Número 614

Environmental Justice and Toxic Releases in Urban Mexico

LOPAMUDRA CHAKRABORTI

NOVIEMBRE 2018

CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONONÓMICAS

Advertencia

Los Documentos de Trabajo del CIDE son una herramienta para fomentar la discusión entre las comunidades académicas. A partir de la difusión, en este formato, de los avances de investigación se busca que los autores puedan recibir comentarios y retroalimentación de sus pares nacionales e internacionales en un estado aún temprano de la investigación.

De acuerdo con esta práctica internacional congruente con el trabajo académico contemporáneo, muchos de estos documentos buscan convertirse posteriormente en una publicación formal, como libro, capítulo de libro o artículo en revista especializada.

Carretera México Toluca 3655, Col. Lomas de Santa Fe, 01210,
Álvaro Obregón, Ciudad de México, México.
www.cide.edu

www.LibreriaCide.com

Oficina de Coordinación Editorial
editorial@cide.edu
Tel. 5081 400

Acknowledgements

I am thankful for the generous comments by anonymous referees that has improved the analysis greatly. I am grateful for the financial support of Mexico's National Council of Science and Technology Conacyt (grant # PNDCPN 2014-01/248906) and guidance at the inception of this research by Jaime Sainz (CIDE), Jay Shimshack, Michael Margolis, and participants at the AERE 4th Annual AERE Summer Conference, San Diego, California, and 90th Western Economics Association International (WEAI) Annual Conference, Honolulu, Hawaii.

Abstract

In this paper we present empirical evidence that major toxic releasing facilities in Mexico pollute more in poorer or marginalized neighborhoods. We are the first to investigate "environmental (in)justice" behavior utilizing a large panel of water pollution data, from 2004 to 2015, reported in the Mexican Pollution Reporting and Transfers database. For our measure of socioeconomic status, we use the Mexican government's "marginalization index" published every five years and available at a disaggregate scale for urban areas in Mexico. We focus on seven toxic substances that pose the greatest health risk and are among the most prevalent water releases in Mexico. Our results show that plants increased their toxics pollution in areas that witnessed a rise in marginalization. For 1% increase in marginalization, an average plant increases their arsenic releases by 0.05 kg/year, cadmium releases by 0.13 kg/year, cyanide releases by 0.12 kg/year, lead releases by 4.1 kg/year, nickel releases by 4.2 kg/year, chromium releases by 1.2 kg/year, and mercury releases by 3.4 kg/year. We interpret these magnitudes as significant as a facility with annual releases of 1 kg/year is classified as a major polluter subject to mandatory reporting for all six metals (except cyanide).

Keywords: a environmental justice; toxic releases in Mexico; self-reported water pollution; informal regulation; community features; local income and unemployment effects

Resumen

En este documento presentamos evidencia empírica que muestra que las instalaciones emisoras de contaminantes más importantes de México contaminan más en comunidades de bajos recursos o marginadas. Esta investigación es la primera en usar un panel de datos extensos sobre la contaminación del agua (2004 a 2015) obtenidos del Registro de Emisiones y Transferencia de Contaminantes (RETC) para investigar la (in)justicia ambiental. Utilizamos el "índice de marginación" del gobierno mexicano como medida de estatus socioeconómico. Este índice es publicado cada cinco años y está disponible a una escala desagregada para áreas urbanas en México. Nos centramos en siete sustancias tóxicas que representan el mayor riesgo para la salud y se encuentran entre las emisiones al agua más frecuentes en México. Nuestros resultados muestran que las instalaciones emisoras aumentaron su contaminación tóxica en áreas urbanas que experimentaron un aumento en marginación. Para un aumento de 1% en la marginación, una instalación promedio aumenta sus emisiones de arsénico en 0.05 kg / año, emisiones de cadmio en 0.13 kg / año, emisiones de cianuro en 0.12 kg / año, emisiones de plomo en 4.1 kg / año, emisiones de níquel en 4.2 kg / año, emisiones de cromo en 1.2 kg / año y emisiones de mercurio en 3.4 kg / año. Estas magnitudes son significativas, ya que una instalación con emisiones anuales de 1 kg / año se clasifica como un contaminador importante y está sujeto a informes obligatorios para los seis metales (excepto el cianuro).

Palabras claves: la justicia ambiental, las emisiones tóxicas en México, los auto-reportes de la contaminación del agua, la regulación informal, características de la comunidad, los efectos de renta y de desempleo local

Introducción

In developed countries, studies show that pollution is higher in poorer or marginalized neighborhoods. The empirical literature has found evidence on two mechanisms that explain this relationship. When industries locate in lower income communities expecting less community pressure or regulatory presence, the direction of causality runs from community characteristics to pollution (Wolverton, 2009). In the literature this mechanism is called the environmental justice principle. On the other hand, higher pollution might lead to the poor moving into dirtier neighborhoods due to lower housing costs etc.; then the causality runs in the opposite direction from pollution to community characteristics (Currie *et al.,* 2015).[1] The latter is called the compensating differential or sorting principle. The objective of this paper is to isolate the evidence on environmental justice in Mexico--higher pollution from toxic releasing facilities associated with marginalized and poorer communities in the neighborhood.

In this paper we present empirical evidence that environmental justice concerns are predominant in developing countries and for Latin America in particular. We utilize the most up-to-date database on toxic releases data from major facilities in Mexico. Most of the environmental justice studies for the developing world, although based in Mexico, are limited to specific border cities and use cross-section sampling to

[1] The authors find that housing values fall after plants releasing toxic pollutants open, within a distance (1/2 to 1 mile) believed to correspond to how far the pollutants travel.

identify association between socioeconomic characteristics and proximity or concentration of heavily polluting industries such as export assembly plants. They find that urban and peri-urban areas with incomes higher than rural regions also have the characteristics of a greater concentration of industries (Grineski, Collins and Aguilar, 2015).

This is also one of the few studies in the environmental justice literature that focuses on toxic discharges into water, unlike most studies in the US and other developed countries that focus on air emissions. The pollutant release data in Mexico allows us to focus on heavy metals that are classified under known or suspected carcinogens, developmental or reproductive toxicants, as well as persistent bioaccumulative and toxic compounds (CEC, 2018). Our decision to study discharges into water was originally guided by a belief that the affected population is more reliably people nearby than is the case for air emissions. These data also appear to be more complete than other releases and transfers.

We present a comprehensive analysis of the relationship between pollution and poverty as captured by the marginalization index published by the Mexican Government's National Population Council (*Consejo Nacional de Población*, Conapo). The index measures aspects of education, housing, and health, but not income per se, which is largely unavailable in Mexican census data. Although we are not the first to use this socioeconomic indicator, we exercise caution in utilizing this information, especially in a panel data framework, as the index published every five years using Census and *Conteo* (count) data is not comparable across the years.[2]

We adopt several empirical strategies to establish that higher pollution is associated with poorer socioeconomic status of the local population. Due to lack of comparable data on the marginalization index, over time, we estimate cross sectional models using a marginalization index from 2000 and 2005 census years, all of them prior to the pollution reported years. To address endogeneity of socioeconomic indicators, at the very least, the local community characteristics need to be from a year prior to the pollution report years (Arora and Cason, 1998). Alternately, we

[2] The *Conteo* databases are actually censuses that collect information on a smaller number of household characteristics compared to the census databases.

utilize the panel nature of our data by looking at change in pollution levels within the same plant. In these models we capture the effect of marginalization after netting out the effect of plant level unobservable factors. In addition, we control for each plants' base levels of pollution. Lastly, the fixed effects model control for all plant specific factors that influence why some polluters pollute more as opposed to others.

The coefficient estimates from the cross sectional models show that higher pollution is associated with higher marginalization index (though seldom statistically significant). The conditional quantile regressions (on the same cross sections) reveal some evidence that facilities that are at higher percentiles of the log pollution distribution (75^{th} and 90^{th} percentiles) increase their toxic pollution even more in response to marginalized communities. In the models exploiting within plant variations in pollution such as the change in pollution levels, heavily polluting plants (in the initial years) are more responsive to local marginalization with a positive sign on the interaction term of initial pollution levels and marginalization index. In the fixed effects models we find that communities that witnessed an increase in marginalization also experienced higher pollution levels, over time.

