Toxic Assets: How the Housing Market Responds to Environmental Information Shocks *

Jeremy G. Moulton † Nicholas J. Sanders ‡ Scott A. Wentland §

October 2018

Abstract

In March 2000, a number of polluting industries, including fossil fuel power plants, were added to the list publicly reporting pollution releases in the Toxics Release Inventory (TRI). Employing microdata from Zillow, which contain information on millions of property transactions and detailed corresponding home characteristics, we examine how housing markets respond to new information about reported toxic pollution by nearby facilities. We investigate this using a regression discontinuity design, which exploits the discrete information shock with fine microdata over time and space. Contrary to prior findings that TRI information does not influence household actions, we find the additional TRI data caused households to revise priors on ambient pollution levels, leading to an immediate reduction in home prices near the most toxic plants after the release. Effects appear isolated to homes within just a few miles of reporting facilities. From a policy standpoint, the results imply that there remains a role for government as provider of information that markets subsequently incorporate into prices.

Keywords: pollution, information disclosure, residential housing, home prices, spillovers, externality, toxics release inventory, regression discontinuity

JEL Classifications: D62, Q50, Q53, H23, G14, R32, R38

Disclaimer: Any views expressed here are those of the authors and not necessarily those of the Bureau of Economic Analysis or the U.S. Department of Commerce. Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group.

*Thanks to Caroline Hoxby, Hilary Hoynes, Christopher Knittel, Jed Richardson, participants in the Stanford Environment and Energy Policy Analysis Center seminar, the Stanford Public/Environmental Seminar, the University of California, Davis Public/Labor Seminar, and the Stanford Institute for Theoretical Economics summer meetings. Special thanks to Tom Blake, who helped substantially with data collection, and to Maggie Goulder for excellent research assistance.

†Department of Public Policy, University of North Carolina — Chapel Hill, moulton@email.unc.edu

‡Policy Analysis & Management, Cornell University, njsanders@cornell.edu

§Office of the Chief Economist, U.S. Bureau of Economic Analysis, scott.wentland@bea.gov

"It's not news that they're polluting, but it is news to the extent that they are polluting." (John F. Sheehan of the Adirondack Council[1])

1 Introduction

In 1998, the Environmental Protection Agency (EPA) instituted a new policy requiring a number of heavily polluting industries to begin publicly reporting annual toxic releases, with data made public in 2000. These industries were added to the Toxics Release Inventory (TRI), data on local toxic pollution released to the public once a year. The addition of these industries drastically shifted reported pollution levels in impacted localities, with increases in the range of 800% in the most extreme cases. Importantly, this was a shift in *reported* releases, not a shift in releases per se, and thus the change in *information* on local amenities was independent of a change in the level of amenities themselves.

This provides a unique opportunity to investigate how additional information on local amenities can be incorporated into local housing markets when the amenity is only partially observable. In this case, the industries added to the TRI included coal and oil power plants. Such factories are large, visibly obvious polluters, which should allow households to establish a perception regarding daily exposure to localized "bads." The TRI serves as an external source of information by which households can update perceptions. Using data from Zillow covering millions of home sales across the United States, we investigate whether specific information provided by the TRI led to shifts in the housing prices above and beyond those correlated with local pollution levels. We find that information revelations result in rapid decrease in housing values, but only in a highly localized space of several miles and primarily near the highest polluters.

Our results add to the literature on environmental information and market solutions to externalities. Several models in economics use market mechanisms to rectify the problem

[1] Hu (May 12, 2000).

of environmental externalities.[2] But for such models to operate efficiently, markets must accurately assess environmental conditions. Households cannot efficiently sort themselves without knowing local externality levels, and imperfect information will result in an equilibrium that is socially inefficient. The role of information and response is equally important for avoidance and mitigation behavior in the face of environmental dangers.[3]

The remainder of this paper is organized as follows. Section 2 describes the TRI in detail, as well as the relevant policy changes used for identification. Section 3 discusses prior findings on TRI information and home prices within the context of the larger TRI literature. Section 4 describes the data used in the analysis. Section 5 describes the methodology. Section 6 presents our primary results and explores various robustness checks. Section 7 discusses our findings in context of prior related work. Section 8 concludes. We also present a discussion of the TRI in the media and media exposure as a potential vector for new information in the Appendix.

This is a work in progress, and a major update from a prior version of the paper, in which analysis was done at the zip code level. As such, all graphs present information at the level of zip code, while regressions use individual home sales.

2 The Toxics Release Inventory

Public Law 99-499 (the "Superfund Amendments and Reauthorization Act of 1986") amended the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 and

[2]Tiebout (1956) proposes a model where individuals sort in communities with their optimal combination of taxes and amenities. This "voting with your feet" can be applied to an environmental context where, rather than government establishing constraints and regulations, firms are allowed to pollute and households sort based on their preferences for environmental quality. Coase (1960) proposes an alternate solution via private bargaining. Property rights are assigned, and households and firms engage in market transactions to find an agreed-upon level of pollution.

[3]Recent research finds when households have information regarding environmental hazards, they adjust behavior in ways that can help offset potential for health consequences. For example, Graff Zivin et al. (2011) find notification of water quality violations leads households to shift consumption from tap to bottled water, and Neidell (2004), Neidell (2009), and Moretti and Neidell (2011) find people adjust behavior to avoid spending time outside on days with dangerous levels of ambient ozone.

created the Toxics Release Inventory. Contained within the Act was the requirement that,

> The owner or operator of a facility subject to the requirements of this section shall complete a toxic chemical release form as published under subsection (g) for each toxic chemical listed under subsection (c) that was manufactured, processed, or otherwise used in quantities exceeding the toxic chemical threshold quantity established in subsection (f) during the preceding calendar year at such facility.
> (Public Law 99-499)

The Act applied to facilities that had 10 or more full-time employees, were within SIC codes 2000 through 3999, and produced or released over a threshold level of specifically noted toxics per year.[4] Data are self-reported, collected by the EPA at the end of each calendar year, and later released to the public as the Toxics Release Inventory report. Due to lags between when data are collected and ultimately released to the public, the full TRI data for any given year become public around 18 months after the end of the relevant reporting year.

A number of studies examine the impact of the early TRI data. Closest to this paper, Bui and Mayer (2003) and Oberholzer-Gee and Mitsunari (2006) both consider how home prices respond to the initial 1989 data release. Neither paper finds any consistent change in home prices when the TRI data first appear. We discuss these results in greater detail in Section 7. More recently, Mastromonaco (2012) considers how a later TRI policy change in 2002 influenced housing prices in a number of California cities. Other work explores how the stock market capitalized information on firm toxic emissions. Hamilton (1995) found stock losses for polluters in the days directly following the initial release, and Konar and Cohen (2006), using 1988 TRI data, find both toxic chemical releases and environmental lawsuits to be associated with negative stock returns. Khanna et al. (1998) found repeated release data had lasting effects on firms already known to be large polluters. Less is known,

[4] In the first reporting year, this threshold was set at 75,000 pounds. This was lowered to 50,000 pounds in the second reporting year, and 25,000 pounds in the third reporting year, and then stabilized for some time. The initial listing of chemicals required to report was a combination of two pre-existing lists of hazardous toxics, the New Jersey Environmental Hazardous Substance List and the Maryland Chemical Inventory Report List. In 1993, the EPA added 23 additional chemicals to the reporting list, with 286 more added in 1994 as the list of who was to report expanded to include all Federal facilities.

however, about the impacts of the large-scale 1998 adjustment to the reporting requirements of the TRI, which were, as we describe below, categorically different than prior releases. Markets may not necessarily react the same to information about heavy polluters as they do to polluters previously reported by the TRI. This raises the empirical question as to whether nonlinearities matter in this context, and whether the market was able to accurately capitalize information for these large polluters into market prices absent the information made public in the TRI.

In the 1998 reporting year, seven industries were added to the list of those required to report information in the TRI: electricity production via coal and oil burning (SIC codes 4911, 4931, and 4939), metal and coal mining (SIC codes 10 and 12), solvent recyclers (SIC code 7389), hazardous waste treatment and disposal facilities (SIC code 4953), chemical distributors (SIC code 5169), and petroleum bulk terminals (SIC code 5171). These industries represented a large share of reported toxic releases, particularly the electricity production sector. As noted in a public statement by then EPA administrator Carol M. Browner upon the release of the new information (emphasis added);

> The new results, when added to the manufacturing sector already reporting, bring the total releases of toxic chemicals reported nationally to 7.3 billion pounds — **nearly triple the previous number**. Americans now will have the best picture ever of the actual amounts of toxic pollution being emitted by industry into local communities [. . .] **For the record, between 1997 and 1998, total releases of toxic pollution for the manufacturing sector continued to decline** — this time by 90 million pounds. Next year, we'll be able to see how all of the combined sectors will "trend" in terms of total emissions and individually [. . .] You have been given press kits today similar to previous years. This time, however, **as a result of the new data being presented, you will notice lists of states and facilities in eight different categories. The categories are the traditional manufacturing sector and the seven new sectors**.[5]
> (Remarks Prepared for Delivery, TRI Announcement, May 11 2000)

In investigating the impact of the policy change, we focus on airborne releases, as they are by

[5] Currently available online at `http://yosemite.epa.gov/opa/admpress.nsf/12a744ff56dbff8585257590004750b6/83c9dac72c1425068525701a0052e3dd!OpenDocument`.

far the largest changes due to the policy.[6] From this point forward, unless otherwise noted, the term "new releases" refers to airborne releases.

