

## Who Lives on the Wrong Side of the Environmental Tracks? Evidence from the EPA's Risk-Screening Environmental Indicators Model

Item Type	Working Paper
Authors	Ash, Michael;Fetter, T. Robert
DOI	<a href="https://doi.org/10.7275/1274557">10.7275/1274557</a>
Download date	2025-06-18 07:11:19
Link to Item	<a href="https://hdl.handle.net/20.500.14394/40204">https://hdl.handle.net/20.500.14394/40204</a>

POLITICAL ECONOMY RESEARCH INSTITUTE

**PERI**

University of Massachusetts Amherst

**Who Lives on the Wrong Side  
of the Environmental Tracks?  
Evidence from the EPA's Risk-Screening  
Environmental Indicators Model**

**Michael Ash and T. Robert Fetter**

2002

**POLITICAL ECONOMY  
RESEARCH INSTITUTE**

10th floor Thompson Hall  
University of Massachusetts  
Amherst, MA, 01003-7510  
Telephone: (413) 545-6355  
Facsimile: (413) 545-2921  
Email: [peri@econs.umass.edu](mailto:peri@econs.umass.edu)  
Website:  
<http://www.umass.edu/peri/>



**WORKINGPAPER SERIES**

Number 50

**Who Lives on the Wrong Side of the Environmental Tracks?**  
**Evidence from the EPA's Risk-Screening**  
**Environmental Indicators Model**

**Michael Ash and T. Robert Fetter\***

December 2002

Program on Development, Peacebuilding, and the Environment  
Political Economy Research Institute  
University of Massachusetts, Amherst

---

\* Michael Ash, mash@econs.umass.edu, is assistant professor of economics and public policy, University of Massachusetts, Amherst. T. Robert Fetter, fettert@saic.com, is an environmental economist with Science Applications International Corporation (SAIC). Ash and Fetter are joint co-authors. We thank Jim Boyce and John Stranlund for their invaluable guidance, and we are deeply grateful to Nick Bouwes and Steven Hassur of the Environmental Protection Agency who have painstakingly developed and graciously provided us with access to the Risk Screening Environmental Indicators data.

**Abstract:** This study analyzes the social and economic correlates of air pollution exposure in U.S. cities using a unique dataset created as a by-product of the EPA's Risk-Screening Environmental Indicators model and finds evidence of disproportionate exposure to environmental hazards in communities with higher concentrations of lower-income people and people of color. We improve on previous studies of environmental inequality in three ways. First, where previous studies focus on the proximity to point sources and the total mass of pollutants released, our measure of toxic exposure reflects atmospheric dispersion and chemical toxicity. Second, we analyze the data at a fine level of geographic resolution. Third, we control for substantial regional variations in pollution, allowing us to identify exposure differences both within cities and between cities. We combine 1998 data on toxicity-adjusted exposure to air pollution with 1990 Census block group data for urbanized areas. We find that blacks tend to live both in more polluted cities in the U.S. and in more polluted neighborhoods within cities. Hispanics live in less polluted cities on average, but they live in more polluted areas within cities. We find an extremely consistent income-pollution gradient, with lower income people significantly more exposed. Our findings highlight the importance of controlling for inter-regional variation in pollution levels in studies of the demographic correlates of pollution.

---

## Introduction

The analysis of environmental justice examines differential availability of environmental amenities or exposure to environmental disamenities on the basis of socioeconomic, ethnic, or racial differences. Depending on how one defines environmental justice, the existence of such inequality may itself constitute a finding of injustice, or interpretations of the cause of the difference may matter as well. Even focusing on the relatively simple question of what demographic attributes correlate with environmental quality, there is little consensus among researchers on methodology or results. Using different indicators of environmental disamenities, units of spatial analysis, explanatory variables, and theoretical frameworks, researchers have found evidence for widely different conclusions.

With respect to the recurrent themes in recent analyses, we contribute three methodological improvements. First, we use a realistic indicator of exposure to pollution which is based upon the U.S. Environmental Protection Agency's Toxics Release Inventory (TRI) but takes into account relative toxicity and chemical fate and transport. Second, we examine correlations at geographically small units, Census block groups, so as to avoid the "ecological fallacy," i.e., reaching conclusions from a large unit of analysis that do not hold at smaller resolution due to spatial heterogeneity. Third, we incorporate regional variations into a national analysis in a novel way.

