The Environmental Kuznets Curve in Small Geographies*

Dakshina G. De Silva† Robert P. McComb‡ Anita R. Schiller§ Aurelie Slechten¶

Abstract

We consider the relationship between location choices of polluting industries and local incomes. Working at the census tract-level, we find these choices are positively correlated with local income over lower income levels and these firms' demands for measures to reduce the likelihood of toxic release, as proxied by waste management firms, show a similar pattern. We find evidence the relative frequency of toxic release is negatively correlated with proximate incomes. We conclude that the Environmental Kuznets Curve is related to profit maximizing behavior of firms seeking to maximize expected profits in recognition of the financial risk associated with toxic release.

JEL Classification: Q2, L6, R1.
Keywords: EKC, Environmental Justice, Agglomeration, Entry and Exit.

*We thank George Deltas, James Hartigan, seminar audiences from INRA in Montpellier and Iowa State University and participants at the 2018 Asian Meeting of the Econometric Society for comments and suggestions. We would also like to thank the Texas Workforce Commission for providing us with fully disclosed Quarterly Census of Employment and Wages data at the establishment level.

†Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK (e-mail: d.desilva@lancaster.ac.uk).

‡Department of Economics, Texas Tech University, MS: 41014, Lubbock, TX 79409-1014, (e-mail: robert.mccomb@ttu.edu).

§Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK (e-mail: anita.schiller@lancaster.ac.uk).

¶Department of Economics, Lancaster University Management School, Lancaster University, Lancaster, LA1 4YX, UK (e-mail: a.slechten@lancaster.ac.uk).

1 Introduction

Local pollution generated by the presence of manufacturing activities has been shown to have a causal impact on human health (Chay and Greenstone, 2003; Graff Zivin and Neidell, 2012). Currie and Schmieder (2009), for example, find that toxic chemicals released by these manufacturing activities have a significant negative effect on birth outcomes. The effects of local pollution on households will of course differ as their proximity to the source of pollution differs. Since these potentially polluting activities are located in industrial concentrations discretely distributed across the landscape, the distribution of polluting activities within a region or country is not uniform and households' exposure differs accordingly.

Industrial localization and concentration has been of considerable interest since, at least, Marshall in 1920 and spawned a broad literature. However, there is the more specific question of whether the localization and distribution of polluting activities in the production of tradeables is the result of firms' strategic decisions based on local demographic characteristics. This more nuanced question falls within the strand of literature in the realm of environmental justice. That is, are low income areas or areas with larger fractions of minorities disproportionately affected by potential health risk associated with exposure to toxic releases (Hamilton, 1995; Levinson,1996; Wolverton, 2009)? Are polluting firms more likely to locate in lower income or higher minority areas? Do households self-select locally or regionally by income due, in part, to a correlation between housing prices and local environmental quality? Or, is there some combination of these circumstances?

This paper addresses the first question by investigating the relationship between the location and toxic waste management choices of potentially polluting firms and local income levels.[1] We first provide a theoretical framework to analyze this relationship. A potentially polluting firm seeking to maximize profits will be concerned about the liability of toxic releases and the threat such releases

[1]A potentially polluting firm is defined as any firm, regardless of size or reporting requirements, in a NAICS code identified by the Environmental Protection Agency's Toxic Release Inventory (TRI), excluding NAICS 562 Hazardous Waste. However, only firms in these NAICS codes that employ at least 10 FTEs and exceed EPA threshold limits in terms of their processing or usage of designated hazardous or toxic chemicals are subject to mandatory reporting within the TRI (https://www.epa.gov/toxics-release-inventory-tri-program). The firms subject to mandatory reporting are denoted in this paper as TRI-reporting firms. TRI-reporting firms responsible for toxic releases that exceed TRI limits are identified and treated in this paper as TRI-polluters. So, TRI-polluters are a subset of TRI-reporting firms, and TRI-reporting firms are a subset of TRI-type (potentially polluting) firms.

pose to its financial results. It will therefore take the residential characteristics of a location into account when valuing the legal costs and compensation it will have to pay in the event of a toxic release (Coase, 1960) and balance these costs with the costs of influencing the likelihood of toxic release through waste management services. Compensation for damages will of course increase as local property values and incomes are higher. Moreover, in the event of a toxic release, the probability of collective action by residents and other businesses may be higher. As proposed by Olson (1965), the expected benefits/awards of such actions are directly related to local income levels.

The combination of these two factors leads to a negative relationship between firm-level pollution and local income, consistent with the environmental justice literature. As the firms' exposure to pollution-related financial risk increases, a prudential response is to manage that risk by employing more waste management and remediation services. Two otherwise identical firms, one in a high income neighborhood and the other in a low income neighborhood, would be expected to demand different levels of waste management and remediation services, positively correlated to the surrounding incomes.

Like any firm choosing a location, a firm engaged in a potentially polluting activity needs to consider the array of attributes of any particular location in terms of their importance in profitability.[2] A potentially polluting firm, however, must also consider the financial risk of release and the necessary costs of managing the likelihood of toxic release. Clearly, as suggested above, a toxic release in a location in an isolated or low-income area will pose less financial loss than would be the case for a firm located in a densely populated or high-income area. And, at the same time, labor costs and rents will be lower in lower-income areas. On the other hand, workforce availability, economic and social infrastructure are generally more attractive in densely populated areas. Firms seeking a high quality workforce or needing well-developed social and economic infrastructure will prefer higher income locations. So, there are, as always, trade-offs and the relationship between local income and the number of firms in a local area might be non-monotonic.

It seems reasonable to assume that the greater the number of potentially polluting firms, ceteris paribus, the greater the risk of pollution release. Recognizing that the incentives faced by a represen-

[2] Economists have long been interested in explaining what factors motivate profit-maximizing firms when the choose to open a new plant or expand an existing facility. There have been studies on the theory of plant location, including the role of taxes and agglomeration economies. Shadbegian and Wolverton (2012) review the theory, evidence, and implications of the role of environmental regulations in plant location decisions.

tative firm result in the realization of a localized aggregation of similar (potentially polluting) firms, we derive some sufficient conditions that yield an inverted U-shaped relationship between local income and the total level of potentially polluting activities in a neighborhood, i.e. a local Environmental Kuznets Curve (EKC). We then investigate empirically the existence of this local EKC and the predictions of our theoretical framework by looking at the relationship between the location choices of potentially polluting firms, pollution hazards and local income levels. In this analysis, we consider the demand for remediation or waste management services as a demand that arises from firms concerned about the liability of toxic release in terms of their expected profits. Although not perfectly correlated with localized firms' demand for environmental quality, waste management/remediation is the only clearly identified industry involved in pollution mitigation in the regional non-tradeables sector for which entry and employment data are available.

By and large, the potentially polluting activities considered here result from industrial activity whose output is not dependent on the local market, i.e., production of tradeables. Firms in these activities are free to choose any location, subject to zoning restrictions. One might naturally think of household demand for localized environmental quality to be expressed collectively through the political process and reflected in a regulatory or statutory framework that restricts the nature, location and technologies of productive activities. By restricting our analysis to a single state, Texas, we control for an otherwise heterogenous regulatory framework.[3] There are few local environmental restrictions imposed in the State of Texas, beyond local zoning laws, and the state itself takes a relatively light hand to regulation. Thus, jurisdictions in the state are largely subject to a practically identical regulatory environment.[4] Differences in local demand for waste management services must arise from the consumers of those environmental services, i.e., local firms responding to localized conditions, and we largely eliminate any localized version of the pollution haven hypothesis due to regulatory

[3] Texas is an attractive setting to consider given its size. It is the second-largest state in the U.S. both geographically and economically (with a gross state product of $1.6 trillion dollars in 2016). Its economy would rank 14th in the world when its gross state product is considered relative to national gross domestic products. It contains significant geo-physical diversity and is home to 25 separate MSAs.

[4] One example of the State's interference in local regulatory efforts was the widely reported lawsuit brought by the State against the City of Denton that banned hydraulic fracturing by referendum with 59% of the vote. The State sought to limit the municipality's ability to regulate oil and gas activities, to allow the State to pre-empt local regulations, and to ensure that all local efforts to impose regulations be "commercially reasonable." See the Texas Tribune, September 18, 2015 for reporting of the city's failure to block oil and gas activities.

heterogeneity.

One additional point needs to be made. Although residents might broadly express their preferences for environmental attributes through the political process, they can also express individual preferences for environmental attributes by moving across jurisdictions (in a Tiebout sense). Thus, a household that has a strong preference for low risk of toxic exposure would choose a locality in which there is little or no presence of potentially polluting activities. This would tend to put upward pressure on property values in the lower risk jurisdictions and to depress property values in the higher risk jurisdictions, assuming that few, if any, households would prefer a more exposed jurisdiction. This leads to reverse causality issues and this type of sorting would produce, through time, a negative simple correlation between the presence of polluting firms and median or average household income.

For this reason, in addition to the analysis of potentially polluting industry localization and household income, we consider the probability of entry of a potential polluter in and across the given geographies. By also focusing on entry, we avoid the question of inter-jurisdictional population sorting that might occur in the years following a potentially polluting firm's entry. Regional sorting by income due to the presence or absence of an environmental hazard, if it occurs, would already be reflected in observed household incomes in the areas proximate or distant to the pre-existing industrial concentrations (demographic characteristics are given at the time of entry). Given the importance of waste management activities to generate a local EKC relationship, we conduct a separate but similar analysis of both localization and likelihood of entry of waste management/remediation firms into those geographies while controlling for the presence of TRI-type firms. In both cases, we control for agglomeration economies that might serve to attract firms into existing industrial concentrations (Glaeser et al., 1992; Henderson et al. 1995; Combes, 2000; Rosenthal and Stange, 2003).

Working at the census tract-level, we estimate the relationship between localization of potentially polluting firms (TRI-type firms) and a set of co-variates including local household income. We find that locations and entry probabilities of potentially polluting firms are positively correlated with local income over only the lower range of income and that these firms' demands for measures to reduce the likelihood of toxic release, as proxied by the presence and entry of waste management/remediation firms, show a similar, but amplified, pattern and are positively correlated with the presence of potentially polluting firms. We also find very persuasive evidence that the relative frequency of toxic release, i.e., the ratio of toxic releases to the number of potentially polluting firms, is negatively correlated with proximate household income.

Both our model's stylized facts and empirical estimation are consistent with a relationship similar to that expressed by the EKC. Our results lead us to conclude that the inverse U-shaped relationship between income and toxic release is, at least partially, the product of potentially polluting firms seeking to maximize expected profits in recognition of the financial risk associated with a toxic release. This empirical finding at small geographical scale is novel in the field of environmental economics and has clear application to both the literature on environmental justice and macro-scale economic analyses of the EKC.

We say this because the environmental justice literature has primarily focused on the relationship between local income and pollution exposure (Arora and Cason, 1999 and, Brooks and Sethi, 1997). While these papers find some empirical evidence of an EKC-type curve, the theoretical relationship between levels of undesirable localized emissions and regional or local income that would generate an inverted U-shaped curve or the role of the remediation industry have not been investigated. Hamilton (1995) and Wolverton (2009) find evidence that polluting firms choose to locate disproportionately in poor areas. However, in their analysis, they consider only firms in the Environmental Protection Agency's Toxic Release Inventory (TRI) which have actually reported a toxic release. While these results might well portray an effective reality, we find that limiting the analysis to TRI firms that have a release on record is overly narrow and may miss a useful, broader picture.

Polluting firms in the TRI are clearly only a subset of the larger universe of firms that can potentially have a toxic release. The question should not revolve around previous releases, assuming post-release remediation measures have been successful, but rather potential releases. Since similar firms in different household income localities may choose different levels of effort to limit releases, firms in lower income or less developed areas may simply be more likely to experience a release because the lower financial consequences of a release do not justify the cost of additional precautionary measures.

Second, with the seminal papers by Grossman and Krueger (1993, 1995) that introduced the construction of the EKC, there was considerable interest in examining the empirical relationship between environmental quality and income (Harbaugh et al., 2002; Stern, 2004). Many theoretical models have tried to offer consistent explanations of the EKC. According to de Bruyn and Heintz (1999), theoretical approaches to explain the EKC fall within the range of five factors: (1) behavioral change and preference (Andreoni and Levinson, 2001; Lieb, 2002); (2) institutional changes (Jones and Maunelli, 2001; Egli and Steger, 2007); (3) technological and organizational changes (Selden and Song, 1995; Stokey, 1998); (4) structural changes (Marsiglio et al, 2016); and (5) international reallocation (Rothman, 1998). These models typically use the representative agent or social welfare framework. They are, thus, focused on the demand side of the equation and not suitable for considering how firms approach production and pollution decisions.