Figure 1 illustrates how newly reported releases from relevant industries compare to releases from earlier reporters. The figure shows all recorded releases in thousands of tons from the 1988 reporting year through 2002 (when the TRI stopped using SIC codes), separated by newly added industries (dashed line) and all other reporting industries (solid line).[7] Total reported releases for the seven impacted industries are effectively zero prior to the 1998 policy change, and after the change, releases from these industries are greater than releases for all other industries combined. Figure 2 shows the number of reporting plants by newly reporting industries (dashed line) and all other industries (solid line). Large releases reported for new industries are not due to a large number of newly reporting firms, but to the average amount of toxics for each firm.

Interestingly, the location of newly reporting industries has almost no relationship to prior levels of reported pollution. That is, there is little correlation between the levels of reported toxics in 1997 (based on the prior year's TRI report) and the level of newly reported toxics in 1998. Appendix Figure A-1 shows a scatter plot of toxics from new industries in 1998 against 1997 reported toxic levels, along with a fitted line to illustrate the lack of correlation. If newly reporting industries locate in areas populated by other manufacturing firms, newly reported releases could be highly correlated with pollution levels in general. This would make it difficult to separate between the impact of new reporting and a change in the view of toxics in general. For example, news reports on new industries might draw new attention to *all* sources of toxics.

Imperfections in data collection make using the TRI an imprecise measure of ambient

[6]This makes our results more comparable to prior findings using the TRI: Bui and Mayer (2003) and Oberholzer-Gee and Mitsunari (2006), for example, focus on airborne releases, as do many of the studies on health using the TRI (Currie and Schmieder, 2009; Currie, 2011; Currie et al., 2011).

[7]Air releases are the sum of stack and fugitive releases, where fugitive releases include equipment leaks, chemical evaporation, etc.

toxics. Firms appear and disappear due to openings/closings, failure to produce the amount of toxics required to report, etc., which can cause year-to-year changes in both number of firms reporting and total emissions. Reported data are often estimates based on production levels rather than directly measured emissions, and while the EPA does enforce reporting, there is no regular verification of reported versus true toxic releases.[8]

Such problems mean the TRI data may be an unreliable measure of exact toxic exposure, which led Currie et al. (2011) to develop an instrumental variable strategy using firm openings and closings.[9] We address this issue by focusing on the addition of large-scale newly recorded releases rather than smaller year-to-year marginal effects. There is also the concern that the general public is unaware of the existence of the TRI, and thus cannot benefit from any expansions of the data available. We show in the Appendix that the media focused on TRI-related stories at the time of each new data release, particularly around the releases impacted by the 1998 policy change. It need not be that households actively sought TRI data, but instead responded to the data provided by the media, or even learned from neighbors who had learned from the media, etc.

Pollution changes and information changes often move together: a toxic event bringing firms to public attention, such as the incident at Three Mile Island (Nelson, 1981; Gamble and Downing, 1982), or newly constructed power plants moving into neighborhoods (Davis, 2011). Our design avoids potential contamination from other factors that move along with changes in environmental quality, such as plants openings/closings, economic development, migration patterns, and emissions regulation. Still, interpreting price changes around the

[8]de Marchi and Hamilton (2006), for example, show that when pollution monitors can be used to examine ambient toxic levels, drops in emissions reported in the TRI are often smaller than those measured by nearby monitors. They further show the distribution of certain reported emissions fails the "Bedford's Law" test for a distribution of "true" data, and in some cases reported numbers appear to suggest "rule of thumb" uses for reporting rather than direct production numbers. Bedford's Law states that in a distribution of data, the first digit of all values is unevenly distributed across the 1-9 spectrum similar to a logarithmic scale, with 1 being represented approximately 30% of the time and each larger number appearing less and less frequently.

[9]When investigating the impact of toxics on infant health, they find no significant effects with OLS and large, significant effects with IV, suggesting measurement error in TRI data is a problem.

time of the TRI release as the result of information requires no other factors correlated with treatment changed due to the policy. For example, if firms that are newly required to report adjust production or employment as a result, there could be economic impacts that, in turn, influence housing prices. Similarly, if firms actively reduce pollution as a result of the policy, information and true pollution levels change simultaneously, making it impossible to separate specific impacts of information disclosure and the willingness to pay to avoid toxics.[10] The lag between when toxics are produced, when data are gathered, and when data become publicly available helps me separate the impact of the information shock from any changes caused within the firms in response to the new reporting regulations. That is, if the policy change itself impacts home prices, changes should occur during the year of toxic production (1998).[11]

Due to an additional policy change in the TRI, we limit analysis to periods just around reporting in 2000. In reporting year 2000, the EPA again expanded the toxics on the reporting list, adding new persistent bioaccumulative toxic (PBT) chemicals and lowering the reporting threshold for certain toxics already on the list, including metals such as lead (100 pound threshold) and mercury (10 pound threshold). Certain dioxins were given low reporting thresholds of anything greater than 0.1 gram of releases.[12] This policy change impacts a number of the same industries. For example, power plants are a large source of both lead and mercury. It impacted a good deal of other dioxin-producing factories as well.[13] Mastromonaco (2012) considers this alternate treatment in greater detail.

[10] Active attempts were made by some firms to reduce emissions after the initial TRI release in the form of the "33/50" plan, where a number of producers aimed to reduce toxic emissions by 33% in 1992 and 50% in 1995 (Environmental Protection Agency, 1999).

[11] As an alternative, there could be a substantial lag between the adjustment actions of the firm in 1998 and the eventual economic effects 18 months later.

[12] For an in depth list of the PBT listing and threshold changes in reporting year 2000, see Chapter 3 of the 2001 Toxics Release Inventory Public Data Release.

[13] Earlier versions of this paper considered how already treated areas saw changes under further treatment. While some negative effects were present, the lack of good controls causes me to omit these results here.

3 Environmental Hedonic Pricing and Prior Evidence From the TRI

Prior studies use changes in the value of homes as a hedonic measure of how households value environmental amenities.[14] Most similar to this work, Bui and Mayer (2003), Oberholzer-Gee and Mitsunari (2006), and more recently Mastromonaco (2012) examine the impact of toxics on home prices using the TRI. Oberholzer-Gee and Mitsunari (2006) use sales records from homes across five Philadelphia counties to investigate how observed prices near TRI facilities changed with the first-ever release of TRI data in 1989. They find home prices decreased across the time of the data release and interpret this change as a revision of the risk expectations of households who, prior to the TRI data release, had underestimated true toxic exposure, though this may be background trends in home prices independent of the TRI period. They also find results are highly sensitive to distance from a TRI facility, with perceptions being revised only in households a quarter to a half-mile away (and zero effect for homes closer to TRI sites). Bui and Mayer (2003) use 231 zip codes in Massachusetts and examine both the impact of the initial data release as well as short-run changes in reported toxics in the years that follow. In both cases, they find no detectable impact on home prices, even in communities with high newspaper readership (as measured by the Audit Bureau of Circulations) taken as a proxy for access to information. And while they find reported releases declined substantially after the first reporting years, the declines did not seem related to political economy, neighborhood influence, or price changes.

[14]For example, Greenstone and Gallagher (2008) examine housing prices near Superfund sites both before and after cleanup, and Gamper-Rabindran et al. (2011) consider how the results in Greenstone and Gallagher (2008) vary with levels of geographic aggregation. Chay and Greenstone (2005) use changes in pollution resulting from the Clean Air Act to show improved air quality was associated with increases in home prices in the impacted regions, and Bento et al. (2011), using the more recent Clean Air Act Amendments, show impacts of air quality improvement vary by spatial aggregation as well. Leggett and Bockstael (2000), using variation in water quality in the Chesapeake Bay, show a positive willingness to pay to avoid exposure to fecal coliform. Studies specifically investigating how power plants influence housing prices include Blomquist (1974), Nelson (1981), Gamble and Downing (1982), and Davis (2011).

Despite no consistent evidence housing prices adjust in response to earlier TRI information, other research finds changes in behaviors correlated with home values. Banzhaf and Walsh (2008) find people "vote with their feet" for environmental quality, using air releases from the TRI as a measure of toxic levels. Using the 2000 policy change, Currie (2011) finds compositional changes in the characteristics of mothers nearby TRI factories when additional information on toxics is provided, and Mastromonaco (2012) finds households in California see a decrease in value. Other work on environmental information further supports that households use such information when it is made available, with accompanying changes in home values. Davis (2004) shows that the proliferation of information on elevated cancer rates in a Nevada county caused a decrease in home prices of almost 16 percent, and Gayer et al. (2000) find the release of risk information about Superfund sites caused households to revise their expected cancer risks.

A possible explanation for the findings of divergent findings is a non-linear market response function regarding information provision. Markets use new information to adjust expectations on things like local amenities, but changing location behavior is expensive. If information causes an update that deviates only slightly from prior expectations, adjustment may be sufficiently costly to prevent response. Given a large enough shift in expectation, markets then respond. As we show in Section 6, imposing a linear response to new information gives me almost the exact same results as those found in Bui and Mayer (2003), while allowing for non-linear response finds something quite different. The next section more formally addresses such a possibility in a theoretical context.