Across all urban areas in the contiguous U.S., we find that block groups with higher proportions of blacks tend to experience higher levels of toxicity-adjusted exposure to air pollution from TRI-reporting facilities than do predominantly white block groups, whereas block groups with higher proportions of Hispanics tend to experience lower levels of pollution. However, a regression model that compares block groups only to other block groups in the same city shows that block groups with more Hispanics and those with more blacks have more pollution on

average. Taken together, these results imply that blacks tend to live in more polluted cities than do whites, and also tend to live in more polluted block groups within cities. While Hispanics tend to live in less polluted cities than do whites, on average they too live in more polluted block groups within cities.

## **Literature Review**

In this section, we survey recent practice in the environmental justice literature, with particular focus on the areas in which we offer methodological improvements. We then briefly address the conventional practices that we follow.

### *Choice of Pollution Indicator*

Many studies of the demographic correlates of pollution use as the dependent variable the presence or absence of a polluting facility, such as a toxic storage and disposal facility (TSDF), in an area (e.g. United Church of Christ 1987, Mohai and Bryant 1992, Anderton et al. 1994, Oakes 1997, Been and Gupta 1997, Boer et al. 1997, Pastor 2001). The presence of a polluting facility within neighborhood boundaries has some validity as an indicator of environmental quality, especially as perceived by residents, but as a sign of actual exposure to hazard its validity is limited. The mass of pollutants reported in the TRI released within community borders is a closer proxy to actual environmental quality, since it represents a fuller set of polluting facilities and accounts for heterogeneous amounts of pollution released. Most studies that use TRI data focus on the mass of pollutants released; see, e.g., Bowen et al. (1995), Kriesel, Centner, and Keeler (1996), and Arora and Cason (1999).

Recognizing that mass of pollutants released within neighborhood boundaries remains a relatively blunt measure of environmental quality, a few researchers have adjusted TRI data for toxicity and atmospheric dispersion. Glickman and Hersh (1995) estimate risks from chronic exposure to industrial facilities in Allegheny County, Pennsylvania. Using TRI and other data, adjusted for toxicity and wind patterns, they find that Census block groups with more blacks, poor people, and people over age 65 face higher risks compared to the rest of the population. Brooks and Sethi (1997) find that U.S. zip code neighborhoods with more blacks experience greater pollution, as measured by an index of TRI releases adjusted for health effects from chronic exposure and distance from pollution source. This relationship holds even when income, education, percent of residents in urban areas, housing value, percent of workforce in manufacturing, and population density are held constant. Using a pollution index based on TRI releases, adjusted for chemical toxicity and accounting for chemical fate and transport using EPA-reviewed models and databases, Bouwes, Hassur, and Shapiro (2001) find that densely populated square-kilometer neighborhoods with more blacks, Hispanics, Asians, and unemployed residents tend to be more polluted than other densely populated neighborhoods.

### *Unit of Spatial Analysis*

A key choice for researchers is the unit of spatial analysis, and this choice may have significant effects on the results of a study. One of the first studies to receive national attention (United Church of Christ 1987) compared the demographics of 5-digit zip codes containing commercial TSDFs to those of zip codes without TSDFs. Hird and Reese (1998) explored demographic correlations with a composite index of 29 indicators of environmental quality at the county level. Brooks and Sethi (1997) use a relatively sophisticated indicator of pollution, but define neighborhoods in terms of 5-digit zip code. Such studies are open to the possibility of ecological fallacy, that is, that the correlations identified at large units of analysis may not hold at finer resolution. Most researchers would probably agree in the abstract that “the area chosen for analysis should correspond to the likely areal distribution of possible harm” (Anderton et al. 1994, 128). Uncertainty about the range of possible harm from point sources makes difficult the precise definition of the appropriate area.

### *Controlling for Regional Characteristics*

A third issue that is increasingly prominent in the literature is what factors should be held constant to control for regional variation. Firms choose the locations of polluting facilities on the basis of linkages to output markets, existing facilities, and transportation networks, which in turn are determined partly by historical patterns of economic development. Moreover, local and state governments conduct much industrial location policy and implement environmental regulations; for these policymakers, the relevant social patterns of pollution are local. National-level analysis may overlook regional variations either with regard to “base” levels of environmental quality or with regard to the structural relationship between race and other variables and pollution.