The structure of the paper is as follows: in section 2, we present our theoretical framework and in section 3 we develop our empirical approach to study the existence of a local EKC-type relationship. In section 4, we explain the patterns of entry and exit in the remediation industry. Section 5 lays out the conclusions.

2 Theoretical analysis

We assume that the expected level of local pollution is largely determined by the number of firms handling hazardous waste (potentially-polluting firms) and the efforts undertaken by those firms to prevent toxic releases through waste management practices such as treatment and recycling. These remediation services are typically supplied by waste management and remediation firms. Local characteristics are likely to affect these two elements. We focus on the role of local income and develop a framework to explain how it can affect both the number of potentially-polluting firms in a locality and their efforts to avoid releases of hazardous waste in the environment. We then use our theoretical structure to motivate our empirical analysis.

2.1 Firm-level toxic releases

A representative potentially polluting firm is located in area l, characterized by a median income m_l and some other local characteristics Z_l. The production process generates or uses some hazardous substance x. The firm has two options: (1) fail to prevent releases of the toxic substance into the local environment, which may pose a threat to the environment and human health; or (2) prevent some or all toxic releases from occurring through waste management/remediation practices.

Let e_l be the amount of the hazardous substance that is released into the local environment and $a_l = x - e_l$ be the amount that is treated or recycled. From an individual polluter's perspective, releasing toxic chemicals is costly because the firm will have to implement a clean-up program, pay penalties and compensate the local residents for damages.[5] We denote this cost by $e_l h(m_l)$. Then $h(m_l)$ is the unit cost of release (e.g. cost of *ex post* remediation per unit of pollutant and/or compensation paid to local residents per unit of pollutant) incurred by the representative firm in a locality characterized by income m_l. As suggested by the Coase Theorem (Hamilton, 1995), $h'(m_l) > 0$.[6] Higher incomes (and associated higher property values) are expected to increase the costs of release in a local area since, in litigation, injured parties recover damages based on reduced property values or, in the case of impacts that limit work or productive ability, lost income.[7] Avoidance of toxic releases does not generate any damage but has a quadratic cost $p(a_l)$ (with $p(0) = 0$, $p'(a_l) > 0$ and $p''(a_l) > 0$ for all $0 \leq a_l \leq x$). The total pollution cost can then be written as:

$$c_l(e_l, m_l) = e_l h(m_l) + p(x - e_l)$$

Conditional on the decision to locate a new plant in area l, a firm will choose a level of releases, $e_l^* = \arg\min_{e_l} c_l(e_l, m_l)$.[8] An interior solution $e_l^* > 0$ will be given by:

$$h(m_l) = p'(x - e_l^*)$$

the optimal pollution cost $c_l(e_l^*, m_l) = c_l^*$ is increasing in income, such that higher income areas imply higher costs of pollution (but lower than if firms did not have the opportunity to avoid releases: $c_l^* < c_l(x, m_l)$). We can also show that firms tend to reduce their releases in higher income locations:[9]

[5] Indeed, most disposal or other release practices are subject to a variety of regulatory requirements designed to minimize potential harm to human health and the environment.

[6] We also assume that $h(0) = \underline{h}$, where $\underline{h}$ is a basic, obligatory clean-up program that must be undertaken by firms, in case of toxic release, regardless of citizens demand for or willingness to pay for a better environment.

[7] Also, a higher-income area might be associated with a higher probability of collective actions by local residents to force the firm to implement a more thorough clean-up program in case of release.

[8] As $p''(a_l) > 0$, the first-order condition is sufficient for a minimum. Moreover, there will be an interior solution to this problem if $p'(0) < h(m_l) < p'(x)$.

[9] In this model, we assume that x (the total amount of hazardous waste generated by the production process) does

$$\frac{de_l^*}{dm_l} = -\frac{xh'(m_l)}{p''(x - e_l^*)} < 0 \qquad (1)$$

This result suggests that lower waste management/remediation levels are undertaken by potential polluters in lower income areas. Consistent with the environmental justice literature, lower income areas will be disproportionately subject to localized releases, although not necessarily more densely populated by TRI-type firms.

2.2 Number of firms

In a profit-maximization framework (Levinson, 1996; Wolverton, 2009), a firm considering the location of a new plant will choose the neighborhood with the attributes (m_l, Z_l) that lead to the highest expected profit. A firm's expected profit in a local area l is given by:

$$\widehat{\pi}_l = F(w_l, y_l, c_l^*, v_l)$$

where w_l is a vector of local input prices, y_l is a vector of local fixed factors (e.g. land, labor), c_l^* is the optimal pollution cost derived in the previous section, and v_l is a vector of other local factors that might affect the expected profits (e.g. minority ratio, residents' propensity to engage in collective actions...). Thus, the geographic distribution of firms/industries depends on the local characteristics that drove their initial location decisions (including existing agglomeration economies) and we can write the equilibrium number of firms in local area l as a function of local characteristics (ignoring the integer constraint on the number of firms for expositional purposes):

$$N^* = g(\widehat{\pi}) = N(m_l, Z_l) \qquad (2)$$

with $sign\left(\frac{dN(m_l,Z_l)}{dm_l}\right) = sign\left(\frac{d\widehat{\pi}_l}{dm_l}\right)$, i.e. the higher the expected profit in a local area, the larger the number of firms located in this area.

One reason why expected profit might be negatively affected by local income is as argued above. That is, a firm's pollution cost c_l^* is increasing, *ceteris paribus*, in local income. There are of course

not vary with income. This is consistent with previous studies (Harrington, 2012, 2013; Khanna et al 2009; Florida and Davison, 2001) which show that local community characteristics, such as local income, have limited impact on pollution prevention activities (P2), i.e. *ex ante* efforts to reduce pollution emissions.

a number of other channels through which local income may affect potential profits. For example, more cohesive and effective collective action in wealthier areas (Olson, 1965) may act as a deterrent to certain firms considering a given location. Some neighborhoods/communities may lobby the local government to adopt stricter environmental standards above any broader regulatory requirements (see Fishel, 2005). Their ability to do so (included in the vector v_l) will depend on their income and willingness to bear the costs of the effort to achieve a cleaner environment. Moreover, higher income areas will be characterized by higher rental costs (w_l), which may reduce the potential profit and the attractiveness of the locality.

Nevertheless, there are, as noted above, some offsetting benefits available in higher income areas that might attract and retain firms. Although wage rates may be higher, higher wage rates reflect higher marginal products of labor (quality of workforce) and, in the case of firms that require high skill labor inputs, reflect the availability of a workforce that matches their hiring needs (y_l). Moreover, higher income neighborhoods tend to have better physical and social infrastructure which may contribute to lower logistical costs and help to retain workers or to attract workers from other areas. Taken together, the elements outlined above suggest that the relationship between the number of firms and local income is difficult to predict on a purely theoretical level. That is, this relationship may be complex and non-monotonic.

2.3 Local pollution and income

We now investigate the relationship between total pollution in area l, $N(m_l, Z_l)e_l^*$, and local income m_l. It will depend on how the functions $N(m_l, Z_l)$ and e_l^* vary with income (i.e. expressions (1) and (2)):

$$\frac{d}{dm_l}\left(N(m_l, Z_l)e_l^*\right) = \frac{dN(m_l, Z_l)}{dm_l}e_l^* + N(m_l, Z_l)\frac{de_l^*}{dm_l}$$

First, note that, if the number of plants in a given location is decreasing in local income (i.e. firms always prefer to locate in low-income areas), total pollution will be decreasing, *ceteris paribus*, for all levels of local income, as predicted by the environmental justice literature.

However, some areas are too undeveloped to attract industrial economic activity. This implies that for relatively low levels of income, the benefits of locating in an area with better quality workforce and physical infrastructure can outweigh the higher costs of land or pollution such that the equilibrium number firms in a local area is increasing with the local income ($\frac{dN(m_l,Z_l)}{dm_l} > 0$). Then, as local incomes increase and pollution, land and labor costs increase *pari passu*, the relationship between the number of industrial firms and income is likely to flatten out or even turn down as the costs tend to increase relative to the benefits ($\frac{dN(m_l,Z_l)}{dm_l} \leq 0$).

If that is the case, we can then derive some sufficient conditions, such that the relationship between local pollution and local income is not monotonic, for the existence of an inverted U-shaped curve (i.e. EKC-type relationship), as shown in the following proposition.

Proposition. If (1) $N(m_l, Z_l)$ is continuous and twice differentiable in m_l, with $\frac{d^2N(m_l,Z_l)}{dm_l^2} < 0$ for all $m_l \geq 0$ and $\frac{dN(\widehat{m},Z_l)}{dm_l} = 0$ for some $\widehat{m} > 0$, (2) $N(0, Z_l) = 0$ and $\frac{dN(0,Z_l)}{dm_l} > 0$, and (3) the unit cost of release $h(m_l)$ is convex in income with $\lim_{m_l \longrightarrow 0} h'(m_l) = 0$,

Then the relationship between total pollution in area l, $N(m_l, Z_l)e_l^*$ and local income will exhibit an inverted U-shape, with a maximum level of total pollution reached at some income level m^*, with $0 < m^* < \widehat{m}$.

Proof. See Appendix.

In addition to the two conditions on the relationship between the number of firms and local income, we also need a third condition on the relationship between the unit cost of release and local income. Note that if condition (3) is satisfied, the optimal amount of releases at the firm-level, e_l^*, will be decreasing and concave in income (see equation (1)).[10] Intuitively, for low levels of income, a small increase in income does not have a big impact on the per unit cost of release and, as a consequence, on the optimal level of release. Therefore, the increase in the number of plants in a location $N(m_l, Z_l)$ more than offsets the reduction in individual firm-level releases. For larger levels of income $m_l > m^*$, as $N(m_l, Z_l)$ flattens out, the increase in the number of plants will be offset by the reduction in individual firm-level releases even as the effect of income on optimal releases increases. This leads to our first testable hypothesis.

Testable Hypothesis 1. When conditions 1-3 are satisfied, higher income neighborhoods will tend to

[10] $\frac{d^2e_l^*}{dm_l^2} = -\frac{xh''(m_l)}{p''(x-e_l^*)} < 0$ (as $p(a_l)$ is a quadratic function).

have more potentially polluting (TRI-type) firms, but those firms will invest more in waste management such that local pollution will exhibit an inverted U-shape with respect to income.

We will check the validity of these first two conditions in the empirical analysis by looking at the relationship between the total number of potentially polluting firms and median income, controlling for the relevant additional factors. We check for the third condition by looking at the relationship between income and property values in a given geography and our data show that the relationship is indeed convex (see Figure A1 in the Appendix). We also verify in our empirical analysis that individual firm-level releases are decreasing and concave in income by using the proportion of TRI-reporting firms that report a release in any given year relative to the set of potentially polluting firms, i.e., TRI-type firms, and estimate how local characteristics, including local income, affect this proportion.[11]

2.4 The Remediation Industry

Both the clean-up in case of release and the treatment/recycling used to avoid release of toxic chemicals are carried out by firms from the *remediation* industry. Thus, waste remediation firms supply pollution risk management services in response to the demand for those services posed by potentially polluting firms. The services supplied by these companies are local and often very specialized (depending on the type of polluting industry, pollutant, etc.) and typically require highly skilled/trained workers. Revenues of the remediation firms are for the most part reflected in costs of the local potentially polluting firms. Recall that the demand for remediation services purchased by polluting firm is partially driven by the income of local residents. As such, we have our second testable hypothesis:

Testable Hypothesis 2. Remediation firms will be more likely to locate in areas with larger numbers of polluting firms and in areas characterized by a higher median income.

3 Testing for a local EKC-type relationship

We now turn to the empirical analysis of the relationship between local income, industrial localization and local environmental quality. For our purposes, a local area is the census tract. Census tracts in

[11] Note that in our data, we only observe toxic releases by firms reporting to the TRI Program and not for all the potential polluters. We use the number of TRI-polluters, not the number of toxic releases. A TRI-polluter could have more than one release per calendar year.

populous areas are relatively small. Thus, it represents the locality closely adjacent to any potentially polluting firm located in the tract. It also closely represents the population that bears the immediate environmental impact in case of toxic release. As previously noted, the analysis is limited to the State of Texas. Aside from the benefit of a homogeneous regulatory environment, we also are able to take advantage of access to detailed establishment-level data from the Texas Quarterly Census of Employment and Wages, as described below. Recognizing that the data are at the establishment level, we use the more usual terminology of the firm to indicate a specific productive facility.