4 Data

4.1 Home Prices

The information shock we examine in this paper is a national one, but likely with very localized or idiosyncratic effects, depending on how far one lives from a major polluter that would receive news coverage. Accordingly, we use national microdata from ZTRAX, a large dataset initially compiled by Zillow that contains transaction data as well as rich individual property characteristics for sales recorded from local tax assessment data[15]. A key advantage of this data is that it allows for both fine-grained analysis of localized effects and large national coverage. Specifically, the coverage of this data is representative of the United States' national housing market, initially containing 374 million detailed records of transactions across more than 2,750 counties, which includes information on each home's sale price, sale date, mortgage information, foreclosure status, and other information commonly disclosed by a local tax assessor's office[16]. We link this data with each home's property characteristics that Zillow also obtained from the local assessors' offices, which typically includes the size of the home (in square feet), number of bedrooms and bathrooms, year built, and a variety of other characteristics of the home[17]. Due to reporting differences across localities, it was a herculean effort by Zillow to accumulate and initially organize this data, which we received in a somewhat raw form, requiring additional cleaning for research purposes.

We scrutinized missing data and extreme values as part of our data cleaning and initial culling of outliers. The raw data included sales of empty plots of land, some commercial property transactions, agricultural sales, and a host of types of properties that are not relevant for the scope of our paper. As a result, we limited the sample to single family homes, townhouses, apartments, condos, and properties that are typically associated with

[15] Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group. Nonproprietary code used to generate the results for this paper is available upon request of the authors.

[16] We note that some states do not require mandatory disclosure of the sale price, so we have limited data for the following states currently: Idaho, Indiana, Kansas, Mississippi, Missouri, Montana, New Mexico, North Dakota, South Dakota, Texas, Utah, and Wyoming.

[17] Zillow's Ztrax data contains separate transaction and assessment files by state, where all transactions need to be linked to corresponding assessment records. With guidance from Zillow, we were able to merge the bulk of the data, but not without some data loss (which figures into the size of our final sample).

the residential market. We cull the top five percent of the lot size distribution (cutting many large farms) and outlier homes that are on the upper tail of the distribution (i.e. they either have more than six bedrooms, more than five bathrooms, or have a garage that holds more than five cars)[18]. We remove homes at the extreme ends of the price distribution for our analysis, which were homes that had a reported sale price of less than $10,000 and greater than $1 million. We cull homes that reported an age of greater than 100 years old (i.e. sale year — year built). While the Zillow data set contains a vast number of property characteristics, in our initial analysis we primarily rely on the variables above that have the most coverage nationally so we limit how much data we would effectively have to throw away[19]. Finally, we exclude California in much of the analysis due to the timing of the policy change we study, which was close in proximity to the timing of a sharp drop in the market valuation of the tech/internet sector that was heavily concentrated in California at the time.

As a check on the quality of the data, we compared our cleaned Zillow sample to the U.S. Census American Community Survey (ACS) to ensure that this administrative data aligned with carefully collected (albeit more limited) survey data provided by the Census. Overall, we found that the limited set of characteristics of homes that were in both the ZTRAX data and the ACS are comparable in terms of their summary statistics. We find that, in untabulated results, the shared characteristics across data sets (number of rooms, bedrooms, year built, acreage, and tax amount) had variable median and mean values that fell within a few percentage points of one another. Our final sample consists of approximately 1.5 million individual home sales that took place in the year surrounding the information shock. Because our analysis centers on only the transactions that occurred within windows around this information shock, our final sample figure above is much smaller than our initial

[18]We also create dummy variables equal to one if the property reported a lot size of zero or there are missing bedrooms or bathrooms.

[19]We conducted a sensitivity analysis in untabulated regressions that incorporates property characteristics to determine whether the results are sensitive to omitted variables for which we can control. Our results are largely robust to omitting variables that have more limited coverage.

data set. We return to this point when we describe our methodology. Table 1 shows the summary statistics for housing data used in our main analysis.

4.2 TRI

Toxic data are from the TRI Basic Data Files on the EPA website, which are annually aggregated by facility and toxic and include information on facility name and location, toxics released, and on- and off-site releases. All data are recorded in pounds until reporting year 2000. After 2000, the majority of data remain in pounds, though dioxins are reported in grams. Also included are the SIC classification codes for each reporting producer, which we use to identify polluters impacted by the policy change.[20]

Prior work on the TRI and home values has separated toxics by categories of potential health damage to test for differences across assessed health risk, and found none (Bui and Mayer, 2003). Our primary models do not separate by toxicity (though results in the earlier version of this paper Sanders (2011) were robust to focusing on toxicity). In future advancements of this work, we plan to separate emissions by whether or not toxics are classified as cancerous.

5 Methodology

Our analysis uses a change to the TRI, a publicly available data set produced by the EPA. Since its first release in 1989, the TRI has focused on the toxic releases produced by the manufacturing sector. It wasn't until a legislative change in 1998 that a number of heavy polluters, including coal power plants, were required to provide information for the TRI. This change provides a unique opportunity to examine the role of changes in amenity information, as the change is entirely independent of any changes to the level of the amenity itself.[21]

[20] In later years, the TRI reports NAICS codes rather than SIC codes.

[21] Most changes in environmental quality, or similar amenities, are accompanied by other, potentially confounding changes such as recessionary periods, economic development, or other such factors that may influence hedonic pricing estimates (e.g., a change in crime drives an increase in focus on crime reporting).

We match TRI facilities to location based on TRI location data. We use reported toxics from newly-reporting SIC codes as the primary information variable. Hereafter, we use the term "emissions" and "pollution" to refer specifically to 1998 TRI-reported toxic emissions coming from firms that had SIC codes impacted by the 1998 policy change unless otherwise specified. Using newly reported pollution is effectively synonymous with using changes in reported pollution between the 1997 and 1998 TRI reports. Figure A-2 in the Appendix shows an almost 1-to-1 relationship between the change in reported toxics for impacted areas and the reported toxics from newly added SIC codes. In zip code level graphical analysis, we aggregate all newly reported emissions to zip code level totals. In regressions done at the home-sale level, we do results location by location. When multiple locations exist within a given distance range, we focus on the largest emitter.

Because we examine an information shock that occurs at a discrete point in time, we follow a regression discontinuity design similar to Moulton, Waller, and Wentland (2018), which is broadly consistent with the event study literature in finance and applied microeconomics more generally. Our goal is to estimate how the home prices responds to this new information; so, we leverage a research design that accounts for portion of a home's price that can be explained by the idiosyncratic characteristics the home, allowing the remaining portion to be explain by time-specific shocks. In particular, the design consists of the combination of a hedonic sale price model and a standard linear spline RD model using the sale week as the running variable, as seen in equation (1)[22]. For this study, we focus on logged home price as the outcome of interest over the 26 weeks prior to and after the disclosure event. Traditionally, hedonic regression analysis has been a commonly used methodology in the

No such concerns are present here, as our identification comes from a change in data rather than a change in exposure.

[22] As our running variable is time, we acknowledge that it may be more precise to refer to this as interrupted time series (ITS). It is not uncommon to use time as a running variable, where the discontinuity is a point in time (for example, Moulton, Waller, and Wentland (2018)). See also Hausman and Rapson (2017) who discuss RD using a time running variable more generally.

housing literature since Rosen (1974); but, more recently the approach has been increasingly coupled with a quasi-experimental framework (for a review, see Parmeter and Pope, 2013).

$$ln(SalePrice_h) = \alpha + \beta_1(SaleWeek_h - C) + \beta_2(SaleWeek_h \geq C) * (SaleWeek_h - C) + \beta_3(SaleWeek_h \geq C) + \beta' X_h + \epsilon_h \quad (1)$$

Logged sale price of each individual house h is the outcome in the model. We re-center the sale week trend around the appropriate cutoff (C) following the policy shock in March 2000, limited the sample to +/- 26 weeks around this policy change. The coefficient β_1 captures the sale price time trend prior to the cutoff. We also include this same re-centered trend interacted with an indicator variable equal to one when the sale week was at or after the policy change cutoff. The coefficient β_2 represents the change in the post-cutoff price time trend, which can be used to determine if any price change following the announcement dissipates or grows over the post-cutoff window. The β_3 coefficient that is associated with an indicator variable equal to one when the sale week is after the policy change, which estimates the difference in the pre- and post-cutoff trends' intercepts at the cutoff. Thus, the estimated intercept difference in this design can be interpreted as the treatment effect of the announcement of new emission reports, which is the key coefficient of interest and is referred to and labeled as "Discontinuity" or "D" in the proceeding tables. X_h represents the following controls common to hedonic price regressions that account for observable characteristics of heterogeneous properties: square footage, number of bedrooms and bathrooms, size of garage (number of cars), acreage, age of the home, whether the home is a single-story ranch, has a pool, has a basement, sale day of the week fixed effects, indicators for no acreage, missing bedrooms or bathrooms, and zip code fixed effects. Standard errors are two-way clustered at the zip code level and by sale week.

