain the stronger price reaction documented above.

Second, there is strong evidence in the regression results to support the conclusion that prices fell after the policy announcement detailing the introduction of the new lead reporting threshold. In contrast, the announcement of the PBT policy appears to have had no affect on prices. Focusing on lead, the descriptive and econometric evidence indicates that the policy’s overall effect on prices was negative. By construction, the treatment houses received only an update in information as the facilities they were proximate to existed as registered polluters with the state of California prior to reporting to the TRI program. While a statistically significant drop in prices may be an intuitive result, the decline in prices after announcement is interesting because it is not intuitive. It does not seem likely that the affected firms would be identified by the market prior to the release of data. The results of this section unfortunately do not add any additional explanations beyond what is offered in Section 4. Rather, they confirm the qualitative price patterns documented in that section. The micro-dynamics of information diffusion in this context is beyond the scope of this paper, and remains a question for future research.

Lastly, the negative price effects come from a set of facilities that have a relatively small amount of emissions relative to the average TRI site in the area. Table 1 notes the average lead treatment site had average annual toxic emissions of 31 pounds, compared to the average annual emissions of 1,155 pounds for the control sites. Prices fell in spite of the new information provided by the TRI data release, which detailed that these facilities had very few emissions. This suggests two conclusions: 1) the housing market cares about the
TRI label and not a firm’s toxic emissions insofar as households need to turn to the data to calculate their exposure to toxic chemicals, or 2) households are concerned about the presence of large quantities of toxic chemicals near their homes. Both of these conclusions support the results of Bui and Mayer (2003) who find no relationship between emissions and housing prices. However, the overall implication of my results for the relationship between prices and the TRI program differs from their analysis.

7 Conclusion

In this paper, I examine the effects that information about the risks posed by facilities reporting to the Toxic Release Inventory program has on the housing market. By utilizing a discontinuous change in the reporting requirements to the TRI program, I find that housing market participants react to new information contained in the TRI data. The firms affected by the policy change have very low emissions however, indicating that market participants internalize the risk of an accidental release, or that TRI status carries a stigma because of perceived risk.

The results indicate that the lead reporting threshold reduction had a more robust negative impact on housing prices than the reporting threshold for PBTs. The relative salience of lead and the dangers it poses, especially to children, could account for the difference in market outcomes. The empirical existence of this difference suggests a level of sophistication on behalf of housing market participants who appear to discern separate implications for the two sets of chemicals.

The findings of this paper suggest that the Toxic Release Inventory program is communicating information to local residents as intended. However, the evidence doesn’t support the claim that households react to the detailed emissions reports, but rather to the knowledge that there are toxic chemicals in significant quantities present in their vicinity, where
“significant” is inferred from the responsibility to report to the TRI. The results also suggest that there may be an alternative mechanism driving the observed reductions in emissions over the lifespan of the TRI program. If local residents are more sensitive to the presence of large quantities of onsite chemicals than to emissions, firms may have an incentive to avoid scrutiny by operating below TRI reporting thresholds. If emissions do not enter household information sets, inclusion of emissions in hedonic models is dubious and violates full information assumptions.

Observed emissions reductions may be in fact driven by public pressure as intended, but only as a byproduct of firms seeking to avoid obligation to report. Under this scenario, emissions may not actually be falling but becoming under reported; an outcome that was probably not the intent of policy makers. This has important implications for the design of right-to-know polices. If markets are affected simply by a firm’s requirement to disclose information, rather than the details of that information, firms only have incentive to avoid having to disclose information. The incentive to improve the details of the information disclosed, for example by reducing reported emissions, might not be strong enough to induce abatement. Unfortunately, without detailed emissions information on firms that do not report, it is difficult to conclusively characterize firm behavior. Regardless, the empirical evidence supports an alternative explanation for the observed reductions in emissions and the literature’s mixed conclusions on the relationship between housing prices and TRI emissions.
References


Figure 1: Bay Area TRI Sites as of 2002

Notes: Data on median home values taken from The Census Bureau, for the year 2000 census. TRI site locations provided by the EPA.
Figure 2: PBT Time Series
(a) Lead Prices