3.1 TRI data

The Toxic Release Inventory (TRI) is a mandatory reporting program managed by the U.S. Environmental Protection Agency for a set of industries that use or produce certain toxic or dangerous chemicals. The EPCRA (Emergency Planning and Community Right to Know Act) Section 313 requires TRI reports to be filed by owners and operators of facilities that meet all of the following criteria:

- The facility has 10 or more full-time employee equivalents;
- The facility is included in a given subset of the North American Industry Classification System (NAICS); and
- The facility manufactures (defined to include importing), processes, or otherwise uses any EPCRA Section 313 chemical in quantities greater than the established threshold in the course of a calendar year.[12]

As pointed out in Footnote 1, we define three sets of TRI-related firms. We refer to all firms located in a NAICS subject to TRI reporting, regardless of whether reporting is mandatory for the firm, as a TRI-type firm or potentially polluting firm but exclude firms in the NAICS 562 Waste Management and Remediation Services (this sub-sector group includes establishments engaged in the collection, treatment, and disposal of waste materials, see details below). Within the category of TRI-type firms are the firms for whom reporting is mandatory. Firms for whom reporting is mandatory

[12] See https://www.epa.gov/toxics-release-inventory-tri-program for details on on NAICS codes, listed chemicals, and chemical thresholds required for reporting.

are called TRI-reporting firms, a subset of TRI-type firms. TRI-reporting firms are not necessarily firms that experienced a toxic release. Firms that actually report toxic chemical releases are treated as a TRI-polluters for the year in which the release is reported. As such, a firm can be a polluter in year t, but not a polluter in year $t+i$. The EPA also provides toxicity weights for each toxic chemical listed in the TRI Program which allows us to compute a tract-level toxicity index by aggregating all firms' releases, measured in pounds of toxicity, within a tract. All other industries are treated as either non-TRI-type industries (i.e. they don't handle any toxic substance listed in the TRI Program) or remediation industries (NAICS 562).

We restrict the waste management/remediation sector to four industry sub-sectors in NAICS 562 Waste Management and Remediation Services. Specifically, we consider establishments in NAICS 562112 Hazardous Waste Collection, 562211 Hazardous Waste Treatment and Disposal, 562910 Remediation Services, and 562920 Materials Recovery Facilities (recycling). We shall refer to these four NAICS codes collectively as either waste management or remediation industries (for our purposes, these two terms refer to the same set of industries).

3.2 Firm-level data

All firm and industry-level data are derived from the Texas Quarterly Census of Employment and Wages (QCEW) for the years 2000–2006 as provided by the Texas Workforce Commission.[13] The QCEW reports data at the establishment level, including exact address, geographical coordinates, age, parent company, monthly employment, and quarterly payroll for all establishments in Texas subject to reporting under the Unemployment Insurance (UI) program. Different establishments within the same firm are identified by unique identification numbers and reported separately.

[13] The main data used in this study were collected and provided by the Texas Workforce Commission. These data are fully-disclosed (tax ID, locations, wages, and employment) and are not available to the general public. We were able to acquire them under the terms of non-disclosure agreement. We can only provide the data under the terms of this agreement either in terms of establishment-level aggregation at the NAICS-6 or some industry aggregation of NAICS-6 establishment-level data at the county level. We can report total county-level data at NAICS-6 if there are at least four establishments in the county with no establishment representing more than 60% of the the total county employment in the given NAICS-6 industry. Interested researchers can contact the Texas Workforce Commission for data requests.

3.3 Census-tract data

Median income and population statistics at the census tract level are taken from the U.S. Census Bureau. For Census-based data, we linearly interpolate co-variates from Census 2000 and Census 2010 to generate yearly observations at the tract level. The realizations of the variables we use from Census 2000 and Census 2010 are of course highly correlated with each other. For example, the correlation between median income in 2000 and 2010 is 0.95. A few Census 2000 tracts are divided in Census 2010. We aggregate variables (or construct population-weighted averages, where appropriate) to obtain corresponding Census 2000 tract information. We also consider some measures of local infrastructure (and, by that proxy, transportation costs) using the number of roads, number of rail roads, road construction expenditures.

Summary statistics are reported in Table 1.[14] We observe that there are only 0.084 incumbent waste remediation firms per tract. For a representative tract, there are 4.018 TRI-type firms. Average median household income for the tracts is about $43,930 and the average wage paid by establishments in each tract is about $38,250 in the sample period.[15] For a given tract, the average population is about 5,088 and the average unemployment rate is about 4.468 percent. The average house value is about $126,000.

3.4 Empirical analysis

To test the predictions of the theoretical analysis (testable hypotheses 1 and 2), we estimate an empirical model that takes the following form:

$$y_{lt} = X_{lt}\mathbf{\Delta} + \tau_t + \mu_{lt} \tag{3}$$

where y, depending on the specification, is total toxicity in pounds in each tract or the total number of firms in the TRI-type sector and in the remediation industry per tract at a given time. $X_l = (m, Z)'_l$ is the tract-l specific characteristics treated as median income m, average wage, college ratio, number of amenity-type establishments, infrastructure, population density, unemployment rate, land area,

[14] A description of these variables is provided in Table A. 1.

[15] The median income refers to the residents of a particular tract, while wage refers to the wage paid by establishments located in this tract to their workers, who are not necessarily living in the same tract.

housing rental ratio, and housing prices in the year, t. Of these variables, median income is of particular interest in this analysis and we employ a cubic specification for income.[16]

Estimation results are reported in Table 2. [17] The results for income are then graphed, *ceteris paribus*, in Figures 2. The familiar inverted U-shaped curve is present for the relationship between total toxicity and median income, peaking at median income of approximately $65,000, as in De Silva, *et al* (2016). These results confirm the presence of an EKC-type relationship between local income and local pollution. Of interest, the total numbers of firms in the TRI-type sectors in a tract increase and then level off at higher income levels, supporting our first testable hypothesis that the EKC is consistent with an increasing number of firms (conditions 1 and 2).[18] The number of firms in the waste remediation industry rise proportionately (slightly) slower than does the TRI-type sector over higher income levels, although they track closely for the most part, indicating that potentially-polluting firms in higher income tracts tend to utilize more remediation services. As one would expect from our second testable hypothesis, there is substantial correlation in Column 4 between the number of remediation firms and the number of TRI-type firms in each tract. These results indicate that, not surprisingly, remediation firms are more likely to locate in areas where a high number of potentially-polluting firms are present. Hence, in the next section, we investigate this result for remediation firms in more detail.

In Table 3, using the observed turning points in Figure 2, we re-estimate these relationships for the same set of correlates and the same three independent variables. However, in these estimations, we identify three intervals for median income in order to estimate linear splines, or piecewise linear relationships between the independent variables and median income. We find a positive relationship over the median income range $0-66,700 for all three specifications, and a weakly significant negative relationship only for the highest range of median income with respect to toxicity. This reinforces the key observation in Figure 2 that a local EKC is consistent with a number of firms in waste remediation

[16]In this setting, we do not include tract-fixed effects because some of our variables don't change over time (e.g. number of roads, number of railroads, etc.). In this context, our model captures a spatial view of the EKC–how firm siting is affected by difference in tract level income. Therefore, we do not consider long run technological changes within tracts.

[17]De Silva et al. (2016) show that high polluting firms locate in high minority areas. As rent and income are correlated with race, we also estimate our models including the minority ratios. Our main findings remain consistent and these results are reported in Tables A2-A3 in Appendix.

[18]We also estimate the effect of local income on the total number of employees in a tract and the results are qualitatively the same as with the number of firms. These results are available upon request.

and TRI-type sectors that first increase with median income and then level off as median income levels increase beyond the threshold value.

Finally, we check for the third condition (in testable hypothesis 1). To this end, we estimate the relative frequency of toxic releases (above the EPA threshold) reported to the TRI Program as a function of local income.[19] Given the nature of our dependent variable, we estimate this fractional model using method proposed by Papke and Wooldridge (1996). This is intended to capture the behavioral outcomes of TRI-type firms as a function of income. In the theoretical analysis, we show that when the number of firms in a tract is increasing with the local income, a condition to obtain an EKC-type relationship between income and pollution is that, for lower-income tracts, firm-level releases should not decrease too rapidly with income. Results in Table 4 confirm that local income has a stronger impact on the frequency of toxic releases in higher-income tracts, indicating that firms tend to spend more on waste management or remediation services only when median income is sufficiently high. This last result is consistent with the environmental justice literature; that is, firms' behavior toward the pollution risk, as well as localization, depends on local income.

3.5 Robustness Checks

3.5.1 Self-sorting by TRI-type firms

One concern might be that this result arises because of self-sorting by TRI-type firms. If those industries that have higher likelihood of release tend to cluster in lower income areas, and those industries that have a lower likelihood of release are more likely to be found in higher income neighborhoods, then this result will emerge. To this point, we might note that this is not inconsistent with our hypotheses and results so far. Rather it would tend to reinforce the environmental justice argument. Nevertheless, we separated three industries at NAICS-3 for individual analysis. Our selection was based on the specific industries shares of all establishments in the TRI-type industries. We chose industries with a high, medium and low proportion of the complete set of TRI-type firms. Specifically, we chose chemicals, fabricated metal, and miscellaneous n.o.s. As reported in Table 5, the relative frequency of firms reporting toxic releases in both the chemicals and fabricated metal sectors retain a similar

[19] As noted, the relative frequency of toxic releases is actually the ratio of TRI- reporting firms that report releases (TRI-polluters) to the total number of TRI-type firms.

appearance to the relative frequency that appears for the TRI-type industries. The relative frequency of miscellaneous is not significantly related to income, perhaps because this industry is too sparsely present in the data.

3.5.2 Causality

A potential issue may be that our demographic variables are contemporaneous and endogenous. It may be the case that a tract's demographic characteristics (including income) change over time precisely because of firms' location decisions: residents express their preferences for environmental attributes by moving across jurisdictions. We deal with this issue in different ways. First, we re-estimate these localization regressions with demographic characteristics lagged by a period. Our results are consistent and robust. However, we do not report these results in order to save space but we can provide them on request.

Additionally, we look at the correlations between ranking of tracts in 2000 and 2006. These results are reported in Table A.5. Our results indicate that ranking of tracts based on income, education, and population are highly correlated between 2000 and 2006 and the correlation coefficient is more than 96 percent for all variables. To the extent that industrial concentrations in 2000 probably represent the cumulative effects of perhaps several decades, we might assume that sorting effects at the beginning or our sample period represent a population location equilibrium, or *ex post* sorting equilibrium.

Another way to deal with this reverse causality issue is to look at firms' entry models. Insofar as entrant's locational calculus is concerned, demographic characteristics of the potential localities are largely given. With this in mind, we utilize an entry model for TRI-type firms by focusing only on new entrants in our sample period (2000-2006). Controlling for other relevant factors, we should observe entrants' preferences with respect to local incomes without the issue of costly relocation that affects incumbent firms. Due to the large number of new entrants (11,752), it is not possible to estimate this entry model using a conditional logit. Instead, we use the Poisson Pseudo Maximum Likelihood (PPML) method with time fixed effects.

Compared to the standard Poisson estimation, the PPML the estimation does not assume that the data are distributed with the mean equal to the variance of the event count. The data need not even come from a Poisson process and may be either under or over-dispersed. However, note that the estimated coefficients are nevertheless identical to the Poisson regression estimates. All that is required for PPML consistency is that the conditional mean function be correctly specified.[20] Here, the dependent variable is the number of TRI-type entrants (y) for a given tract (l) for a given year (t). Tract-level independent variables are as described before. The basic model is as follows:

$$E[y_{lt}|X_{lt}] = \exp(X_{lt}^{'}\psi + \tau_t) \tag{4}$$

These entry results are presented in the Table 6. We observe a similar correlation between TRI-type entrants and income (and other tract characteristics) as in our localization models. This exercise further supports our theoretical prediction that local income drives firms' location decision and pollution behavior.

3.5.3 Multiple establishments

Additionally, one might be concerned that firms with multiple establishments may locate their manufacturing plant in a low income neighborhood while their sales office may be located in an affluent area. This does not seem to be the case in our sample as the average firm has only 1.3 branches.[21]

3.5.4 Non-TRI-type firms

It is natural to wonder if this inverted U-shaped relationship is not also observed in the case of non-TRI-type firms. That is, is this result specific to TRI-type firms or does industrial localization more generally exhibit this same pattern? To answer this question, we conducted a similar regression to analyze the localization of non-TRI-type firms. As can be seen in Table A.4, there is a positive relationship between income and localization of firms, ceteris paribus, over the two lower income splines and then the relationship flattens out (column 1).

[20] For a more detailed discussion of this reasoning, see Gourieroux, Monfort, and Trognon (1984) and Santos Silva and Tenreyro (2006, 2011).

[21] This number is computed excluding industries as such as retail gasoline, commercial printing, and food processing which represent less than 5 percent of the sample. If we include these industries this number is about 2.6.