THE SELF-REPORTED, MANDATORY TOXIC RELEASES DATABASE

As part of a side accord of the North American Free Trade Agreement (NAFTA), the three countries of Canada, Mexico and US have collaborated to address environmental challenges and promote public disclosure of toxics pollution by major facilities. Following the Emergency Planning and Community Right-to-Know Act (EPCRA), the US adopted the Toxics Release Inventory (TRI) in 1987. In 1993, Canada adopted its National Pollutant Release Inventory (NPRI). Following a 2001 amendment to Mexico´s General Law of Ecological Equilibrium and Environmental Protection (*Ley General de Equilibrio Ecológico y Protección al Ambiente*) a rule adopted in June of 2004, the Pollutant Release and Transfer Register (*Registro de Emisiones y Transferencias de Contaminantes* (RETC) was adopted by Mexico. The Commission for Environmental Cooperation of North America (CEC), an intergovernmental agency, with the task of harmonizing environmental reporting by all three countries archives all data with a couple of years' lag. The respective pollutant registers with information on releases of listed substances into air, land, water, and transfers (e.g., sent to recycling or sewerage) are updated annually.

The CEC in its Taking Stock reports highlight that differences in the national registers means lack of comparability across nations. As of the 2013 reporting year, the three PRTRs (Pollutant Release and Transfers Register) differ in coverage of pollutants with TRI 675, NPRI 346 and RETC 104 with only about 60 pollutants common to all three (CEC, 2018). The NPRI covers all facilities manufacturing or using a listed substance, the TRI covers most manufacturing and federal facilities like electric utilities and hazardous waste management while RETC covers 11 industrial sectors that are subject to federal jurisdiction. [3] [4] In addition, the RETC applies to any facility that handles hazardous waste or discharge pollutants into national water bodies. Facilities under NPRI and TRI are required to report if they meet or exceed the "activity" and employee thresholds while RETC facilities are subject to either "activity" or "release" thresholds with no restrictions on number of employees.

According to the Mexican Secretariat for the Environment and Natural Resources *(Secretaria de Medio Ambiente y Recursos Naturales,* or Semarnat) learning-by-doing of facilities is a common feature of any self-reported pollution data, which is not atypical to Mexico. It seems likely, then, that the later cross sections are more representative, not to mention that there is considerable variation among the sample of facilities reporting even in the most recent years.[5] We focus on seven pollutants that are fairly common and pose the greatest threat to health from exposure (CEC, 2009). The metals are arsenic, cadmium, chromium, lead, mercury, nickel and their compounds; and cyanide (organic and inorganic). These seven toxics were among the top 25 pollutants for on-site water releases in Mexico.[6] Our dataset on discharges into

---

[3] Notably, facilities that fall under state or municipal jurisdiction are not included in the National RETC database.

[4] These sectors are: petroleum, chemicals, paints and ink manufacturing, primary and fabricated metals, automotive, pulp and paper, cement/limestone, asbestos, glass, electric utilities, and hazardous waste management.

[5] Results are similar upon dropping the first 3-year period of 2004-2006, throughout the various specifications; even though sample sizes drop by more than half.

[6] http://www.cec.org/Page.asp?PageID=749&SiteNodeID=1215&BL_ExpandID=754

water actually shrinks over time. In Section 4 we explore some empirical strategies to address the unbalanced nature of this panel.

By law, firms are required to report on these toxic substances if the total amount manufactured, processed, or otherwise used exceeds 5 kg per year, or the amount emitted exceeds 1 kg. For cyanide, the thresholds are 2,500 kg used or 100 kg emitted. Facilities use a variety of methods to measure or estimate their emissions, including emission factors, mass balance, engineering calculation, stack testing and direct measurement. When reporting under their annual operation certificates (*Cedula de Operacion*, COA), facilities include information about the type of method used. However, this information is not recorded in the RETC database. Also, simple reporting error appear to be common--according to a Semarnat official, these include many "errors in the conversion of units and errors in the selection of the appropriate substance for report (substances with similar names are often interchanged)" (Eicker et al 2010, p11-12).[7]

Examination of individual plants turns up other likely symptoms of reporting errors. There are some very improbable fluctuations that could be attributed to data entry errors. One such example is for cadmium discharges at one plant of 0.37 tons in 2007, 441 tons in 2008 and 0.05 tons in 2009. About a quarter of the pollution reports have duplicates at the same plant out to five or six significant digits. For example, one plant reports "161.88409" kg of lead for the years 2010, 2011, 2012 and 2013. Another example is a plant reporting "0.0005" kg for arsenic, cadmium, chromium, nickel and lead for the year 2006. The latter example most likely correspond to the detection threshold on some monitoring device. However, the former can hardly be interpreted as anything other than failure to take new measurements.

The inaccuracies in the pollution reports appear to result from the lack of updating of pollution reports because of poor enforcement and limited budgets of implementing agencies. This is because pollution reports albeit mandatory are not subject to standard monitoring and enforcement protocols and hence they appear not to result from any attempt to deceive regulators or the public. Since random

[7] Two such obvious cases of data entry error were dropped from the sample, where units reflect tons rather than kilograms or even grams.

measurement error biases regression coefficients towards zero (i.e., attenuation bias), the results below most likely understate the actual association between marginalization and pollution.

Along with the pollution reports, the RETC database contain information on each facility or business establishment such as name, address, and geographic location. However, for purposes of pooling the pollution data across the years we faced the problem of the same physical plant appearing with different identification codes. For example, each time the plant underwent a change in name, ownership, sector designation, etc., it was assigned a new identification code in the RETC system. If the same facility had multiple activities (e.g., generation of electricity and treatment of toxic residuals), it appeared with different identification codes within the same year. Hence, we had to manually consolidate the number of unique RETC facilities (i.e., the same physical plant/business) across the different years in the database.

Finally, similar to the TRI literature that mention inaccuracies in facility location records, we suspect data errors in the geographical coordinates in the RETC database. Our final sample of georeferenced facilities with water pollution data is 2,443 RETC facilities. We used Mexico's Statistical Agency's (INEGI's) National Directory and Statistics on Economic Units (Denue) database and mapping software to validate the location of each polluter.

REVIEW OF ENVIRONMENTAL JUSTICE LITERATURE

In the developed world, environmental justice variables, as captured by socio-economic characteristics of the population affected, have been associated with higher industrial pollution. Most of the empirical literature stemmed from Hamilton's (1995) classification of discrimination based on race/ethnicity, economic vulnerability, and willingness to engage in collective action. The literature on the Toxics Release Inventory (TRI) which serves as a mandatory public information disclosure mechanism in the U.S. has focused primarily on the environmental justice perspective. Higher pollution is linked to lower income levels in Brooks and Sethi, (1997) on air emissions; Arora and Cason (1998) on aggregate emissions; Helland and Whitford

(2003) on emissions into air, water, and land treated separately), although not in all (Gray, Shadbegian and Wolverton, 2012).

The US EPA's Risk-Screening Environmental Indicators (RSEI) Model has been utilized in studies such as Bouwes, Hassur and Shapiro (2003), Ash and Fetter (2004), and more recently in Zwickl, Ash and Boyce (2014). These papers model exposure to toxics by weighing the "dose" from toxics released by TRI facilities with toxicity of chemicals and population exposed to get precise values for potential health risks. More recently, formal investigations into environmental justice by modeling toxicity adjusted province level exposures to toxics releases were conducted in the European Union economies such as Germani, Morone, and Testa (2014). Increased political participation as captured by voter turnout has been shown to lower potential risks from toxic chemicals pollution in Shapiro (2005). Saha and Mohr (2013) find that media attention has a deterring effect on pollution from toxics releasing facilities in the US.

Empirical evidence on the relation between pollution and income (or poverty), from developing countries, and Latin America in particular, is both sparse and ambiguous. Early evidence on informal regulation or community pressure for organic water pollutants comes from Indonesia (Pargal and Wheeler, 1996). More recently, this relationship is put to scrutiny in emerging economies notably in China (Ma, 2010). In Latin America, Seroa de Motta (2006) uses survey data of large manufacturing plants in Brazil, for the year 1997, to conclude that pressure from communities and NGOs are important factors in explaining environmental performance of manufacturing firms in Brazil. In Brazil too, Feres and Reynaud (2012) find that informal regulation and more specifically community pressure influence formal regulatory actions such as inspections and sanctions. They capture informal regulation by using the count of citizen complaints about pollution at a given plant or in a given location and the number of meetings organized with environmental NGOs, local community or political representatives. In Bogota, Colombia, Cruz (2004) finds that higher income and education are significant in explaining complaints filed by the local population. By contrast, Dasgupta, Lucas and Wheeler (2002) find that particulate matter emissions are higher in higher wage (urban) municipalities in Brazil.

