Press and Leaks: Do Newspapers Reduce Toxic Emissions?

Pamela Campa[*]

IIES, Stockholm University
JOB MARKET PAPER
November 2012

Abstract

I investigate whether media coverage induces firms to reduce toxic emissions. I first develop a simple model where newspapers inform consumers about the level of toxic emissions of a firm, potentially lowering demand. Firms react to this "threat of coverage" by proactively adjusting emissions. I test this using data on emissions from 25,523 plants in 1998-2009 from the Toxic Release Inventory of the US Environmental Protection Agency, coupled with data on location and content of newspapers. I find that an increase in *Newspapers Density*, that is the number of newspapers nearby the plant, raises the press coverage of the plant's toxic emissions and reduces the amount of these emissions. If a plant were to move from the fifth to the ninety-fifth percentile of *Newspapers Density*, its emissions would be 15% lower. I show that this association is unlikely to be driven by selection on unobservables, and is larger in industries exposed to consumer pressure and in counties subjected to extreme negative health outcomes. My estimates suggest that aggregate toxic emissions from plants producing final goods would be 11% larger if there were no newspapers within a 20-mile radius from the plants, a number that doubles for plants located in counties that have experienced an extreme number of cancer deaths.

1 Introduction

This paper investigates how media coverage shapes corporate decisions on environmental issues, above and beyond regulatory compliance. Anecdotal evidence suggests that the diffusion of information through newspapers may substantially influence corporate environmental decisions. For example, in 2003, toxic emissions at a Chevron oil refinery in Contra Costa county, in California, were nearly equal to one million pounds. Emissions in the same year at another Chevron oil refinery in Jackson County, in Mississippi, were slightly larger, all according to the data reported in the Toxic Release Inventory (TRI). The former oil refinery was the second biggest polluter in its county, the latter was the third biggest. Within two days from the publication of the TRI data, five articles were published in local newspapers in California, which featured the Chevron oil refinery in Contra Costa county as one of the top polluters in the Bay Area. Ten more articles covering toxic emissions at this plant were written in local newspapers between 2004 and 2008 in the following three days. In comparison, there was no press coverage of toxic emissions at the refinery in Jackson county between 2003 and 2008. Emissions at the Chevron plant in Contra Costa county declined by 50% between 2003 and 2009. Over the same period, emissions at the Jackson county plant went down by only 7%. The Contra Costa plant is located near several newspapers, whereas the Jackson plant has one only newspaper within a 20-mile radius.

To what extent can the differential exposure to newspapers explain the different speed at which emissions decreased at the two Chevron refineries between 2003 and 2009? Does the presence of newspapers nearby a plant make the plant accountable for its environmental performance, solving a problem of asymmetric information on pollution between firms and residents? Answering this question matters, given the effects of pollution on human health (Chay and Greenstone, 2003; Currie and Schmieder, 2009; Agarwal et al., 2010) and on cognitive abilities (Nilsson, 2009; Sanders, 2012).

In this paper, I show that the characteristics of the local newspaper market affect firms' decisions on legal toxic emissions. I first develop a simple model of firm accountability to consumers, mediated by newspaper coverage (accountability through investors' behavior is not modeled, but it is discussed in the implications of the model). Local consumers, the "constituents", gain utility from a firm's local production, through jobs and economic spillovers. However, they also bear the health and environmental costs of local industrial production. The local consumers do not directly observe the level of emissions from the plant, but may get this information from local newspapers. The higher the density of newspapers in the vicinity of the plant, the larger the probability of newspaper coverage. If a newspaper article reports high emissions, consumers decrease their demand for the good produced at the plant to minimize the health impact of local industrial production. The firm trades-off the expected loss of demand following bad press against the cost of using a clean production technology. Since a higher density of newspapers increases the expected loss it lowers the probability that the plant pollutes.

I test these hypotheses using data from the emissions of about 16,000 plants in 1998-2009 from the Toxic Release Inventory (TRI), which is the *Pollutant Release and Transfer Register* for the US, and which is produced by the Environmental Protection Agency. I combine this with data

on the location of about 1,500 local newspapers, and data on the newspaper coverage of the TRI announcements of any plant that was ever among the top 20 polluters.

I first analyze the determinants of newspaper coverage of emissions. I provide graphical evidence that the probability that a plant's emissions are featured in a nearby newspaper is approximately inverse to its distance from the newspaper. Total expected coverage in any newspaper is then approximately equal to what I call *Newspapers Density:* the number of newspapers in 20 mile rings around the plant, weighted by the inverse of their distance to the plant. Holding state and industry-sector fixed, a higher *Newspapers Density* raises the probability that a plant is covered in newspapers located within its respective ring. To guard against any spurious correlation between *Newspapers Density* and a plant's newsworthiness, I control for coverage in newspapers outside the rings, which should react to a plant's newsworthiness as much as those inside the rings. According to my estimates, a plant five miles away from a newspaper has a 0.004 lower probability of coverage than a plant five miles away from three newspapers. This is equal to 16% of the average predicted probability of coverage in the sample.

I then analyze the effect of *Newspapers Density* on toxic emissions. Comparing plants in the same industry group, county and year shows that plants in areas with a higher *Newspapers Density* have lower toxic emissions. I show that this relationship is not explained by selection on observables and is unlikely to be explained by selection on unobservables (Altonji et al., 2005). According to my estimates, if a plant were to move from the 5th percentile of *Newspapers Density* to the 95th percentile, its emissions would be 15% lower. An indication that the effect goes through consumer influence is that the effects are largest for plants in subsectors that mostly sell final goods. Moreover, the association of *Newspapers Density* with emissions is larger in counties that have experienced extreme negative health outcomes in the recent past, and where consumer awareness is likely to be high. This is consistent with the model's assumption that the constituents pressure firms because of concerns about the health effects of toxic emissions.

There is no evidence that multi-plant firms move emissions from one plant to another, in response to high *Newspapers Density*. This has relevant implications for the interpretation of the estimates in aggregate terms, analogously to the discussion of leakage in response to carbon taxes.

I finally estimate the fall in aggregate toxic emissions if there were no newspapers close to any plant, so that *Newspapers Density* was zero for all plants. In this counterfactual scenario, aggregate emissions would be 3% higher. Moreover, the emissions from plants in industries that sell final goods would be 11% higher. Finally, emissions from plants located in areas that have experienced extreme infant mortality and mortality from cancer would be, respectively, 17% and 22% higher.

This paper contributes to the literature on the determinants of corporate social responsibility and, more specifically, of corporate environmentalism. Papers in this literature have analyzed both theoretically and empirically what motivates corporate pro-active behavior in terms of social and environmental outcomes (Arora and Cason, 1995; Hamilton, 1995; Konar and Cohen, 1997; Khanna and Damon, 1999; Hamilton, 1999; Maxwell et al., 2000; Harrington et al., 2008). According to

Kitzmueller and Shimshack (2012), the empirical evidence points toward a prominent role of public politics and *private politics*.[1] However, the role of exposure to mass media in fostering pro-active behavior has not been investigated. This paper fills this gap, studying whether newspapers create incentives for firms to pollute less than their industry-group counterparts that are less exposed to press coverage, increasing the threat of public and private politics.[2]

This paper also contributes to the literature on the effects of mass-media on policy and political outcomes (Prat and Strömberg, 2011) and on corporate decisions (Dyck et al., 2008; Guerrero, 2012). In this literature, Dyck and Zingales (2002) show that the cross-country variation in firms' environmental responsiveness is partly explained by the diffusion of the press. With respect to their study, I exploit within-country variation, using plant-level data on emissions and proximity to newspapers. I directly address the role of the geography of the newspaper market in shaping incentives for newspapers to cover firms, deriving implications on the type of policies that would increase citizens' information on corporate behavior.

The rest of the paper is organized as follows. Section 2 presents the conceptual framework. Section 3 describes the data used. Section 4 presents the estimation strategy and the empirical results. Section 5 investigates the existence of spillovers and computes aggregate effects. Section 6 concludes the paper and draws policy implications.

2 Conceptual Framework

I develop a simple model that studies why the density of newspapers near firms' plants induces firms to decrease toxic emissions. There are two types of agents: a continuum of firms that produce a locally demanded good and a group of consumers-residents (people who consume the local good and live in the vicinity of the firms' plants) that I will refer to as "constituents". In the interaction between a firm and its constituents, moral hazard arises because of an unobservable action, i.e. whether the firm's plant pollutes or not. Local newspapers can mitigate the moral hazard problem.

Each firm sells the locally demanded quantity y_l produced at a fixed cost c_j. The fixed cost varies with the production technology, which can be clean, e_c, or dirty, e_d. The clean technology is more expensive, $c_c > c_d$, but also produces a larger health and environmental cost, h, to the consumers:

$$h(y_l, e_c) = 0,$$
$$h(y_l, e_d) = \overline{h}(y_l) > 0,$$

where $\overline{h}$ increases with y_l, at an increasing rate. The firm decides whether to use the technology e_c or e_d with the objective of maximizing profits.

The constituents demand the quantity y_l. They act as a group.[3] Their utility of y_l is made up of three components. First, it is linearly increasing in the consumption of y that they can buy from local producers (y_l) and in the national market (y_{nl}); $y = y_l + y_{nl}$. Second, they get some direct utility from the presence of each local plant, $V(y_l)$, because of the creation of jobs and indirectly through economic spillovers. $V(y_l)$ is increasing in y_l at a decreasing rate. Third, they bear the health and environmental cost h, caused by the production of y_l. h can take the form of increased incidence of certain diseases, dirty local rivers or lakes, high level of dust in the neighborhood etc. The constituents do not observe h, because they do not observe what technology is adopted at the plants located in their neighborhood. They only see the aggregate environmental and health damage H, which can be caused by different plants as well as by mobile sources. Moreover, they know how the choice of technology affects their health, i.e. the function $h(y_l, e_j)$.

The constituents are endowed with wealth W. I assume that for any value y_l larger than W, the marginal utility from the production of y_l is lower than its marginal cost when the dirty technology is used, so that:

Assumption 1. $V'(y_l) < \overline{h}'(y_l) \quad \forall y_l \geq W$

This occurs when plants are located in relatively rich places, where the marginal value of job creation and economic growth is low, so that it is more than offset by the health and environmental cost of economic activity. I further assume that the value of job creation is higher than the health cost at zero production, $V(0) \geq h(0)$.

The constituents maximize their portion of utility of good y as follows:

$$\max_{y_l, y_{nl}} \quad R = E\left[y_l + y_{nl} + V(y_l) - h(y_l, e_j)\right]$$

$$s.t. \quad y_l + y_{nl} \leq W.$$

The constituents first choose the quantity y^* that they want to consume and then how to allocate this demand between local and non local producers. Their problem involves setting up the optimal "demand strategy" of y_l, once y^* has been chosen. y_{nl} is then residually determined. Therefore, the constituents' problem boils down to maximizing the expected $V(y_l) - h(y_l, e_j)$ subject to the constraint that y_l must be between zero and W. If they observed e_j, they would adopt the following demand strategy:

$$y_l = \begin{cases} W & \text{if } e = e_c \\ \underline{y_l} < W & \text{if } e = e_d. \end{cases} \tag{1}$$

---

[3]One way of thinking about the constituents acting as a group is to think, as in Aghion et al. (2012), of a rule set by the individuals belonging to the group which, if followed by every group member, maximizes the utility of the group.

where

$$\underline{y_l} = argmax[V(y_l) - \overline{h}(y_l)].$$

See Proof in Appendix A.

However, the constituents cannot directly observe what technology is used at the local plant. They can only learn that the plant uses a dirty technology through reports featured in newspapers. I assume that if one newspaper publishes news about a local plant, all constituents become informed.[4] Even though, in reality, the constituents could access information on the type of technology used at the firm directly, in practice they seem unlikely to do so; collective action problems and ignorance of the availability of data on emissions are likely to explain this behavior.[5]

Newspapers access information on whether or not the plant uses a dirty technology. Then, they write a story, $s = s_d$, with a probability that depends on their location with respect to the plant. If the plant uses a clean technology, there is no coverage in newspapers, i.e. $s = \oslash$. There are N newspapers, located within a distance D to the plant,which face a demand for news about toxic emissions from the plant. I call these "relevant" newspapers.[6]

Let p_i denote the probability that a relevant newspaper i writes an article, given that the plant uses the dirty technology, $p_i = Pr(s_i = s_d|e_d)$. I assume that p_i falls with the distance d_i from the plant. Two aspects motivate this assumption: first, the likely increasing (in distance) cost of coverage, which includes the cost of traveling to the plant to get more information and interview the employees, and the utility cost for the journalist of writing an article;[7] second, the decreasing (in distance) demand for news about the plant, as readers further away from the plant are less affected by its toxic emissions. The probability p_i is assumed to be zero for non relevant newspapers, i.e. newspapers at a distance larger than D from the plant.

The probability P that any newspaper writes a story about the plant after the news $e = e_d$ is out is equal to:

$$P = 1 - \prod_{i=1}^{N}(1 - p_i),$$

---

[4]This simply requires allowing constituents to communicate among themselves. Given that this communication could induce some moral hazard, possibly leading to a situation in which the constituents do not read newspapers, I assume that constituents read newspapers not only to become informed but also because they intrinsically enjoy reading news. Note that the results hold if only a share of constituents becomes informed.