4 Entry and exit patterns in the remediation industry

In the previous section, we have shown that a local EKC-type relationship is consistent with an increasing number of TRI-type firms coupled with increasing expenditures of those firms in waste management activities. As mentioned above, this is our proxy for local demand for environmental quality. Waste remediation firms supply pollution risk management services in response to the demand for those services posed by potentially polluting firms. The services supplied by these companies are local and often very specialized (depending on the type of polluting industry, pollutant, etc.) and typically require highly skilled/trained workers. Given the importance of these activities to obtain the EKC-type relationship, we investigate further the structure of the industry supplying these remediation/waste management services.

4.1 Entry

In this section, we estimate the entry process of waste remediation establishments by census tract in Texas on an annual basis over the years 2000-2006. Establishment entry in any year is defined as the appearance (initial UI liability) of a new Enterprise Identification Number (EIN). We estimate the number of entrants in a particular location (tract) as a function of location characteristics. These characteristics include the number of remediation firms already present in that county (localization effects), the number of establishments in both TRI and non-TRI-related industries in that tract, median personal income, level of education attainment, amenities, infrastructure, population density, unemployment rate, and controls for housing ownership. We also include a dummy variable to control for tracts that are in counties bordering nearby states and Mexico. We present the distributions of incumbent and entrant firms for the waste remediation industry in Table 7.

The most common year of firm entry was 2002. This coincides with a TRI rule making which lowered reporting thresholds for lead and lead compounds.[22] De Silva *et al.*, (2016) show that, in 2002, the number of TRI reports involving lead or lead compounds increased six-fold compared to 2001. Also note that the aggregate announced toxic weight increased by a factor greater than four

[22] See Title 40, Part 372 of the Code of Federal Regulations which is summarized in volume 66, number 11 of the Federal Register.

in 2002 relative to 2001. We argue that, with this threshold change, demand for waste remediation increased in 2002 and, hence, we see an uptick in entrants in 2002. There is an average of about 0.013 remediation entrants per tract over the entire period of the sample.

We empirically model a firm's (i) location (l - tract) choice of entry at time t in order to maximize expected profits using a conditional Logit model (see McFadden, 1974). Each firm's after-entry profit from location l, π_{ilt}, can be written as follows:

$$\pi_{ilt} = A'_{lwt}\gamma + A'_{lpt}\lambda + X'_{lt}\beta + \epsilon_{ilt} \tag{5}$$

A_{lwt} is the number of incumbent waste remediation firms that are in a tract l (which is a proxy for the total investment in prevention e). A_{lpt} is the number of incumbent firms that are TRI-type in tract l, and X_l is the tract l specific characteristics, as mentioned above, in a given year, t. These two variables, A_{lwt} and A_{lpt}, together capture the agglomeration effects in tract l. We assume that the disturbance, ϵ_{ilt}, is independent and identically distributed.

In order to have a closed form expression for a firm's choice probabilities, we also assume that the ϵ_{ilt} are distributed with a Type 1 extreme value distribution. We further assume that each firm knows its private costs and expected profits. This asymmetric information assumption enables us to convert the discrete actions of competitors into continuous location choice probabilities. Then, we can specify the conditional logit model as follows:

$$\Pr(E_{ilt} = 1 | A_{lwt}, A_{lpt}, X_{lt}, \tau_t) = \Pr(\pi_{ilt} > \pi_{ikt} \text{ for all } l \neq k) \tag{6}$$

The dependent variable E_{ilt} equals 1 if a firm i chooses location l and 0 otherwise. A firm i will choose location l when $\pi_{il} > \pi_{ik}$ for all $k \neq l$. Therefore, conditional on the decision to open a new plant, the probability that firm i will choose a particular location l can be written as follows:

$$\Pr(E_{ilt} = 1) = \frac{\exp(A'_{lwt}\gamma + A'_{lpt}\lambda + X'_{lt}\beta)}{\sum_{k=1}^{m} \exp(A'_{kwt}\gamma + A'_{kpt}\lambda + X'_{kt}\beta)} \tag{7}$$

Results for the likelihood of entry are reported in Table 8. In all model specifications, the localization of remediation activity is important, all else equal. The estimated coefficient of the variable that uses the number of existing remediation firms as a measure of industrial concentration is positive and significant at the .01 level. In other words, the presence of incumbent remediation firms has a positive impact on the likelihood of additional entry of remediation firms. This is consistent with the presence of localization economies, or economies of scale from industrial concentration, that enhance the attractiveness of a given location for a start-up or relocating establishment. Not surprisingly, the presence of TRI-type firms matters with consistently and highly significant coefficient estimates across all models.

This indicates that industries with a history of polluting firms are an important factor in the presence of remediation firms. The estimated coefficient on median income is also significant and positive and there is some evidence that the second derivative on the median income variable may be negative or the relationship is concave. We also conclude that population matters, as does land area. The signs of the coefficients for both variables, *ceteris paribus*, imply that greater population density, as would be expected, results in a higher likelihood of entry of remediation establishments. Of further interest is the estimated coefficient on the ratio of college-educated residents. These results suggest that higher levels of education correlate to lower remediation firm entry probabilities.

As a robustness check, we estimate the entry process using a count data model with time fixed effects, specifically a Poisson model estimated by Pseudo Maximum Likelihood (PPML). Our dependent variable is the number of waste management entrants (y) for a given tract (l) for a given year (t). The basic model is as follows:

$$E[y_{lt}|A_{lwt}, A_{lpt}, X_{lt}] = \exp(A_{lwt}^{'}\nu + A_{lpt}^{'}\omega + X_{lt}^{'}\psi + \tau_{lt}) \qquad (8)$$

Estimation results for these PPML regressions are contained in Table 9. No qualitative differences are observed between these two models. As before, there is evidence of localization effects and demand-side factors associated with a larger TRI-type firm sector on the likelihood of remediation firm entry. In summary, the above findings support our earlier conjecture that remediation firms will locate closer to TRI-type firms.

4.2 Entry at random locations

As an additional robustness check, we look at entry by establishments in the remediation industries into random locations that are not dependent on legal jurisdictional boundaries. The locations are defined as non-overlapping rings of one-mile radius centered on establishments that are not in the defined set of remediation industries -that is, establishments in either a TRI-type or a non-TRI-type industry, excluding the remediation sector.

This brings an additional level of spatial acuity into the analysis. We center the rings on existing establishments because we want to limit the analysis to areas where there is commercial activity in order to ensure that the chosen areas are actually potential choices for locating a new establishment. Not doing so might result in choosing locations in which there is virtually no population, such as remote rural or agricultural land with no industrial or commercial infrastructure, or no commercial or industrial activity due to, say, zoning restrictions. Any non-remediation industry establishment that existed at any point during the time frame of the study is a potential center point, thus allowing for the possibility of new areas of commercial activity that came into existence during the course of the analysis and the possibility that a remediation firm is the first to enter the area.

By maximizing the number of potential rings while imposing the non-overlapping condition, we get 8,142 rings. Table 10 provides summary statistics for these non-overlapping rings. We see 231 out of 395 entrants enter into these random locations. Figure 3 shows these locations.

The specific industry containing the firm that is used to center the random location ring does not matter –it only serves to locate a ring in an area that allows commercial activity. It is, rather, the industrial content captured in the ring that matters. In this analysis, the variables of interest are remediation industry establishments and TRI-type firms contained in the random rings. We are unable to measure non-establishment variables, such as household income or other population characteristics, at the spatial division of the rings. Hence, the other variables in the model are still measured at the census tract level and reflect the census tract in which the ring center is located In general, this represents measures for the census tract variable values that reflect the tract in which the majority of the area of the ring is located.

The results of the PPML estimation in Table 11 are quite similar to those based on establishments at the census tract. That is, localization is, as before, an important determinant of the likelihood of remediation establishment entry. Further, the presence of TRI-type establishments in the rings is an important factor in location decisions of remediation industry establishments. Of interest, higher wages in the surrounding census tract are significantly associated with higher entry probabilities which can reflect the industry's demand for higher skilled labor.

4.3 Exit

Thinking about establishment exits, we have in mind a theoretical model involving a threshold rule that is analogous to the profit maximization problem considered in our entry analysis. In this case, if firms do not make a sufficient level of profit, they choose to exit the industry. In order to consider the question of remediation industry establishment exit, we estimate logit models in which the dependent variable is coded one if the firm exits during a given period, or as zero if the firm continues operation through the entire period.

Exit is identified as having taken place if the firm (EIN) disappears from the data set at the outset of a given calendar year and is treated as having occurred during the last year in which the firm is present in the data (year previous to disappearance). The time to failure is relative to the year in which the firm entered the market. That is, we only consider firms that enter during the time frame of the analysis, i.e., entry in years 2000-2006, and consider the exit event relative to the number of years since entry. The observed number of years to failure, therefore, ranges from a low of one (failure in year of entry) to a high of seven (no failure observed) across the establishments in the sample. Table 12 illustrates exit patterns. To interpret Table 12, note that, in year 2002, there were 98 entrants. By 2006, or over the course of the following five years, 69 of them exited the market.

Overall, there were 395 entrants of which 214 of them had exited by 2006. Table 13 provides establishment-level summary statistics. About 36% of the establishments have past experience in the market. On average, these establishments have an additional branch or are part of a multi-establishment firm. Additionally, these establishments pay about $49,000 per year in wages and employ about 19 workers. This indicates that this labor force is highly skilled and limited to a given area. Hence, we conjecture that firms will compete for the same resources and this will affect the survival rate. In this case, we expect agglomeration to increase the likelihood of exit.

Estimation results are reported in Table 14. In the case of localization effects, the presence of other remediation firms increases the likelihood of failure for these new establishments. On the other hand, the numbers of local area TRI-type firms appear to have no influence on exit probabilities. Not surprisingly, like firms in most industries, the age of the firm and firm size (employment) are negatively correlated with probability of failure in any given period. As with the other estimations above, local income appears to have no statistically significant effect on exit.

5 Conclusion

We have employed detailed data for small geographies to analyze the posited theoretical relationship between the localization of potentially polluting firms, toxic releases, and local-area incomes. Our model suggests that profit-maximizing, potentially polluting firms behave rationally toward the financial risk inherent in a toxic release. Our conjecture is that, as risk exposure increases with incomes within spatial proximity to those firms, the firms will take measures to manage that risk. We utilized the localization of the waste management industry as a means of observing evidence of demand for risk-reducing options. In this context, firms can vary their utilization of waste management services as an additional means of managing pollution risk across the spectrum of locations in lower income, lower risk exposure areas to location in higher income, higher risk exposure areas.

We find our results to be persuasive. We find evidence consistent with our conjecture, both in terms of potentially polluting firms' localization, the localization of waste remediation firms, and in terms of the relative frequency of these firms' reported releases. Since the analysis was made within a single state, we have largely controlled for heterogeneity in regulatory structures that, under the traditional pollution haven argument, would lead to a similar result if polluting firms were to exploit opportunities to locate in lax regulatory environments, hungry for economic development. While both explanations will lead to a similar outcome, ours is driven by non-policy related incentives, although it does depend on enforcement of a common set of laws. That is to say, the traditional economic paradigm of profit-maximizing behavior in the presence of risk can explain the phenomenon of the

EKC independent of any differences across regulatory regimes.

Our results bear implications for policies that aim to enhance environmental justice, but also speaks to policies that can exploit the incentives inherent in the standard paradigm of profit-maximization. One of the objectives of the TRI Program was to create a strong incentive for companies to take measures to reduce their toxic release and be good neighbors in their communities. During our sample period 2000–2006, overall toxic releases in the US decreased by about 34% with a further decline of about 21% since 2006. However, our paper shows that this decline might not be uniformly distributed because firms respond to local demographic characteristics (including local income). In trying to implement the lowest-cost response to the publication of the TRI data, firms will tend to reduce releases through waste management when local opposition is the highest.

Combined with our finding that economic activity and local income are positively correlated, this implies that the population group most affected by toxic releases will be a working-class population located in an industrial area. Our analysis suggests that without further actions, the disparities in exposure to toxic release faced by certain population groups will not be reduced by simply requiring that firms report their releases. If the attainment of greater environmental justice across population groups is a policy goal, serious thoughts should then be given to the regulations on compensation schemes, designed to offset the costs of potential environmental risk.

References

[1] Andreoni, J., and Levinson, A. (2001). The simple analytics of the environmental Kuznets curve, *Journal of Public Economics*, 80 , 269-286.

[2] Arora, S., and Cason, T. N. (1999). Do community characteristics influence environmental outcomes? Evidence from the Toxic Release Inventory, *Southern Economic Journal*, 65(4), 691-716.