In Mexico, a 1995 confidential survey of 236 major polluters found that only about a fourth of them consider pressure from the neighboring community a significant factor in environmental decisions (Dasgupta, Hettige, and Wheeler, 2000). Also in Mexico, Blackman, Batz, and Evans (2004) reported that in Ciudad Juarez, the export-assembly plants known as *maquiladoras* contribute substantially to air pollution, but they conclude that the poor are not disproportionately affected. However, as pointed out before, none of the prior studies actually consider plant level pollution data rather the focus was on concentration of high risk industries and/or perceptions of the polluting industries themselves.

EMPIRICAL APPROACH

Our main variable of interest is self-reported toxic pollution discharged into water. Our empirical specifications, however, consider direct discharges into water and indirect discharges through sewerage system.[8] The physical nature of metals is such that metals disposed of directly into water bodies or through sewerage are not likely to be destroyed and hence can be treated as releases that eventually enter the environment (CEC, 2011).We adopt this strategy as a small proportion of the plants 'switch' their pollution media from water to land, or water to sewage, or both. Most of the land and sewage pollution are available from 2010-2013 and 2009-2013, respectively. This change in reporting system is a puzzle considering that there was no change in the regulation or any overarching changes in policy or economic conditions that might explain this phenomenon. We exercise caution in aggregating across pollution media as exposure to water pollution (either directly or indirectly through the sewerage system) are distinct from exposure to land pollution, which primarily infiltrates into the ground.

We access the latest update to the RETC database with pollution data until 2015. Our main models make use of pollution levels averaged across four, three-year

[8] Results are similar when considering only water pollution.

periods for the 12 years of pollution data available (2004-2006, 2007-2009, 2010-2012, and 2013-2015) of annual toxics reported by each facility in the RETC database. We do not exploit the annual nature of the pollution reports due to potential issues with the data such as apparent lack of updates in the pollution reports across the years and/or identical pollution reports across toxics within a year. We do, however, disaggregate on one important dimension. In all the models, each substance is treated separately since it is likely that their health impacts, public perception, and awareness differ.

The Census of Population and Housing do not directly ask questions on the income or poverty status of households. We capture socioeconomic status with an index of urban marginalization available from Mexico's National Population Council (Conapo) website. The data is available from 1990 onwards but at the level of municipalities for the first census year reported (reporting at a more disaggregate level of AGEBs [*Área Geoestadística Básica Urbana*] began in 2000).[9] It is calculated as the first principal component of indicators of education, social status and demographics (see Table 1 for details). The values are categorized into five classes of "very high," "high," "medium," "low," and "very low," with positive numbers classified as highly marginalized and negative numbers as less marginalized.[10] However, the set of variables utilized to calculate the marginalization index varied across the years. The indices calculated every five years (based on Census and *Conteo* data) are not comparable over time as they are scaled against the corresponding socioeconomic variables in each year.

We adopt a threefold strategy to investigate the relation between pollution and marginalized neighborhoods. The first is to simply investigate correlations between pollution and marginalization indices in a cross sectional framework, similar to the reduced form bivariate analysis of Chakraborti and Margolis (2017). The second is to model changes in pollution, over time, controlling for base levels of pollution of each

[9] AGEBs are roughly comparable to urban census tracts in the United States. AGEBs are fairly small urban areas with more than 2,500 inhabitants and relatively homogeneous socioeconomic characteristics. To compare AGEBs versus municipalities, the state of Aguascalientes, Mexico, comprises of only 6 municipalities while it has about 300 AGEBs.

[10] Upon rescaling the index to take only positive values, all our results yield very similar coefficients.

facility and explore its interaction with the marginalization index i.e. whether more polluting facilities increase their rate of toxic releases if community experiences a rise in marginalization index (worse off). The third is an attempt to exploit the panel nature of the pollution data by estimating a fixed effects model for the four time periods. Noteworthy, facility level fixed effects control for all time invariant factors that cannot or have not been explicitly modeled in the specifications.

For the cross-sectional associations, we regress toxic releases of plants on IMU of AGEBs within 1km of each plant.[11] We utilize marginalization indices based on 2000 census and 2005 count data. Separate estimations are carried out for each of the four time periods with the corresponding IMU always from a year prior to the pollution report years. To avoid endogeneity of current socioeconomic characteristics, i.e. current pollution influencing move in by the poor (in response to lower housing prices etc.) we consider data from the two earliest available years with IMU at the disaggregate, AGEB level.

As is common practice in the literature, we include population density as an additional measure of local community characteristics. The empirical evidence is not clear on its expected sign— if denser neighborhoods are exposed to greater pollution burden (per person) then its sign is positive; alternately, if local officials and communities exert more pressure on the plants to reduce exposure to pollution, the sign would be negative. In keeping with our choice of IMU, we calculate population density for each plant by aggregating underlying population data reported at the level of AGEBs from the same census years.

Equation (1) below presents the cross-sectional specification where our dependent variable is the three-year average toxic emissions of each substance. $lPoll_{ist}$ is the log of average toxic emission of substance s, by plant i and in time t, our dependent variable, IMU_{it_0} is the marginalization index constructed for each plant, based on the underlying data of "Urban Marginalization Index" published by the Conapo at the AGEB level. The time index t_0 refers to the census/count year with

[11] In addition, we consider the marginalization index at different spatial scales. We try alternative radii of AGEBs within 2 km and 5 km of each plants. We notice that the coefficient falls in magnitude and significance when moving from the 1 km to 2 km and 5 km measures. This pattern is consistent with the 1 km distance criterion used in other spatial environmental justice studies.

available data, prior to the time t when the pollution is reported. The variable $popden_{it_0}$ is the population density of the surrounding AGEBs (within 1 km) calculated from the same census year 2000 and count 2005. State-level dummy variables $state_i$ controls for differences in regulatory pressure and other differences such as political participation.[12] [13]

$$lPoll_{ist} = \alpha + \beta IMU_{it_0} + \delta popden_{it_0} + \theta_i state_i + \varepsilon_{ist} \qquad \text{...... (1)}$$

where, $t = \{2004 - 2006, 2007 - 2009, 2010 - 2012, 2013 - 2015\}$

and $t_0 = \{2000, 2005\}$

Our second approach is to model change in pollution within the same plant and from one time-period to the next. We add an important explanatory variable in this model: the base level of pollution of each facility which is captured by the level of releases reported in the 2004-2006 period. Our priors are that past levels of pollution will be an important determinant of changes in pollution levels, as we expect pollution levels for the same plant to be correlated over time. Inclusion of this variable also allows us to control for some of the plant specific attributes that explain why certain plants are more polluting than others.

$$lPoll_{ist} - lPoll_{is(t-1)} = \alpha + \theta(lBasepoll_{is}) + \rho IMU_{it_0} + \sigma\left[lBasepoll_{is} * IMU_{it_0}\right]$$
$$+\delta popden_{it_0} + \mu\left[lBasepoll_{is} * popden_{it_0}\right] + \vartheta_i toxic_i + \varepsilon_{ist}$$
$$\text{............(2)}$$

Equation (2) above presents the change in pollution model where the dependent variable is the annualized growth rate or change in log pollution from one 3-year average to the next time period. Our expectation on the interaction term between the baseline level of pollution and marginalization index is that heavily polluting plants increase their pollution if communities experience an increase in

[12] Voter turnout data at the municipality level did not yield significant results.
[13] Results similar without state dummies.

marginalization index and hence the sign on the coefficient would be positive. Population density is also interacted with base pollution as it is included as a local community characteristic. The interpretation of the uninteracted marginalization index (or population density) is the effect on change in pollution in communities that experience '0' baseline levels of pollution; while the interpretation of the uninteracted baseline level of pollution is the effect on change in pollution in communities with neutral marginalization index (or population density) with value of '0'. Toxic substance dummies are included as pollutants are pooled.

The third approach adopted to establish the relationship between pollution and marginalized neighborhoods is to exploit the true panel nature of the data. We pool the four time periods; however, given lack of time variation in our independent variable (IMU) we assign a "marginalization score" to each plant utilizing the underlying socioeconomic variables that are also available in the same Conapo databases. Our marginalization score is time varying and calculated as the first principal component of all seven socioeconomic characteristics of AGEBs that are available in the 2000 and 2005 publication years (see Table 1). These are: population between 6 and 14 that do not attend school, population without access to health services, child mortality rate of women between ages 15 and 49, percent of houses without piped water, percent of houses with overcrowding, and percent of houses without refrigerator. Table 1 below shows the correlation coefficients between the seven variables and the first principal component are very similar in pattern to the correlation coefficients between the published IMU indices and its corresponding socioeconomic variables. "Marginalization score" captures about 63 percent of the total variance of the seven underlying variables.