[5]Analyzing the lessons that can be learned from the publication of data on toxic emissions in the US, Hamilton (2005) emphasizes the important role of information intermediaries, given the apparent unwillingness of citizens to gather the information on their own from the direct source of data dissemination (i.e. the Environmental Protection Agency's - EPA - website and publications). Similarly, commenting on the spread of TRI data in the years 1987-1992, Bui and Mayer (2003) observe that the primary source of TRI information for communities was not the raw data release, but rather the media accounts.

[6]Being within a certain distance to the plant makes a newspaper "relevant", because the constituents and the local newspaper's readers substantially overlap. The assumption that newspapers sell in an area nearby their headquarters is especially plausible for local newspapers. In Section 4, the empirical analysis focuses on the US, where local newspapers represent the largest share of the newspaper market.

[7]I assume that the utility cost of coverage is lower if the journalist has a direct interest in shaming a polluting plant, as might be the case if the journalist works and lives in the same area as the plant.

which, for very small p_i, can be approximated to:

$$P \approx \sum_{i=1}^{N} \frac{\partial P}{\partial p} \Delta p_i = \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} p(d_i). \tag{2}$$

As in Besley and Prat (2006), three assumptions are implicit in this setup: news cannot be fabricated, only negative stories about a firm are news, and all newspapers have access to the same type of information. The second assumption could easily be relaxed by imposing that newspapers are more likely to write negative rather than positive stories, and the third assumption reflects exactly the type of information flow that I exploit in the empirical application in Section 4.[8]

2.1 Equilibrium

The constituents make their decision only based on s, choosing:

$$y_l^* = argmax[V(y_l) - (1 - P(e_c|s))\overline{h}(y_l)],$$

where the probability that the plant uses the clean technology is computed using Bayes' rule[9],

$$P(e_c|s) = \begin{cases} \approx Pr[(\overline{y}_l - \underline{y}_l) \sum_{i=1}^{N} p(d_i) \geq c_c - c_d] & \text{if } s = \oslash \\ 0 & \text{if } s = s_d. \end{cases} \tag{3}$$

Given these beliefs, the constituents' best response is

$$y_l^* = \begin{cases} \overline{y}_l > \underline{y}_l & \text{if } s = \oslash \\ \underline{y}_l & \text{if } s = s_d. \end{cases} \tag{4}$$

See the Proof in Appendix A. Given this best response, the firm faces expected profits equal to:

$$\pi_c = \overline{y}_l - e_c$$

and

$$\pi_d = (1 - P)(\overline{y}_l - e_d) + P(\underline{y}_l - e_d).$$

---

[8]In the empirical application, I look at emissions as reported in the EPA's administered Toxic Releases Inventory. This is a database with information on plant-level toxic emissions, which can be freely accessed through different media.

[9]See Appendix A for the derivation of the beliefs with the Bayes' rule; $Pr[(\overline{y}_l - \underline{y}_l) \sum_{i=1}^{N} p(d_i, \alpha) \geq c_c - c_d]$ is the probability that in equilibrium a firm chooses not to pollute.

Therefore, the firm decides what technology to use as follows:

$$e_j = \begin{cases} e_c & \text{if } (\bar{y}_l - \underline{y}_l) \sum_{i=1}^{N} p(d_i, \alpha) > e_c - e_d \\ e_d & \text{otherwise.} \end{cases} \tag{5}$$

See Proof in Appendix A.

2.2 Testable Implications

In the theoretical framework, I assume that the probability that a polluting plant is featured in a newspaper is decreasing with the distance of the plant to the newspaper. This implies that the plant's probability of coverage is increasing in $\sum_{i=1}^{N} p(d_i)$, where p is decreasing in d_i. In order to take this relationship to the data, I need to make a functional form assumption on p. In Figure 1, I use data on coverage of toxic emissions and the distance of plants from newspapers in the US, and I show that the relationship between the probability that a plant's toxic emissions are featured in a newspaper and its distance from the newspaper is approximately inverse.[10] Given this evidence, I specifically assume that

Assumption 2. $p_i = \frac{1}{d_i}$.

Assumption 2 and Equation 2 imply that the probability of news coverage equals

$$P \approx \sum_{i=1}^{N} \frac{1}{d_i} = Newspaper\ Density, \tag{6}$$

where the last equality is a definition. Let $\eta = (\bar{y}_l - \underline{y}_l)/(e_c - e_d)$. We now have the following proposition.

Proposition 1. *If Newspaper Density $\geq \eta$, then plants do not pollute, and there is no coverage of toxic emissions. If Newspaper Density $< \eta$, then plants pollute and each plant gets featured in the news with probability equal to Newspaper Density.*

I will test these implications in the empirical analysis in Section 4.2. Further, note that since plants only pollute when *Newspaper Density* is sufficiently low, the probability that coverage of a polluting plant arises must be close to zero:

Lemma 1. *In equilibrium, coverage of polluting plants is close to zero.*

Before concluding the discussion of the conceptual framework, it is worth emphasizing that, although I focused on pressures made by local consumers, the threat of *private politics* (Baron, 2001, 2003) that affects firms' decisions may come from national or international consumers. For example, local newspapers might "pull the alarm" for national newspapers: once local newspapers

---

[10]In Figure 1 the dots are means, and the line connects the fitted values from the regression of coverage on the inverse of distance.

have met the initial cost of writing an article, the additional cost of reporting for other newspapers is arguably low; then the news might spread nationwide. Moreover, if consumers value the health of other citizens and environmental conditions at any location, local newspapers could provide input for boycotts that potentially spread nation-wide, or even internationally. Finally, other instances of private politics, such as massive sales of stock market shares, or lawsuits and pressures from the civil society for more stringent regulation, can also cause a reduction in expected profits, inducing the pro-active behavior modeled in this section.[11]

3 Data and Variables Description

Toxic Emissions I use data on toxic releases collected by EPA's TRI program. Starting in 1989, every year plants with more than 10 employees that operate in certain sectors (primarily in manufacturing) and that manufacture, process or otherwise use each of about 650 toxic substances above certain thresholds must report the quantity of each toxic substance released into the air, the water and the land. [12] Firms can update their data on emissions when they discover mistakes in previous reporting.[13] However, for the purpose of the current research, I use original data on emissions. These data were provided by EPA for the years 1996 to 2009. In 1998, there was a major change in the program, with new sectors being added among those that were required to report their emissions. Therefore, I limit my sample to plants observed in the years 1998-2009. I use information on total plant-level emissions, the industry in which a plant operates, and its exact geographic location (latitude and longitude).

This is an unbalanced panel; the number of plants is 25,523, for a total of 154,587 plant-by-year observations.

---

[11]In an extension of the simple model presented in this section, I allow for the possibility that firms "capture" newspapers. The strategic interaction between firms and newspapers is a variant of Besley and Prat (2006), modified to allow for the role of geography (i.e. distance) in determining coverage. As in the model presented in this section, relevant newspapers have incentives to cover a plant because of their proximity to the plant. However, firms can capture these newspapers through advertisements and bribes (this is what Ellman and Germano (2009) call the "regulatory view" on advertising, as opposed to the "liberal view"; see Gambaro and Puglisi (2010) for empirical evidence on the "regulatory" role of firms' advertising in newspapers). Newspapers thus decide whether to cover a firm depending on their distance from the plant, conditional on not being captured. Distance plays a role in determining the probability that a relevant newspaper is captured through its effect on the cost of coverage: given that the firm has to compensate the newspapers for the forgone net revenues from not publishing the news, the lower the cost of coverage, the higher these foregone net revenues. Therefore, in this model, distance affects the probability of coverage through an additional channel. The firm captures the newspaper only if the cost of capture is lower than the cost of shifting to the clean technology. The cost of capture is increasing in the number of newspapers that are relevant for the firm, because, for the equilibrium of the bargaining game to be robust to deviations, each newspaper has to be compensated as if it were a monopolist in the market, i.e. as if its readers are all the constituents of the plant. The main difference between the extended model and the model presented in this section is thus that, in the former, the number of relevant newspapers matters because the probability that the firm bribes the newspapers is decreasing in this number. This extension has implications for the relationship between advertisement, pollution and newspaper market that are not testable with US data because of the lack of information on firm-newspaper specific advertisement relationships (and possibly bribes).

[12]If the quantity released is lower than 500 pounds, firms can choose not to report the quantity released, and will just submit a form that certifies that they manufacture some of the listed toxic substances; for these cases, emissions are set to zero. This exception does not hold for Persistent, Bioaccumulative and Toxic (PBT) chemicals.

[13]Updated data on emissions from 1989 to 2009 can be accessed on the EPA website.

The main dependent variable is total toxic emissions, Y_{pcit}, from plant p in county c and industry i at time t. All released amounts are reported in pounds except dioxin which is reported in grams. The mean is 164,481, with a standard deviation of 2,465,825.

A potential concern is that TRI data are self-reported. However, this is unlikely to be a major issue for several reasons. First, EPA's Office of Enforcement and Compliance Assistance (OECA), in conjunction with EPA's Regional Offices, conducts inspections and audits, and sanctions plants that misreport their toxic emissions.[14] OECA also queries big polluters and some facilities whose numbers change drastically over a short time period to check that the numbers are correct. EPA provides compliance incentives, i.e. "policies and programs that reduce or waive penalties under certain conditions for business, industry, and government facilities that voluntarily discover, promptly disclose, and expeditiously correct environmental problems".[15] Second, accounts in previous papers that use TRI data are reassuring on data quality. Currie and Schmieder (2009) quote several studies according to which general compliance in the TRI is high. Moreover, they find that reported emissions of TRI chemicals that are expected to impair birth outcomes do indeed have a negative effect on birth weight. Agarwal et al. (2010) point out that, although quality assurance tests carried out by EPA annually on a very small sample of facilities (< 100) document a large degree of error, there is not currently any evidence of systematic over- or under-reporting; they thus argue that the measurement error is not correlated with their dependent variable, which is infant mortality. In the empirical analysis of this paper, toxic emissions are the dependent variable; therefore, if the measurement error is not correlated with the independent variable of interest, *Newspapers Density*, it would only increase the estimator's variance; if it is instead correlated with *Newspapers Density*, it would be a source of bias, in a direction that cannot be determined a priori. However, given the lack of evidence on systematic over- or under-reporting, and given that EPA inspects plants that have more incentives to under-report in response to high *Newspapers Density*, i.e. big polluters, I assume that the error is not correlated with *Newspapers Density*. Finally, Hamilton (2005) reports the results of an enforcement program launched in Minnesota in 1991 to identify facilities that should have reported. The facilities identified turned out to be very small polluters, and the reason for non compliance was ignorance rather than intentional concealment. Ignorance of the reporting requirements or of the correct estimation method is likely larger in the first years of implementation of the program, and therefore it is less of a concern in my analysis where I look at emissions reported between 2001 and 2009.[16]

Newspapers Density To measure plant level *Newspapers Density,* defined in equation (6), I use a dataset reporting the name of almost all US newspapers with their city of location, the year they were founded and the year they were closed, if relevant.[17] This dataset is based on information

---

[14]Different examples of plants that are fined for failure to report can be found at http://www.epa.gov/tri/stakeholders/enforcement/enforce.html

[15]See http://www.epa.gov/compliance/incentives/index.html. The other information on compliance and enforcement reported here has been acquired through visits to EPA's website and emails with EPA employees.

[16]Emissions reported in the TRI need not be measured exactly but may be estimated.

[17]Unfortunately, more precise information on newspapers' location is not available; therefore I assume that a newspaper is located at the centroid of its city. While this certainly induces some error in the measure of distance,

published on the website *Chronicling America*. (In total, there are about 1,500 newspapers every year in my sample.). The key independent variable, $NewspapersDensity_{pt}$, is computed as the sum of all newspapers within rings of 20 miles radius around plant p at time t, weighed by their distance from the plant. The mean of *Newspapers Density* is 0.52, and the standard deviation is 0.78.

Newspaper Coverage I collected data on coverage of plants' toxic emissions in newspapers as follows. Among the plants that appear in the TRI dataset, I select those that were among the state top-20 polluters at least once in the period analyzed (i.e. 1998-2009). I look for articles covering these plants between 2000 and 2008. These plants reported emissions between 1998 and 2006, since TRI data are released by EPA with a two-year lag. I use a program called Imacros to iteratively search Newslibrary, an archive of editions of a large number of US newspapers, for each plant-by-year observation.[18] [19] The algorithm delivers articles published in newspapers within five days from the release of the TRI data, mentioning a short form of the name of the plant, the word EPA, and the name of the city or the county where the plant is located. Figure 2 in Appendix B shows an example of the output of this search. I manually verify that the articles identified cover the plant's yearly toxic emissions recorded in the TRI; hence, while undercounting may be an issue, overcounting is unlikely. I then check whether the articles selected are published in newspapers that are located within or outside the 20 miles rings around the plant. I finally construct the variable indicating whether each newspaper covers each plant in each year.