[3] Brooks, N., and Sethi, R.(1997). The distribution of pollution: community characteristics and exposure to air toxics, *Journal of Environmental Economics and Management*, 32, 233-250.

[4] Chay, K., and Greenstone, M. (2003). The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession, *The Quarterly Journal of Economics*, 118(3), 1121-1167.

[5] Coase, R. H. (1960), The Problem of Social Cost, *The Journal of Law & Economics*, 3, 1-44.

[6] Combes, P. (2000). Economic structure and local growth: France 1984-1993, *Journal of Urban Economics*, 47 , 329-355.

[7] Currie, J., and Schmieder, J. F. (2009). Fetal Exposures to Toxic Releases and Infant Health, *American Economic Review*, 99 (2): 177-83.

[8] de Bruyn, S. M., and Heintz, R. J. (1999). The Environmental Kuznets Curve Hypothesis. Chapters,in: Handbook of Environmental and Resource Economics, chapter 46. Edward Elgar Publishing.

[9] De Silva, D. G., Hubbard T., and Schiller, A. R. (2016). Entry and Exit Patterns of 'Toxic' Firms, *The American Journal of Agricultural Economics*, 98 (3), 881-909.

[10] Egli, H., and Steger, T. M. (2007). A Dynamic Model of the Environmental Kuznets Curve: Turning Point and Public Policy, *Environmental & Resource Economics*, 36, 15-34.

[11] Fischel, W.A. (2005). The Home voter Hypothesis: How Home Values Influence Local Government Taxation, School Finance, and Land-Use Policies. Harvard University Press.

[12] Florida, R. and Davison, D. (2001). Why do firms adopt (advanced) environmental management practices (And do they make a difference?). In Regulating from the Inside, ed. C. Coglianese and J. Nash, RFF Press: Washington D.C.

[13] Glaeser, E. L., Kallal, H. L., Scheinkman, J. A., Schliefer, A. (1992). Growth in cities, *Journal of Political Economy*, 100 , 1126-1152.

[14] Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson Models, *Econometrica*, 52, 701-720.

[15] Graff Zivin, J., and Neidell, M. (2012). The Impact of Pollution on Worker Productivity. *American Economic Review*, 102 (7), 3652-73.

[16] Grossman, G. M., and Krueger, A. B. (1993). Environmental impacts of a North American free-trade agreement, in *The U.S.-Mexico Free Trade Agreement*, P. Garber, (ed.), MIT Press, Cambridge, MA.

[17] Grossman, G. M., and Krueger, A. B. (1995). Economic growth and the environment, *Quarterly Journal of Economics*, 110, 353-378.

[18] Hamilton, J. H. (1995). Testing for environmental racism: Prejudice, profits, political power? *Journal of Policy Analysis and Management*, 14, 107-132.

[19] Harbaugh, W. T., Levinson, A., and Wilson, D. M. (2002). Reexamining the Empirical Evidence for an Environmental Kuznets Curve, *Review of Economics and Statistics*, 84 (3), 541-551.

[20] Harrington, D.R. (2012). Two-stage Adoption of Different Types of Pollution Prevention Technologies, *Resource and Energy Economics,* 34(3), 349-373.

[21] Harrington, D.R. (2013). Effectiveness of State Pollution Prevention (P2) Programs and Policies, *Contemporary Economic Policy*, 31(2), 255-278.

[22] Henderson, V., Kuncoro, A., and Turner, M. (1995). Industrial development in cities, *Journal of Political Economy*, 103, 1067-1090.

[23] Jones, L. E., and Manuelli, R. E. (2001). Endogenous Policy Choice: The Case of Pollution and Growth, *Review of Economic Dynamics*, 4, 369-405.

[24] Khanna, M., Deltas, G., and Harrington, D.R. (2009). Adoption of Pollution Prevention Techniques: The Role of Management Systems, Demand-Side Factors and Complementary Assets. *Environmental and Resource Economics*, 44(1), 85-106.

[25] Levinson, A. (1996). Environmental regulations and manufacturers' location choices: Evidence from the Census of Manufactures, *Journal of Public Economics*, 62 (1–2), 5-29.

[26] Lieb, C. M. (2002). The environmental Kuznets curve and satiation: a simple static model, *Environment and Development Economics,* 7, 429-448.

[27] Marsiglio, S., Ansuategi, A., and Gallastegui, M. C. (2016). The Environmental Kuznets Curve and the Structural Change Hypothesis, *Environmental & Resource Economics*, 63, 265-288.

[28] McFadden, D. L., (1974). Conditional logit analysis of qualitative choice behavior, *Frontiers in Economics*, P. Zarembka (ed.), Academic Press: New York, 105-142.

[29] Olson, M. (1965). The logic of collective action. Cambridge, MA: Harvard.

[30] Papke, L. E., and Wooldridge, J. M. (1996). Econometric Methods for fractional response variables with an application to 401(K) plan participation rates, *Journal of Applied Econometrics*, 11, 619-632.

[31] Rosenthal, S. S., and Strange, W. C. (2003). Geography, industrial organization, and agglomeration, *Review of Economics and Statistics*, 85, 377-393.

[32] Rothman, D. S. (1998). Environmental Kuznets curves - real progress or passing the buck? A case for consumption-based approaches, *Ecological Economics*, 25, 177-194.

[33] Santos Silva, J.M., and Tenreyro, S. (2006). The Log of Gravity, *Review of Economics and Statistics,* 88 (4), 641–58.

[34] ——, (2011). Further Simulation Evidence on the Performance of the Poisson Pseudo-Maximum Likelihood Estimator, *Economics Letters* 112 (2), 220–2.

[35] Selden, T. M. and Song D. (1995). Neoclassical growth, the J-curve for abatement, and inverted U curve for pollution, *Journal of Environmental Economics and Management*, 29, 162-168.

[36] Shadbegian, R., and Wolverton, A. (2010). Location Decisions of U.S. Polluting Plants: Theory, Empirical Evidence, and Consequences, *International Review of Environmental and Resource Economics*, 4(1), 1-49.

[37] Stern, D. I. (2004). The rise and fall of the environmental Kuznets curve, *World Development*, 32(8), 1419-1439.

[38] Stokey, N. L. (1998). Are there limits to growth? *International Economic Review*, 39, 1–31.

[39] Wolverton, A. (2009). Effects of Socio-Economic and Input-Related Factors on Polluting Plants' Location Decisions. *The B.E. Journal of Economic Analysis & Policy,* 9, 1-32.

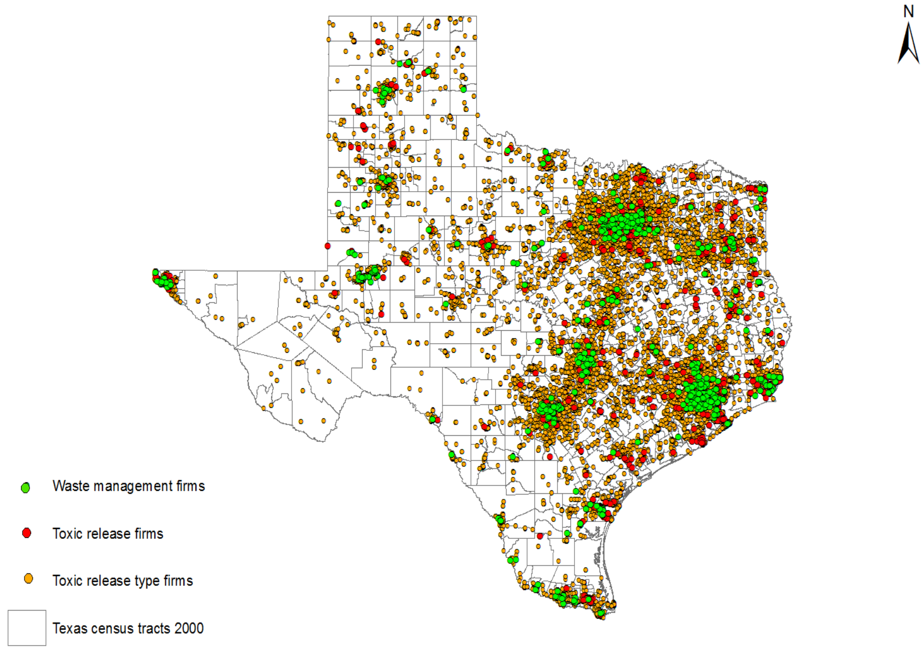


Figure 1: Locations of waste remediation and TRI firms

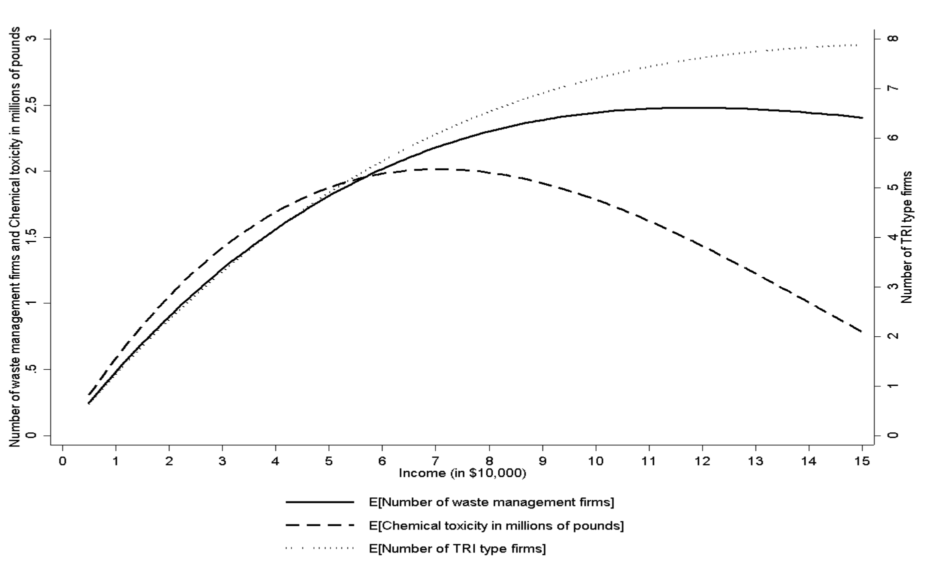


Figure 2: Estimated cubic functions relating number of firms to median income

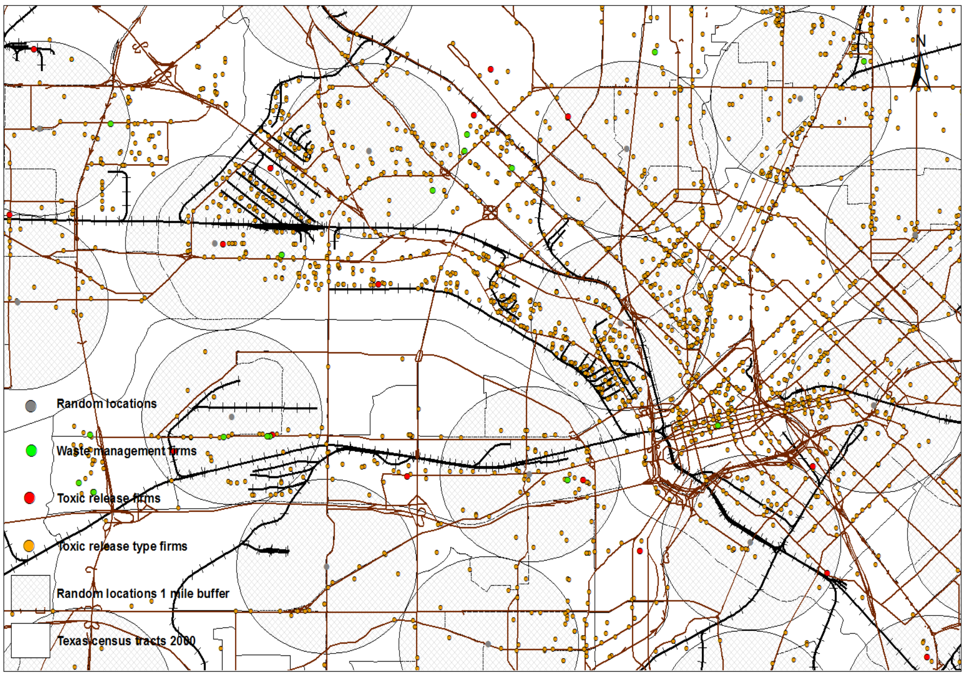


Figure 3: Non-overlapping one mile rings – Dallas area

Table 1: Summary statistics by tract

Variable	Mean (Standard deviation)
Average toxicity in pounds$_{lt}$ (in millions)	0.010
	(0.180)
Average number of environmental remediation firms$_{lt}$	0.084
	(0.342)
Average number of TRI type firms$_{lt}$	4.018
	(9.994)
Median income (in $10,000)$_{lt}$	4.393
	(2.278)
Average wage (in $10,000)$_{lt}$	3.825
	(4.470)
College ratio$_{lt}$	0.094
	(0.078)
Number of amenity establishments$_{lt}$	5.478
	(12.878)
Number of roads$_{lt}$	13.150
	(12.022)
Number of rail roads$_{lt}$	2.153
	(4.154)
Unemployment rate$_{lt}$	4.468
	(3.204)
Population (in 1,000)$_{lt}$	5.088
	(2.884)
Land area (in 100 in square miles)$_{lt}$	0.622
	(2.200)
Housing rental ratio$_{lt}$	0.315
	(0.201)
TxDOT expenditures (in $1,000,000)$_{lt}$	9.068
	(23.076)
Average house value (in $10,000)$_{lt}$	11.423
	(9.619)