Table 1. Correlation Coefficient of Urban Marginalization Index (IMU), Marginalization Score (MS) and Socioeconomic Variables for AGEBs, 2000-2005

Category	Indicators (percent, unless otherwise indicated)	IMU, 2000	IMU, 2005	MS, 2000-2005
Education	Population between ages 6 and 14 who do not attend school	0.6816	0.5626	0.6652
	Population 15 years and above without post-primary education	0.9028		
	Population 15 years and above without secondary education		0.8460	
Status	Population employed earning no more than twice the minimum wage	0.8624		
	Houses without drainage	0.7709	0.7318	0.7784
	Houses without piped water	0.8517	0.8225	0.8455
	Houses without proper roofing material	0.7291		
	Houses without refrigerator	0.9065	0.8674	0.9080
	Houses with overcrowding	0.8825	0.8556	0.8382
	Houses without septic connection		0.9184	
	Houses with mud floor		0.8068	
Demographic	Child mortality rate of women between ages 15 and 49	0.7204	0.6253	0.7110
	Population without access to health services	0.8173	0.7413	0.7786
	Women between 12 and 17 years of age with at least one live childbirth	0.2929		

Equation (3) below presents the multivariate, panel data model estimated to answer whether changes in "marginalization score" was associated with changes in pollution from local facilities and in a consistent direction as the cross section models.

$$lPoll_{ist} = \alpha_i + \gamma MS_{it_0} + \delta popden_{it_0} + \mu_t period_t + \epsilon_{ist} \quad \dots\dots\dots \quad (3)$$

The dependent variable that we construct is the 3-year average toxic releases of seven substances where t refers to the four periods of 2004-2006, 2007-2009, 2010-2012, and 2013-2015. Our main explanatory variable MS_{it_0} refers to the time varying measure of marginalization "score" based on seven socioeconomic variables from the census 2000 and count 2005 data on AGEBs within 1km of each plant. We include population density drawn from 2000 census and 2005 count years, respectively.

Inclusion of time varying independent variables allows us to control for plant specific fixed effects α_i. Notably, this dummy would subsume all time-invarying aspects such as size of operation, type of manufacturing facility and regulatory presence in the state. Dummy variables for each 3-year average period, $period_t$ captures time period specific effects such as variations based on each three-year period. In particular, these control for the attrition in the pollution reports observed in the water pollution data. Finally, we cluster standard errors at the level of municipality to control for spatial correlation of plants located near each other.[14]

REGRESSION SAMPLE

Our dataset has 4,321 RETC facilities that report at least once, over 2004-2015, its pollutant emissions or transfers on any one of the 7 toxic substances. Of these 4,321 facilities, we could verify the location of 2,443 facilities with water and sewage pollutant release data (roughly 56 percent match). Majority of these plants are located in the State of Mexico (18 percent), Tamaulipas and Nuevo Leon (10 percent), Mexico City (8 percent), and Jalisco (7 percent). Note that the three biggest metropolitan statistical areas (MSAs) of Mexico City, Monterrey and Guadalajara are included in these states covered. Majority of these plants are in the chemical (24 percent), automobiles, metallurgy and others (11 percent), electronics (6 percent), and petroleum and petrochemicals (5 percent).

We present pollution data summed across direct discharges into water and indirect discharges into water through the sewerage system as part of our main

[14] Clustering standard errors by plant as we expect pollution levels from the same plant will be correlated over time would address arbitrary serial correlation.

results. Table 2 shows a more balanced sample of all seven pollutants and for the four, 3-year averages in contrast to only water. It shows a lot of variabilities depending on the toxic substance as well as the period under consideration. We then take the logarithm as toxic releases reported exhibit considerable variation both within and between different polluting facilities.

Table 2. Average Water and Sewage Pollution 2004-2015

Variable	Obs.	Mean	Standard deviation	Min	Max
Arsenic, 2004-2006	998	15.87275	272.9713	5.63e-13	8210
Arsenic, 2007-2009	851	18.98477	270.6547	7.00e-10	6996.607
Arsenic, 2010-2012	639	9.892782	126.5048	5.26e-09	2432.106
Arsenic, 2013-2015	411	53.83123	710.3534	2.00e-09	14022
Cadmium, 2004-2006	893	6.581314	44.34351	2.00e-10	666.8361
Cadmium, 2007-2009	830	89.673	1338.272	5.00e-10	30420
Cadmium, 2010-2012	624	41.31844	423.995	3.00e-08	8826.217
Cadmium, 2013-2015	416	37.16841	226.2592	1.00e-08	2821
Chromium, 2004-2006	902	43.46531	513.2508	4.80e-10	11620
Chromium, 2007-2009	838	317.9041	5391.58	2.70e-09	114075
Chromium, 2010-2012	635	268.5724	2669.348	1.20e-07	40948.49
Chromium, 2013-2015	447	183.9916	2160.134	1.00e-08	43909.78
Cyanide, 2004-2006	1,005	4.869878	38.51604	5.63e-13	883.008
Cyanide, 2007-2009	874	73.33431	1000.977	1.00e-09	19421.97
Cyanide, 2010-2012	634	19.3203	221.4729	4.00e-09	5218.367
Cyanide, 2013-2015	389	73.12594	897.9623	4.73e-12	17237.56
Lead, 2004-2006	975	43.84804	506.1654	2.80e-09	13200
Lead, 2007-2009	899	221.1405	3744.506	1.30e-09	82242
Lead, 2010-2012	675	3482.796	84651.66	1.20e-07	2198218
Lead, 2013-2015	457	161.7345	2119.394	1.00e-07	43915.43
Mercury, 2004-2006	976	12.28124	270.7388	1.00e-10	8210
Mercury, 2007-2009	835	1041.457	29841.55	1.00e-10	862312.5
Mercury, 2010-2012	602	71.26333	1540.71	8.00e-10	37710.83
Mercury, 2013-2015	375	553.9914	10548.6	5.10e-13	204259.4
Nickel, 2004-2006	963	34.07277	231.7184	2.50e-09	5274.903
Nickel, 2007-2009	889	411.2412	6201.55	1.30e-09	137070
Nickel, 2010-2012	681	3006.851	72776.4	1.53e-07	1898465
Nickel, 2013-2015	488	254.7132	2386.393	1.00e-08	43915.43

Note: All pollution data are in kilograms.

To get our regression sample, we had to spatially join the geocoded RETC facilities with the AGEB boundaries layer of the census files. Our final sample size is 1,600 RETC facilities that was within 1km of an urban boundary (AGEB) for the census 2000 data. For the count 2005 data, we could match 1,653 facilities that were within 1km of an AGEB boundary. Note that our regression sample considers AGEBs that are exposed to non-zero level of toxic pollution over the years considered in the dataset. This implies that our results cannot be generalized to all urban areas in Mexico; rather they are applicable to urban communities with marginalization indices that are within range of the summary statistics table presented below. Table 3 below presents the dependent variables for the cross-sectional, panel data, and change in pollution models, and the marginalization indices and population density variables used in the regressions. The community characteristics assigned to each plant is the simple average of those for AGEBs with centroids within one kilometer of the location of each plant.

Table 3. Summary Statistics for Regression Sample

Variable	Obs.	Mean	Standard deviation	Min	Max
Dependent Variables					
log water and sewage pollution, 2004-2015	20,201	-2.787009	4.552728	-28.30437	14.6032
Change in water and sewage pollution, 2004-2015	9,333	.2066684	3.870614	-19.35677	26.2696
Community Characteristics					
Urban Marginalization Index, 2000	1,600	-1.425966	1.553065	-4.25317	5.40725
Urban Marginalization Index, 2005	1,653	-.4937153	.5511779	-1.359009	2.71063
Marginalization Score, 2000-2005	3,253	-.9918163	1.234029	-2.890151	6.31086
Population Density, 2000	1,600	43.40724	41.21065	.1672506	275.927
Population Density, 2005	1,653	45.43644	40.05436	.1966283	246.125

RESULTS

First, we present the cross-sectional reduced form models with marginalization index and population density as our measures of local community characteristics, regressed on each 3-year average pollution levels (log transformed) and state level controls. The marginalization index and population density based on the location of each plant is drawn from the years 2000 and 2005 in separate regressions, with results shown in Table 4. Overall, we find that plant-level toxic releases into water (direct and indirect discharges) are higher in more marginalized areas. However, this coefficient is rarely statistically significant.[15] By contrast, population density around the plant exerts a statistically significant negative effect on pollution releases. We infer that population density captures willingness to engage in collective action in other words higher exposure puts a downward pressure on pollution.