(b) Lead Transaction Volume

Figure 3: Lead Time Series
Figure 4: Local Linear Regression Coefficients vs. Distance

Notes: Markers represent the coefficient from a local linear regression of log price residuals on distance from the nearest TRI facility. The residuals are calculated from a regression of the natural logarithm of price on housing attributes, time and neighborhood dummies.
Table 1: TRI Sites: Average Annual Releases (lbs), 2001 - 2009

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Zero Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRI Control Sites</td>
<td>353</td>
<td>1,155</td>
<td>3,946</td>
<td>0</td>
<td>39,693</td>
<td>27.2</td>
</tr>
<tr>
<td>PBT Treatment Sites</td>
<td>2</td>
<td>2.661</td>
<td>3.721</td>
<td>0.0300</td>
<td>5.292</td>
<td>0</td>
</tr>
<tr>
<td>Lead Treatment Sites</td>
<td>21</td>
<td>30.53</td>
<td>137.0</td>
<td>0</td>
<td>628.3</td>
<td>42.9</td>
</tr>
</tbody>
</table>

Notes: Statistics are derived from the annual facility level on-site emissions for facilities that were active at least one year from 2001 to 2009. Means and standard deviations taken over active years only. Annual emissions are reported in pounds.
Table 2: Housing Samples

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Sites</td>
<td>Control</td>
<td>Treatment</td>
<td>Lead Treatment</td>
</tr>
<tr>
<td>Lead Treatment Site</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PBT Treatment Site</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>TRI Site</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.949</td>
</tr>
<tr>
<td>Age</td>
<td>29.21</td>
<td>36.57</td>
<td>37.36</td>
<td>28.65</td>
</tr>
<tr>
<td>Lotsize (Sq. Feet)</td>
<td>7,296</td>
<td>5,084</td>
<td>1,791</td>
<td>4,578</td>
</tr>
<tr>
<td>No. Sq Feet</td>
<td>1,840</td>
<td>1,489</td>
<td>1,347</td>
<td>1,512</td>
</tr>
<tr>
<td>No. Bathrooms</td>
<td>2.232</td>
<td>1.917</td>
<td>1.789</td>
<td>2.088</td>
</tr>
<tr>
<td>No. Bedrooms</td>
<td>3.235</td>
<td>2.963</td>
<td>2.710</td>
<td>3.013</td>
</tr>
<tr>
<td>No. Rooms</td>
<td>6.709</td>
<td>6.296</td>
<td>5.661</td>
<td>6.341</td>
</tr>
<tr>
<td>Price ($)</td>
<td>486,070</td>
<td>399,596</td>
<td>397,855</td>
<td>442,793</td>
</tr>
<tr>
<td>Obs.</td>
<td>315,804</td>
<td>337,439</td>
<td>2,134</td>
<td>57,438</td>
</tr>
</tbody>
</table>

Notes: “Lead Treatment Site”, “PBT Treatment Site”, and “TRI Site” provide the fraction of houses in each sample that are within three kilometers of at least one site of the specified type. Prices are normalized to year 2000 prices to control for inflation.
Table 3: Main Results - Proximity