Table 2: Explaining Variation in the Number of Toxic Pounds, Remediation Firms, and TRI Type Firms

Variable	Toxicity	Number of firms in		
	in pounds$_{lt}$	TRI type$_{lt}$	Remediation$_{lt}$	
	(1)	(2)	(3)	(4)
Median income (in \$10,000)$_{lt}$	0.642***	1.091***	0.515***	0.464***
	(0.114)	(0.148)	(0.086)	(0.084)
Median income (in \$10,000)$^2_{lt}$	-0.060***	-0.071***	-0.034***	-0.031***
	(0.015)	(0.020)	(0.010)	(0.010)
Median income (in \$10,000)$^3_{lt}$	0.001***	0.001*	0.001*	0.001*
	(0.000)	(0.001)	(0.000)	(0.000)
Number of TRI type incumbent establishments$_{lt}$				0.027***
				(0.005)
Number of TRI type incumbent employees$_{lt}$				
Average wage (in \$10,000)$_{lt}$	0.033***	0.079***	0.024***	0.021***
	(0.006)	(0.012)	(0.004)	(0.004)
College ratio$_{lt}$	-5.773***	-10.447***	-3.989***	-3.437***
	(0.951)	(0.887)	(0.641)	(0.627)
Number of amenity establishments$_{lt}$	-0.008**	0.110***	0.011***	0.007***
	(0.004)	(0.013)	(0.001)	(0.001)
Number of roads$_{lt}$	0.014***	0.167***	0.024***	0.015***
	(0.002)	(0.025)	(0.002)	(0.003)
Number of rail roads$_{lt}$	0.071***	0.402***	0.058***	0.043***
	(0.009)	(0.038)	(0.006)	(0.006)
Unemployment rate$_{lt}$	0.016***	0.010	0.021***	0.023***
	(0.005)	(0.016)	(0.007)	(0.007)
Population (in 1,000)$_{lt}$	0.004	0.113***	0.034***	0.032***
	(0.008)	(0.019)	(0.009)	(0.009)
Land area (in 100 in square miles)$_{lt}$	-0.045***	-0.335***	-0.110***	-0.082***
	(0.013)	(0.055)	(0.025)	(0.021)
Housing rental ratio$_{lt}$	0.387***	2.803***	1.216***	1.045***
	(0.144)	(0.352)	(0.181)	(0.175)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Number of observations	30,114	30,114	30,114	30,114
Log likelihood	-3,302.000	-109,680.000	-9,644.000	-9,517.000
Uncensored observations	833	29,968	2,104	2,104

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Explaining Variation in the Number of Toxic Pounds, Remediation Firms, and TRI Type Firms – alternate specification

Variable	Toxicity	Number of firms in		
	in $pounds_{lt}$	TRI $type_{lt}$	$Remediation_{lt}$	
	(1)	(2)	(3)	(4)
Median income \$0 – $\$66,700_{lt}$	0.206***	0.566***	0.308***	0.278***
	(0.031)	(0.046)	(0.033)	(0.032)
Median income >\$66,700 – $\$100,000_{lt}$	0.013	0.107	-0.040	-0.054
	(0.068)	(0.118)	(0.058)	(0.056)
Median income $>\$100,000_{lt}$	-0.679*	-0.093	0.025	0.035
	(0.385)	(0.061)	(0.052)	(0.048)
Number of TRI type incumbent $establishments_{lt}$				0.027***
				(0.005)
Number of TRI type incumbent $employees_{lt}$				
Average wage (in $\$10,000)_{lt}$	0.033***	0.079***	0.024***	0.021***
	(0.006)	(0.012)	(0.004)	(0.004)
College $ratio_{lt}$	-6.062***	-10.477***	-4.118***	-3.566***
	(0.986)	(0.920)	(0.641)	(0.626)
Number of amenity $establishments_{lt}$	-0.008**	0.110***	0.011***	0.007***
	(0.004)	(0.013)	(0.001)	(0.001)
Number of $roads_{lt}$	0.015***	0.168***	0.024***	0.015***
	(0.002)	(0.025)	(0.002)	(0.003)
Number of rail $roads_{lt}$	0.070***	0.400***	0.057***	0.043***
	(0.009)	(0.038)	(0.006)	(0.006)
Unemployment $rate_{lt}$	0.014***	0.004	0.020***	0.022***
	(0.005)	(0.016)	(0.007)	(0.007)
Population (in $1,000)_{lt}$	0.001	0.114***	0.033***	0.031***
	(0.008)	(0.019)	(0.009)	(0.009)
Land area (in 100 in square $miles)_{lt}$	-0.040***	-0.332***	-0.108***	-0.080***
	(0.013)	(0.055)	(0.024)	(0.020)
Housing rental $ratio_{lt}$	0.358**	2.753***	1.229***	1.061***
	(0.145)	(0.361)	(0.182)	(0.175)
Border county $effects_{lt}$	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Number of observations	30,114	30,114	30,114	30,114
Log likelihood	-3,312.000	-109,684.000	-9,636.000	-9,508.000
Uncensored observations	833	29,968	2,104	2,104

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Relative frequency of toxic release firms reported

Variable	(TRI polluters / TRI–type firms)$_{lt}$		
	(1)	(2)	(3)
Median income \$0 – \$66,700$_{lt}$	-0.036***	-0.033***	-0.031**
	(0.011)	(0.012)	(0.014)
Median income >\$66,700 – \$100,000$_{lt}$	-0.038	-0.034	-0.044
	(0.040)	(0.040)	(0.042)
Median income >\$100,000$_{lt}$	-0.259*	-0.260*	-0.289*
	(0.138)	(0.135)	(0.154)
Average wage (in \$10,000)$_{lt}$			0.014***
			(0.002)
Number of amenity establishments$_{lt}$			-0.020***
			(0.003)
Number of roads$_{lt}$			0.003**
			(0.001)
Number of rail roads$_{lt}$			0.026***
			(0.002)
Unemployment rate$_{lt}$			0.003
			(0.005)
Population (in 1,000)$_{lt}$		-0.013*	-0.002
		(0.007)	(0.006)
Land area (in 100 in square miles)$_{lt}$		0.003	-0.016*
		(0.005)	(0.009)
Housing rental ratio$_{lt}$			-0.029
			(0.110)
Year effects	Yes	Yes	Yes
Border county effects$_{lt}$	Yes	Yes	Yes
Number of observations	30,114	30,114	30,114
Log likelihood	-1,215.000	-1,214.000	-1,178.000

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level.

Table 5: Relative frequency of toxic release firms reported by industry

Variable	(Chemical polluters / total chemical firms)$_{lt}$	(Fabricated metal polluters / total fabricated metal)$_{lt}$	(Miscellaneous manuf. polluters / total miscellaneous manuf.)$_{lt}$
	(1)	(2)	(3)
Median income \$0 – \$66,700$_{lt}$	-0.005	0.013	0.012
	(0.012)	(0.013)	(0.029)
Median income >\$66,700 – \$100,000$_{lt}$	-0.026	-0.213***	-0.034
	(0.038)	(0.046)	(0.053)
Median income >\$100,000$_{lt}$	-0.541***	-0.132**	0.028
	(0.190)	(0.059)	(0.057)
Average wage (in \$10,000)$_{lt}$	0.018***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)
Number of amenity establishments$_{lt}$	0.001	0.001***	0.002***
	(0.001)	(0.000)	(0.001)
Number of roads$_{lt}$	0.006***	0.009***	0.010***
	(0.001)	(0.001)	(0.002)
Number of rail roads$_{lt}$	0.048***	0.024***	0.011***
	(0.003)	(0.003)	(0.004)
Unemployment rate$_{lt}$	0.023***	0.013***	-0.011
	(0.003)	(0.004)	(0.020)
Population (in 1,000)$_{lt}$	0.020***	0.035***	0.040***
	(0.005)	(0.005)	(0.007)
Land area (in 100 in square miles)$_{lt}$	-0.019*	-0.068***	-0.038
	(0.010)	(0.021)	(0.024)
Housing rental ratio$_{lt}$	0.058	-0.361***	0.591***
	(0.087)	(0.097)	(0.143)
Year effects	Yes	Yes	Yes
Border county effects$_{lt}$	Yes	Yes	Yes
Number of observations	30,114	30,114	30,114
Log likelihood	-3,334.000	-2,649.000	-992.700

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level.

Table 6: Models of aggregate entry counts for TRI type firms

Variable	Number of entrants						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of TRI type incumbents$_{ilt}$	0.013***	0.012***	0.012***	0.012***	0.012***	0.013***	0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Median income \$0 – \$66,700$_{lt}$		0.067***		0.138***	0.216***	0.189***	
		(0.013)		(0.017)	(0.020)	(0.022)	
Median income >\$66,700 – \$100,000$_{lt}$		-0.145***		-0.089**	-0.033	-0.002	
		(0.038)		(0.041)	(0.041)	(0.042)	
Median income >\$100,000$_{lt}$		-0.044		-0.040	-0.036	-0.027	
		(0.031)		(0.030)	(0.029)	(0.030)	
Average wage (in \$10,000)$_{lt}$		0.018***	0.019***	0.019***	0.019***	0.019***	0.019***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
College ratio$_{lt}$			-0.973***	-2.272***	-3.417***	-4.058***	-0.898***
			(0.220)	(0.363)	(0.395)	(0.410)	(0.277)
Number of amenity establishments$_{lt}$				0.005***	0.004***	0.005***	0.004***
				(0.000)	(0.000)	(0.001)	(0.000)
Number of roads$_{lt}$		0.002	0.002	0.001	0.001		-0.000
		(0.002)	(0.002)	(0.002)	(0.002)		(0.002)
Number of rail roads$_{lt}$		0.028***	0.026***	0.026***	0.026***		0.026***
		(0.003)	(0.003)	(0.003)	(0.003)		(0.003)
Unemployment rate$_{lt}$					0.004	-0.002	0.007
					(0.004)	(0.005)	(0.005)
Population (in 1,000)$_{lt}$		0.052***	0.056***	0.043***	0.041***	0.053***	0.056***
		(0.006)	(0.005)	(0.006)	(0.006)	(0.007)	(0.005)
Land area (in 100 in square miles)$_{lt}$					0.021***	0.032***	0.015**
					(0.006)	(0.005)	(0.006)
Housing rental ratio$_{lt}$					0.733***	0.818***	0.116
					(0.120)	(0.121)	(0.099)
TxDOT expenditures (in \$1,000,000)$_{lt}$						0.001*	
						(0.001)	
Average house value$_{lt}$							-0.000
							(0.002)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs.	30,114	30,114	30,114	30,114	30,114	30,114	30,114
Log likelihood	-23,848.000	-23,027.000	-23,061.000	-22,885.000	-22,816.000	-23,072.000	-22,974.000

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level. The dependent variable is the number of environmental remediation firms in a tract.