Studies have controlled for broad regional variations by allowing the base level of pollution to differ by urban or rural status, or by performing separate regressions by region, for example, South versus non-South (Arora and Cason 1999), South, West, and remainder of U.S. (Hird and Reese 1998), or the nine EPA regions (Anderton et al. 1994). Other studies have controlled for variations by choosing the comparison population carefully. Anderson et al. (1994) compared demographics between Census tracts containing TSDFs and tracts without TSDFs within the same Metropolitan Statistical Area (MSA) or rural county. They reasoned that only tracts in the same MSA or rural county as a facility could serve as possible alternative sites for the same market. Their choice of the comparison population contrasts with that of the United Church of Christ (1987), which compared demographics between tracts containing TSDFs and all tracts in the U.S. without TSDFs. As Been and Gupta (1997) note, the choice of Anderson et al. (1994) eliminated about 18,000 non-host tracts from the comparison population and reduced the observed differences between the racial and ethnic composition of the host and non-host tracts.

Kriesel, Centner and Keeler (1996) report a positive correlation between the total mass of TRI releases within one mile of a block group and the percent of nonwhite residents in the block group in both Georgia and Ohio, holding constant the percent of the population in poverty and the percent voting in the 1992 national election. When they add six so-called “non-discriminatory industrial location factors” – number of manufacturing employees, presence of an interstate highway in the county, county’s average weekly manufacturing wage, percent of adults

with high school degrees (to proxy labor productivity), average house value (to proxy the cost of living), and population density – the positive association with percent nonwhite disappears.

The inclusion of supposedly non-discriminatory industrial location factors as control variables in a regression is itself problematic, for at least two reasons. First, it is difficult to identify variables that proxy the desired information without being causal factors or effects of facility location in their own right. For example, in addition to affecting productivity, educational level could affect the propensity of a community to accept or resist polluting industrial facilities; or housing values could be an effect rather than a cause of facility location. Indeed, Kriesel, Centner, and Keeler found unexpected signs on two of the six “industrial location factors” in the Georgia regression (percent of adult high school graduates, average weekly manufacturing wage) and on four of the six factors in the Ohio regression (percent of adult high school graduates, average weekly manufacturing wage, size of manufacturing workforce, population density). Second, so-called “non-discriminatory” factors may not actually be non-discriminatory. For instance, Rabin (1989) found evidence in the zoning histories of twelve cities that local planners changed the zoning of residential land occupied at the time mainly by low-income blacks to industrial or commercial. Rabin found that this practice often led blacks to leave their old neighborhoods, but in many cases some continued to live side-by-side with new industrial land uses. If policy makers created conditions in the mid-twentieth century that encouraged industrial development in and near black neighborhoods, then factors such as the proximity of interstate highways and housing value may not, in fact, be non-discriminatory.

#### *Additional Control Variables*

Virtually every major environmental justice study has included median household or family income as an explanatory variable. Incomes tend to be lower among blacks and Hispanics than among whites in the U.S. To the degree that income correlates with housing value, land values are likely to be lower in more polluted neighborhoods, either because a polluting facility reduces environmental quality or because a firm siting a new facility will choose an area with lower land costs if other factors are equal. Areas with low-income residents may not be able to mount effective citizen opposition to polluters. For these reasons, areas with lower incomes are likely to be more polluted. If correlation between ethnicity or race and pollution disappears when income is held constant, then we have not found environmental racism, *per se*, but have identified environmental inequality and, some would argue, environmental injustice. Researchers who use a linear specification in income tend to find a negative correlation between median income and pollution, but several researchers have observed a quadratic relationship between pollution and income. Of these, most find that median income is positively correlated with pollution at low levels of income, but as income rises, the correlation becomes negative (e.g. Been and Gupta 1997, Boer et al. 1997, Brooks and Sethi 1999). The dominant explanation for this inverse-U shape is that in the poorest neighborhoods there is insufficient economic activity to generate pollution, and in the richest neighborhoods people have the financial or political power to obtain high environmental quality; thus most pollution falls on people in the middle.