The coefficient in column (3) of Table 4 can be interpreted as a one-unit increase in marginalization index is associated with 26% higher chromium pollution during 2004-2006. A one-unit increase in marginalization represents about 70% increase for the average community, in 2000. In terms of elasticity, a 1% increase in marginalization in 2000 is associated with a 0.37% increase in chromium pollution from average levels of pollution at 43.47 kg/year i.e. 0.16 kg/year. For all the other individual toxic substances we find a positive coefficient on marginalization index for the 2004 to 2006 time period (except arsenic). The respective estimates represent an increase of 0.17% for mercury, 0.18% for cadmium and cyanide, 0.21% for nickel and 0.22% for lead. Expressed as a percentage of average levels of pollution, these estimates represent 0.02 kg/year for mercury and cadmium, 0.01 kg/year for cyanide, 0.07 kg/year for nickel and 0.1 kg/year for lead.

For the 2007-2009, we obtain much smaller but positive estimates across all (except cadmium) with elasticities ranging between 0.02% for nickel and 0.11% for chromium. For 2010-2012 pollution, in elasticity terms the magnitude of coefficients on marginalization index (from 2005) varies from 0.03% for lead, 0.07% for cadmium, 0.1% for arsenic, 0.12% for chromium, 0.15% for mercury and 0.20% for cyanide.

[15] Results are similar when summing over all six metals weighted by corresponding health risks.

Evaluated at the sample average levels of pollution, the corresponding estimates represent an increase of 1.2 kg/year for lead, 0.03 kg/year for cadmium, 0.01 kg/year for arsenic, 0.31 kg/year for chromium, 0.1 kg/year for mercury, and 0.04 kg/year for cyanide. Hence, the coefficient on marginalization index vary greatly depending on the pollutant and time period considered; but the sign is consistently positive except for the most recent time period (2013-2015) with the smallest sample size.

The coefficient on population density in column (3) of Table 4 can be interpreted as an increase in density by one person per square km is associated with a decline in chromium pollution by 0.8%. We interpret the coefficient on the other community variable considered as follows-- a one-unit increase in density represents about 2.3% increase for the average community in 2000. In terms of elasticity, a 1% increase in population density in 2000 is associated with a 0.35% decline in chromium pollution for the average plant i.e. a magnitude of 0.15 kg/year. We infer that perhaps environmental justice concerns i.e. higher pollution in more marginalized areas are being offset by community's willingness to engage in collective action as captured by the population density measure. However, marginalization index is seldom statistically significant.

Table 4. Log Water and Sewage Pollution 2004-2006, 2007-2009, 2010-2012 and 2013-2015 on Urban Marginalization Index (IMU) and population density, from 2000 and 2005

Period	Variables	(1) Arsenic	(2) Cadmium	(3) Chromium	(4) Cyanide	(5) Lead	(6) Mercury	(7) Nickel
2004-2006	IMU, 2000	-0.018	0.126	0.261**	0.130	0.156	0.116	0.150
		(0.115)	(0.121)	(0.122)	(0.115)	(0.113)	(0.116)	(0.114)
	Density, 2000	-0.006	-0.010**	-0.008*	-0.005	-0.008*	-0.005	-0.006
		(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)
	R^2	0.09	0.05	0.07	0.07	0.05	0.06	0.05
	N	998	893	902	1,005	975	976	963
2007-2009	IMU, 2000	0.030	-0.036	0.077	0.042	0.051	0.065	0.016
		(0.117)	(0.121)	(0.114)	(0.111)	(0.108)	(0.116)	(0.113)
	Density, 2000	-0.010**	-0.012**	-0.014***	-0.010**	-0.009**	-0.007	-0.009**
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
	R^2	0.12	0.05	0.08	0.10	0.05	0.06	0.05
	N	851	830	838	874	899	835	889
2010-2012	IMU2005	0.211	0.139	0.237	0.404	0.070	0.297	-0.019
		(0.339)	(0.346)	(0.334)	(0.337)	(0.325)	(0.348)	(0.328)
	Density, 2005	-0.010*	-0.007	-0.010**	-0.010**	-0.007	-0.011**	-0.010**
		(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
	R^2	0.18	0.07	0.12	0.13	0.06	0.11	0.08
	N	639	624	635	634	675	602	681
2013-2015	IMU2005	-0.311	-0.017	-0.204	-0.215	-0.333	0.007	-0.163
		(0.426)	(0.409)	(0.366)	(0.446)	(0.354)	(0.451)	(0.354)
	Density, 2005	-0.015**	-0.006	-0.018***	-0.009	-0.006	-0.009	-0.014**
		(0.007)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)
	R^2	0.20	0.10	0.17	0.17	0.12	0.16	0.12
	N	411	416	447	389	457	375	488

Note: Standard errors errors in parentheses; * p<0.1; ** p<0.05; *** p<0.01; All regressions include state dummy variables.

Robustness Check: Conditional Quantile Regressions

In this section, we estimate conditional quantile regressions to investigate whether the responsiveness of polluters varies at different quantiles of the log pollution distribution, after controlling for state level variations such as regulatory stringency or political participation. Similar to the OLS results, we estimate separate regressions for each time period and utilizing 2000 and 2005 socioeconomic information. We adopt this strategy as the underlying pollution data spans a very great range, from micrograms to tons. At the lower end, a pollution increase of, say, 300% is a change from one innocuous level to another; at the upper end, it is immense. Thus, we do not wish to treat the same log-difference as the same amount in these two cases. We believe that our quantile regressions results are more robust to outliers as we estimate the regressions at different percentiles of the log pollution distribution.

Tables 5 and 6 present the conditional quantile regression results for the earliest two time periods—2004 to 2006 and 2007 to 2009. The marginalization index and population density are calculated from AGEB-level data for the two available years 2000 and 2005. In these regressions, we observe that the coefficient on marginalization index increases in magnitude (and significance) at higher percentiles of the log pollution distribution (75^{th} and 90^{th} percentiles). We infer that plants that are the higher percentiles of the log pollution distribution increase their pollution discharges by an even higher magnitude in more marginalized urban communities, in contrast to plants that are at the lower percentiles of the log pollution distribution. For example, at the 75^{th} percentile, the coefficient for all pollutants and total are higher than estimating the relationship at sample mean levels. A one percent increase in marginalization index in 2000 is associated with an increase in 0.2% for arsenic, 0.21% for cyanide, 0.22% for lead and mercury, 0.28% for cadmium, 0.43% for nickel and 0.44% for chromium. However, expressed as magnitudes of increase at the 75^{th} percentile pollution levels, these coefficients represent an increase of only 0.0002 kg/year for arsenic, 0.001 kg/year for cyanide, 0.005 kg/year for lead, 7.3e-05 kg/year for mercury, 0.002 kg/year for cadmium, 0.02 kg/year for nickel and 0.005 kg/year for chromium. The considerable variability in the pollution data implies that the mean varies significantly from the median or the quartiles.

For population density, plants at the middle percentiles of 50th and 75th have a significant negative coefficient—implying that potential higher burden of exposure is likely to be associated with lower pollution levels for plants within this range of the log distribution.

For Table 6, statistical significance for marginalization index is seen only at the 90th percentile for 2007-2009 pollution levels; while, population density statistically significant predominantly from the 50th percentile. Tables 7 and 8 show the quantile results estimated for the recent two time periods. Because there is a substantial loss of statistical power we present only the results with state dummy variables dropped from these regressions.

Change in Pollution Models

Next we investigate our second empirical strategy i.e. the change in pollution models. We utilize the marginalization index and population density from census 2000 and count 2005 in separate regressions, as consistency checks. Drawn from either year (2000 or 2005) local community variables are calculated prior to the change in pollution levels as the first observation is generated from the difference in log pollution in 2007-2009 and the log pollution in 2004 to 2006. As mentioned before, these regressions allow us to control for baseline levels of pollution (2004 to 2006) for each plant. We interact both measures of community characteristics with base levels of pollution.