I estimate the effect of *Newspapers Density* on coverage in states with some coverage of toxic emissions during the sample period. The dependent variable in this analysis is C_{psit}, which is a dummy taking the value of 1 if plant p in State s and 2-digit industry i is covered in year t. Summary statistics for this sample are shown in Table 1. Figure 3 in Appendix B shows the number of articles by state during the sample period (the names of no coverage states are labeled in red). Around 5.3% of the plant-by-year observations in the sample receive some coverage. Thus, coverage is a relatively rare event. The average number of newspapers in the rings is nearly 2, and, among the newspapers in the rings, the average distance from a plant is about 10 miles.

For the analysis of the effect of *Newspapers Density* on toxic emissions, I consider the entire sample of plants. In different parts of the analysis, I distinguish between states where there is coverage ("coverage states"), and states where there is no coverage ("no coverage states"). Given that data on emissions are published with a two-year lag from the reporting year, and that I look at the effect of the one-year lag of *Newspapers Density*, in practice I analyze the determinants of emissions between 2001 and 2009. [20] Table 2 shows summary statistics for the estimation sample,

---

in a randomly selected sample of twenty newspapers whose location is searched in Google Maps, only three have a distance from the centroid (which presumably nearly coincides with the city center) larger than 1.5 miles, and of these only one is located in one of the three largest cities in this sample.

[18] Newslibrary archives articles from 4,508 titles. However, these are not all newspapers, as they include TV channels, radio channels, newswires and transcripts. After the articles have been extracted, I check which ones are published in newspapers, and I keep these.

[19] I limit the search to plants that reported emissions between 1998 and 2006, because the data on emissions (and therefore the articles) are published two years after the reporting year until 2008, whereas in 2009 data for more than one year are released by EPA.

[20] I drop plants that do not have a match in the respective county of location in the year in which they are observed

distinguishing between "coverage states" and "no coverage states."

University Density The variable *University Density* is constructed in the same way as *Newspapers Density*. I use data published by the US National Center for Education Statistics, listing all colleges and universities offering programs of two years or more, with the respective zip codes.[21]

Other controls The control variables that I use in the analysis are based on Census data. I download block-group data for each of the variables of interest.[22]. Using Geographic Information Systems as ArcGIS and Geospatial Modeling Environment, I split the US territory into cells with areas equal to one square kilometer (≈ 0.39 square miles). Every cell gets the value of the census block group that has its maximum area in the cell itself. Then, I calculate the average across these cells in rings with a radius of five miles from the plant. I use Census data for the years 2000 and 2010, and I linearly interpolate the values in between.

4 Results

4.1 Newspapers Density and Coverage

I first analyze how the probability that a plant is covered by any newspaper is affected by *Newspapers Density*. I test that implication of Proposition 1 by estimating the following equation:

$$C_{psit} = \alpha_0 + \alpha_1 NewspapersDensity_{pt} + \mu_{st} + \eta_{i2} + \epsilon_{psit} \tag{7}$$

where C_{psit} is a dummy taking the value of 1 if plant p in State s and two-digit industry i is covered in year t, and φ_{st} and η_{i2} are, respectively, state-by-year and 2-digit Primary NAICS (industry sector) fixed effects. The parameter estimates are reported in Table 3. *Newspapers Density* is positively correlated with the probability that a plant is covered in newspapers within its respective ring (Column 1).

However, within a state and an industry, plants that are located in areas with a high concentration of newspapers can be different from those in areas with a low concentration along several dimensions. For example, large plants could be both more well-known and "newsworthy" and located in areas with high *Newspapers Density*. I address this selection issue by separately studying how *Newspapers Density* correlates with coverage in newspapers inside and outside the 20-mile rings.

---

because I estimate a model with county-by-year effects; plants that do not have a match in the same industry-group (4-digit Primary NAIC), because I estimate industry-group effects; and plants located in counties whose area is above the 90th percentile of the sample distribution, where the heterogeneity across location with different levels of newspapers density is likely to be very large.

[21] The list of colleges is updated to the date at which I accessed the website of the *National Center for Education Statistics* (June 2012). Therefore, measurement error occurs if some universities were opened or closed in recent years. However, I assume that this is not a frequent event. For the purpose of the analysis in Section 4, the mere fact that a college opening or closure is approaching is sufficient to capture the characteristics of the location that I am interested in.

[22] A census block group is a cluster of census blocks that contains between 600 and 3,000 people, with an optimum size of 1,500 people.

Figure 4 shows the newspaper market for plants in Oregon, US. The black dots are plants that report to the TRI, the red triangles are newspapers, the larger circles are the 20-mile rings in which I measure *Newspapers Density*, and the smaller circles are the rings in which I measure the control variables. *Bhelen Manufacturing Co.* has one newspaper within its ring, whereas *Ash Grove Cement Co.*, *Grant Western Lumber Co.* and *Praire Wood Prods.* have none. If *Bhelen Manufacturing Co* is more likely to be covered because it is more newsworthy, and newsworthiness correlates with *Newspapers Density*, then *Newspapers Density* should also be positively associated with coverage in newspapers outside the *Bhelen Manufacturing Co*'s ring. If *Newspapers Density* is not a proxy for newsworthiness, its effect on coverage in newspapers outside the rings should be smaller. This is indeed what is shown in Table 3, Column (2): the correlation between *Newspapers Density* and probability of coverage outside the rings is positive, but it is lower than the coefficient in Column (1), and not different than zero at conventional levels. [23]

A limitation of the coefficient estimate in Column (1) in Table 3 is that the probability of coverage in newspapers within the ring is mechanically zero for plants with *Newspapers Density* equal to zero. Thus, I restrict the sample to plants in rings where there is at least one newspaper. In Column (3) I show that the effect of *Newspapers Density* on the probability that the plant is covered in newspapers within its ring is not purely mechanical. Moreover, the estimate of α_1 is relatively stable when I control for a measure of plant's newsworthiness, i.e. a dummy for coverage in newspapers outside the rings (Column 4), and when I control for a full set of demographics measured in the area around the plant (Column 5).[24]

According to the estimates in Columns (4) and (5), if *Newspapers Density* increases by one unit, the probability of coverage of toxic emissions increases by 0.009. If we compare a plant that is five miles away from one newspaper with another that is five miles away from three newspapers, then the probability of coverage in the relevant newspapers for the latter is larger by 0.0036. This is equal to nearly 16% of the average predicted probability of coverage of toxic emissions in the sample. Although the overall probability of coverage of toxic emissions is low, this is a large effect.

4.2 Newspapers Density and Toxic Emissions

I now turn to the main question, and test the implication of the model that plants located in areas with higher *Newspapers Density* pollute less. I estimate the following equation:

$$Y_{pcit} = \beta_0 + \beta_1 NewspapersDensity_{pt} + \beta_2 X_{pt} + \mu_{ct} + \eta_{i4} + \epsilon_{pcit} \tag{8}$$

where Y_{pcit} are toxic emissions from plant p in county c and industry i at time t, X_{pt} is a full set of controls measured at the plant level, μ_{ct} are county-by-year effects, η_{i4} are 4-digit Primary

---

[23]The positive association between $Pr(cov\ out)$ and *Newspapers Density* is not surprising because the newspapers within the ring may "pull the alarm" for the other newspapers, which will find it cheaper to write a story once the initial cost of writing an article has been met.

[24]The demographic controls are: log population density, log income, share of African-American, share with high school diploma or some college, share with associate diploma or more, share aged below 20, share aged above 65, unemployment in 2000.

NAICS (industry-group) effects, and *NewspapersDensity* is the measure described in section 3. In practice, I compare plants that operate in the same county, industry-group and year.[25] In Figure 4, for instance, I would compare emissions at *Behlen Manufacturing* and at *Ash Grove Cement*, provided that they operate in the same industry-group.

I estimate equation (8) using plant-level data on toxic emissions in states where there was some coverage during the sample period. The distribution of emissions is highly skewed to the right, so I perform the analysis on log emissions. Given that emissions can be equal to 0, I add a constant equal to 1 to the original value of emissions.

The estimation results are shown in Table 4. Column (1) reports the baseline estimate of Equation (8). As predicted in Section 2, *Newspapers Density* is negatively correlated with toxic emissions.

The estimate in Column (1) is not necessarily causal. I now discuss robustness to the inclusion of more controls and to sample selection. In section 4.3, I discuss this issue further.

In order to address selection, I restricted the sample to counties whose area is lower than the 90th percentile in the sample distribution.[26] Moreover, I control for the most likely correlates of emissions and *Newspapers Density*, using demographic variables measured near each plant. The selection of these controls is first of all guided by the historical account reported in Gentzkow et al. (2010), according to which population and income explain most of the variation in newspapers' entry and exit in the US. I thus control for the logarithms of population density (*Log Pop Density*) and of income per capita (*Log Income*) within a five-mile radius from the plant. I also allow for non-linear effects of these two variables, including dummies for realizations of log population density and log income within different percentiles. As shown in Column (2), the coefficient is still negative and statistically significant, although it decreases in magnitude.

Beside population density and income, other local characteristics might be correlated with emissions and *Newspapers Density*. First, emissions and *Newspapers Density* likely correlate with local preferences for environmental amenities. I proxy these local preferences with the share of people with tertiary education (*Education - associate or more*) in the five-mile ring around the plant.[27] This evidence suggests that *Education - associate or more* is a reasonable proxy for the unobserved variable. The point estimate of β_1 is remarkably robust to including the proxy in the controls. The share of people with tertiary education is negatively associated with toxic emissions.

In Column (4) I control for a number of additional variables. The literature on environmental

---

[25]Notice that there is only a limited time-variation in *Newspapers Density* generated by newspapers entry and exit. Therefore, I do not exploit the panel dimension of the dataset.

[26]Locations within large counties are most likely to be different in ways that cannot be controlled for.

[27]The standard environmental Kuznets Curve argument, according to which the quality of the environment improves with income growth above a certain level of income, has been contested on the ground that income is correlated with other variables that play a role in shaping preferences for environmental outcomes, among which education (Bimonte, 2002). A cross-tabulation of responses of US residents to the World Value Survey shows that higher levels of education are associated with more willingness to take actions in favor of the environment (see Figures 5 - 7 in Appendix B). This is in line with other survey evidence, as in Fleishman-Hillard and the National Consumers League (2007), that "environmental preferences in a society strongly depend on demographic characteristics such as education or technological development" (Kitzmueller and Shimshack, 2012).

justice suggests an association between toxic emission and the share of minorities in a neighborhood of the plant (Bullard, 1990): Thus, I control for the share of African-Americans living in the five-mile rings centered at the plant (*Share Black*). Moreover, people who live in the neighborhood of a plant gain utility from the creation of jobs and from the economic spill-overs associated with local production. Consequently, in a context of high unemployment, the trade-off between cost-effectiveness and environmental impact could be more easily resolved in favor of cost-effectiveness, so I control for the unemployment levels in the five-mile rings around a plant in 2000.[28] I also include two measures of the demographic composition of the population, *Share Younger 20* and *Share Older 65*, and the share of people with high school education and some college, *Education - high school or some college*, to account for the observation that a more educated workforce may make it easier to adopt a new clean technology.

The estimate of β_1 is stable to these additional controls. Using a measure developed by Altonji et al. (2005), and adapted to the continuous case by Bellows and Miguel (2009), I conclude that the selection on unobservables should be 2.5 times as large as the selection on observables in order to drive the point estimates for β_1 to zero. This is a large number, if one considers that I included the most powerful predictors of newspaper entry and exit, that the coefficient is remarkably stable to the inclusion of observables, and that I attempted to account for unobserved selection through a proxy variable approach.

The estimated effects are sizable. According to the estimates in Column (4), if we compare a plant that is five miles away from one newspaper, to one that is five miles away from three newspapers, the latter reports 3.8% lower toxic emissions. If a plant moved from the fifth percentile of *Newspapers Density* (0) to the ninety-fifth percentile (1.55), its emissions would be 15% lower.

In Section 5, I will discuss the implications of this effect in aggregate terms. While it might appear small, in Section 4.4.3 I show that this is the average of very heterogeneous effects across locations and industries, and this effect is much larger for some subsamples of plants. Moreover, accounting for industry-group effects, I might underestimate the effect of *Newspapers Density* on toxic emissions at a given locality, because I do not capture the effect of *Newspapers Density* on the location decision of firms in heavily polluting industry-groups.

4.3 Robustness Checks

The estimated coefficient on *Newspapers Density* in Column (4) of Table 4 is identified under the assumption that *Newspapers Density* is uncorrelated with the error term, conditional on the other included variables. This means that the controls should capture the county-year and industry-group differences that are correlated with emissions and *Newspapers Density*. The fact that the estimate of β_1 in Column (2) is robust to the introduction of additional controls suggests that this assumption is plausible. In this section, I will provide further evidence to corroborate this claim using different approaches. I first perform two placebo tests. The first tests whether *Universities*

---

[28]Census-block-group data on unemployment are not available for recent years. In a robustness check in Section 4.3.4, I address the potential issue arising from mis-measurement of unemployment in the latest years.