Methodologically, the controls serve a number of purposes. We are examining a large cross-section of homes over time, and these homes are heterogeneous along numerous key dimensions. While aggregation across a large national data set may allay compositional concerns, controlling for arguably the most important determinants of a home's price (size, bedrooms, bathrooms, age, location, etc.) allows for a more straightforward apples-to-apples comparison of a cross-section of homes within a given period. That is, a handful of characteristics and location explain most of the time-invariant variation in home prices. Hence, the estimated emission announcement effect comes from the *variation in price not explained by these factors*, just as the "event" in the finance literature explains the variation in a firm's equity price or return not already explained by quantifiable firm-specific factors that make up the fundamentals of its valuation.

We first estimate the model in equation 1 separately for homes that are situated in concentric circles that are 0 to a half mile, a half mile to a mile, one mile to two miles, two to five miles away, and more than five miles away from emitting plants. We expect that homes closest to the polluting plants will be most affected by the policy. Our default specification restricts the sample to only those homes that are located near plants emitting 100 or more thousands of tons, but also provide results for plants emitting more than 30, 60, or 80 thousands of tons. We hypothesize that the estimated effect should be larger for homes near higher emitting plants, as the highest polluting plants tended to be most noticeable by their neighbors and tended to receive the most news coverage during the time period around the information release. Alternatively, rather than stratifying the sample by the concentric circles, we modify equation 1 to include an interaction of each concentric circle with the discontinuity and linear spline trend variables, with the homes that are five or more miles away from the pollution emitting plants serving as the reference group. The discontinuity coefficient associated with each concentric circle estimates the difference in discontinuities, or the extent to which home prices discontinuously changed in relation to homes significantly far enough away to serve as a control or placebo.

$$
\begin{aligned}
ln(SalePrice_h) = \alpha + \beta_1(SaleWeek_h - C) * (DistanceCircle) +) \\
\beta_2(SaleWeek_h \geq C) * (SaleWeek_h - C) * (DistanceCircle) \\
+\beta_3(SaleWeek_h \geq C) * (DistanceCircle) + \beta' X_h + \epsilon_h
\end{aligned} \tag{2}
$$

6 Results

In an earlier version of this work, Sanders (2011) found no price response for newly added facilities with fewer than 100,000 tons of emissions at the zip code level. In the following analysis we similarly focus on only high emitters. Table 2 shows our results by distance range and a similar cutoff of over 100,000 tons. A home sale is considered treated if it falls within the stated distance range of a newly reporting facility with at minimum 100,000 tons of reported emissions. Using a range of 0-0.5 miles, we find that introduction of a relevant TRI site reduced housing sales prices by approximately 4.5%. However, this sample is very limited, as only around 3,300 homes fall within such a close range of the relevant TRI sites. When we expand this circle to cover 0.5-1 miles, we see an *increase* in home prices, though the result is not statistically significant. For the remaining distance circles, we find similar results: a positive but statistically insignificant increase in home prices. Figure 3 provides a graphical examination of our results from Table 2. In the 0-0.5 mile range, there is a general upward trend in prices, with a break around the time of TRI information release. However, for any distance donuts larger than 0.5 miles and less than 5 miles, we observe a small jump in prices followed by a general upward trend.

None of these regressions involve a control group — they are single difference, pre/post regressions, comparing housing prices for sales before the new data release to those after within the specified distance range. We next add an additional difference by comparing houses within the specified range pre/post to houses beyond 5 miles pre/post. We also expand our treatment group to cover more of the reporting distribution (e.g., plants over the 90th, 95th, 97th, and 98th percentile of reporting emissions, corresponding to approximately 30,000, 60,000, 80,000, and 100,000 tons), we see a more stable pattern in effects. Column 1 of Table 3 shows that, looking at homes near sites that reported above the 90th percentile, houses within 0.5 miles fall in price by approximately 3.5% relative to the omitted group furthest away. The effect jumps around as we increase our distance donut, but is largely stable around 3%. A key empirical result of the paper is that as we increase the cutoff to over the 98th percentile, we find estimates are larger, with losses around 8% for houses within a half mile, decreasing to around 3% for houses 2-5 miles away (where all houses beyond 5 miles serve as the control group). We note here that our coefficient on "post" is around 4%, even after controlling for a running variable in time. This suggests houses beyond 5 miles see an increase in value that simple time trends do not explain. Understanding what drives this effect is an avenue for our future research.

6.1 The Financial Impact of Information Capitalization

To place our findings within the context of similar studies on environmental bads, it is useful to consider prior environental hedonic estimates using housing values. Davis (2011) finds the construction of new natural gas power plants reduced home values within 2 miles of plants by 4.1-7.1 percent.[23] Chay and Greenstone (2005) find that Clean Air Act total suspended particulate reductions during the 1970s increased home values by 2-3.5 percent, and Bento et al. (2011) find similarly sized county-level results for the later 1990 Amendments. Gamper-Rabindran et al. (2011) find cleanup of Superfund sites raised highly localized housing values by up to 19 percent, though at a different level of aggregation Greenstone and Gallagher

[23] Davis (2011) shows many new plants opened in 2000, but notes almost all new plants were natural gas plants, which are exempt from reporting to the TRI. For our results to be due to newly constructed power plants, new natural gas plants would have to have opened in the same areas as already existing impacted industries at the same time as the new TRI release. Davis (2011) also finds the effects of being close to a power plant fade within approximately 2 miles, meaning the probability of a treatment zip code in our analysis being close to a treatment area from that analysis is relatively small.

(2008) find cleanups to have no effect.

More directly related to the dissemination of environmental information, Davis (2004) finds that the increased information on cancer clusters dropped home values by 14 percent, while Gayer et al. (2000) find increased information on Superfund hazardous waste risk shifted risk expectations downward but still led to a home price decrease of approximately 1 percent.[24,25] As noted in Section 3, most prior works on the TRI and housing values find no consistent change in home prices due to new TRI information (Bui and Mayer, 2003; Oberholzer-Gee and Mitsunari, 2006), but most recently, Mastromonaco (2012) finds the 2000 TRI policy change regarding lead and PBTs was associated with a value drop of up to 8.6 percent for nearby homes in California.

7 Discussion

7.1 Prior findings from the TRI and housing markets

Prior work by Bui and Mayer (2003) and Oberholzer-Gee and Mitsunari (2006) on TRI information and housing prices finds no consistent effects of newly reported TRI releases. We note there is a large difference in the size of the information shock regarding changes in reported emissions. Figure 4 reports average releases per zip code for impacted zip codes from 1988 through 2002. Conditional on non-zero releases, the average reported releases in the first TRI data were less than 200 tons (mean non-zero releases were around 200 tons in Bui and Mayer (2003) as well). In 2000, however, average non-zero newly reported emissions level for treatment zip codes was almost 2 million pounds.

Our geographic variation is also larger, covering millions of sales across multiple states, and the world in which TRI data were released for the first time is different from that in

[24]This is calculated using their reported price drop of \$661 divided by the mean housing value of \$74,176 (in 1996 dollars).

[25]A recent example of the effect of information in non-environmental literature is Linden and Rockoff (2008) who find that releasing information on sexual offenders in the neighborhood lowers home values by approximately 4 percent.

which TRI data are updated.[26] The initial 1989 data release, for example, did not have the advantage of the Internet, and households had to seek out hard copies of the TRI if they wanted information. Data are now available online, news outlets have expanded both in number and scope of coverage, and additional information is more readily available on the dangers of environmental toxics. Communication was more costly in the past, so dissemination of information across households and neighborhoods is now higher.

Finally, in the first years of TRI data reporting, the housing market may not have held solid priors before the first TRI data were released, and it may have taken time before people knew how to interpret toxic data. By the time of the 2000 data release, the TRI had been around for over a decade. If the market believed the prior TRI releases were an accurate reflection of true ambient toxics, priors would have been more solidified, and thus their response would be greater with the 2000 data release.

8 Conclusion

We show how housing markets respond to increased information on local amenities, using changes in reporting policies as a shock to perception. The Toxics Release Inventory, a publicly released annual report of pollution produced by the manufacturing sector, served as a manner for households to assess local environmental amenities. Some of the largest polluters, such as power plants, did not report pollution information in the TRI until 1998. The addition of the relevant industries resulted in a large increase in publicly reported pollution, upwards of 800% in some areas.

We find the appearance of new data led to statistically and economically significant decreases in home sales prices that are non-linear in nature. At higher emissions levels, we see find decreases in home sales price of approximately 8% in areas very close to large polluters, relative to the reference group of homes further away, which experienced a small

[26] Bui and Mayer (2003) use 231 zip codes in one state, and Oberholzer-Gee and Mitsunari (2006) have data from 5 counties.

positive bump in prices after the information release. Information on how households view toxic pollution from fossil fuel power plants is important given recent potential expansions of coal power plant regulation expected to reduce mercury releases by approximately 90 percent (approximately 44 tons), and cost $10.9 billion in the year 2016.[27] Our results also speak to the role of market forces in the task of dealing with environmental externalities. Market mechanisms exist that, in theory, achieve socially efficient equilibria, but they require all markets to be fully informed about the size of the externality. At least in the case of environmental toxics, household perceptions of the externality are imperfect, which makes Pareto optimal free-market solutions unlikely.