There is no clear expectation about whether population density should correlate positively or negatively with pollution. One might expect a positive correlation because there is likely to be more economic activity and thus more pollution in areas with more people. On the other hand,

planners and public health officials likely work to reduce pollution in densely populated places. Boer et al. (1997) find that population density is not an important predictor once they control for the proportion of land devoted to industry, utilities, transportation, and communication, and suggest that in most studies (which lack land-use data) population density stands in for industrial land use. Unfortunately, data on either land use or zoning are generally unavailable or expensive. In any case, controlling for population density is a frequent if not universal practice in environmental justice studies. If denser neighborhoods are also neighborhoods with more people of color, then a finding of disparate pollution burdens in minority neighborhoods when population density is held constant offers, if nothing else, some additional information about what is not the underlying cause of the correlation.

Many studies include additional variables as demographic and socioeconomic correlates. Although there is little consensus about which variables are important to include – either as controls, to narrow the possible reasons for racial and ethnic patterns that may be observed, or as variables whose correlations with pollution are interesting in their own right – several variables are commonly used, including education, voter turnout, the percent of owner-occupied housing units, and the percent of vacant housing units. For the most part, the direction of correlation for these other variables is difficult to predict, especially in a multivariate model with controls for income. As noted already, education levels could be positively or negatively correlated with pollution. The expected correlation between pollution and the percent of vacant housing units may also be argued both ways: a positive correlation may result because neighborhoods with more vacant housing might have less social cohesion and thus pose less opposition to polluting facilities; on the other hand, high vacancy rates may reflect low economic activity that may be associated with low pollution.

Hamilton (1993) observed a negative correlation between facility expansion plans within a county and that county's voter turnout in the 1980 presidential election. Since Hamilton's study, most of the researchers who have examined correlations between pollution and voter turnout (e.g. Brooks and Sethi 1997, Arora and Cason 1999) have found the same negative correlation. The prevailing explanation is that either higher voter turnout influences decision makers, or that voter turnout is a proxy for collective civic spirit or "social capital," and thus reflects the likelihood and capacity of a community to engage in collective action to resist polluting facilities. Studying the capacity expansion plans of commercial hazardous waste facilities, most researchers who include a variable for "collective action potential" follow this lead, using voter turnout. In contrast to the usual results, Kriesel, Centner, and Keeler (1996) find a positive correlation between voting and pollution and conclude that "industrial locations which generate toxic releases are perceived as positive outcomes" (496). All else equal, however, most people likely prefer industrial economic activity that generates fewer toxic releases.

## **Methodology**

In this section we develop our multivariate model of the social and demographic correlates of TRI pollution exposure. The measure of exposure to environmental hazards used in the present study captures great detail about the dispersion and toxicity of pollution. Our pollution measure



also has very fine geographic resolution, which allows us to use the smallest unit at which most Census socioeconomic data are available, the block group. By analyzing exposure rather than proximity to source and by using Census block groups as the unit of analysis, we address the problem posed by Anderton et al. (1994) of determining the “areal distribution of possible harm”; in addition, we avoid the ecological fallacy to the extent feasible. We control for regional variation in a novel way, by including fixed effects for Census-defined urbanized areas. In so doing, we allow for different base levels of pollution in individual cities, thereby controlling for city-specific historical patterns of economic development. At the same time, by restricting the regression coefficients to be identical across cities, we are able to pose the question of whether – at a national level – there are common demographic characteristics of “the wrong side of the environmental tracks.”

Although many studies group all racial and ethnic minorities together, we include Hispanics, non-Hispanic blacks, and non-Hispanic Asians and Pacific Islanders as our separate categories and find that different groups have different patterns of exposure. We do not include Native Americans, who have very low representation in U.S. cities, or the residual “other race” category in the Census.

Based on the discussion of explanatory variables in the review of the literature, we include the percent of residents with less than a high school education, the percent of housing units that are vacant, population density, and both median household income and its square. While most studies compute voter turnout as voters divided by registered voters, we follow Arora and Cason (1999) in computing voter turnout as the number of people who voted divided by the total population as the measure of mobilization. This variable is available only at the county level, so we assign the county-level voter turnout to each block group within the county; for all other variables, we use black-group level data.