Table 5. Conditional Quantiles of Log Water and Sewage Pollution 2004-2006 on Urban Marginalization Index (IMU) from 2000

IMU, 2000	(1) Arsenic	(2) Cadmium	(3) Chromium	(4) Cyanide	(5) Lead	(6) Mercury	(7) Nickel
10% Quantile	-0.201	0.117	0.182	0.114	0.406	0.052	0.295
	(0.301)	(0.387)	(0.306)	(0.282)	(0.362)	(0.347)	(0.275)
25% Quantile	-0.000	0.205	0.112	0.043	0.293	0.126	0.123
	(0.156)	(0.257)	(0.153)	(0.192)	(0.216)	(0.169)	(0.188)
50% Quantile	0.019	0.100	0.202*	-0.021	0.097	0.110	0.051
	(0.101)	(0.131)	(0.109)	(0.124)	(0.120)	(0.109)	(0.113)
75% Quantile	0.141	0.193*	0.310***	0.148	0.157	0.151	0.301***
	(0.098)	(0.105)	(0.112)	(0.125)	(0.119)	(0.110)	(0.107)
90% Quantile	0.090	0.183	0.279*	0.231*	0.190	0.015	-0.006
	(0.190)	(0.126)	(0.150)	(0.129)	(0.116)	(0.151)	(0.105)
Density, 2000							
10% Quantile	-0.016	-0.017	-0.012	-0.014	-0.009	-0.002	-0.007
	(0.016)	(0.017)	(0.011)	(0.015)	(0.014)	(0.016)	(0.012)
25% Quantile	-0.006	-0.002	-0.008**	0.000	-0.001	-0.006	-0.006
	(0.009)	(0.006)	(0.003)	(0.007)	(0.011)	(0.005)	(0.006)
50% Quantile	-0.005	-0.012***	-0.011**	-0.008**	-0.007*	-0.007**	-0.010***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)
75% Quantile	-0.003	-0.009*	-0.006	-0.007	-0.006**	-0.005	-0.006*
	(0.005)	(0.005)	(0.007)	(0.005)	(0.003)	(0.004)	(0.003)
90% Quantile	0.009	-0.007	0.001	-0.005	-0.004	-0.006	-0.005
	(0.008)	(0.005)	(0.008)	(0.006)	(0.006)	(0.006)	(0.005)
N	998	893	902	1,005	975	976	963

Note: Standard errors in parentheses; * p<0.1; ** p<0.05; *** p<0.01; Includes state dummy variables

Table 6. Conditional Quantiles of Log Water and Sewage Pollution 2007-2009 on Urban Marginalization Index (IMU) from 2000

IMU, 2000	**(1) Arsenic**	**(2) Cadmium**	**(3) Chromium**	**(4) Cyanide**	**(5) Lead**	**(6) Mercury**	**(7) Nickel**
10% Quantile	-0.089	-0.205	-0.168	-0.046	0.243	0.050	0.329
	(0.310)	(0.316)	(0.312)	(0.316)	(0.327)	(0.344)	(0.221)
25% Quantile	-0.009	0.018	0.003	-0.038	0.016	0.088	0.071
	(0.195)	(0.204)	(0.139)	(0.169)	(0.115)	(0.149)	(0.155)
50% Quantile	0.041	-0.052	0.064	0.137	0.067	0.063	-0.001
	(0.064)	(0.100)	(0.136)	(0.094)	(0.087)	(0.124)	(0.139)
75% Quantile	0.120	0.005	0.096	0.059	0.034	-0.016	-0.061
	(0.148)	(0.082)	(0.155)	(0.095)	(0.078)	(0.082)	(0.177)
90% Quantile	0.266	0.344***	0.442***	0.282**	0.076	0.272	0.033
	(0.233)	(0.130)	(0.158)	(0.138)	(0.123)	(0.197)	(0.172)
Density, 2000							
10% Quantile	-0.011	-0.032***	-0.013	-0.018**	-0.007	-0.013	-0.001
	(0.009)	(0.010)	(0.011)	(0.007)	(0.010)	(0.011)	(0.013)
25% Quantile	-0.017***	-0.009	-0.018***	-0.011*	-0.006	-0.005	-0.005
	(0.006)	(0.009)	(0.003)	(0.006)	(0.004)	(0.006)	(0.004)
50% Quantile	-0.004	-0.007**	-0.013***	-0.011***	-0.008*	-0.005	-0.011**
	(0.006)	(0.004)	(0.005)	(0.003)	(0.005)	(0.004)	(0.004)
75% Quantile	-0.007	-0.010**	-0.012**	-0.012**	-0.009***	-0.008	-0.014**
	(0.006)	(0.004)	(0.006)	(0.005)	(0.002)	(0.007)	(0.006)
90% Quantile	-0.011	-0.011	-0.021***	-0.007	-0.010	-0.011	-0.013**
	(0.011)	(0.008)	(0.005)	(0.005)	(0.007)	(0.012)	(0.006)
N	851	830	838	874	899	835	889

Note: Standard errors in parentheses; * p<0.1; ** p<0.05; *** p<0.01; Includes state dummy variables.

Table 7. Conditional Quantiles of Log Water and Sewage Pollution 2010-2012 on Urban Marginalization Index (IMU) from 2005

IMU, 2005	(1) Arsenic	(2) Cadmium	(4) Chromium	(5) Cyanide	(6) Lead	(7) Mercury	(8) Nickel
10% Quantile	-1.735	-1.021	0.850	-0.207	0.239	-0.432	-0.282
	(1.374)	(0.994)	(1.374)	(1.322)	(0.824)	(1.620)	(1.093)
25% Quantile	-0.467	0.435	0.344	-0.340	-0.083	0.282	-0.015
	(0.392)	(0.377)	(0.482)	(0.591)	(0.398)	(0.563)	(0.538)
50% Quantile	-0.085	0.059	0.065	0.086	-0.236	0.207	-0.215
	(0.343)	(0.253)	(0.381)	(0.294)	(0.367)	(0.260)	(0.318)
75% Quantile	0.611**	0.072	0.559	0.971***	0.178	0.349	0.476**
	(0.270)	(0.206)	(0.345)	(0.209)	(0.368)	(0.314)	(0.238)
90% Quantile	0.315	0.425	0.848*	0.477	0.299	0.923*	0.524
	(0.450)	(0.584)	(0.478)	(0.340)	(0.501)	(0.477)	(0.389)
Density, 2005							
10% Quantile	-0.025**	-0.006	-0.012	-0.010	-0.015	-0.020	-0.022
	(0.012)	(0.013)	(0.008)	(0.009)	(0.013)	(0.016)	(0.015)
25% Quantile	-0.024***	-0.008*	-0.014***	-0.017***	-0.005*	-0.011	-0.010***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.007)	(0.003)
50% Quantile	-0.022***	-0.009**	-0.017***	-0.017***	-0.005	-0.014***	-0.015***
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.002)	(0.004)
75% Quantile	-0.024***	-0.011***	-0.013**	-0.016***	-0.009**	-0.016***	-0.013***
	(0.004)	(0.003)	(0.005)	(0.003)	(0.004)	(0.003)	(0.003)
90% Quantile	-0.031***	-0.012***	-0.013	-0.017*	-0.002	-0.017***	-0.012***
	(0.007)	(0.004)	(0.008)	(0.009)	(0.005)	(0.005)	(0.005)
N	639	624	635	634	675	602	681

Note: Standard errors in parentheses; * p<0.1; ** p<0.05; *** p<0.01.

Table 8. Conditional Quantiles of Log Water and Sewage Pollution 2013-2015 on Urban Marginalization Index (IMU) from 2005

IMU, 2005	(1) **Arsenic**	(2) **Cadmium**	(3) **Chromium**	(4) **Cyanide**	(5) **Lead**	(6) **Mercury**	(7) **Nickel**
10% Quantile	-0.464	0.218	0.200	-0.064	-0.148	0.267	-0.244
	(0.862)	(0.690)	(0.397)	(1.160)	(0.549)	(1.355)	(0.531)
25% Quantile	-0.683	0.476	0.219	0.084	-0.258	0.638	-0.105
	(0.516)	(0.469)	(0.457)	(0.421)	(0.431)	(0.784)	(0.353)
50% Quantile	0.256	0.357	0.212	0.045	-0.384	0.591	0.192
	(0.449)	(0.420)	(0.334)	(0.614)	(0.373)	(0.691)	(0.260)
75% Quantile	0.073	0.167	0.468*	0.293	0.229	0.341	0.215
	(0.419)	(0.432)	(0.239)	(0.420)	(0.246)	(0.586)	(0.327)
90% Quantile	-0.493	0.352	0.343	0.104	0.581	0.030	0.502
	(0.693)	(0.695)	(0.466)	(0.429)	(0.433)	(0.421)	(0.604)
Density, 2005							
10% Quantile	-0.035***	-0.014	-0.030***	-0.023	-0.013	-0.027*	-0.018
	(0.012)	(0.017)	(0.011)	(0.015)	(0.014)	(0.015)	(0.013)
25% Quantile	-0.028***	-0.009	-0.017***	-0.015***	-0.009	-0.014*	-0.017***
	(0.009)	(0.006)	(0.006)	(0.005)	(0.006)	(0.007)	(0.006)
50% Quantile	-0.025***	-0.012**	-0.019***	-0.013**	-0.009**	-0.014***	-0.012***
	(0.007)	(0.005)	(0.003)	(0.005)	(0.004)	(0.004)	(0.005)
75% Quantile	-0.027***	-0.011**	-0.016***	-0.015***	-0.010***	-0.018**	-0.014***
	(0.007)	(0.005)	(0.003)	(0.005)	(0.004)	(0.008)	(0.005)
90% Quantile	-0.034***	-0.015**	-0.027***	-0.028***	-0.014***	-0.029***	-0.022***
	(0.008)	(0.006)	(0.008)	(0.005)	(0.005)	(0.011)	(0.008)
N	411	416	447	389	457	375	488

Note: Standard errors in parentheses; * p<0.1; ** p<0.05; *** p<0.01.