Density, computed analogously to Newspaper Density, correlates with emissions. The second tests whether Newspaper Density correlates with emissions in states which have no actual emissions coverage. Then, I analyze the relationship between emissions and newspaper density using only within-firm variation. Finally, I analyze robustness to using different samples, measures of the dependent variable and estimation methods.

4.3.1 Universities Density and Toxic Emissions: Placebo

Universities and colleges, like newspapers, are arguably more likely located in urban areas where the flow of knowledge and information is faster, where there are more residential amenities etc. Therefore, if the negative association of *Newspapers Density* with toxic emissions is due to non-controlled-for characteristics of urban locations, a measure of *Universities Density*, constructed as the main independent variable *Newspapers Density*, should be correlated with toxic emissions, when county-by-year and industry-group variation is exploited, and all controls are included. I test this using a placebo regression. Using the specification in equation (8), but replacing $NewspapersDensity$ by $UniversitiesDensity$, I get

$$Y_{pcit} = \varphi_0 + \varphi_1 UniversitiesDensity_{pt} + \varphi_2 X_{pt} + \mu_{ct} + \eta_{i4} + \epsilon_{pcit}. \tag{9}$$

According to the estimate in Table 5, we cannot reject the hypothesis that φ_1 is equal to zero, and the point estimate is substantially smaller than the estimates of β_1 shown in Table 4. If many of the characteristics of a location associated with the presence of newspapers overlap with those associated with the presence of universities, this makes a good case for a causal effect of *Newspapers Density* on emissions.

4.3.2 Newspapers Density and Toxic Emissions in States with no Coverage

Factors such as journalists' distaste for pollution or the occurrence of important events, such as local elections, at the date the EPA data are released, could drive the probability of coverage at some localities to zero, independently of the characteristics of the newspaper market. In these areas, there is no news coverage, no matter how close the newspapers are, and therefore closeness to a newspaper should not have any impact. On the other hand, if the correlation between Newspaper Density and pollution is driven by other factors, such as unobserved plant type, then the correlation would still be there also in these non-coverage areas.

I evaluate the identifying assumption estimating the effect of *Newspapers Density* for plants in states where newspapers never cover emissions. In the empirical analysis so far, I focused on states with some coverage during the sample period. However, as shown in Figure 3, for eight states the text search did not find any articles on toxic emissions published between 2000 and 2008. In these states, there should not be any effect of *Newspapers Density* on toxic emissions, because of the lack of coverage and threat of coverage. I test this hypothesis by estimating Equation 8 in "no coverage" states. In Table 5, Column (2), I report the result of this estimation.

The effect of *Newspapers Density* in states where there is no coverage cannot be distinguished from zero. To guard against the possibility that there is not enough variation, in this smaller sample, to estimate a precise coefficient, I also estimate Equation (8) with industry sub-sector effects (3-digit Primary NAICS), using the entire sample of counties (i.e. not excluding those with areas above the 90th percentile), or exploiting within state-year variation: also, in these regressions the effect of *Newspapers Density* in states where there is no coverage is virtually zero (results not shown and available upon request), although the large confidence intervals call for cautiousness in interpreting this result. Specifically, one can interpret the estimate in Column (2) either as a result of insufficient variation in the data, or as a signal of no clear association between *Newspapers Density* and toxic emissions. The latter interpretation, if valid, rules out an alternative channel of influence between the two variables. Namely, emissions may be lower in areas with larger *Newspapers Density* because residents, reading more newspapers, are more informed about environmental issues and as such look independently for more information on toxic emissions. However, if this were the case then *Newspapers Density* should be negatively associated with toxic emissions also in states where TRI statistics are not covered in newspapers.

4.3.3   Within-firm variation

A possible concern is that the *Newspapers Density* of plant p is correlated with unmodeled characteristics of its parent company, which affect firm-level decisions on toxic emissions. Some of the big firms in the dataset, for example, could be more concerned about reputation because they have been targeted by activists in the past. Coca-Cola, to name one, might pollute less because of these reputation-related concerns, and might locate its plants in areas dense with newspapers so that it can more easily influence journalists with its public relation activities. In this section, I try to address this concern by exploiting within-firm variation in *Newspapers Density* and toxic emissions.

Plants that report their toxic emissions to the TRI can be part of multi-plant firms. The TRI dataset records the parent company name and/or Dun and Bradstreet number (D&B number) for most of the plants in the estimation sample. I identify these plants as linked to the same firm if they are associated with the same parent company name or Dun and Bradstreet number.[29] 80% of the plants with parent company information are linked to a firm that has more than one plant reporting to the TRI. I use these plants to estimate a variant of Equation (4), which includes firm-by-year fixed-effects (μ_{ft}):

$$Y_{pcit} = \pi_0 + \pi_1 NewspapersDensity + \pi_2 X_{pt} + \mu_{ft} + \eta_{i4} + \epsilon_{pcit} \tag{10}$$

---

[29]The match is mostly based on the parent company name, because, in an audit study conducted by D&B for EPA, the submitted TRI D&B numbers matched the D&B U.S Domestic Ultimate Parent Number only in 31% of cases (information provided by EPA employee through email communication). Matching on the parent company name has also some issues, because several variants of a name can be used (for instance, ARCELORMITTAL INC and ARCELORMITTAL USA INC); this implies that the results in this section, although suggestive, should be interpreted with some caution.

Table 5, Column (3) shows the point estimate for π_1 in Equation (10). This is the effect of newspapers density on toxic emissions, when comparing plants operated by the same firm in the same 4-digit industry and year. The coefficient is negative, although the standard errors are larger than in the specification with county-by-year effects, due to the smaller sample size. In Column (2), I exclude plants linked to firms that only count two observations in a given year. The effect of *Newspapers Density* for the remaining sample is larger, very similar to that shown in Table 4, Column (4), and marginally significant. The point estimate for π_1 does, in fact, increase as the sample is restricted to larger and larger firms (result not shown).[30] This suggests that the effect of *Newspapers Density* on toxic emissions is increasing with the size of the parent company. Different hypotheses are consistent with this finding: larger firms are more exposed to reputation-related concerns, they are more resourceful and thus more capable of investing in corporate environmentalism, and could reduce emissions at a lower cost, due to economies of scale in the adoption of new technologies. Overall, these results indicate that *Newspapers Density* explains variation in emissions also when we compare plants operated by the same firm, thus addressing the concern that the effect estimated in section 4.2 is driven by different location decisions across firms.[31]

4.3.4 Additional Robustness Checks

I show that the estimates presented in Table 4 are robust to certain departures from the main sample and estimation method.

Table 6 Column (1) shows that the estimated effect of *Newspapers Density* on toxic emissions is not driven by the higher concentration of newspapers in big cities compared to other areas in the same county. When I exclude the twenty-five most populated cities in the sample, the coefficient is almost the same as that estimated on the main estimation sample.

Another concern is that the bulk of the effect comes from the comparison between plants that are not surrounded by any newspapers and those that are surrounded by some newspapers. The conceptual framework and the analysis in Section 4.1 show that in areas with at least one newspaper, coverage is increasing in the number of newspapers and in their respective proximity to the plant. If newspaper density affects emissions through the threat of coverage, we should then expect this to be true also when comparing plants that have at least one newspaper in their 20-mile rings. This is indeed shown in Column (2) of Table 6.

In estimating equation (8), I use as the dependent variable the natural logarithm of toxic emissions because the distribution of emissions is highly skewed to the right. This has the drawback that I have to add a constant equal to 1 to the original value of emissions due to the presence of plants that have emissions equal to 0. In Column (3) of Table 6, I show that *Newspapers Density* is negatively associated with *log Emissions* when I exclude the observations that have zero emissions and therefore do not have to perform any transformation on the dependent variable. According to

---

[30]I assume that firms that list more plants in the TRI dataset are larger.

[31]The low precision of the estimates is not surprising, given the noise in the data created by the difficulty in identifying plants linked to the same firm using the parent company name.

the point estimate in Column (3), a one unit increase in *Newspapers Density* would reduce emissions by 6%.[32]

As pointed out in Section 3, when the quantity of a toxic substance released is lower than 500 pounds, firms can choose not to report the quantity released and instead submit a form that certifies that they manufacture some of the listed toxic substances; for these cases, emissions are set to zero. I choose to treat these values as zeros in the main specification because I study the effect of *Newspaper Density* on reported toxic emissions, as they are presented to the public; emissions being a continuous variable, I run a standard OLS regression. However, I also estimate a model that takes into account the fact that whenever the plant reported what is called the "Form A," i.e. a form that only certifies that emissions of a particular substance are below 500 pounds, the quantity actually released could be any value in the interval [0,499]. I thus run an interval regression, using the lower and upper bounds implied by this reporting technique.

I define the true value of emissions as y_{pt}^*. The observed value is:

$y_{pt} = 0$ if $0 \leq y_{pt}^* < k$, $y_{pt} = y_{pt}^*$ if $y_{pt}^* \geq k$

Assuming that $y^*|x \sim \mathcal{N}(x\beta, \sigma^2)$, where $\sigma^2 = Var(y^*|x)$ is assumed to be independent of x, the parameters β and σ can be estimated by maximum likelihood, defining the likelihood as follows (Wooldridge, 1995):

$$\log\left[L(\beta, \sigma)\right] = \sum_i \left(1\left[y_i \geq k\right]\right) log \left\{\sigma^{-1}\phi\left[\frac{(Y_i - X_i\gamma)}{\sigma}\right]\right\} +$$
$$\left(1\left[0 \geq y_i > k\right]\right) log \left\{\Phi\left[\frac{(k - X_i\gamma)}{\sigma}\right] - \Phi\left[\frac{-X_i\gamma}{\sigma}\right]\right\}$$

Due to computational limitations, I estimate a model with county and year fixed effects rather than with county-by-year effects, and 3-digit industry effects rather than 4-digit industry effects. The estimated coefficient is reported in Table 6, Column (4). The point estimate of β_1 is very close to that obtained with the OLS regression.

Given that unemployment is not necessarily persistent over time and that substantial economic shocks could quickly change the economic landscape of a location, the unemployment level in 2000 might not be a good measure for unemployment in recent years. Therefore, I first run the regression estimated in Table 4, Column (4) without including the control for *Unemployment 2000* and limiting the sample to the years 2001-2004. The point estimate for β_1 in equation (8) is -0.084, and it is significant at the 5% level. Then, I run this regression on the same sample, adding a control for unemployment in 2000. As shown in Column (5), the coefficient is stable to the introduction of this further control, even in this sub-sample for which *Unemployment 2000* should be a good measure of the actual level of unemployment.

---

[32]The effect estimated in this sample is, not surprisingly, smaller than that estimated on the full sample: the effect of *Newspapers Density* on the decision to not emit toxic substances is not captured in this specification.

Finally, in Column (6) I drop from the sample all plants that are located at a linear distance of five or less miles from the border with Canada or Mexico or from the coastline. The independent variables are measured with error for these plants because Census-block-group data are not available for a portion of their ring. If this error is correlated with *Newspapers Density*, the estimated effect of *Newspapers Density* will be downward biased. In Column (6) I show that, excluding these plants, the effect of *Newspapers Density* is larger, but broadly similar to that estimated in the full sample.

4.4 Mechanism

To shed light on the mechanism behind the effect of *Newspapers Density* on toxic emissions, I now analyze heterogeneous effects. I first investigate whether *Newspapers Density* affects more emissions of substances that have strong documented negative effects on humans, which should be the case if the effects go through some public reaction to negative health outcomes (rather than say a correlation between unobserved plant type and *Newspapers Density*). Next, I analyze whether emissions from firms producing final goods respond more to higher *Newspapers Density*, consistently with the model of response to potential consumer reactions. I finally check whether plants in communities that have a prior history of negative health outcomes, and for this reason are probably more aware of the problems associated with pollution, react more to higher *Newspapers Density*; this should be the case if the effects go through health concerns and are driven by locals.

4.4.1 Substances with strong health effects

In this section, I focus on *Newspapers Density* and emissions of two substances, dioxin (and dioxin-like compounds) and lead. The TRI is a register of emissions of 650 toxic substances that have different levels of danger for human health and for the environment. Among these substances, EPA pays special attention to the so-called Persistent and Bioaccumulative Substances (PBTs) because they tend to persist in the environment and to bioaccumulate into food chains. Dioxin and lead belong to this category. The reason for focusing on these substances is that their negative effects on humans and the environment have been extensively analyzed, so that it is not implausible that journalists and/or readers are more aware of the danger they pose.[33]

I analyze how coverage and emissions react to *Newspapers Density*, controlling for total plant-level emissions, by estimating the following regressions:

$$C_{psit} = \theta_0 + \theta_1 \text{PBT}_{p,t-2} + \theta_2 emissions_{p,t-2} + \theta_3 Newspapers Density_{pt} + \theta_4 X_{pt} + \mu_{st} + \eta_{i2} + \epsilon_{psit}, \ (11)$$

---

[33]In the epidemiological literature, many studies on the health effects of dioxin exploit the "natural experiment" created by the Seveso disaster in Italy in 1976. The most recent of these studies (Pesatori et al., 2009) confirmed an excess risk of lymphatic and hematopoietic tissue neoplasms in the most exposed zones, and uncovered a high risk of breast cancer in women contaminated 15 years since the accident. Previous studies established a relationship between exposure to dioxin and, among others, chloracne (Caramaschi et al., 1981), diabetes (Pesatori et al., 1998) and sex ratio (Mocarelli et al., 2000). In the economics literature, Nilsson (2009) shows that early exposure to lead impairs scholastic performance, cognitive abilities and labor market outcomes.

$$\text{PBT}_{pcit} = \vartheta_0 + \vartheta_1 NewspapersDensity_{pt} + \vartheta_2 emissions_{pt} + \vartheta_3 X_{pt} + \mu_{ct} + \eta_{i4} + \epsilon_{pcit}, \quad (12)$$

where PBT is an index for emissions of one of the PBT's substances on which I focus, and $emissions$ is the total level of emissions reported by the plant.[34] [35] If the correlation between emissions and *Newspapers Density* arises because of newspaper coverage, a high ϑ_1 estimate should be associated with a high θ_3 estimate.