We also include non-English speaking households as an independent variable and predict a positive relationship with pollution for two reasons: first, neighborhoods with fewer English speakers are less likely to oppose an environmental hazard effectively; second, even controlling for income, non-English speakers may face greater constraints in the credit and housing markets. We include the percentage of households with capital income, intended as a proxy for the relative wealth of neighborhoods. We note, however, that this variable may be a weak indicator of wealth because it measures only the percent of households receiving these types of income and not the average amount per household. Lastly, we include the percent of housing units that are owner-occupied. For this variable, we predict a negative correlation with pollution: Residents of neighborhoods with a high share of rental units are likely to be more transient, have less social cohesion, and therefore be less effective in resisting the siting of polluting facilities.

We estimate three multivariate specifications of the pollution exposure model. In the equations below, the  $\beta$ s are, throughout, conforming vectors of coefficients;  $i$  indexes block group,  $c$  indexes county, and  $a$  indexes urbanized area. By including fixed effects for the 393 urbanized areas in some specifications, we control for the component of the error term associated with each urbanized area. We thereby identify the demographic correlates of pollution between neighborhoods within urbanized areas. The fixed effect captures the base level of all variables

for the area; hence, in the fixed-effect specification, the coefficients are identified on the basis of variation within each area.

The first specification uses only ethnicity and race as independent variables: the percent of Hispanics, percent of non-Hispanic blacks, and percent of non-Hispanic Asian-Americans. Formally, the model can be expressed as:

$$\text{SCORE}_{ica} = \beta_0 + \text{MINORITY}_{ica} \beta_{\text{MINORITY}} + \delta_a + \varepsilon_{ica},$$

where SCORE equals the measure of pollution or, in the case of the dichotomous models described below, an indicator for membership in the relevant quantile of each city, and MINORITY is a vector with the percent of Hispanic, non-Hispanic black, and non-Hispanic Asian residents. The first component of the error term,  $\delta_a$ , is a fixed effect for the entire urbanized area, and the second component,  $\varepsilon_{ica}$ , is a standard white-noise error for the block-group neighborhood.

In the second specification, we incorporate median household income as a control variable in quadratic form:

$$\text{SCORE}_{ica} = \beta_0 + \text{MINORITY}_{ica} \beta_{\text{MINORITY}} + \text{INCOME}_{ica} \beta_{\text{INCOME}} + \delta_a + \varepsilon_{ica}$$

where INCOME is a vector of median household income and its square.

The third model adds six additional explanatory variables at the block-group level designated by the vector  $X_{ica}$ : population density (persons per square kilometer); the percent of households in which English is not the primary language; the percent of adults without a high school diploma or the equivalent; the percent of households receiving interest, dividend, or net rental income; the percent of vacant housing units; and the percent of occupied housing units that are occupied by the owner. The third specification also includes  $\text{VOTE}_{ca}$ , the percent of the total population that voted in the 1992 national election in the county which contains the block group. We therefore estimate:

$$\text{SCORE}_{ica} = \beta_0 + \text{MINORITY}_{ica} \beta_{\text{MINORITY}} + \text{INCOME}_{ica} \beta_{\text{INCOME}} + X_{ica} \beta_x + \text{VOTE}_{ca} \beta_{\text{VOTE}} + \delta_a + \varepsilon_{ica}$$

For some of these additional variables the correlation with pollution is of interest, but our principal interest is whether the correlation between the racial/ethnic variables and pollution remain stable when additional variables are added.

## Data

### *Pollution Data*

The pollution data for this paper are derived from the EPA's Risk-Screening Environmental Indicators model. Because it incorporates detailed data on the toxicity and dispersion of chemical releases, the Indicators model gives more realistic information on potential human health effects from air pollutants than has been available for most previous studies. The output of the Indicators model is a unitless measure, intended for relative comparisons, rather than a physically denominated measure of the quantitative risk or exposure potential in a particular location.