Table 7: Entry patterns

Entry year	Remediation firm	
	Entrants	Incumbents
2000	43	203
2001	47	241
2002	98	277
2003	54	362
2004	58	367
2005	50	368
2006	45	375

Table 8: Conditional logit results for firm entry at tract level

Variable	Firm entry						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of environmental remediation incumbents$_{lt}$	1.048***	1.008***	1.027***	0.980***	0.960***	0.990***	1.019***
	(0.059)	(0.063)	(0.063)	(0.064)	(0.066)	(0.062)	(0.064)
Number of TRI type incumbents$_{ilt}$	0.008***	0.006***	0.005***	0.005***	0.004**	0.008***	0.004**
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Median income \$0 – \$66,700$_{lt}$		0.160***		0.265***	0.391***	0.376***	
		(0.042)		(0.052)	(0.061)	(0.062)	
Median income >\$66,700 – \$100,000$_{lt}$		-0.084		0.001	0.088	0.077	
		(0.091)		(0.095)	(0.098)	(0.098)	
Median income >\$100,000$_{lt}$		0.001		0.005	-0.001	0.015	
		(0.089)		(0.088)	(0.086)	(0.087)	
Average wage (in \$10,000)$_{lt}$		0.013*	0.016**	0.015**	0.013*	0.014**	0.014**
		(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
College ratio$_{lt}$			0.080	-3.326***	-5.050***	-5.307***	-0.444
			(0.706)	(1.082)	(1.215)	(1.221)	(0.880)
Number of amenity establishments$_{lt}$				0.004***	0.003***	0.003***	0.003***
				(0.001)	(0.001)	(0.001)	(0.001)
Number of roads$_{lt}$		0.004	0.002	0.003	0.006		0.004
		(0.004)	(0.004)	(0.004)	(0.004)		(0.004)
Number of rail roads$_{lt}$		0.019**	0.018**	0.017*	0.019**		0.020**
		(0.009)	(0.009)	(0.009)	(0.009)		(0.009)
Unemployment rate$_{lt}$					0.013	0.013	0.021*
					(0.011)	(0.011)	(0.013)
Population (in 1,000)$_{lt}$		0.052***	0.068***	0.041***	0.038**	0.044***	0.067***
		(0.015)	(0.014)	(0.016)	(0.016)	(0.015)	(0.014)
Land area (in 100 in square miles)$_{lt}$					-0.052	-0.023	-0.076*
					(0.037)	(0.032)	(0.042)
Housing rental ratio$_{lt}$					1.211***	1.135***	-0.028
					(0.335)	(0.341)	(0.274)
TxDOT expenditures (in \$1,000,000)$_{lt}$						-0.004	
						(0.004)	
Average house value$_{lt}$							0.010**
							(0.005)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of entrants	395	395	395	395	395	395	395
Number of tratcs	4,302	4,302	4,302	4,302	4,302	4,302	4,302
Log likelihood	-3,180.000	-3,157.000	-3,165.000	-3,149.000	-3,137.000	-3,141.000	-3,155.000
χ^2	249.100	293.100	280.600	312.500	336.400	327.400	298.900

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level. The dependent variable takes the value of 1 for entry and 0 otherwise.

Table 9: Models of aggregate entry counts

Variable	Number of entrants						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of environmental remediation	1.048***	1.008***	1.027***	0.980***	0.960***	0.990***	1.019***
incumbents$_{lt}$	(0.059)	(0.063)	(0.062)	(0.065)	(0.067)	(0.063)	(0.065)
Number of TRI type incumbents$_{ilt}$	0.008***	0.006***	0.005***	0.005***	0.004**	0.008***	0.004**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Median income \$0 – \$66,700$_{lt}$		0.160***		0.265***	0.391***	0.376***	
		(0.045)		(0.059)	(0.062)	(0.063)	
Median income >\$66,700 – \$100,000$_{lt}$		-0.084		0.001	0.088	0.077	
		(0.096)		(0.106)	(0.107)	(0.107)	
Median income >\$100,000$_{lt}$		0.001		0.005	-0.001	0.015	
		(0.087)		(0.085)	(0.083)	(0.083)	
Average wage (in \$10,000)$_{lt}$		0.013***	0.016***	0.015***	0.013***	0.014***	0.014***
		(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)
College ratio$_{lt}$			0.080	-3.326**	-5.050***	-5.307***	-0.454
			(0.761)	(1.378)	(1.392)	(1.408)	(0.942)
Number of amenity establishments$_{lt}$				0.004***	0.003***	0.003***	0.003***
				(0.001)	(0.001)	(0.001)	(0.001)
Number of roads$_{lt}$		0.004	0.002	0.003	0.006		0.004
		(0.004)	(0.004)	(0.004)	(0.004)		(0.004)
Number of rail roads$_{lt}$		0.019*	0.018*	0.017*	0.019*		0.020*
		(0.010)	(0.010)	(0.010)	(0.011)		(0.010)
Unemployment rate$_{lt}$					0.013	0.013	0.021
					(0.013)	(0.014)	(0.018)
Population (in 1,000)$_{lt}$		0.052***	0.068***	0.041***	0.038**	0.044***	0.067***
		(0.015)	(0.013)	(0.016)	(0.015)	(0.014)	(0.014)
Land area (in 100 in square miles)$_{lt}$					-0.052	-0.023	-0.076
					(0.039)	(0.030)	(0.049)
Housing rental ratio$_{lt}$					1.211***	1.135***	-0.027
					(0.367)	(0.373)	(0.327)
TxDOT expenditures (in \$1,000,000)$_{lt}$						-0.004	
						(0.004)	
Average house value$_{lt}$							0.010**
							(0.004)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs.	30,114	30,114	30,114	30,114	30,114	30,114	30,114
Log likelihood	-1,987.000	-1,964.000	-1,972.000	-1,956.000	-1,944.000	-1,948.000	-1,962.000

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level. The dependent variable is the number of environmental remediation firms in a tract.

Table 10: Summary statistics for randomly chosen non-overlapping locations

Variable	Non-overlapping locations
Number of non-overlapping locations	8,142
Number of environmental remediation entrants	231
Average number of environmental remediation entrants$_{lt}$	0.004
	(0.072)
Average number of environmental remediation incumbents$_{lt}$	0.024
	(0.173)
Average number of TRI type firms$_{lt}$	1.212
	(5.708)

Standard deviation are in parentheses.

Table 11: Aggregate entry counts at a random location

Variable	Number of entrants		
	(1)	(2)	(3)
Number of environmental remediation	1.142***	1.083***	1.138***
incumbents within 0-1 mile$_{lt}$	(0.117)	(0.116)	(0.117)
Number of TRI incumbents within 0-1 mile$_{lt}$	0.012***	0.011***	0.012***
	(0.002)	(0.002)	(0.002)
Median income \$0 – \$66,700$_{lt}$	0.407***	0.309***	
	(0.084)	(0.087)	
Median income >\$66,700 – \$100,000$_{lt}$	0.254**	0.193	
	(0.128)	(0.127)	
Median income >\$100,000$_{lt}$	-0.075	-0.079	
	(0.244)	(0.242)	
Average wage(in \$10,000)$_{lt}$	0.023***	0.022***	0.022***
	(0.008)	(0.008)	(0.007)
College ratio$_{lt}$	-1.753	-1.038	2.544*
	(1.595)	(1.617)	(1.304)
Number of amenity establishments$_{lt}$	0.004***	0.003**	0.004***
	(0.001)	(0.001)	(0.001)
Number of roads$_{lt}$	-0.009		-0.009
	(0.006)		(0.006)
Number of rail roads$_{lt}$	0.013		0.015
	(0.012)		(0.011)
Unemployment rate$_{lt}$	0.066***	0.065***	0.075***
	(0.014)	(0.014)	(0.015)
Population (in 1,000)$_{lt}$	-0.009	0.002	0.030*
	(0.018)	(0.017)	(0.017)
Housing rental ratio$_{lt}$	4.430***	3.799***	3.307***
	(0.322)	(0.367)	(0.313)
TxDOT expenditures (in \$1,000,000)$_{lt}$		-0.030**	
		(0.012)	
Average house value$_{lt}$			0.020***
			(0.007)
Border county effects$_{lt}$	Yes	Yes	Yes
Year effects	Yes	Yes	Yes
Number of obs.	56,994	56,994	56,994
Log likelihood	-1,364.000	-1,351.000	-1,378.000

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level. The dependent variable is the number of environmental remediation firms in a random location (within 0-1 mile.)

Table 12: Exit patterns

Entry year	Total entrants	Exit year 2000	2001	2002	2003	2004	2005	2006	Total
2000	43	0	2	2	5	6	3	2	20
2001	47		5	6	16	6	3	3	39
2002	98			4	29	18	11	7	69
2003	54				7	11	10	5	33
2004	58					13	10	9	33
2005	50						9	7	16
2006	45							5	5
Total	395	0	7	12	57	54	46	38	214

Table 13: Establishment level summary statistics for exit

Variable	Mean (Standard deviation)
Establishments with past experience	0.364
	(0.481)
Average number of branches	1.021
	(2.334)
Age (in months)	42.702
	(22.532)
Average wage (in $10,000)	4.908
	(11.619)
Size	18.872
	(37.543)

Table 14: Logit results for exit

Variable	Exit					
	(1)	(2)	(3)	(4)	(5)	(6)
Number of environmental remediation	0.076***	0.075***	0.077***	0.078***	0.078***	0.078***
incumbents$_{ilt}$	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Number of TRI type incumbents$_{ilt}$	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Establishments with past experience$_i$	-0.016	-0.017	-0.016	-0.017	-0.018	-0.018
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Number of branches$_{it}$	-0.002	-0.002	-0.002	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age$_{it}$	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Median income \$0 – \$66,700$_{lt}$	-0.005		-0.006	0.006	0.006	
	(0.007)		(0.009)	(0.011)	(0.011)	
Median income >\$66,700 – \$100,000$_{lt}$	-0.003		-0.003	0.003	0.005	
	(0.013)		(0.014)	(0.014)	(0.015)	
Median income >\$100,000$_{lt}$	0.015		0.015	0.016	0.016	
	(0.012)		(0.012)	(0.012)	(0.012)	
Average wage (in \$10,000)$_{lt}$	-0.003	-0.003	-0.003	-0.002	-0.003	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Size$_{it}$	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
College ratio$_{lt}$		-0.016	0.019	-0.204	-0.190	-0.260
		(0.128)	(0.171)	(0.211)	(0.209)	(0.200)
Number of amenity establishments$_{lt}$			-0.000	-0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.000)	(0.000)
Number of roads$_{lt}$	-0.000	-0.000	-0.000	-0.000		-0.000
	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)
Number of rail roads$_{lt}$	-0.002	-0.002	-0.002	-0.001		-0.001
	(0.001)	(0.001)	(0.001)	(0.001)		(0.001)
Unemployment rate$_{lt}$				-0.002	-0.002	-0.002
				(0.003)	(0.003)	(0.003)
Population (in 1,000)$_{lt}$	0.002	0.001	0.002	0.001	0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Land area (in 100 in square miles)$_{lt}$				-0.009	-0.014	-0.009
				(0.011)	(0.012)	(0.011)
Housing rental ratio$_{lt}$				0.090	0.098*	0.106*
				(0.056)	(0.058)	(0.054)
MSA	-0.005	-0.008	-0.004	-0.017	-0.001	-0.019
	(0.036)	(0.036)	(0.036)	(0.039)	(0.047)	(0.039)
TxDOT expenditures (in \$1,000,000)$_{lt}$					0.001	
					(0.001)	
Average house value$_{lt}$						0.003
						(0.002)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	1,346	1,346	1,346	1,346	1,346	1,346
Log likelihood	-420.972	-422.477	-420.822	-419.178	-419.228	-419.759

*** denotes statistical significance at the 1 percent level, ** denotes statistical significance at the 5 percent level and *denotes statistical significance at the 10 percent level. The dependent variable takes the value of 1 for exit and 0 otherwise.

Appendix

Proof for the proposition in the theoretical analysis

Under conditions (1) and (3), total pollution in area l, $N(m_l, Z_l)e_l^*$, will be concave in income for $0 \leq m_l < \widehat{m}$. Indeed, by differentiating twice $N(m_l, Z_l)e_l^*$ with respect to m_l, we get:

$$\frac{d^2}{dm_l^2}\left(N(m_l, Z_l)e_l^*\right) = \frac{d^2N(m_l, Z_l)}{dm_l^2}e_l^* + 2\frac{dN(m_l, Z_l)}{dm_l}\frac{de^*}{dm_l} + N(m_l, Z_l)\frac{d^2e_l^*}{dm_l^2}$$

which is negative for all $0 \leq m_l < \widehat{m}$, as $\frac{dN(m_l,Z_l)}{dm_l} > 0$ and $\frac{d^2N(m_l,Z_l)}{dm_l^2} < 0$ (condition 1), and $\frac{d^2e_l^*}{dm_l^2} < 0$ (condition 3).