As shown in Columns (1) and (2) of Table 7, higher emissions of dioxin trigger coverage, whereas higher emissions of lead do not, controlling for total plant-level emissions. If firms are worried about negative coverage when located in areas with a high concentration of newspapers, the level of dioxin that they emit should then respond to *Newspapers Density* more than the level of lead.

As shown in Table 7, Columns (3) and (4), emissions of dioxin are higher in areas with less *Newspapers Density*. This negative association does not hold when the dependent variable measures emissions of lead. Although silent on the reasons for which dioxin triggers coverage in newspapers while lead does not, this test lends further support to the thesis that the association between the density of newspapers and toxic emissions is caused by actual and potential coverage of pollution statistics in nearby newspapers.

4.4.2 Consumer Pressure

In the model, when faced with a high concentration of newspapers at their plants' locations, firms have incentives to adopt good environmental practices because they want to avoid drops in demand due to bad press. Gupta and Innes (2011) show that high emitters of toxic substances are more likely to be targeted for a boycott, and that firms are more likely to adopt an environmental management system and an environmental protocol after being subjected to a boycott. Innes and Sam (2008) show that the threat of boycott is positively associated with corporate environmentalism.

Firms in certain industries are more exposed to boycotts than others. Gupta and Innes (2011), for instance, conjecture that firms that sell final goods are more likely to be targeted for a boycott.[36] If the characteristics of the newspaper market play a substantial role in giving (or threatening to give) relevant information to organized groups of consumers, then the effect of *Newspapers Density* should be larger for firms that sell final goods. In Table 13 I test for this implication by estimating

---

[34]The index takes the value of 0 if the plant-level emissions of the PBT substance are 0, and a number from 1 to 4 depending on to which quartile of the positive part of the sample distribution that the plant's emissions belong. The choice of the index is motivated by the fact that the percentage of plants having non-zero emissions of dioxin and lead is very low (5% and 12%, respectively, in the full sample). The categoric variable is used instead of a dummy for emissions above zero to exploit the information from the few plants that emit relatively high quantities of these substances; the results hold up when using a dummy as dependent variable, however.

[35]The second lag of emissions is relevant for coverage at time t, because, as already explained, data on emissions are published with a two-year lag.

[36]Although in their data they only find a positive but not significant effect of selling final goods on the probability that a firm is the target of a boycott, Gupta and Innes (2011) do not interpret the result as evidence that selling final goods has no effect on the probability of boycott; they rather think that the insignificant effect is due to the introduction of industry controls, that remove much of the available variation, and to the limitations of their main measure of environmental constituency.

the following equation:

$$Y_{pt} = \lambda_0 + \lambda_1 NewspapersDensity_{pt} + \lambda_2 Final_{i3} + \lambda_3 X_{pt} + \lambda_4 NewspapersDensity_{pt} * Final_{i3}$$
$$+ \mu_{ct} + \eta_{i4} + \epsilon_{pt}$$

$$(13)$$

In practice, I test whether the effect of *Newspapers Density* estimated in Table 4 is larger for plants in industry-subsectors that mainly sell final versus intermediate goods.

I construct the variable *Final* as follows: I look at the 6-digit Primary Naics included in each industry-subsector and classify a subsector as mainly selling final goods if most of its 6-digit Primary Naics sell final goods. Table 8 lists the subsectors classified as selling final goods and those classified as selling intermediate goods.[37] To make some examples of firms in this category, the industry "Utilities" lists companies as American Electric Power and Edison, "Food Manufacturing" lists companies as Kellogg and Kraft, "Beverage and Tobacco Product Manufacturing" includes Coca-Cola and Pepsi. These are all companies with high proximity to consumers.

In Column (1) of Table 9 I show that the effect of *Newspapers Density* on emissions is significantly larger in subsectors that mostly sell final goods.[38] In fact, most of the effect of *Newspapers-Density* shown in Table 4, Column (4), seems to be estimated in these subsectors. This result is robust to excluding subsectors whose composition of Primary Naics in terms of type of good sold is very heterogeneous (result not shown but available upon request).[39]

4.4.3 Local Health Salience

I now test whether firms react more to *Newspapers Density* in places that are more vulnerable to adverse health shocks. The conceptual framework is based on the assumption that the constituents observe H, i.e. the aggregate environmental and health damage at their place of residence, but they do not know the contribution of each local plant to h. Press coverage of pollution can fill this informational asymmetry. Therefore, the larger is H the more likely it is that constituents both ask for information about toxic emissions and respond to coverage taking action against the firm. In both cases, if the firm also knows H, the presence of newspapers nearby the plant creates larger incentives to cut emissions when the health and environmental outcomes in the area around the plant are more negative.

Given the unavailability of data on local health outcomes, I use county-level data, provided by the *US Department of Health and Human Services* in the *Area Resource Files*. I test the hypothesis

---

[37]I excluded plants in the "Public Administration" sector.

[38]Note that the main effect of producing final goods is absorbed by the 4-digit industry fixed effects.

[39]Note that the effect of *Newspapers Density* on toxic emissions could be larger in subsectors selling final goods for a reason different than consumer pressure; namely, plants that are "compliers", i.e. that change their emissions because of *Newspapers Density*, are mostly in subsectors selling final goods. The "compliers" are most likely plants in industries that are traditionally highly polluting. However, while mean emission are slightly larger in the category "final goods", the standard deviation and the 99th percentile of the distribution of emissions is larger in the sample of plants selling intermediate goods; in fact, log emissions are larger in the sample of plants selling final goods. This suggests that there are also "compliers" in this sample.

that the effect of *Newspapers Density* is larger in counties where there are extreme negative health outcomes by estimating the following equation:

$$Y_{pcit} = \delta_0 + \delta_1 NewspapersDensity_{pt} + \delta_2 X_{pt} + \delta_3 NewspapersDensity_{pt} * \widetilde{H}_c + \mu_{ct} + \eta_{i3} + \epsilon_{pcit} \quad (14)$$

$\widetilde{H}_c$ is a dummy for extreme negative health outcome, equal to 1 if the level of infant mortality and or mortality from cancer is larger than the 99th percentile in the respective cross-county distributions.[40] I choose to measure the extreme negative health outcome with this dummy because I intend to capture a situation of "alarm" that is most likely perceived by people living in the county. Given this definition, relatively few plants are located in counties that are classified as having experienced extreme health outcomes (see Figure 8). For this reason, I control for industry-subsector, rather than industry-group, fixed effects, to increase the total amount of variation used to estimate δ_3.

The estimates of δ_3 in Table 9 show a pattern that is consistent across the two different measures of $\widetilde{H}_c$. As hypothesized, the effect of *Newspapers Density* on toxic emissions is larger in counties that have experienced some extreme negative health outcomes in the recent past. In light of the theoretical and empirical results presented in this paper, this suggests that polluting firms may face a higher risk from proximity to newspapers where there is a "health alarm", because the "threat of coverage" may be larger and also, holding the "threat of coverage" constant, because the constituents might be more likely to take initiatives against firms spotted as big polluters.

5   Spillovers and Aggregate Effects

The aggregate effect of newspapers on emissions depends crucially on the degree to which multi-plant firms move emissions from one plant to another, in response to high *Newspapers Density*. Emissions at different plants are different inputs, whose cost depends on the level of *Newspapers-Density* of each plant, and the amount of input needed to produce a certain level of output varies with the technology used (clean or dirty). Depending on the degree of substitutability between the two inputs, the firm might find it convenient to shift part of the production from a place where the input is more costly (i.e. with large *Newspapers Density*) to one where it is less costly (i.e. with low *Newspapers Density*), or rather to change the technology used where emissions are more costly.

I check for the existence of substitution patterns by regressing emissions on *Newspaper Density* and the aggregate *Newspapers Density* of all other plants belonging to the same firm and 4-digit

---

[40]*Infant Mortality* is the number of deaths of infants aged below one per 1,000 live births; *Mortality from Cancer* is the share of deaths caused by cancer. Figure 8 in Appendix B shows the distributions of *Infant Mortality* and *Mortality by Cancer* in the estimation sample.

industry:

$$Y_{pcit} = \sigma_0 + \sigma_1 NewspapersDensity_{pt} + \sigma_2 X_{pt} + \sigma_3 NewspapersDensityFirm_{pt} +$$
$$\mu_{ct} + \eta_{i4} + \epsilon_{pt} \tag{15}$$

where $NewspapersDensityFirm_{pt}$ is defined as follows. Let me call a *group* the set of plants linked to the same firm and that operate in the same 4-digit industry. Indexing by p' any plant in the same *group* as plant p, and that is different than p, I define:

$$NewspapersDensityFirm_{pt} = \sum_{p' \neq p} NewspapersDensity_{p't}$$

This is a measure of the cost of emissions at the other plants in the same firm and industry-group as plant p. Thus, I implicitly assume that firms have an opportunity to substitute emissions if they have more than one plant in the same industry-group. If a firm manages plants operating in different industry-groups, then the mechanism of substitution described above is unlikely to hold, because it is unlikely that the firm treats emissions in different industry-groups as substitutes. In estimating Equation 15, I limit the sample to plants belonging to a *group* that has at least two elements. If a firm substitutes emissions across plants, depending on the relative levels of *Newspapers Density*, then σ_3 should be positive.

Moreover, if a firm shifts a part of its emissions to the plant with the minimum relative cost of pollution, i.e. with the minimum realization of *Newspapers Density*, then emissions should be increasing in a dummy (*Min Density*) that takes the value of one if plant p has the lowest level of *Newspapers Density* across the other plants in its *group*. Having the minimum value of *Newspapers Density* is correlated with the number of plants in the same *group*, which might also be correlated with emissions, and therefore this variable needs to be controlled for. I test these hypotheses as follows:

$$Y_{pcit} = \rho_0 + \rho_1 NewspapersDensity_{pt} + \rho_2 X_{pt} + \rho_3 MinDensity_{pt} + \rho_4 Nplants_{fit} +$$
$$\mu_{ct} + \eta_{i4} + \epsilon_{pcit} \tag{16}$$

where

$$MinDensity_{pt} = 1 \quad if \quad NewspapersDensity_{pt} = min(NewspapersDensity_{pfit})$$

and *N plants* is the number of plants in the same *group* as plant p. The average value of $MinDensity_{pt}$ in the sample is 0.12.

In equations (15) and (16), I exploit variation across plants in the same county and industry-

group. Thus, I answer the following question: when we compare plants in the same county, industry-group and year, and once *Newspapers Density* and other characteristics of the area where the plant is located are accounted for, does the relative level of *Newspapers Density* at plant p, with respect to other plants in the same *group*, explain its toxic emissions?

Table 10, Columns (1) and (2), provide no evidence for the existence of substitution patterns within firms. The estimates of σ_3 and ρ_3 are highly imprecise and negative. The coefficient of *Newspapers Density* is the same as in previous specifications, although not precisely estimated due to the smaller sample size. The lack of within-firm substitution has important implications for computing the aggregate effect of newspapers on total emissions. In order to compute these aggregate effects, I conduct a counterfactual analysis, based on that in Yanagizawa-Drott (2010).

I use different estimates from the previous sections: the effect of *Newspapers Density* on toxic emissions estimated in equation (4), the effect of *Newspapers Density* on toxic emissions for plants in industries that sell final goods estimated in equation (13), and the effect of *Newspapers Density* on toxic emissions in counties that have experienced extreme negative health outcomes in the recent past estimated in equation (14). As in Yanagizawa-Drott (2010), I define the notion of counterfactual outcome y_c as the outcome that we would observe had *Newspapers Density* been zero for any observation in the sample. For example, this means that in equation (4),

$$y_c = exp(ln(y_{pt} + 1) - \beta_1 Newspapers Density_{pt}) - 1 \tag{17}$$

To account for the variance in the estimated parameters, I first generate a normal random variable for each coefficient of interest, with a mean equal to the point estimate and a standard deviation equal to the standard error. [41] Then, I extract a value from these distributions for each coefficient of interest and I calculate the y_c's. Summing over all the y_c's in the sample of interest I get Y_c, which is the aggregate counterfactual. This is a measure of what the total level of toxic emissions would be in the sample of interest had *Newspapers Density* been zero for any observation. For each coefficient of interest, I repeat this procedure 500 times, deriving a mean and a standard deviation for Y_c. Comparing the mean of Y_c and the true aggregate emissions Y in the sample of interest, we can interpret the estimates in this paper in aggregate terms.