[27]From the EPA mercury and air toxics fact sheet available at `http://epa.gov/airquality/powerplanttoxics/pdfs/proposalfactsheet.pdf`.

References

Atlas, M. (2007): "TRI to Communicate: Public Knowledge of the Federal Toxics Release Inventory," *Social Science Quarterly*, 88, 555–572.

Banzhaf, H. S. and R. P. Walsh (2008): "Do People Vote With Their Feet? An Empirical Test of Tiebout's Mechanism," *The American Economic Review*, 843–863.

Bento, A., M. Freedman, and C. Lang (2011): "Spatial and Social Disparities in the Benefits from Air Quality Improvements," Mimeo.

Blomquist, G. (1974): "The effect of electric utility power plant location on area property value," *Land Economics*, 50, 97–100.

Bui, L. T. and C. J. Mayer (2003): "Regulation and capitalization of environmental amenities: Evidence from the toxic release inventory in Massachusetts," *Review of Economics and Statistics*, 85, 693–708.

Chay, K. Y. and M. Greenstone (2005): "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy*, 113, 376–424.

Coase, R. (1960): "The problem of social cost," *Journal of Law and Economics*, 3, 1–44.

Currie, J. (2011): "Inequality at Birth: Some Causes and Consequences," *American Economic Review: Papers & Proceedings*, 101, 1–22.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2011): "Toxic Releases and Health at Birth: Evidence from Plant Openings and Closings," Mimeo.

Currie, J. and J. F. Schmieder (2009): "Fetal Exposures to Toxic Releases and Infant Health," *American Economic Review: Papers & Proceedings*, 99, 177–183.

Davis, L. W. (2004): "The Effect of Health Risk on Housing Values: Evidence From a Cancer Cluster," *American Economic Review*, 94, 1693–1704.

——— (2011): "The Effect of Power Plants on Local Housing Values and Rents," *Review of Economics and Statistics*, 93, 1391–1402.

de Marchi, S. and J. T. Hamilton (2006): "Assessing the accuracy of self-reported data: an evaluation of the Toxics Release Inventory," *Journal of Risk and Uncertainty*, 32, 57–76.

Environmental Protection Agency (1999): "33/50 Program: The Final Record," .

Gamble, H. and R. Downing (1982): "Effects of Nuclear Power Plants on Residential Property Values," *Journal of Regional Science*, 22, 457–478.

Gamper-Rabindran, S., R. Mastromonaco, and C. Timmins (2011): "Valuing the Benefits of Superfind Site Remediation: Three Approaches to Measuring Localized Externalities," NBER Working Paper 16655.

Gayer, T., J. T. Hamilton, and K. Viscusi (2000): "Private Values of Risk Tradeoffs at Superfund Sites: Housing Market Evidence on Learning about Risk," *Review of Economics and Statistics*, 82, 439–451.

General Accounting Office (1991): "Toxic Chemicals: EPA's Toxic Release Inventory Is Useful But Can Be Improved," Tech. rep.

Graff Zivin, J., M. Neidell, and W. Schlenker (2011): "Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption," *American Economic Review: Papers & Proceedings*, 101, 448–453.

Greenstone, M. and J. Gallagher (2008): "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program*," *Quarterly Journal of Economics*, 123, 951–1003.

Hamilton, J. T. (1995): "Pollution as News: Media and Stock Market Reactions to the Toxics Release Inventory Data," *Journal of Environmental Economics and Management*, 28, 98–113.

Hebert, H. J. (May 12, 2000): "Mining, power plants lead in releasing toxic pollutants," Pittsburgh Post-Gazette.

Hu, W. (May 12, 2000): "Estimate of Toxic Chemical Emissions Is Tripled," The New York Times.

Khanna, M., W. R. H. Quimio, and D. Bojilova (1998): "Toxics Release Information: A Policy Tool for Environmental Protection," *Journal of Environmental Economics and Management*, 36, 243–266.

Konar, S. and M. A. Cohen (2006): "Does the Market Value Environmental Performance?" *The Review of Economics and Statistics*, 83, 281–289.

Leggett, C. G. and N. E. Bockstael (2000): "Evidence of the Effects of Water Quality on Residential Land Prices," *Journal of Environmental Economics and Management*, 39, 121 – 144.

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

Mastromonaco, R. (2012): "Do Environmental Right-to-know Laws Affect Markets? Capitalization of Information in the Toxic Release Inventory," Mimeo.

Moretti, E. and M. Neidell (2011): "Pollution, Health, and Avoidance Behavior," *Journal of Human Resources*, 46, 154–175.

Murray, S. D. (May 12, 2000): "BGE plants near top for toxins Two facilities release 11.5 million pounds in 1998, EPA report says," The Sun.

Neidell, M. (2009): "Information, avoidance behavior, and health," *Journal of Human Resources*, 44, 450–478.

Neidell, M. J. (2004): "Air pollution, health, and socio-economic status: the effect of outdoor air quality on childhood asthma," *Journal of Health Economics*, 23, 1209–1236.

Nelson, J. P. (1981): "Three Mile Island and residential property values: Empirical analysis and policy implications," *Land Economics*, 57, 363–372.

Oberholzer-Gee, F. and M. Mitsunari (2006): "Information regulation: Do the victims of externalities pay attention?" *Journal of Regulatory Economics*, 30, 141–158.

Sanders, N. J. (2011): "Toxic Assets: How the Housing Market Responds to Environmental Information Shocks," Mimeo.

Schlenker, W. and J. Scorse (2011): "Does Being a "Top 10" Worst Polluter Affect Facility Environmental Releases? Evidence from the U.S. Toxic Release Inventory and Scorecard," Mimeo.

Tiebout, C. M. (1956): "A pure theory of local expenditures," *Journal of Political Economy*, 64, 416–424.

Table 1
Summary Statistics for All Homes with Sales in 52 Week Period Around the TRI Announcement

Price	141,667
	(77,817)
Age	28.13
	(26.32)
Acreage	0.27
	(0.31)
No Acreage	0.15
	(0.35)
Square Foot	1729
	(734)
Ranch	0.45
	(0.50)
Pool	0.05
	(0.22)
Basement	0.22
	(0.42)
Garage Size	1.13
	(1.01)
Bathrooms	1.92
	(0.71)
Bedrooms	3.06
	(0.82)
Missing Bed or Bath	0.31
	(0.46)
N	1,494,239

Notes: Summary statistics for home information used in main analysis. Data derived from Zillow information on recorded home sales.

Table 2
Impact of Being Near a Newly Reporting TRI Facility by Distance

Distance	0 to 0.5mi Log(Price) (1)	0.5 to 1mi Log(Price) (2)	1 to 2mi Log(Price) (3)	2 to 5mi Log(Price) (4)	More 5mi Log(Price) (5)
Discontinuity	-4.48*	2.84	0.51	1.34	0.84
	(2.63)	(1.76)	(1.15)	(1.09)	(0.87)
Trend	0.18	-0.02	-0.03	-0.07	-0.05
	(0.13)	(0.08)	(0.06)	(0.05)	(0.03)
Post-Trend	0.23	0.20*	0.29***	0.37***	0.34***
	(0.18)	(0.09)	(0.07)	(0.06)	(0.04)
N	3,258	14,634	62,872	332,989	1,080,156
R2	0.78	0.75	0.71	0.67	0.64

Notes: Primary results of the impact of TRI information revelation on local housing prices. Includes only home transactions within the specified distance of TRI facilities that began to report as a result of the 1998 legislation and reported over 100,000 tons of emissions, in a range of +/- 26 weeks around the reporting date. Regressions use equation 2, allowing for differential trends in housing prices by distance donut both before and after the policy. "D" is an indicator for being after March 2000 when the first new TRI data were reported. Main outcome is log of housing price, so estimates are effectively percentage changes. N represents the number of observed home sales. All regressions focus on the 26 weeks before and after the 2000 data release.