The pollution sources considered in this paper are those facilities that report emissions to the EPA's Toxics Release Inventory (TRI). The provisions of the 1986 Emergency Planning and Community Right-to-Know Act (EPCRA) related to the TRI now require operations engaged in manufacturing (i.e., Standard Industrial Classification (SIC) codes 20 through 39), metal and coal mining, hazardous waste treatment and disposal, and solvent recovery, as well as electrical utilities, chemical distributors, petroleum terminals, and federal facilities, to report their releases of designated pollutants to air, water, and land if the operations exceeded specified thresholds of employee size and chemical use. For air releases, TRI guidelines for the 1998 reporting year required facilities to report emissions of 604 different chemicals and chemical categories (Office of Pollution Prevention and Toxics 1999). Facilities must estimate and report both intentional ("stack") and unintentional ("fugitive") releases, with some monitoring and oversight by EPA. In 1998, 23,396 facilities reported direct releases of 2.1 billion pounds of chemicals to air, 4.3 billion pounds to land, 0.3 billion pounds to underground injection, and 0.2 billion pounds to surface water (EPA, 2000). In addition, facilities reported 0.4 billion pounds of "off-site" releases, meaning that chemicals were transported to another facility and released from there. Most off-site releases in 1998 were to land.

The EPA uses TRI for a variety of purposes, including prioritization of cleanup efforts. EPA also prepares an annual public data release, which has been used by federal, state and local agencies, public interest groups, businesses, health professionals, students and educators, journalists and others – including community-based environmental justice activists – to influence environmental outcomes (Bouwes, Hassur, and Shapiro 2001). Firms that report to the TRI often take the public response seriously, and some evidence suggests that the largest polluters in the TRI suffer stock price declines following the annual public data release (Hamilton 1995a).

The TRI data have been used in a number of national and smaller-scale analyses of the distribution of pollution. However, the TRI data themselves do not include information on the toxicity of the various chemicals nor on their dispersion once released. Most previous studies have simply used proximity to a TRI facility, or the total number of pounds of all TRI pollutants emitted by the facility, as indicators of pollution.

The EPA's Risk-Screening Environmental Indicators model adds toxicity and dispersal information to the TRI data. The data used in this paper differ from the data available in the current public release of the Indicators model in three main ways. First, the data used here are organized by small-area exposure, rather than by TRI facility. These data were generated as a by-product of the Indicators modeling procedure and are not available in the public release. Second,

the data used here consider only exposure to air pollution via inhalation while the published data also consider water and ground pollution and multiple pathways (e.g., ingestion, direct skin contact). Finally, the data in the public release include a population-weighting term used in calculating a risk-related measure, while the data used here do not.<sup>1</sup>

According to the EPA databases used in constructing the Indicators model, the 604 chemicals and chemical categories listed in the TRI vary in toxicity by up to eight orders of magnitude. For example, exposure to one part per trillion of a highly toxic chemical such as methyl isocyanate is toxicologically equivalent to one hundred million parts per trillion of the least toxic chemicals, such as ethylene or ammonium nitrate (Bouwes, Hassur, and Shapiro 2001; Bouwes and Hassur 1997). Toxicity is considered in terms of both cancer and non-cancer effects: if a chemical has both cancer and non-cancer effects, the higher of the cancer or non-cancer weights is assigned as the final toxicity weight. The data on chemical toxicity come primarily from the EPA's Integrated Risk Information System (IRIS), which has been peer-reviewed and represents Agency-wide judgments. For chemicals where IRIS data are not available, the model uses data from the Health Effects Assessment Summary Tables, constructed for use in Superfund and the Resource Conservation and Recovery Act; data from the Office of Pesticide Programs Reference Dose Tracking Report, or alternative data sources, all of which have been reviewed by Office of Pollution Prevention and Toxics scientists (Bouwes and Hassur 1999b).

The Indicators model incorporates facility- and chemical-specific data relevant to potential human exposure. Facility-specific factors include wind speed, direction, and turbulence, and stack heights and exit gas velocities that are either facility-specific (where such data are available) or based on median values for the facility's Standard Industrial Classification (SIC) code (Bouwes and Hassur 1997, 1999b). Chemical-specific factors include the rate of decay and deposition in air and sunlight. The fate and transport model used is the Industrial Source Complex Long-Term (ISCLT3) model, developed by EPA's Office of Air Quality Planning and Standards (Bouwes and Hassur 1999b; U.S. EPA 1995).