The results of this exercise are reported in Table 11. In the full sample, the total lower emissions accounted for by *Newspapers Density* are 3% of the actual emissions. Once we focus on samples where we expect the effect to bite the most, newspapers seem to have an economically important impact on the environmental outcome studied. Without any newspapers in the 20-mile radius around the plants, toxic emissions from plants that produce final goods would have been 11% larger, toxic emissions in counties with extreme infant mortality in the recent past would have been 17% larger, and toxic emissions in counties with extreme mortality from cancer would have been 22% larger.

---

[41]The point estimates in the interaction models are given by the linear combination of the point estimate for *Newspapers Density* and that for the interaction term; the standard errors are calculated using the estimated variance and covariance matrix for the vector of estimates in the respective equation.

6 Conclusion and Policy Implications

This paper assesses how newspapers induce firms to reduce toxic emissions at their plants. I show that the probability that a newspaper covers a plant's toxic emissions falls approximately inversely to its distance from the plant. Specifically, using newly-collected data on the coverage of toxic emissions in US newspapers, I show that the higher is a plant's *Newspapers Density*, the more likely it is that some articles about its toxic emissions are written in nearby newspapers. A simple model predicts that plants located in areas with a higher *Newspapers Density* would pollute less. The empirical evidence in this paper is consistent with this prediction: if a plant moved from the 5th percentile of *Newspapers Density* to the 95th percentile, its toxic emissions would be 15% lower.

I show that this association is unlikely to be explained by selection on unobservables, and that it is only valid in states where some coverage of toxic emissions has been observed during the sample period. Moreover, this effect is larger in industries mainly selling final goods, revealing an interaction of newspaper coverage and consumer pressure in shaping corporate environmental decisions. The effect is also larger in counties that have recently experienced extreme negative health outcomes, suggesting that the probability that citizens take action against firms shamed as polluters is higher under such circumstances. A counterfactual exercise suggests that, if there were no newspapers within a 20-mile radius from the plants that report emissions to the TRI, these emissions would be 3% larger. For plants in industry-sectors that mainly sell final goods, emissions in the counterfactual scenario would be 11% larger, a number that doubles when looking at plants in counties exposed to an extreme incidence of mortality from cancer.

The substantial effect of the presence of newspapers on reported emissions has several policy implications. First, it gives an important insight into the effectiveness of the principle of "Regulation through Revelation " (Hamilton, 2005). Pollutant Release and Transfer Registers are currently used in several countries and regions as an environmental policy tool aimed at reducing pollution. The Kiev Protocol on Pollutant Release and Transfer Registers states that "although regulating information on pollution, rather than pollution directly, the Protocol is expected to exert a significant downward pressure on levels of pollution, as no company will want to be identified as among the biggest polluters (UNECE, 2003)."[42] The findings in this paper show that the availability of information is not a sufficient condition to make people informed. The lower levels of emissions reported by plants located nearby newspapers, paired with the documented effect of *Newspapers Density* on the probability of receiving bad press, reveals that people are more likely to be informed when there are some *intermediaries* that *lower the costs to the public of public information* (Hamilton, 2005). Absent these intermediaries, information provision programs might not reap their potential benefits, *de facto* depriving public policy of a potentially successful policy tool. While the widespread adoption of Pollutant Release and Transfer Registers addresses the problem of informational asymmetries as a possible determinant of polluting behavior on the firms'side, it

---

[42]The Protocol was adopted on May 2003, following the 1998 *Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters*, and establishes the first Pollutant Release and Transfer Register with international scope

does not provide a sufficient regulatory framework to overcome these asymmetries. A more appropriate approach would require a focus on the incentives for information providers to disseminate the information in the Registers.

More generally, these findings provide further evidence, beyond that summarized in Prat and Strömberg (2011), on the large set of outcomes that mass-media can affect. Some of the pollutants reported to the TRI are carcinogens that cause an excess risk of mortality from cancer if released in large quantity. Moreover, Currie and Schmieder (2009) show that specific chemicals in the TRI impair birth outcomes, and Agarwal et al. (2010) estimate a negative effect of TRI releases on infant mortality.[43] These facts, and the magnitude of the aggregate effects for some categories of plants, as presented in Table 11, suggest that it is not implausible that newspapers ultimately affect human health. In the words of Hamilton (2005), "if the scrutiny generated by the TRI results in fewer emissions of air carcinogens or lower releases of toxics into waterways, the TRI can (for a set of people who may never realize it) improve human health." The policy implications following from this are tied to the other finding in this paper, that geography creates incentives for newspapers to write about a plant. This finding is similar in spirit to that in Snyder Jr and Strömberg (2010), that congruence between newspapers readers and a congressman's constituents increases the coverage given to the congressman. In the case of physical plants, the constituents are identified as the people who live nearby the plant. As in Snyder Jr and Strömberg (2010), the relationship between the geography of the newspaper market and coverage has policy implications in terms of press regulation.

Given the importance of citizens' information for firms' environmental performance, it is crucial to understand the implications of the spread of Internet for environmental outcomes. As observed in Snyder Jr and Strömberg (2010), the potential Internet audience being global, the incentives to cover local plants might be low especially when it comes to small firms that produce at few localities relative to global firms. Given the potential high costs of bad reputation for firms coming from consumer boycotts, loss of brand value, risk of litigation, employee productivity and cost of capital (Heal, 2008), this could make the incentives of big corporations increasingly aligned to those of their constituents whereas local firms would not internalize the external costs of their decisions. As suggested in Snyder Jr and Strömberg (2010), the same reasoning also applies when it comes to the trend in newspaper concentration, and in broadcast media substituting newspapers. The findings in this paper suggest that these trends could decrease residents' information about local plants' environmental performance, ultimately removing the incentives for many firms to control toxic emissions that are detrimental for human health.

The results in this paper point out several directions for further research. The coefficient on *Newspapers Density* is very stable to the introduction of controls for the main determinants of the presence of newspapers at a locality. The different tests presented in this paper are consistent with a causal interpretation of the point estimate of β_1. However, in the absence of experimental

---

[43]They estimate that the average county-level decrease in the concentration of different TRI toxics observed between 1998 and 2002 saved 13,800 infant lives.

or quasi-experimental variation, we cannot rule out the possibility that the results are driven by selection on unobservables. While it is difficult to instrument for the presence of newspapers, the selection problem could be addressed by exploiting time variation. In particular, did the passage of the *Community Right to Know Act* give incentives for plants that were more exposed to the monitoring of newspapers to further reduce emissions? While data on emissions are only available since the passage of the law, one could look at other outcomes, such as health outcomes linked to TRI chemicals, ambient concentration data, and location decisions of plants in traditionally "dirty" industries.

I currently impose a radius of 20 miles for the rings in which I measure *Newspapers Density*, as coverage in rings tends to go to zero as the radius approaches 20 miles. However, I plan to use concentric rings (Rosenthal and Strange, 2008) or maximum likelihood estimation to estimate the optimal radius of the rings. This extension would shed some light on the relevant geographical dimension within which newspapers can work as "watchdogs" for polluting firms. It may also address the identification issue if the optimal radius of the ring creates a discontinuity for the relationship between *Newspapers Density* and toxic emissions.

I currently pursue these new directions.

References

Agarwal, N., C. Banternghansa, and L. Bui (2010). Toxic exposure in america: Estimating fetal and infant health outcomes from 14 years of tri reporting. *Journal of health economics 29*(4).

Aghion, P., T. Persson, and D. Rouzet (2012). Education and military rivalry. *NBER Working Paper*.

Altonji, J., T. Elder, and C. Taber (2005). Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools. *Journal of Political Economy 113*(1).

Arora, S. and T. Cason (1995). An experiment in voluntary environmental regulation: Participation in epa' s 33/50 program. *Journal of environmental economics and management 28*(3).

Baron, D. (2001). Private politics, corporate social responsibility, and integrated strategy. *Journal of Economics & Management Strategy 10*(1).

Baron, D. (2003). Private politics. *Journal of Economics & Management Strategy 12*(1).

Bellows, J. and E. Miguel (2009). War and local collective action in sierra leone. *Journal of Public Economics 93*(11).

Besley, T. and A. Prat (2006). Handcuffs for the grabbing hand? the role of the media in political accountability. *American Economic Review 96*(3).

Bimonte, S. (2002). Information access, income distribution, and the environmental kuznets curve. *Ecological economics 41*(1).

Bui, L. and C. Mayer (2003). Regulation and capitalization of environmental amenities: Evidence from the toxic release inventory in massachusetts. *Review of Economics and statistics 85*(3).

Bullard, R. (1990). *Dumping in Dixie: Race, class, and environmental quality*, Volume 3. Westview Press Boulder, CO.

Caramaschi, F., G. Del Corno, C. Favaretti, S. Giambelluca, E. Montesarchio, and G. Fara (1981). Chloracne following environmental contamination by tcdd in seveso, italy. *International journal of epidemiology 10*(2).

Chay, K. and M. Greenstone (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).

Currie, J. and J. Schmieder (2009). Fetal exposures to toxic releases and infant health. *The American Economic Review*.

Dyck, A., N. Volchkova, and L. Zingales (2008). The corporate governance role of the media: Evidence from russia. *The Journal of Finance 63*(3).

Dyck, A. and L. Zingales (2002). The corporate governance role of the media. *The Right to Tell: The Role of Mass Media in Development*.

Ellman, M. and F. Germano (2009). What do the papers sell? a model of advertising and media bias*. *The Economic Journal 119*(537).

Fleishman-Hillard and the National Consumers League (2007). Rethinking corporate social responsibility. `http://fleishmanhillard.com/wp-content/uploads/2007/05/csr_white_paper.pdf`.

Gambaro, M. and R. Puglisi (2010). What do ads buy? Daily coverage of listed companies on the italian press. *RSCAS Working Papers 2010/26, European University Institute*.

Gentzkow, M., J. Shapiro, and M. Sinkinson (2010). The effect of newspaper entry and exit on electoral politics. *American Economic Review 101*(7).

Guerrero, S. (2012). Who is selling you chiquilitros of gasoline? evidence from a public disclosure policy. *Journal of Public Economis, forthcoming*.

Gupta, S. and R. Innes (2011). Private politics and environmental management. *Working Paper, U. of Florida and UC, Merced*.

Hamilton, J. (1995). Pollution as news: Media and stock market reactions to the toxics release inventory data. *Journal of environmental economics and management 28*(1).

Hamilton, J. (1999). Exercising property rights to pollute: Do cancer risks and politics affect plant emission reductions? *Journal of Risk and Uncertainty 18*(2).

Hamilton, J. (2005). *Regulation through revelation: the origin, politics, and impacts of the Toxics Release Inventory Program*. Cambridge University Press.

Harrington, D., M. Khanna, and G. Deltas (2008). Striving to be green: the adoption of total quality environmental management. *Applied economics 40*(23).

Heal, G. (2008). *When principles pay: corporate social responsibility and the bottom line*. Columbia Business School Publishing.

Innes, R. and A. Sam (2008). Voluntary pollution reductions and the enforcement of environmental law: An empirical study of the 33/50 program. *Journal of Law and Economics 51*(2).

Khanna, M. and L. Damon (1999). Epa's voluntary 33/50 program: Impact on toxic releases and economic performance of firms. *Journal of Environmental Economics and Management 37*(1).

Kitzmueller, M. and J. Shimshack (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature 50*(1).

Konar, S. and M. Cohen (1997). Information as regulation: The effect of community right to know laws on toxic emissions. *Journal of environmental Economics and Management 32*(1).

Maxwell, J., T. Lyon, and S. Hackett (2000). Self-regulation and social welfare: The political economy of corporate environmentalism. *Journal of Law and Economics 43*(2).

Mocarelli, P., P. Gerthoux, E. Ferrari, D. Patterson, S. Kieszak, P. Brambilla, N. Vincoli, S. Signorini, P. Tramacere, V. Carreri, et al. (2000). Paternal concentrations of dioxin and sex ratio of offspring. *The Lancet 355*(9218).

Nilsson, J. (2009). The long-term effects of early childhood lead exposure: Evidence from the phase-out of leaded gasoline. *Working Paper, Standford University*.

Pesatori, A., D. Consonni, M. Rubagotti, P. Grillo, P. Bertazzi, et al. (2009). Cancer incidence in the population exposed to dioxin after the Seveso accident: twenty years of follow-up. *Environ Health 8*.

Pesatori, A., C. Zocchetti, S. Guercilena, D. Consonni, D. Turrini, and P. Bertazzi (1998). Dioxin exposure and non-malignant health effects: a mortality study. *Occupational and Environmental Medicine 55*(2).

Prat, A. and D. Strömberg (2011). The political economy of mass media. *CEPR Discussion Paper No. DP8246*.

Rosenthal, S. and W. Strange (2008). The attenuation of human capital spillovers. *Journal of Urban Economics 64*(2).