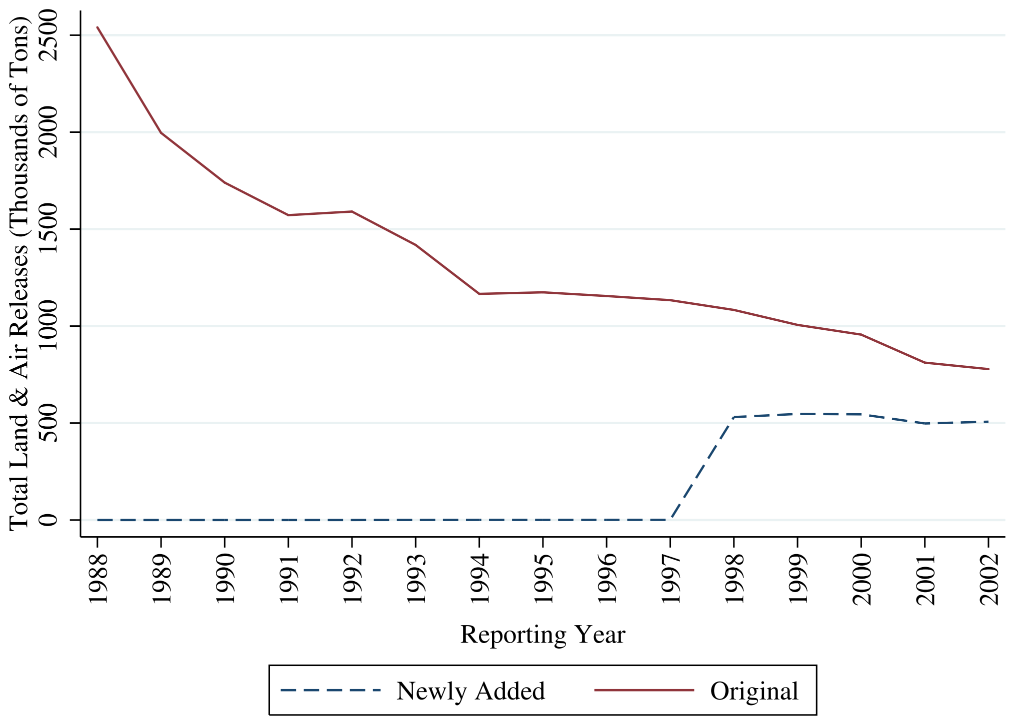
Table 3
Impact of Being Near a Newly Reporting TRI Facility by Range of Newly Reported Emissions

Percentile (for non-zero)	90th	95th	97th	98th
Emissions >	30	60	80	100
	Log(Price)	Log(Price)	Log(Price)	Log(Price)
	(1)	(2)	(3)	(4)
D×(0 to 0.5mi)	-3.44*	-6.51***	-7.21***	-7.92***
	(1.93)	(1.70)	(1.95)	(2.44)
D×(0.5 to 1mi)	-1.90*	-2.90**	-1.72	-1.20
	(1.09)	(1.43)	(1.67)	(1.96)
D×(1 to mi)	-3.14***	-3.08***	-3.16***	-3.47***
	(0.99)	(1.10)	(1.07)	(1.17)
D×(2 to 5mi)	-2.91***	-2.78***	-2.99***	-2.88***
	(0.89)	(0.91)	(0.97)	(0.96)
D	4.20***	4.19***	4.24***	4.30***
	(0.55)	(0.56)	(0.56)	(0.56)
N	1,494,239	1,494,239	1,494,239	1,494,239
R2	0.66	0.66	0.66	0.66

Notes: Primary results of the impact of TRI information revelation on local housing prices. Includes only home transactions within the specified distance of TRI facilities that began to report as a result of the 1998 legislation and reported over the indicated tons of emissions, in a range of +/- 26 weeks around the reporting date. Regressions use equation 2, allowing for differential trends in housing prices by distance donut both before and after the policy. "D" is an indicator for being after March 2000 when the first new TRI data were reported. Main outcome is log of housing price, so estimates are effectively percentage changes. N represents the number of observed home sales. All regressions focus on the 26 weeks before and after the 2000 data release.

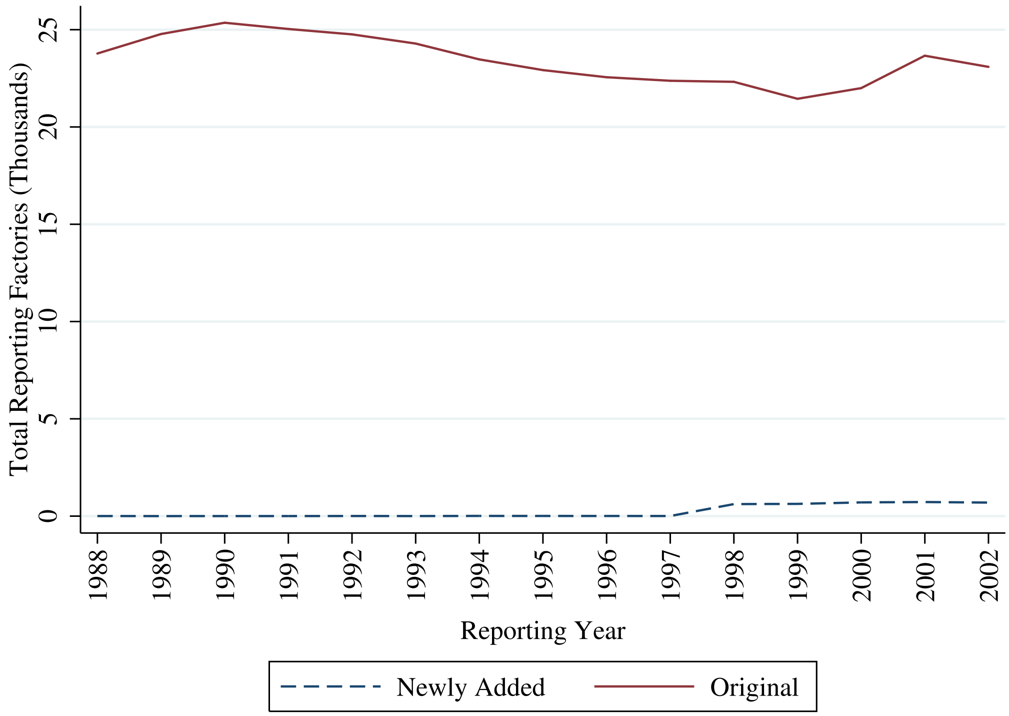
9 FIGURES

Figure 1
Total Toxic Releases Reported for Impacted vs. Not Impacted Industries



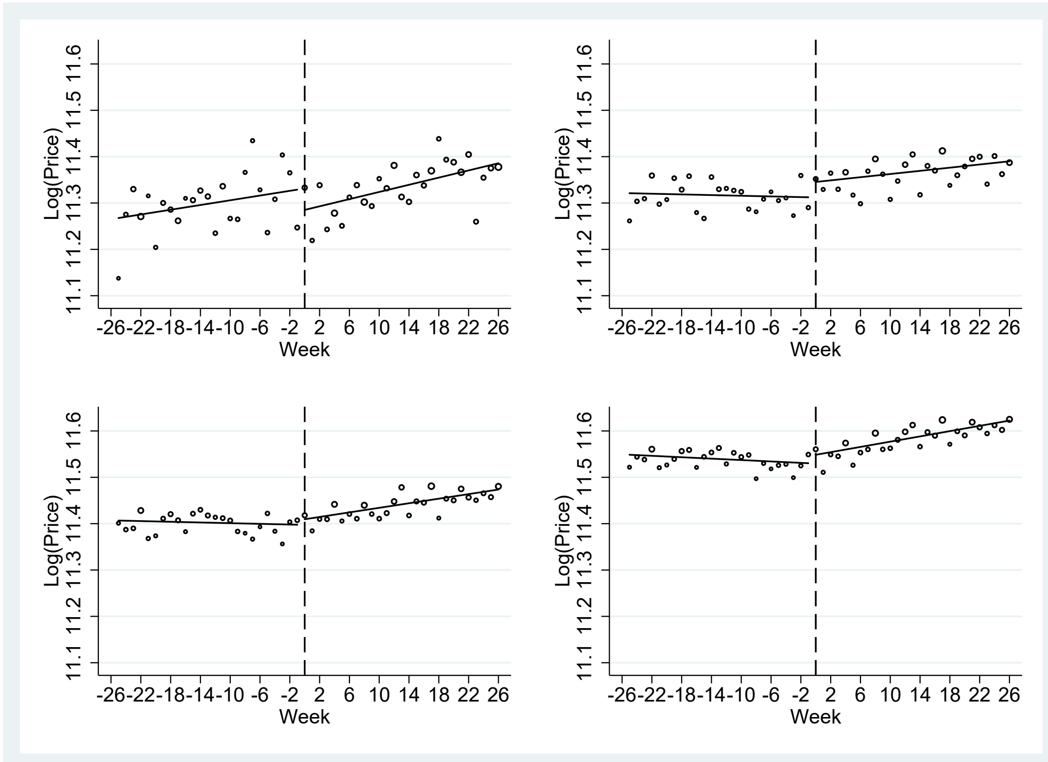
Notes: Toxics are the sum of all land and air releases, in thousands of tons, across all toxics recorded as reported in the Toxics Release Inventory. "Newly Added" indicates firms classified under SIC codes 10, 12, 4911, 4931, 4939, 4953, 5169, 5171, and 7389 (see Section 2). "Original" includes all other industries. Data are from all available TRI locations and are not restricted to the zip codes used in the primary analysis.