Based on these data, the Indicators model estimates ambient concentrations of each TRI pollutant. Concentrations are determined for each square kilometer of the 101-km by 101-km grid in which the facility is centered (the footprint of the facility is assumed to be one square kilometer). After calculating ambient concentrations, the model uses standard assumptions about human exposure to derive a surrogate dose, an estimate of the amount of chemical contacted by an individual per kilogram of body weight per day. In the public data release, these exposure-related estimates are aggregated across cells for each facility, and a single score is reported for each facility. Thus, the Indicators scores for individual cells are created as a necessary step towards building the public data release but are not available in that release. They were made available for this analysis by special arrangement.

---

<sup>1</sup> The EPA's screening method identifies priorities for clean-up based on overall environmental danger or damage, which increases with the exposed population. For our purposes of identifying the demographic factors that correlate with increased individual exposure to pollution, we do not consider a more populated area, given the same ambient concentrations of pollutants, to be more polluted than a less populated one.

The Indicators model combines the chemical-specific toxicity weights with the surrogate dose delivered by each release to obtain a grid cell element for each square kilometer cell that represents the toxicity-adjusted chronic human exposure potential. The grid cell elements resulting from releases at different facilities are added within each grid cell. The following equation shows how the components are combined to obtain the score for each cell:

$$Score_g = \sum_f \sum_c Toxicity\ Weight_c \times Surrogate\ Dose_{cfg}$$

for square kilometer cell  $g$  where  $c$  and  $f$  index chemical  $c$  released by facility  $f$ . We can compare scores across cells to evaluate the relative potential for chronic human health effects.

Compared to the measures used in most previous studies of the distribution of pollution, the data used in this paper have two major advantages. First, the detailed information on chemical toxicity allows a much more realistic measure of the potential human health effects arising from pollution. Second, the data used here are based on a realistic representation of exposure, incorporating chemical fate and transport, as well as stack heights and exit gas velocities. Most previous studies have used a single threshold distance or a simple distance-decay function to approximate the dispersion of pollution, ignoring site-specific characteristics. Moreover, the model achieves fine geographic resolution of pollution risk-related impacts, allowing the use of correspondingly fine units for demographic data without the danger of mismatched areal units.

The data used in this paper also have several important limitations. First, the risk-related impacts are based entirely on TRI reporting, which is self-reported and estimated with limited EPA oversight; firms may misreport or inaccurately estimate their releases. Second, the data used here represent only chronic health effects from inhalation of the 604 TRI-listed chemicals released by TRI reporting facilities. Third, these data are model-based: for instance, although the Indicators model incorporates more site-specific data on stack height and exit gas velocity than any other dataset used in national-level pollution distribution studies, some generalizations are made about these and other variables.<sup>2</sup>

### *Demographic Data*

We use data from the block group level from the 1990 Census of Population and Housing to represent neighborhoods. We use three racial and ethnic characteristics of a population: the percent of Hispanics, non-Hispanic blacks, and non-Hispanic Asians and Pacific Islanders, with whites as the omitted category.<sup>3</sup> The other neighborhood characteristics taken from the Census are median household income, population density, the percent of households with a primary

---

<sup>2</sup> See Bouwes and Hassur (1997, 1998, 1999a, 1999b) for more information about sensitivity analysis and “groundtruthing” of the Indicators model.

<sup>3</sup> Because Hispanic is considered an ethnicity rather than a distinct racial category, the racial categories used in the Census (black or African-American, Asian/Pacific Islander, Native American, and “other race”) are not mutually exclusive with Hispanic origin. For instance, Cuban persons of African descent would be counted as both Hispanic and black for Census purposes. In order to avoid double-counting, this study uses for each block group the percent of Hispanic residents and the percent of non-Hispanic blacks.

language other than English, percent of adults without a high school degree, percent of households receiving interest, dividend, or net rental income, percent of vacant housing units, and the percent of occupied housing units that are occupied by the owner. The percent of people voting in the 1992 national election is available only at the county level (U.S. Department of Commerce 1994a).