Sanders, N. (2012). What doesn't kill you makes you weaker prenatal pollution exposure and educational outcomes. *Journal of Human Resources 47*(3).

Snyder Jr, J. and D. Strömberg (2010). Press coverage and political accountability. *Journal of Political Economy 118*(2).

UNECE (2003). Kiev protocol on pollutant release and transfer registers. `http://www.unece.org/env/pp/prtr.html`.

Wooldridge, J. (1995). *Econometric Analysis of Cross Section and Panel Data*, Volume 5. The MIT Press.

Yanagizawa-Drott, D. (2010). Propaganda and conflict: Theory and evidence from the rwandan genocide. *Working Paper, Harvard University*.

Appendix A: Proofs of Results in Section 2

6.1 Proof of result in Equation 1

Trivially, the utility of the consumer is maximized at $y_l = W$ when $e = e_c$, because the consumer bears no cost from industrial production.

When $e = e_d$, the utility is maximized at $\underline{y}_l$ s.t.:

$$V'(\underline{y}_l) = \overline{h}'(\underline{y}_l)$$

$y_l^* < W$ from Assumption 1

6.2 Proof of results in equations 3 and 4

Define γ as $Pr[(\overline{y} - \underline{y}) \sum_{i=1}^{N} p(d_i, \alpha) \geq c_c - c_d]$. The beliefs are formed following Bayes' rule as follows:

$$Pr(e_c | s = \oslash) = \frac{Pr(s = \oslash | e_c) Pr(e_c)}{Pr(s = \oslash)}$$
$$= \frac{1 * \gamma}{\gamma + (1 - \gamma)(1 - P)} \approx \gamma$$

where the approximation comes from the fact that, when a plant chooses $e = e_d$, it is because P is very small.

and

$$Pr(e_c | s = s_d) = \frac{Pr(s = s_d | e_c) Pr(e_c)}{Pr(s = s_d)}$$
$$= \frac{0 * \gamma}{(1 - \gamma)(P)} = 0$$

Given these beliefs, the constituents choose:

$$\underline{y}_l = argmax[V(y_l) - \overline{h}(y_l)] \quad \text{if } s = s_d \tag{18}$$

and

$$\overline{y}_l = argmax[V(y_l) - (1 - \gamma)\overline{h}(y_l)] \quad \text{if } s = \oslash \tag{19}$$

I prove that $\overline{y}_l > \underline{y}_l$ by Contradiction.

Suppose that $\overline{y}_l \leq \underline{y}_l$. Then $V'(\overline{y}_l) \geq V'(\underline{y}_l) = \overline{h'}(\underline{y}_l) \geq \overline{h'}(\overline{y}_l)$.

Therefore,

$$V'(\overline{y}_l) \geq \overline{h'}(\overline{y}_l)$$

but also, from FOC of the optimization problem in 19,

$$V'(y_l) = (1 - \gamma)\overline{h'}(y_l) \qquad \implies V'(y_l) > \overline{h'}(y_l) \quad \text{for } \gamma > 0$$

Contradiction.

6.3 Proof of result in equation 5

Given the consumer demand strategy, the manager chooses e_c if:

$$\overline{y}_l - e_c \geq (1 - P)(\overline{y}_l - e_d) - P(\underline{y}_l - e_d)$$

Rearranging:

$$(\overline{y}_l - \underline{y}_l)\sum_{i=1}^{N} p(d_i, \alpha_i) \geq e_c - e_d.$$

Appendix B: Tables and Figures

Figure 1
Distance from plant and probability of coverage in newspaper
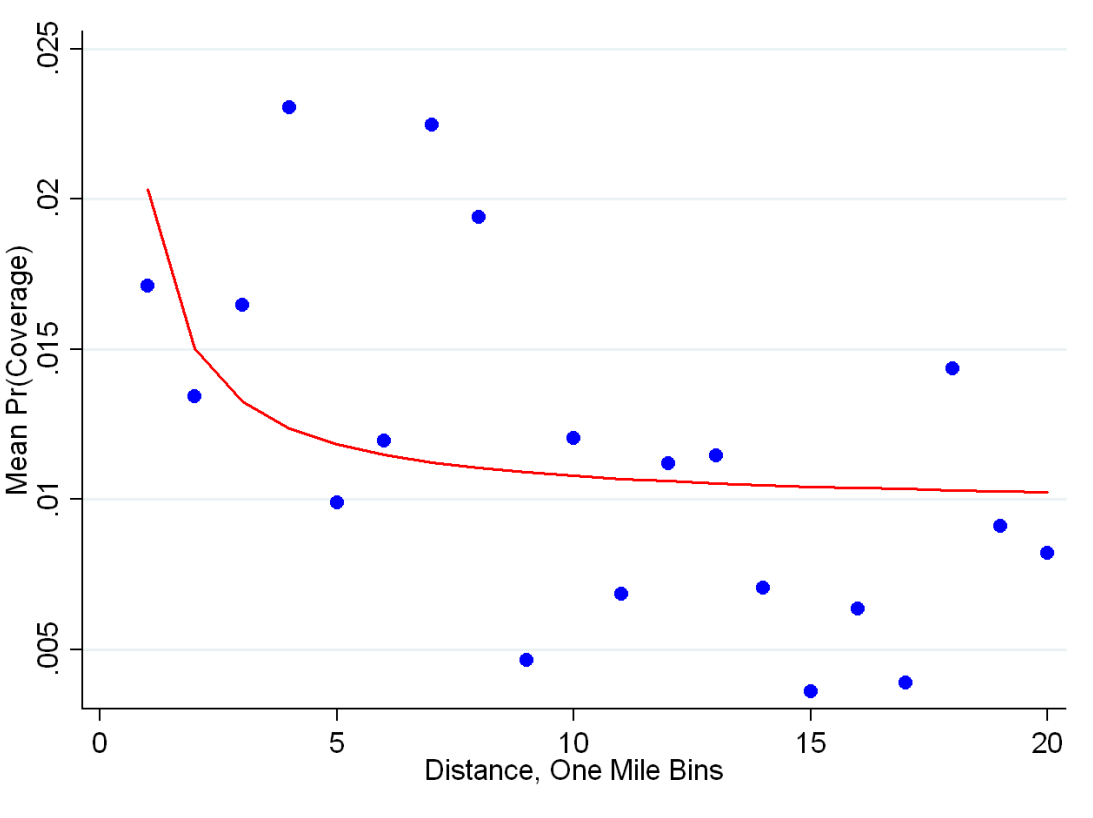
Figure 2

An example of article search output in Newslibrary

Figure 3
Articles Published by State, 2000-2008
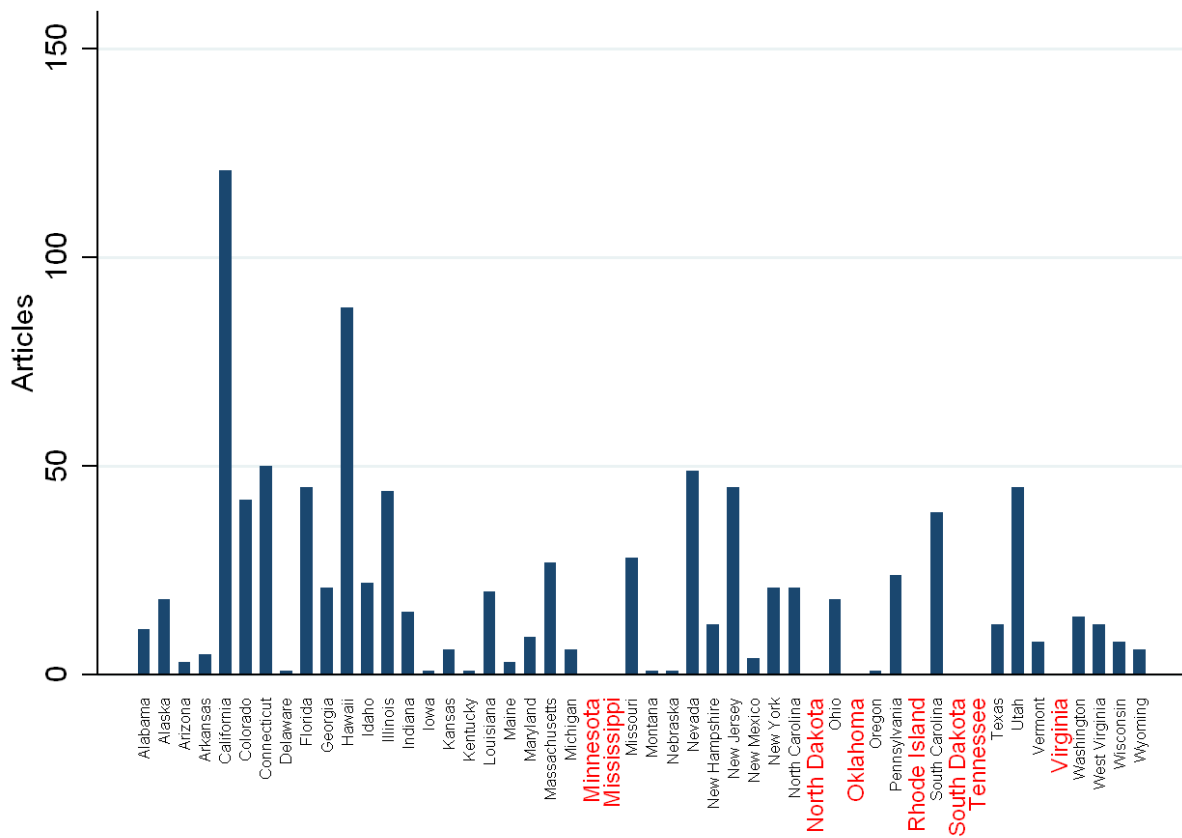
Figure 4
Newspapers Density in Oregon, US

Figure 5
Education and Preferences for Clean Environment, I
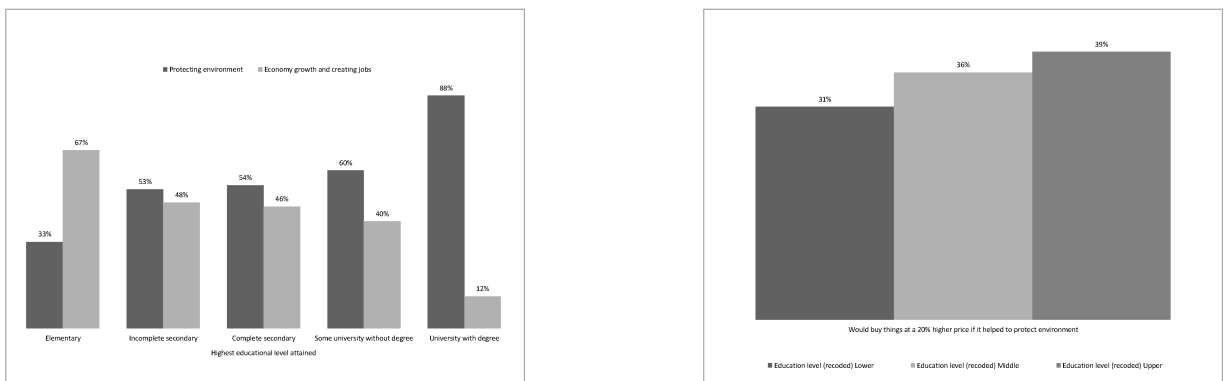
Figure 6
Education and Preferences for Clean Environment, II

57% 76% 81%

Chosen products that are better for environment

■ Education level (recoded) Lower ■ Education level (recoded) Middle ■ Education level (recoded) Upper

9.30% 13.50% 26.40%

Attend meeting, signed petition

■ Education level (recoded) Lower ■ Education level (recoded) Middle ■ Education level (recoded) Upper
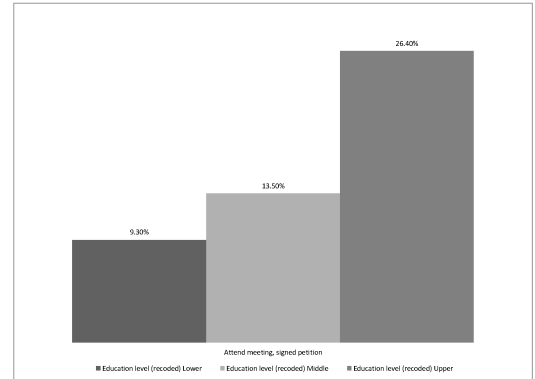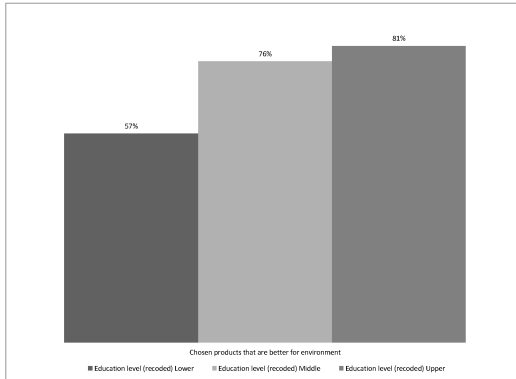
Figure 7
Education and Preferences for Clean Environment, III