Figure 2
Number of Reporting Facilities for Impacted vs. Not Impacted Industries



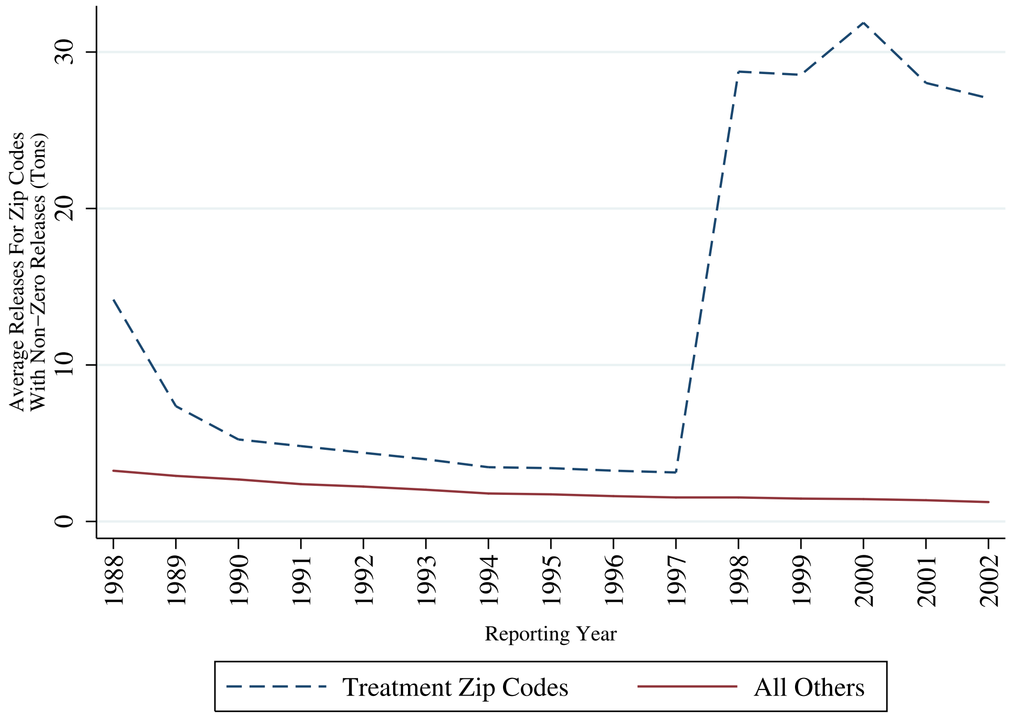
Notes: Count of total reporting firms reporting any non-zero land and air releases to the Toxics Release Inventory. "Newly Added" indicates firms classified under SIC codes 10, 12, 4911, 4931, 4939, 4953, 5169, 5171, and 7389 (see Section 2). "Original" includes all other industries. Data are from all available TRI locations and are not restricted to the zip codes used in the primary analysis.

Figure 3
Graphic Illustration of Results from Table 2



Notes: Graphs correspond to columns 1, 2, 3, and 4 of Table 2 (from top left to bottom right).

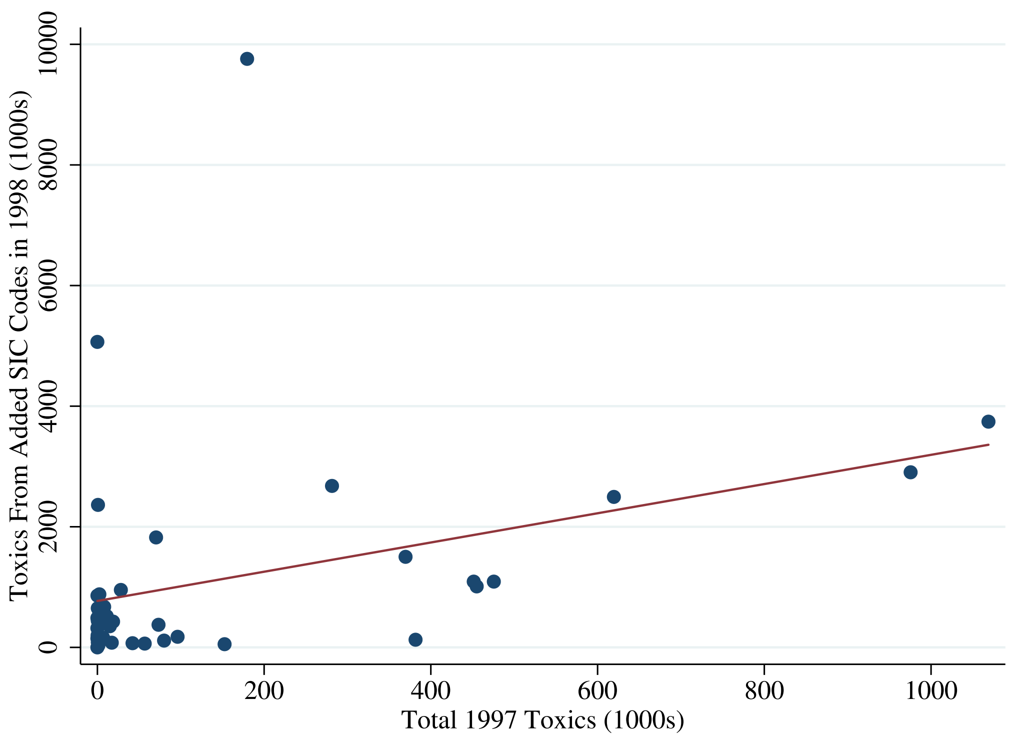
Figure 4
Average Reported Toxics Released per Zip Code by Treatment



Notes: Toxics are the sum of all land and air releases, in thousands of tons, across all toxics recorded as reported in the Toxics Release Inventory. Average per zip is calculated by dividing total releases by number of zip codes in each group. "Treatment" is classified by zip code and based on the amount of new SIC code reported toxics. Zip codes are classified as treated if the inverse hyperbolic sine of 1998 TRI releases from SIC codes 10, 12, 4911, 4931, 4939, 4953, 5169, 5171, or 7389 was above 13. Includes data from all 2,842 zip codes (2,796 control, 46 treatment) used in the primary analysis.

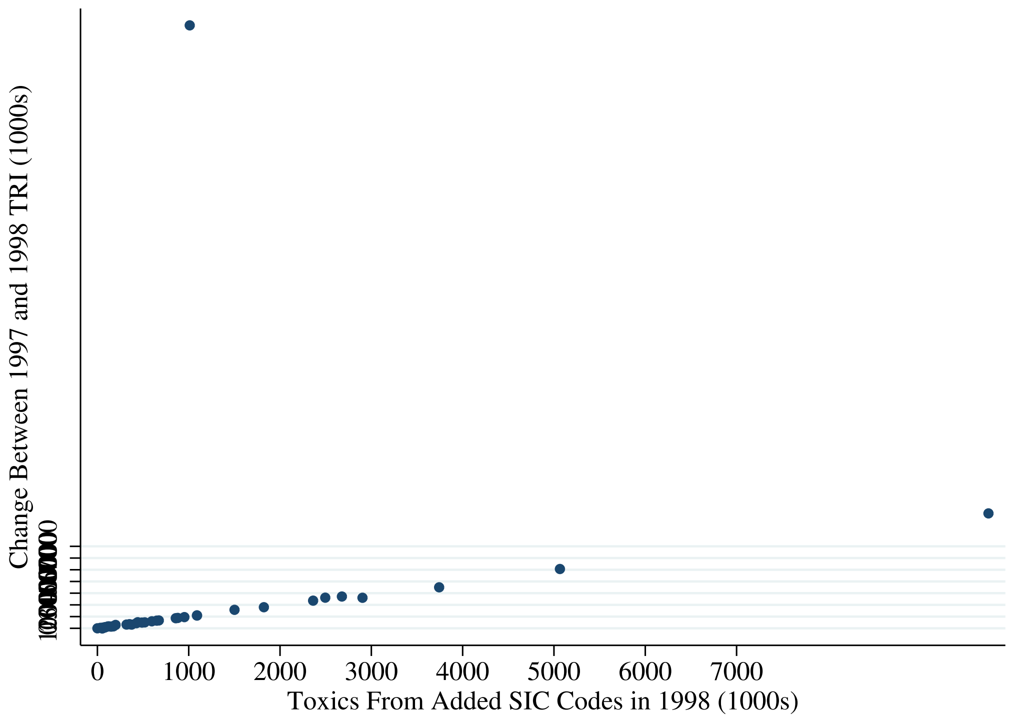
A Appendix A

Figure A-1
Scatter Plot of 1997 Toxic Levels and 1998 Toxics in New SIC Code Categories



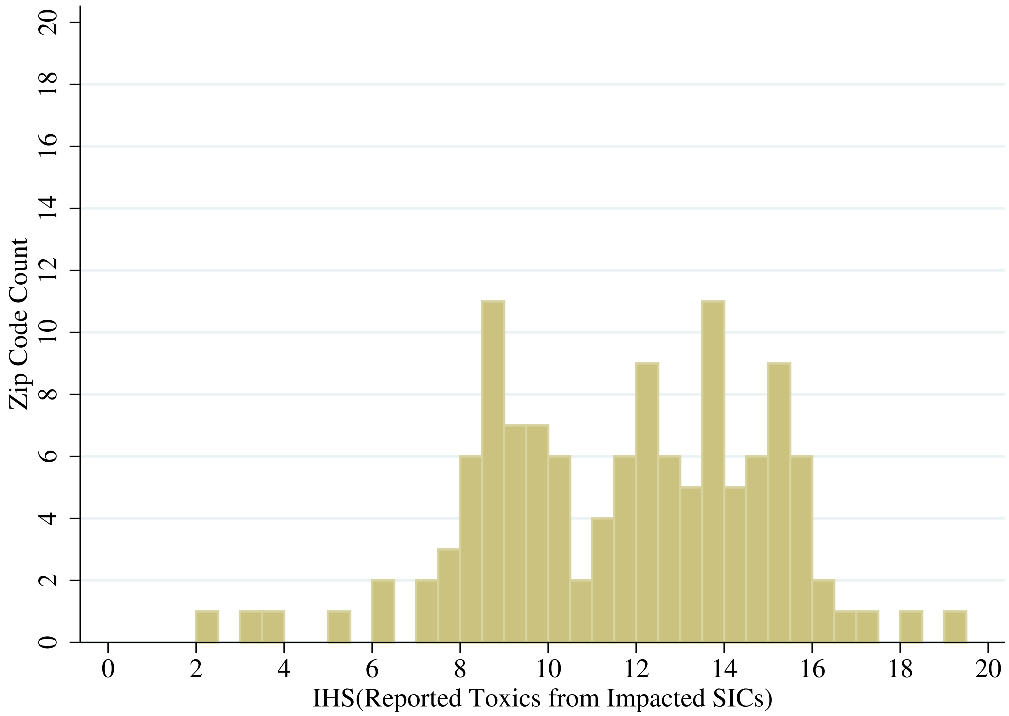
Notes: Horizontal axis is all land and air toxics reported in the 1997 TRI release, in thousands of pounds. Vertical axis is the 1998 TRI reported toxics for only new SIC codes 10, 12, 4911, 4931, 4939, 4953, 5169, 5171, and 7389 (see Section 2), in thousands of pounds. Scatter plot only includes zip codes with non-zero toxics for newly impacted SIC codes used in the primary analysis (130 total zip codes, see Section 5). Graph includes linear predicted fit of new 1998 releases using 1997 totals.