Moreover, under conditions (1)-(3), $N(m_l, Z_l)e_l^*$ will be increasing for low-income levels: at $m_l = 0$, we have $N(0, Z_l) = 0$ (condition 2) and $e_l^* > 0$ (interior solution), so:

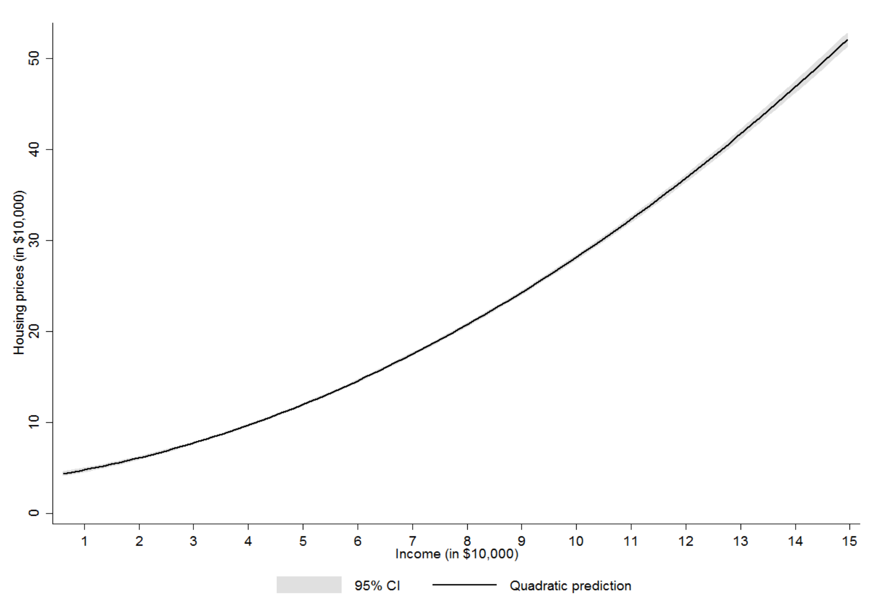
$$\frac{d}{dm_l}\left(N(0, Z_l)e_l^*\right) = \frac{dN(0, Z_l)}{dm_l}e_l^* > 0$$

Finally, under condition (1), $N(m_l, Z_l)e_l^*$ will be decreasing for high-income levels: at $m_l = \widehat{m}$, $\frac{dN(m_l,Z_l)}{dm_l} = 0$, which implies that

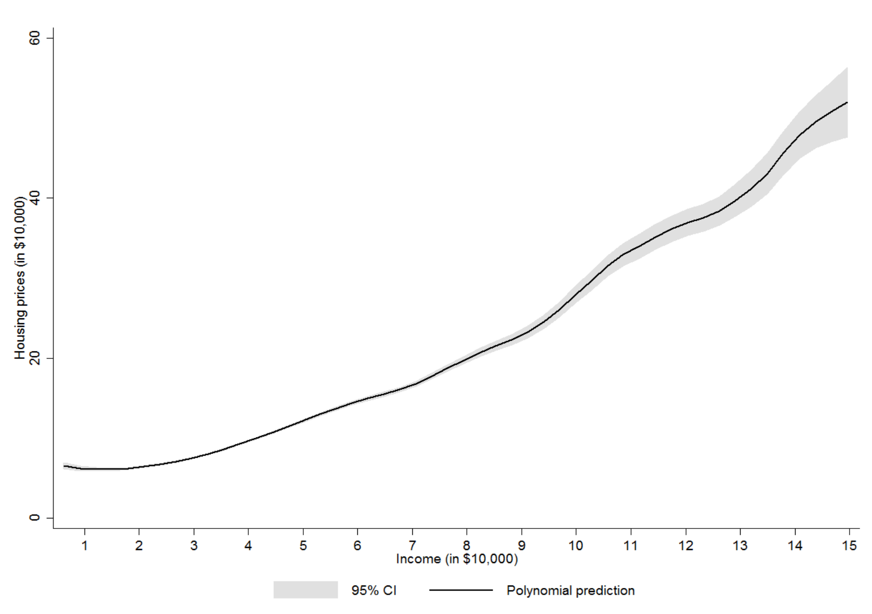
$$\frac{d}{dm_l}\left(N(\widehat{m}, Z_l)e_l^*\right) = N(\widehat{m}, Z_l)\frac{de_l^*}{dm_l} < 0$$

Therefore $N(m_l, Z_l)e_l^*$ must reach a maximum at some $0 < m^* < \widehat{m}$.

Figures and Tables



Panel A



Panel B

Figure 1: Relaionship between housing prices and median income

Table A.1: Variable Descriptions

Variable	Description
Number of environmental remediation firms$_{lt}$	Number of environmental remediation per tract per year. NAICS codes: 562112, 562211, 562910, and 562920
TRI type establishments$_{lt}$	Number of TRI type firms per tract per year.
Tract-level toxicity (in pounds)	The EPA provides toxicity weights for each chemical reported in the TRI which allows for aggreating heterogeneous releases. We then aggregate total release per tract for a given year.
Age$_{it}$	Establishment's age in months
Number of employees$_{it}$	Establishment or tract level number of employees per year.
Employement ratio$_{it}$	This is the establishment's employment divided by the total employees in the industry in TX at a given year.
Wage$_{it}$	Establishment level wage per year.
Plant with past experience$_{i}$	This is a plant with past experience in the industry
Number of branches in TX$_{it}$	Number of branches in TX
Number of roads$_{lt}$	We use the U.S. Census Bureau's Census Feature Class Codes (CFCC) to identify roads. These road maps are provided by ESRI Data & Maps (2000) at census tract level. We use all major highways to small roads that provide access to businesses, facilities, and rest areas along limited-access highways number of roads.
Number of rail roads$_{lt}$	As in roads we use the U.S. Census Bureau's Census Feature Class Codes (CFCC) and ESRI Data & Maps (2000) to identify rail roads. We use all major and minor rail tracks identified by ESRI Data & Maps.
Median income (\$)$_{lt}$	Census tract level median income.
Poverty ratio$_{lt}$	Census tract level percentage of the population under the poverty rate.
College ratio$_{lt}$	Census tract level college graduates as percentage of the population.
Number of amenity establishments$_{lt}$	To measure the relative local presence of amenities, we compute the tract level number of establishments in NAICS 71, Arts, Entertainment, and Recreation, and NAICS 721110 (hotels and motels), 722110 (full service restaurants), and 722410 (drinking places, alcoholic beverages) as reported in the QCEW data.
Housing rental ratio$_{lt}$	Tract level percentage of hosing stock rented.
Unemployment rate$_{lt}$	Census tract level unemploymentrate.
Population$_{lt}$	Census tract level total population.
TxDOT expenditures$_{lt}$(in \$1,000,000)	We construct tract level road construction expenditures by weighting county totals by tact level population.
Average house value$_{lt}$	Census tract level average house value.
Percent nonwhite residents	Census tract-level share of nonwhite population per year.

Table A.2: Explaining Variation in Toxic Pounds, TRI Type Firms, and Remediation Firms

Variable	Toxicity in pounds_{lt}	Number of firms in TRI type_{lt}	Number of firms in Remediation_{lt}	
	(1)	(2)	(3)	(4)
Median income (in \$10,000)$_{lt}$	0.792***	1.302***	0.517***	0.450***
	(0.127)	(0.138)	(0.088)	(0.085)
Median income (in \$10,000)$^2_{lt}$	-0.076***	-0.093***	-0.034***	-0.030***
	(0.016)	(0.019)	(0.011)	(0.010)
Median income (in \$10,000)$^3_{lt}$	0.002***	0.002***	0.001*	0.001*
	(0.001)	(0.001)	(0.000)	(0.000)
Number of TRI type incumbent establishments$_{lt}$				0.027***
				(0.005)
Number of TRI type incumbent employees$_{lt}$				
Percentage nonwhite residents$_{lt}$	1.028***	2.760***	0.018	-0.178
	(0.191)	(0.532)	(0.199)	(0.198)
Average wage (in \$10,000)$_{lt}$	0.033***	0.077***	0.024***	0.021***
	(0.006)	(0.012)	(0.004)	(0.004)
College ratio$_{lt}$	-5.325***	-8.741***	-3.978***	-3.543***
	(0.916)	(0.988)	(0.648)	(0.633)
Number of amenity establishments$_{lt}$	-0.007*	0.111***	0.011***	0.007***
	(0.003)	(0.014)	(0.001)	(0.001)
Number of roads$_{lt}$	0.015***	0.171***	0.024***	0.015***
	(0.002)	(0.025)	(0.002)	(0.003)
Number of rail roads$_{lt}$	0.072***	0.400***	0.058***	0.043***
	(0.009)	(0.037)	(0.006)	(0.006)
Unemployment rate$_{lt}$	0.014***	-0.005	0.021***	0.023***
	(0.005)	(0.017)	(0.007)	(0.007)
Population (in 1,000)$_{lt}$	-0.001	0.106***	0.034***	0.032***
	(0.008)	(0.019)	(0.009)	(0.008)
Land area (in 100 in square miles)$_{lt}$	-0.034***	-0.315***	-0.110***	-0.083***
	(0.013)	(0.052)	(0.025)	(0.021)
Housing rental ratio$_{lt}$	-0.070	1.591***	1.208***	1.120***
	(0.178)	(0.334)	(0.200)	(0.192)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Number of observations	30,114	30,114	30,114	30,114
Log likelihood	-3,283.000	-109,657.000	-9,644.000	-9,516.000
Uncensored observations	833	29,968	2,104	2,104

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.3: Explaining Variation in Toxic Pounds, TRI Type Firms, and Remediation Firms – alternate specification

Variable	Toxicity in pounds$_{lt}$	Number of firms in TRI type$_{lt}$	Number of firms in Remediation$_{lt}$	
	(1)	(2)	(3)	(4)
Median income \$0 – \$66,700$_{lt}$	0.244***	0.621***	0.309***	0.274***
	(0.034)	(0.048)	(0.033)	(0.033)
Median income >\$66,700 – \$100,000$_{lt}$	-0.017	0.047	-0.041	-0.051
	(0.069)	(0.113)	(0.058)	(0.056)
Median income >\$100,000$_{lt}$	-0.628*	-0.088	0.025	0.036
	(0.359)	(0.062)	(0.052)	(0.048)
Number of TRI type incumbent establishments$_{lt}$				0.027***
				(0.005)
Number of TRI type incumbent employees$_{lt}$				
Percentage nonwhite residents$_{lt}$	0.868***	2.633***	0.046	-0.141
	(0.180)	(0.538)	(0.196)	(0.195)
Average wage (in \$10,000)$_{lt}$	0.033***	0.077***	0.024***	0.021***
	(0.006)	(0.012)	(0.004)	(0.004)
College ratio$_{lt}$	-5.735***	-8.845***	-4.090***	-3.650***
	(0.960)	(1.021)	(0.648)	(0.632)
Number of amenity establishments$_{lt}$	-0.007*	0.111***	0.011***	0.007***
	(0.003)	(0.014)	(0.001)	(0.001)
Number of roads$_{lt}$	0.015***	0.172***	0.024***	0.015***
	(0.002)	(0.025)	(0.002)	(0.003)
Number of rail roads$_{lt}$	0.071***	0.398***	0.057***	0.043***
	(0.009)	(0.037)	(0.006)	(0.006)
Unemployment rate$_{lt}$	0.012**	-0.012	0.020***	0.022***
	(0.005)	(0.017)	(0.007)	(0.007)
Population (in 1,000)$_{lt}$	-0.002	0.107***	0.033***	0.031***
	(0.008)	(0.019)	(0.009)	(0.009)
Land area (in 100 in square miles)$_{lt}$	-0.031**	-0.312***	-0.107***	-0.081***
	(0.012)	(0.052)	(0.024)	(0.021)
Housing rental ratio$_{lt}$	-0.028	1.580***	1.209***	1.121***
	(0.177)	(0.339)	(0.200)	(0.193)
Border county effects$_{lt}$	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes
Number of observations	30,114	30,114	30,114	30,114
Log likelihood	-3,298.000	-109,663.000	-9,636.000	-9,508.000
Uncensored observations	833	29,968	2,104	2,104

Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.4: Explaining Variation of Non–TRI Type Firms

Variable	Number of firms$_{lt}$
	(1)
Median income \$0 – \$66,700$_{lt}$	0.910***
	(0.088)
Median income >\$66,700 – \$100,000$_{lt}$	1.561***
	(0.161)
Median income >\$100,000$_{lt}$	0.006
	(0.135)
Average wage (in \$10,000)$_{lt}$	0.161***
	(0.024)
College ratio$_{lt}$	33.129***
	(2.359)
Number of amenity establishments$_{lt}$	0.362***
	(0.053)
Number of roads$_{lt}$	0.239***
	(0.018)
Number of rail roads$_{lt}$	0.051**
	(0.023)
Unemployment rate$_{lt}$	0.092**
	(0.038)
Population (in 1,000)$_{lt}$	0.884***
	(0.037)
Land area (in 100 in square miles)$_{lt}$	-0.287***
	(0.044)
Housing rental ratio$_{lt}$	15.099***
	(0.636)
Border county effects$_{lt}$	Yes
Year effects	Yes
Number of observations	30,114
Log likelihood	-113,216.000
Uncensored observations	0

Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: Tract ranking

Variable		2006		
		Income	Education	Population
	Income	0.964		
2000	Education		0.990	
	Population			0.960