Table 9. Change in Pollution, 2004-2015, on Baseline Pollution, Urban Marginalization Index (IMU) and population density, from 2000 and 2005

Variables	(1) Pooled*	(2) Arsenic	(3) Cadmium	(4) Chromium	(5) Cyanide	(6) Lead	(7) Mercury	(8) Nickel
Base pollution	-0.319***	-0.319***	-0.312***	-0.297***	-0.320***	-0.326***	-0.357***	-0.272***
	(0.014)	(0.038)	(0.042)	(0.037)	(0.037)	(0.040)	(0.044)	(0.037)
Base*IMU	0.020***	0.029*	0.030*	0.027*	0.014	0.027*	0.024	0.013
	(0.006)	(0.016)	(0.017)	(0.015)	(0.015)	(0.016)	(0.018)	(0.015)
density	-0.004***	-0.007*	-0.001	-0.005	-0.003	-0.002	-0.002	-0.006**
	(0.001)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)
Base* density	0.000	0.000	0.000	-0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
IMU, 2000	0.051	0.100	0.079	0.105	0.107	-0.014	0.143	-0.061
	(0.032)	(0.098)	(0.096)	(0.081)	(0.082)	(0.071)	(0.125)	(0.068)
R^2	0.17	0.17	0.16	0.18	0.17	0.17	0.17	0.14
N	7,809	1,137	1,041	1,053	1,145	1,186	1,046	1,201
Base pollution	-0.325***	-0.325***	-0.324***	-0.293***	-0.328***	-0.335***	-0.365***	-0.279***
	(0.014)	(0.038)	(0.041)	(0.035)	(0.036)	(0.038)	(0.042)	(0.036)
Base*IMU	0.051***	0.084*	0.071	0.091**	0.026	0.072	0.058	0.033
	(0.016)	(0.046)	(0.050)	(0.042)	(0.042)	(0.045)	(0.049)	(0.044)
density	-0.004***	-0.008*	-0.002	-0.005	-0.004	-0.002	-0.002	-0.006**
	(0.001)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.003)
Base* density	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
IMU, 2005	0.178**	0.309	0.255	0.344	0.263	-0.017	0.527	-0.157
	(0.089)	(0.288)	(0.264)	(0.225)	(0.233)	(0.200)	(0.340)	(0.196)
R^2	0.16	0.17	0.16	0.18	0.17	0.17	0.17	0.14
N	7,961	1,161	1,062	1,070	1,164	1,210	1,070	1,224

Note: Baseline pollution represents log pollution levels in 2004-2006; Standard errors in parentheses; * p<0.1; ** p<0.05; *** p<0.01; Includes pollutant dummy variables.

Results presented in Table 9 above shows that as expected baseline levels of pollution as captured by pollution reports from the first time period (2004-2006) have a statistically significant influence in explaining change in pollution levels within the same plant, but in subsequent time periods. This leads us to believe that concerns on data validity and under reporting might not be so pervasive as suspected initially. Plants that reported high pollution levels in the beginning years of mandatory reporting are reducing their pollution in more recent years. This variable allows us to control for all plant specific factors that explain high (or low) pollution reports of each plant.

We interpret the positive coefficient on the interaction term of baseline pollution and marginalization index as consistent with our previous evidence on environmental justice concerns. The positive sign on the coefficient can be interpreted as the relationship between marginalization index and pollution becomes stronger or more positive for plants that are at the upper end of the pollution distribution. The coefficient on the un-interacted marginalization index in the pooled model +0.051 (column 1) can be interpreted as the relevant effect where the community factors are zero. At the median of the baseline pollution distribution, the marginal impact remains in the neighborhood to +0.051. However, at the third quartile of the baseline pollution distribution, the marginal impact increases to +0.070; while at the 90th percentile for baseline pollution, the marginal impact increases by almost five folds to +0.245.

We interpret the results for the individual pollutants; in general, the results are weaker than the pooled model. For arsenic releasing plants at the 75th and 90th percentile of baseline, the marginal impact of marginalization increases from +.10 to +.11 and +.19 respectively. For cadmium releasing plants, at the 75th and 90th percentile of baseline, the marginal impact increases from +0.079 to +.10 and +0.32 respectively. For chromium plants, at 75th and 90th percentile of baseline, the marginal impact increase from +.105 to +.132 and +.45 respectively. For cyanide plants, at 75th and 90th percentile of baseline, the marginal impact increase only from +.107 to +.12 and +.193 respectively. For lead plants, at 75th and 90th baseline levels, the marginal effect increases from -.014 to as much as +.07 and +.70 respectively. For mercury plants, at 75th and 90th baseline, the marginal effect increase only from +.143 to +.144 and +.155 respectively. For nickel plants, at 75th and 90th baseline, the effect of marginalization increase from -0.061 to as much as +.006 and +0.53 respectively. Again, we see a great variability in the exact magnitudes of the coefficients for individual pollutants; however, the direction of total impact of marginalization remains consistently positive even when considering changes in pollution within plants.

As for population density, we find that this variable is not a significant factor in explaining changes in pollution levels within plants; especially after controlling for baseline pollution levels of each plant. We infer that for plants that are heavily polluting in the initial years, environmental justice concerns i.e. marginalization index might be a predominant factor in contrast to population density. This result holds for the 2005 marginalization index presented in the bottom panel of Table 9.

Panel Data Models

Our third and last empirical strategy is to estimate a panel data model with fixed effects as this captures all other plant specific effects that influence pollution levels of plants. Table 10 below shows that increase in marginalization is associated with an increase in pollution, after we control for plant fixed effects. Standard errors are clustered at the level of municipalities to address spatial correlation issues of plants that are located nearby or in the same municipality. The second panel of Table 10 presents the results estimated on the sample of 1,595 facilities with marginalization data for both years 2000 and 2005. Results are very similar to the overall sample; however, given the considerable unbalanced nature of the pollution data we could not estimate these models on a true balanced panel.

Table 10. Fixed Effects of Log Water and Sewage Pollution, 2004-2015 on Marginalization Score, 2000 and 2005

Variables	**(1) Arsenic**	**(2) Cadmium**	**(3) Chromium**	**(4) Cyanide**	**(5) Lead**	**(6) Mercury**	**(7) Nickel**
Marginalization	0.258	0.305	0.619**	0.327	0.465	0.839**	0.501
Score, 2000-2005	(0.323)	(0.405)	(0.305)	(0.313)	(0.395)	(0.346)	(0.329)
Density, 2000-2005	0.033	0.022	-0.002	-0.003	0.003	-0.008	0.009
	(0.025)	(0.018)	(0.020)	(0.020)	(0.018)	(0.023)	(0.015)
R^2	0.01	0.02	0.03	0.00	0.01	0.01	0.02
N	2,899	2,763	2,822	2,902	3,006	2,788	3,021
Sample of 1,595 plants							
Marginalization	0.258	0.305	0.619**	0.328	0.465	0.840**	0.501
Score, 2000-2005	(0.323)	(0.406)	(0.305)	(0.313)	(0.395)	(0.346)	(0.329)
Density, 2000-2005	0.033	0.022	-0.002	-0.003	0.003	-0.008	0.009
	(0.025)	(0.018)	(0.020)	(0.020)	(0.018)	(0.023)	(0.015)
R^2	0.01	0.02	0.03	0.00	0.01	0.01	0.02
N	2,853	2,716	2,774	2,857	2,958	2,745	2,972

Note: Clustered standard errors within municipality in parentheses; * p<0.1; ** p<0.05; *** p<0.01; All regressions include time period dummy variables.