Figure A-2
Scatter Plot of Change in TRI Toxics from 1997 to 1998 and 1997 Toxics in New SIC Code Categories



Notes: Horizontal axis is the 1998 TRI reported toxics for only new SIC codes 10, 12, 4911, 4931, 4939, 4953, 5169, 5171, and 7389 (see Section 2), in thousands of pounds. Vertical axis is the change in all reported toxics between 1997 and 1998 TRI releases. Scatter plot only includes zip codes with non-zero toxics for newly impacted SIC codes used in the primary analysis (130 total zip codes, see Section 5).

Figure A-3
Histogram of Inverse Hyperbolic Sine-adjusted New SIC Code Releases



Notes: Histogram of land and air releases reported in the 1998 TRI for all new SIC codes 10, 12, 4911, 4931, 4939, 4953, 5169, 5171, and 7389 (see Section 2), normalized using the inverse hyperbolic sine function (see Section 5). Includes only zip codes with non-zero toxics for newly impacted SIC codes used in the primary analysis (130 total zip codes, see Section 5). Vertical axis shows count of total zip codes in each bin. Bin width is 0.5.

B-1 Appendix C: Third-Party Sources of TRI Information

TRI data can only change behavior if information is accessed and used in the household decision process. Early TRI data were available in hard copy from the EPA, and eventually on compact data disc. Later, a number of sources made data publicly available online, through venues such as the EPA website or the Right-to-Know network (www.rtknet.org). For a period in the late 1990s and early 2000s, the website Scorecard (scorecard.goodguide.com) provided rankings of the worst polluters by area, which Schlenker and Scorse (2011) use to identify the effects of being a "Top 10" polluter on later firm releases. All require active decisions to seek out data, but Atlas (2007) found that in a survey of approximately 1,300 people, few individuals knew about the TRI or the names of TRI facilities in their area. A report by the United States General Accounting Office found that "more than half of the residents in three counties with high levels of emissions were unaware that the data were available to the public" (General Accounting Office, 1991). This raises questions for this and any analysis considering the response to specific TRI information.

One information vector is the popular media, which brought the TRI to the attention of households around each new data release. The media paid particular attention when power plants were added to the list of reporters. A survey of news stories from LexisNexis®, finds stories with "Toxics Release Inventory" in the headline or opening paragraphs occur with high frequency every year around March through June when the EPA releases new data. One of the largest spikes occurs in mid 2000, when the EPA press release specifically notes press packets have information on new, highly polluting sectors. As examples of how the media relayed this information, I include below text from three articles released on May 12th, 2000, that note specific locations recently targeted as high polluters. As early as March of 2000, state-level EPA departments had begun producing press releases and public reports on some of the worst newly reporting polluters, and there are a number of articles starting in early March detailing local pollution levels.

As a measure of when and to what extent the TRI is discussed in the media, Panel A of Figure B-1 shows, by month, the number of articles on LexisNexis® that mention the TRI in the headline or leading paragraph. Panel B shows counts for occurrences of the words coal, oil, and electricity within the TRI articles, which would have been most relevant given fossil fuel power plants were a major factor in the increase in reported levels. Dashed lines indicate the annual official Federal EPA data releases. Almost all stories on the TRI occur just around the annual releases, and a particularly large number of articles appear in 2000 and 2002, the releases corresponding to the data impacted by the 1998 and 2000 reporting policy changes, respectively. The coal/oil/electricity graph shows the substantial increase in articles regarding power plants after 1998. The increase in 1999 is a combination of news articles discussing that the 1998 data were now gathered and 2000 data will include coal power plants, and some plants making data public information on their own just before providing it to the EPA in July of 1999, which may have sparked media interest.

B-1 Example TRI news articles

> Two Baltimore Gas and Electric Co. power plants in Anne Arundel County released 11.5 million pounds of toxic chemicals into the air in 1998, ranking them first in the state and 11th in the nation for toxics, the U.S. Environmental Protection Agency said yesterday.[28] (*The Sun*)

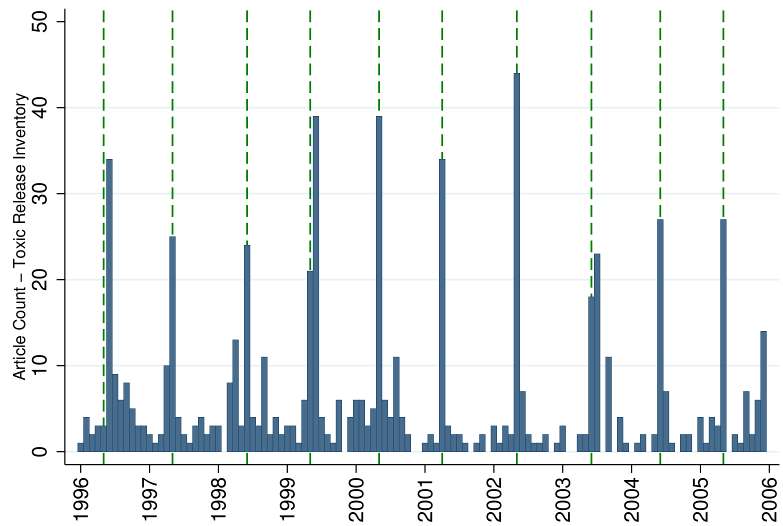
> The heaviest polluters were the 27 power plants in Ohio, which emitted 113.9 million pounds of toxic chemicals in 1998. In comparison, the 34 power plants in New York released 18.7 million pounds, and the 15 power plants in New Jersey released 8 million pounds.[29] (*The New York Times*)

> The report shows that two of the state's coal power plants, Sithe Energy's Keystone plant in Armstrong County and Edison Mission Energy Inc.'s Homer City plant in Indiana County are among the top 20 power plants in the nation, releasing a combined 18.5 million pound (sic) of toxic chemicals into the air, water, and land.[30] (*Pittsburgh Post-Gazette*)

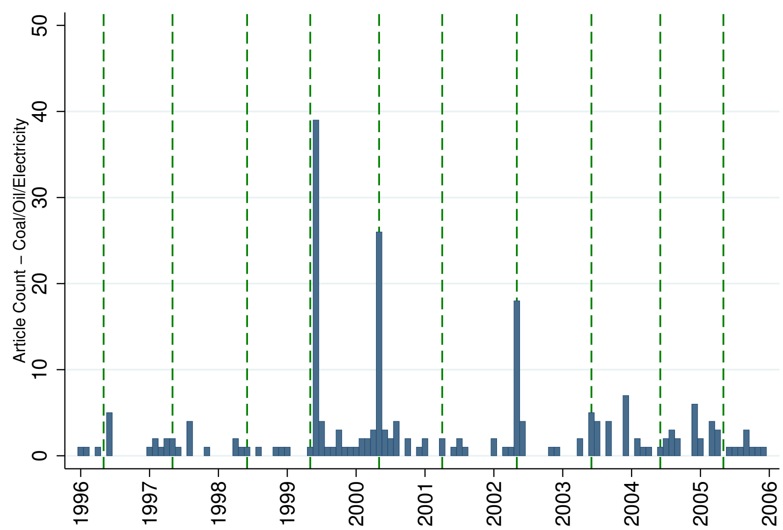
[28] Murray (May 12, 2000)
[29] Hu (May 12, 2000)
[30] Hebert (May 12, 2000)

Figure B-1
Number of Articles Archived on LexisNexis® Containing Selected Keywords

Panel A: Occurrence of "Toxics Release Inventory"



Panel B: Occurrence of "Coal", "Oil", and/or "Electricity"



Notes: Counts of news stories archived on LexisNexis® containing particular text in specifically the headline or opening paragraph, by month and year. Relevant texts are shown on y-axis labels. Dashed lines mark the annual Federal public release time of newest TRI data, as specified by the EPA website.