77% 87% 90%

Recycle

■ Education level (recoded) Lower ■ Education level (recoded) Middle ■ Education level (recoded) Upper

13.9% 21.8% 34.9%

Contributed to environmental organization

■ Education level (recoded) Lower ■ Education level (recoded) Middle ■ Education level (recoded) Upper
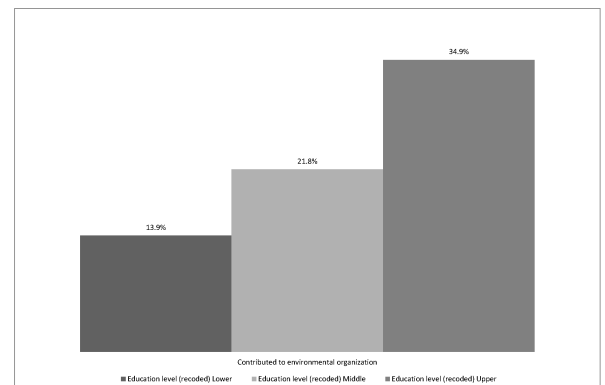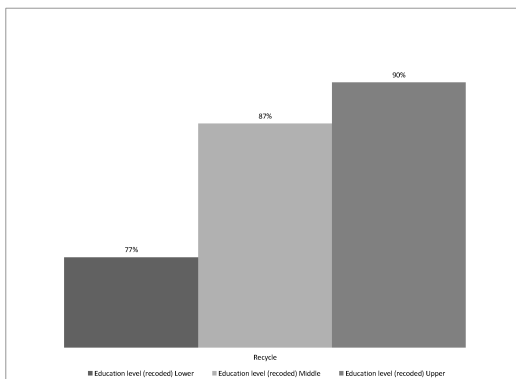
Figure 8
Distribution of county-level Health Outcomes across plants
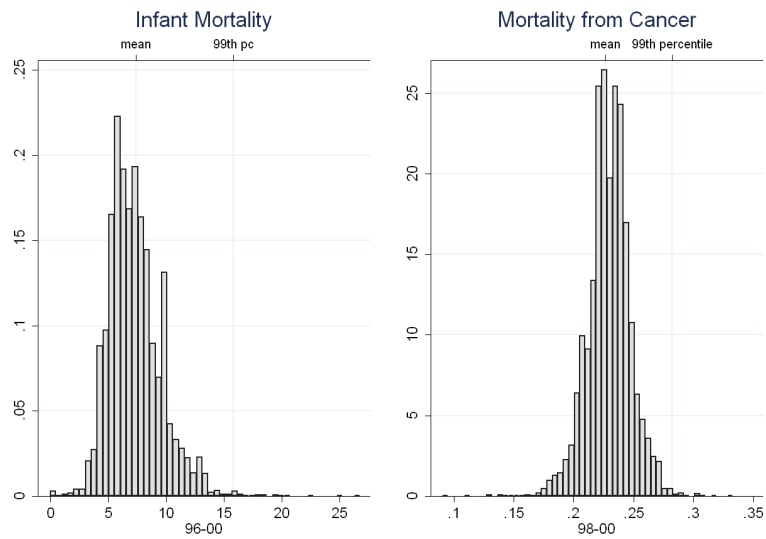
Table 1
Summary statistics, state top 20 polluters

Variable	Mean	Std. Dev.	N
Newspapers Density	0.402	0.665	12772
Covered	0.053	0.224	12772
Number Newsp in Rings	2.215	2.866	12772
Avg Distance Newsp in Rings	9.837	4.863	10064

Table 2
Summary statistics, by coverage in State

	Coverage in State	No Coverage in State
Toxic Emission (in 000)	166 (2586)	154 (1232)
Lag 1 Newspapers Density	0.53 (0.80)	0.41 (0.59)
Pop Density	1378 (2003)	897 (1100)
Income pc	24444 (7055)	23052 (6241)
Education - high school or some college	53 (8)	53 (6)
Education - associate or more	29 (11)	28 (11)
Share Black	11 (14)	14 (18)
Share Younger 20	28 (3)	28 (3)
Share Older 65	13 (3)	13 (3)
Universities Density	3 (4)	2 (3)
Unemployment 2000	5 (3)	5 (2)

Means. Standard Deviations in parenthesis

Table 3

Newspapers Density and probability of coverage of toxic emissions

Dep. var.	(1) Pr(cov ring)	(2) Pr(cov out)	(3) Pr(cov ring) Density>0	(4) Pr(cov ring) Density>0	(5) Pr(cov ring) Density>0
Newspapers Density	0.013**	0.006	0.010*	0.009*	0.009*
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Covered Out				0.082***	
				(0.026)	
Controls					X
Observations	12,772	12,772	10,064	10,064	9,967
Adjusted R-squared	0.073	0.136	0.093	0.102	0.098
$Pr(\overline{X}\widehat{\beta})$	0.0186	0.0387	0.0232	0.0232	0.0234

Standard errors clustered by county. State-by-year and Industry-subsector FE included in all the regressions. The controls included in Column (5) are: log pop density, log income, share black, share with high school education or some college, share with college education ore more, share aged less than 20, share aged more than 65, share unemployed in 2000 included in Column (5)

Table 4

Newspapers Density and toxic emissions

Dep Variable: ln *Emissions*	(1)	(2)	(3)	(4)
Lag 1 Newspapers Density	-0.136***	-0.092**	-0.095**	-0.097**
	(0.041)	(0.041)	(0.041)	(0.041)
Log Pop Density		-0.233***	-0.186**	-0.192**
		(0.078)	(0.080)	(0.087)
Log Income		-0.895***	0.191	0.282
		(0.279)	(0.448)	(0.486)
Education - associate or more			-0.027***	-0.028**
			(0.010)	(0.014)
Share Black				-0.002
				(0.005)
Share Younger 20				-0.029
				(0.018)
Share Older 65				-0.031
				(0.022)
Education - high school or some college				-0.000
				(0.012)
Unemployment 2000				0.014
				(0.028)
Observations	154,587	154,582	154,582	154,582
Adjusted R-squared	0.272	0.273	0.273	0.274

Standard Errors clustered by County. County-year and industry-group fixed effects included in all the regressions. Dummies for values of ln pop density and ln income < the 10th, between the 10th and the 25th,between the 25th and the 50th, and between the 50th and the 75th percentile included in Columns (2)-(4).

Table 5

Placebo, states with no coverage and within-firm variation

Dep Variable: ln *Emissions*	(1)	(2)	(3)	(4)
	Placebo	*No coverage states*	*Within-firm*	
			> 1 plant	> 2 plants
Universities Density	-0.012			
	(0.008)			
Lag 1 Newspapers Density		0.027	-0.068	-0.080*
		(0.171)	(0.047)	(0.048)
County-by-year FE	X	X		
Firm-by-year FE			X	X
Observations	154,582	20,661	90,245	75,805
Adjusted R-squared	0.273	0.281	0.501	0.506

Standard Errors clustered by County. Industry-group fixed effects included. Controls for demographics and dummies for values of ln pop density and ln income < the 10th, between the 10th and the 25th,between the 25th and the 50th, and between the 50th and the 75th percentile included.

Table 6
Alternative samples and estimation methods

Dep Var: ln *Emissions*	(1) No Large Cities	(2) Density >0	(3) Emissions>0	(4) Interv Reg	(5) 2001-2004	(6) Not at border
Lag 1 Newspapers Density	-0.106**	-0.100**	-0.066*	-0.095***	-0.085*	-0.121***
	(0.043)	(0.042)	(0.040)	(0.015)	(0.043)	(0.044)
Unempl 2000					0.007	
					(0.031)	
Observations	146,512	138,439	120,569	154,475	72,145	131,998
Adjusted R-squared	0.277	0.263	0.280		0.256	0.270

Standard Errors clustered by County. County-year and industry-group fixed effects included in Columns (1), (2), (3), (5) and (6). County, industry-subsector and year fixed effects included in Column (4). Controls for demographics and dummies for values of ln pop density and ln income < the 10th, between the 10th and the 25th, between the 25th and the 50th, and between the 50th and the 75th percentile included in all the regressions.

Table 7
Heterogeneous effects by substance

	(1) Pr(cov ring)	(2) Pr(cov ring)	(3) emission dioxin	(4) emissions lead
Dep. var.				
emissions dioxin	0.004**			
	(0.002)			
emissions lead		-0.003		
		(0.003)		
Lag 1 Newspapers Density			-0.010**	0.003
			(0.004)	(0.007)
State-by-year FE	X	X		
2-digit industry FE	X	X		
County-by-year FE			X	X
4-digit industry FE			X	X
Observations	9,967	9,967	154,582	154,582
Adjusted R-squared	0.108	0.107	0.364	0.117

Total emissions, demographics, and dummies for values of ln pop density and ln income < the 10th, between the 10th and the 25th, between the 25th and the 50th, and between the 50th and the 75th percentile are included in all the regressions. Control for *Newspapers Density* included in Columns (1) and (2). *emissions dioxin* and *emissions lead* are categoric variables that take value 0 if the respective emissions are zero, and values from 1 to 4 depending if emissions are, respectively, before the 1th, between the 2nd and the 3rd, between the 3rd and the 4th, and above the 4th quartile, in the part of the distribution for which emissions are larger than zero.

Table 8

Industry subsectors, by type of good or service

Final
Utilities
Construction of Buildings
Food Manufacturing
Beverage and Tobacco Product Manufacturing
Textile Product Mills
Apparel Manufacturing
Petroleum and Coal Products Manufacturing
Computer and Electronic Product Manufacturing
Furniture and Related Product Manufacturing
Miscellaneous Manufacturing
Building Material and Garden Equipment and Supplies Dealers
Food and Beverage Stores
Health and Personal Care Stores
Nonstore Retailers
Air Transportation
Publishing Industries (except Internet)
Rental and Leasing Services
Educational Services
Hospitals
Repair and Maintenance
Personal and Laundry Services

Intermediate
Crop production
Animal farming
Forestry and Lodging
Support Activities for Agriculture and Forestry
Oil and Gas Extraction
Mining
Support Activities for Mining
Specialty Trade Contractors
Textile Mills
Leather and Allied Product Manufacturing
Wood Product Manufacturing
Paper Manufacturing
Printing and Related Support Activities
Chemical Manufacturing
Plastics and Rubber Products Manufacturing
Nonmetallic Mineral Product Manufacturing
Primary Metal Manufacturing
Fabricated Metal Product Manufacturing
Machinery Manufacturing
Electrical Equipment, Appliance, and Component Manufacturing
Transportation Equipment Manufacturing
Merchant Wholesalers, Durable Goods
Merchant Wholesalers, Nondurable Goods
Wholesale Electronic Markets and Agents and Brokers
Support Activities for Transportation
Warehousing and Storage
Real Estate
Professional, Scientific, and Technical Services
Management of Companies and Enterprises
Administrative and Support Services
Waste Management and Remediation Services

Table 9

Heterogeneous effects by consumer pressure and incidence of extreme negative health outcomes

Dep Var: ln $Emissions$	(1)	(2)	(3)
	By Consumer Pressure	By Incidence of Negative Health Outcomes	
Lag 1 Newspapers Density	-0.057	-0.084**	-0.085**
	(0.044)	(0.042)	(0.042)
Lag 1 Newspapers Density*Final Good	-0.260**		
	(0.104)		
Lag 1 Newspapers Density*Extreme Infant Mortality 97-01		-0.564***	
		(0.144)	
Lag 1 Newspapers Density*Extreme Mortality from Cancer 98-00			-3.549***
			(1.067)
Industry Effects	4-digit	3-digit	3-digit
Observations	153,479	154,582	154,548
Adjusted R-squared	0.275	0.180	0.181

Standard Errors clustered by County. County-year fixed effects included. Controls for demographics and dummies for values of ln pop density and ln income < the 10th, between the 10th and the 25th, between the 25th and the 50th, and between the 50th and the 75th percentile included.

Table 10
Within-firm substitution

Dep Var: $\ln Emissions$	(1)	(2)
		> 1 plant per firm-ind group
Lag 1 Newspapers Density	-0.099	-0.100
	(0.089)	(0.090)
Min Density		-0.016
		(0.130)
N plants Firm-Industry		-0.007
		(0.005)
Newspapers Density Firm	-0.087	
	(0.122)	
Observations	65,995	65,546
Adjusted R-squared	0.385	0.384

Standard Errors clustered by County. County-by-year and Industry-group effects included. Controls for demographics and dummies for values of ln pop density and ln income < the 10th, between the 10th and the 25th,between the 25th and the 50th, and between the 50th and the 75th percentile included.

Table 11
Aggregate effects

	Total Emissions, Actual	Total Emissions, Counterfactual	Lower Emissions due to Newspapers Density	Lower Emissions due to Newspapers Density, %
Full Sample	25,600,000	25,400,000	800,000	3%
Industries selling final goods	9,460,000	10,500,000	1,040,000	11%
Counties with extreme infant mortality	31,300	36,500	5,200	17%
Counties with extreme mortality from cancer	79,800	97,800	18,000	22%

Means, in thousands. Standard deviations in parenthesis. See Section 5 for a description of how the counterfactuals are calculated.