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MAXDIST</strong></td>
<td></td>
<td>PBT</td>
<td>Lead</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 km</td>
<td>5 km</td>
<td>3 km</td>
<td>5 km</td>
</tr>
<tr>
<td><strong>PBT × ANN(β₁)</strong></td>
<td>-0.012</td>
<td>0.005</td>
<td>-0.034***</td>
<td>-0.019***</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.009)</td>
<td>(0.004)</td>
<td>(0.002)</td>
</tr>
<tr>
<td><strong>PBT × POLICY(β₂)</strong></td>
<td>-0.007</td>
<td>-0.025***</td>
<td>-0.053***</td>
<td>-0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.038)</td>
<td>(0.009)</td>
<td>(0.004)</td>
<td>(0.002)</td>
</tr>
<tr>
<td><strong>PBT(α)</strong></td>
<td>0.02</td>
<td>0.001</td>
<td>-0.006</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.011)</td>
<td>(0.004)</td>
<td>(0.002)</td>
</tr>
<tr>
<td><strong>LEAD × ANN(β₁)</strong></td>
<td>-0.034***</td>
<td>-0.019***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LEAD × POLICY(β₂)</strong></td>
<td>-0.053***</td>
<td>-0.028***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LEAD(α)</strong></td>
<td>0.006</td>
<td>-0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TRI(γ)</strong></td>
<td>-0.007**</td>
<td>-0.007**</td>
<td>-0.007**</td>
<td>-0.008***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td><strong>ANN(λ₁)</strong></td>
<td>0.045***</td>
<td>0.045***</td>
<td>-0.021</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.007)</td>
<td>(0.013)</td>
<td>(0.013)</td>
</tr>
<tr>
<td><strong>POLICY(λ₂)</strong></td>
<td>0.092***</td>
<td>0.093***</td>
<td>0.037***</td>
<td>0.047***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td><strong>POLICY – ANN</strong></td>
<td>-0.019</td>
<td>-0.031***</td>
<td>-0.019***</td>
<td>-0.009***</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Observations</td>
<td>339,500</td>
<td>339,500</td>
<td>394,781</td>
<td>394,781</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.596</td>
<td>0.596</td>
<td>0.588</td>
<td>0.589</td>
</tr>
</tbody>
</table>

Notes: Treatment status is a function of house proximity to facilities. Proximity of house i to site j is defined as \( PROX = \max(0, MAXDIST − d_{ij}) \). \( MAXDIST \) is set to three kilometers in Columns (1) and (3), and five kilometers in Columns (2) and (4). Standard errors are clustered at the house level and in parenthesis below each coefficient estimate. Significance is denoted: *** 1% level; ** 5% level; * 10% level.
Table 4: Main Results - Indicators

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radius</td>
<td>2 km</td>
<td>3 km</td>
<td>2 km</td>
<td>3 km</td>
</tr>
<tr>
<td>(PBT \times ANN(\beta_1))</td>
<td>-0.003</td>
<td>0.005</td>
<td>(0.078)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>(PBT \times POLICY(\beta_2))</td>
<td>-0.041</td>
<td>-0.064**</td>
<td>(0.077)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>(PBT(\alpha))</td>
<td>0.075</td>
<td>0.005</td>
<td>(0.086)</td>
<td>(0.034)</td>
</tr>
<tr>
<td>(LEAD \times ANN(\beta_1))</td>
<td>-0.047***</td>
<td>-0.055***</td>
<td>(0.008)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>(LEAD \times POLICY(\beta_2))</td>
<td>-0.084***</td>
<td>-0.080***</td>
<td>(0.007)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>(LEAD(\alpha))</td>
<td>0.013</td>
<td>0.002</td>
<td>(0.009)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>(TRI(\gamma))</td>
<td>-0.005***</td>
<td>-0.005***</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>(ANN(\lambda_1))</td>
<td>0.045***</td>
<td>0.044***</td>
<td>-0.015</td>
<td>-0.017</td>
</tr>
<tr>
<td>(POLICY(\lambda_2))</td>
<td>0.094***</td>
<td>0.092***</td>
<td>0.047***</td>
<td>0.041***</td>
</tr>
<tr>
<td>(POLICY - ANN)</td>
<td>-0.038</td>
<td>-0.069***</td>
<td>-0.037***</td>
<td>-0.025***</td>
</tr>
<tr>
<td>Observations</td>
<td>229,659</td>
<td>339,500</td>
<td>256,622</td>
<td>394,781</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.574</td>
<td>0.596</td>
<td>0.572</td>
<td>0.589</td>
</tr>
</tbody>
</table>

Notes: Treatment status is a function of house proximity to facilities. Proximity of house \(i\) to site \(j\) is defined as \(1[d_{ij} < RADIUS]\). \(RADIUS\) is set to two kilometers in Columns (1) and (3), and three kilometers in Columns (2) and (4). Standard errors are clustered at the house level and in parenthesis below each coefficient estimate. Significance is denoted: *** 1% level; ** 5% level; * 10% level.