The coefficient in column (3) can be interpreted as increase in marginalization by one unit leads to increase in chromium pollution by almost 62%. A one-unit increase in marginalization score represents more than doubling in terms of its average level (101%). In elasticity terms, a one percent increase in marginalization leads to 0.61% of average levels of chromium pollution (i.e. an increase of 1.22 kg/year on average for 2004-2015). The coefficients for the other pollutants vary greatly with elasticities ranging between 0.26% for arsenic, 0.3% for cadmium, 0.32% for cyanide, 0.46% for lead, 0.5% for nickel, and 0.83% for mercury. In general, the magnitudes are larger than the cross sectional estimates presented before. In terms of average pollution levels over the entire time period, these estimates represent 0.05 kg/year for arsenic, 0.13 kg/year for cadmium, 0.12 kg/year for cyanide, 4.1 kg/year for lead, 4.2 kg/year for nickel, and 3.4 kg/year for mercury. Considerable variability in pollution data means that the sample means for the entire 2004-2015 period are at times higher than the means for each single 3-year period (nickel, lead and mercury e.g.).

Population density, on the other hand, does not exhibit a statistically significant robust influence on pollution levels over time. We speculate that inclusion of plant level fixed effects might subsume the influence of variations in density surrounding each plant; especially if it is a slow moving variable.

Conclusions

In this paper, we show that major toxic releasing plants in urban areas in Mexico pollute more in marginalized neighborhoods. We address reverse causality of pollution influencing local socioeconomic characteristics (through the channel of residential sorting), by incorporating marginalization and population density measures from years prior to the pollution reports. We build a detailed dataset linking plant level pollution to surrounding community characteristics such as marginalization and density. We are one of the few studies in environmental justice literature to focus on water pollution. We find robust evidence on marginalization employing both cross-sectional (mean and quantile regressions), and panel data techniques by exploiting within plants variations in pollution levels.

We estimate these models for individual pollutants in separate regressions. We believe that the strongest evidence on higher pollution in more marginalized communities come from the change in pollution levels within plants and the fixed effect models. In the change in pollution models we find a positive interaction of baseline pollution levels and marginalization index. We interpret this sign as the marginal impact of marginalization becomes stronger with increase in baseline levels of pollution. With an increase in baseline pollution from 0 to the 90th percentile, the marginal impact of marginalization increases from +.143 to +.155 for mercury releasing plants, from +.107 to +.193 for cyanide releasing plants, from +.10 to +.19 for arsenic releasing plants, from +0.079 to +0.32 for cadmium releasing plants, from +.105 to +.45 for chromium releasing plants, from -0.061 to as much as +0.53 for nickel releasing plants, and from -.014 to as much as +.70 for lead releasing plants. We infer that heavily polluting plants are more responsive to local marginalization index than population density.

The fixed effects models show that an increase in marginalization index by one percent is associated with an increase in toxics releases that vary between 0.26% for arsenic, 0.3% for cadmium, 0.32% for cyanide, 0.46% for lead, 0.5% for nickel, 0.61%

for chromium, and 0.83% for mercury pollution. These elasticities exhibit great variability which justifies disaggregation of toxics. However, expressed as a magnitude of increase from average pollution levels for the entire time period 2004 to 2015, these estimates represent an increase of 0.05 kg/year for arsenic, 0.13 kg/year for cadmium, 0.12 kg/year for cyanide, 1.22 kg/year for chromium, and as much as 4.1 kg/year for lead, 4.2 kg/year for nickel, and 3.4 kg/year for mercury. We interpret these coefficients as significant given that releases of 1 kg/year mandates pollution reporting under the law.

In conclusion, our qualitative result that higher toxics pollution is witnessed in more marginalized neighborhoods is supported by the quantitative estimates found in the various empirical specifications

References

Arora, S. and T. N. Cason (1998). Do community characteristics influence environmental outcomes? *Journal of Applied Economics, I* (2), 413-453.

Ash, M. and T.R. Fetter (2004). Who lives on the wrong side of the environmental tracks? Evidence from the EPA's Risk-Screening Environmental Indicators Model. *Social Science Quarterly*, 85 (2), 441-462.

Blackman, A., M. Batz, and D. Evans (2004). Maquiladoras, air pollution, and human health in Ciudad Juárez and El Paso. *RFF Discussion Paper*, 3–18, Resources for the Future, Washington, DC.

Bouwes, N.W., S.M. Hassur, and M.D. Shapiro (2003). Information for empowerment: the EPA's Risk-Screening Environmental Indicators Project. J.K. Boyce, B.G. Shelley (eds.), *Natural Assets: Democratizing Environmental Ownership*, Princeton Island Press, Washington D.C.

Brooks, N. and R. Sethi (1997). The distribution of pollution: Community characteristics and exposure to air toxics. *Journal of Environmental Economics and Management, 32*(2), 233-250.

Chakraborti, L. and M. Margolis (2017). Do industries pollute more in poorer neighborhoods? Evidence from toxic releasing plants in Mexico. *Economics Bulletin*, AccessEcon, vol. 37(2), 853-870.

CEC (2018). *Taking Stock: North American Pollutant Releases and Transfers*, Vol. 15. Montreal, Canada: Commission for Environmental Cooperation. 154 pp.

CEC (2011). *Taking Stock: North American Pollutant Releases and Transfers*, Vol. 13 Montreal, Canada: Commission for Environmental Cooperation.

CEC (2009). *Taking Stock: 2005 North American Pollutant Releases and Transfers*, Montreal, Canada: Commission for Environmental Cooperation, June 2009.

Cruz, G. (2004). *El efecto del regulador y de la comunidad sobre el desempeño ambiental de la industria en Bogotá, Colombia*. *Revista Desarrollo y Sociedad*, 54 (September 2004) https://doi.org/10.13043/dys.54.6.

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

Dasgupta, S., H. Hettige, and D. Wheeler (2000). What improves environmental compliance? Evidence from Mexican industry. *Journal of Environmental Economics and Management, 39*(1), 39-66.

Dasgupta, S., R. Lucas, and D. Wheeler (2002). Plant size, industrial air pollution, and local incomes: Evidence from Mexico and Brazil. *Environment and Development Economics 7*(2), 365-381.

Eicker, M., T. Ruddy, R. Zah, and H. Hurni (2010). Potentials of Latin American pollutant release and transfer registers as a source of local data for environmental assessments. Paper presented at the Annual Meeting of the International Association for Impact Assessment (IAIA), Geneva, Switzerland.

Feres, J. and A. Reynaud (2012). Assessing the Impact of Formal and Informal Regulations on Environmental and Economic Performance of Brazilian Manufacturing Firms. *Environmental and Resource Economics* 52(1): 65–85.

Germani, A.R., P. Morone and G. Testa (2014). Environmental justice and air pollution: A case study on Italian provinces. *Ecological Economics*, Volume 106, Pages 69-82.

Gray, W. B., R. J. Shadbegian, and A. Wolverton (2012). Environmental justice: Do poor and minority populations face more hazards? *Oxford Handbook of the Economics of Poverty.* Philip N. Jefferson (Ed.). Oxford University Press.

Grineski, S. E., T. W. Collins, and M. L. R. Aguilar (2015). Environmental injustice along the US-Mexico border: Residential proximity to industrial parks in Tijuana, Mexico. *Environmental Research Letters, 10 (9)* http://iopscience.iop.org/1748-9326/10/9/095012.

Hamilton, J. (1995). Testing for environmental racism: prejudice, profits, political power? *Journal of Policy Analysis and Management, 14*(1),107-132.

Helland, E. and A. B. Whitford (2003). Pollution incidence and political jurisdiction: Evidence from the TRI. *Journal of Environmental Economics and Management, 46*(3), 403-424.

Ma, C. (2010). Who bears the environmental burden in China—An analysis of the distribution of industrial pollution sources? *Ecological Economics*, Volume 69, Issue 9, pp. 1869-1876.

Pargal, S. and D. Wheeler (1996). Informal Regulation of Industrial Pollution in Developing Countries: Evidence from Indonesia. *Journal of Political Economy*, 1996, vol. 104, issue 6, pp. 1314-27.

Saha, S. and R.D. Mohr (2013). Media attention and the Toxics Release Inventory. *Ecological Economics* 93:284–291.

Seroa da Motta, R. (2006). Analyzing the environmental performance of the Brazilian industrial sector. *Ecological Economics* 57(2):269–281.

Shapiro, M. D. (2005). Equity and Information: Information Regulation, Environmental Justice, and Risks from Toxic Chemicals. *Journal of Policy Analysis and Management* Vol. 24, No. 2, pp. 373-398.

Wolverton, A. (2009), "Effects of Socio Economic and Input Related Factors on Polluting Plants' Location Decisions." *The B.E. Journal of Economic Analysis & Policy (Advances)* 9(1):132.

Zwickl, K. M. Ash and J. K. Boyce (2014). Regional variation in environmental inequality: Industrial air toxics exposure in U.S. cities. *Ecological Economics*, Volume 107, pp. 494-509.

.