We limit our analysis to Census-designated urbanized areas in the 48 contiguous United States. An urbanized area is a continuously built-up area with a population of 50,000 or more (U.S. Department of Commerce 1994b). Urbanized areas correspond better than do larger metropolitan statistical areas to an intuitive idea of city boundaries. An appendix, available from the authors, lists the 393 urbanized areas used in this analysis, along with their populations, total number of block groups, average pollution level and other characteristics. In all, 66 percent of the population in the contiguous United States lived in urbanized areas in 1990.<sup>4</sup>

With the help of local committees, the Census Bureau defines Census block groups to correspond to neighborhoods; block groups typically contain between 250 and 550 households and fully partition census tracts, which contain between 2,500 and 8,000 residents. The block group is the smallest aggregation for which the Census Bureau publicly releases detailed social and demographic data.

### *Merging Datasets*

The Census data and Indicators model data are well matched in geographic precision, but are not in the same geographic format. The Indicators model divides the continental U.S. into approximately eleven million square-kilometer cells, of which about 2.2 million have positive Indicator scores (for 1998 TRI releases).<sup>5</sup> Census block groups generally have irregular boundaries, and can be either larger or smaller than one square kilometer.

To take full advantage of the fine geographic resolution of the Indicators data, we merge the pollution and Census data by Census *blocks*, a finer level of resolution than the block group (an average block group contains about 30 blocks), and then aggregate block-level scores to the block group level. After converting Census latitude and longitude to the geography of the grid-score lattice, we assign each Census block the score of the square kilometer cell in the grid-score

---

<sup>4</sup> About 2,500 block groups were removed from the analysis because they reported zero population, zero land area, zero median household income, or because zero people in the block group reported their race, ethnicity, education levels, or other characteristics. One explanation for non-reporting of income and demographic characteristics is that the population was sufficiently small that accurate conclusions about the population could not be drawn based on the small number of people who received the sample (“long form”) questionnaire used for data in Summary Tape File 3.

<sup>5</sup> Our analysis excludes Alaska and Hawaii. We smoothed the data to provide scores, either zero or positive, for all 11 million cells using an interpolated distance-weighted smoothing routine. An alternative approach of simply assigning zero to cells without scores gives very similar results.

lattice in which is located the Census-reported internal point of the block.<sup>6</sup> Then we compute block group score as an average of the scores of its component blocks.

## **Descriptive Statistics and Modeling Techniques**

Table 1 reports descriptive statistics for the dependent variable (the Indicators score) and the independent variables. About 1 percent of the observations, or approximately 1,400 block groups, have an Indicators score below one. Figure 1 shows a histogram of the Indicators score variable (for values up to 500). The dependent variable has a large mass below one and a very long tail; the distribution is clearly non-normal. Both the histogram and a comparison of the mean and median show that Indicators score has a strong right skew. The 95<sup>th</sup> percentile of the dependent variable is around 2,100, but scores range up to 1.7 million. The skew of the dependent variable suggests the use of limited dependent variable estimation rather than ordinary least squares.<sup>7</sup>

To motivate the censored approach of Tobit, we posit the existence of a latent variable representing the incapacity to avoid pollution; negative values of this latent variable would reflect progressively higher levels of capacity to secure environmental quality. But since the observed pollution cannot fall below zero, we observe scores of essentially zero for many block groups. Because several observations with very high values of Indicator score variable could overwhelm the results, for the Tobit analysis we impose upper-tail censoring of the dependent variable at 2,300, slightly above the 95<sup>th</sup> percentile, to limit the influence of several very high scores but not to discard those block groups altogether. That is, the dependent variable for the Tobit is both lower- and upper-censored with three ranges, i.e., below one, continuous between one and 2,300, and larger than 2,300.

Consistency in the Tobit model depends on the assumption of normal and homoskedastic errors (Greene 1993). Furthermore, Tobit estimates may be inconsistent in models with fixed effects even when other assumptions are satisfied (Honoré 1992). However, because our data include many block groups within regions, we consider Tobit estimates in the fixed-effects model reliable.<sup>8</sup>

We also apply a dichotomous estimation technique. We estimate the probability that a neighborhood is in the more polluted half and most polluted tenth of its city. The quantiles are

---

<sup>6</sup> The internal point is the location of the Census landmark (e.g., street intersection) closest to the block centroid and inside the block polygon. We experimented with alternative assignments, e.g., directly assigning to each block group the score associated with the block-group internal point.

<sup>7</sup> Applied to censored data, ordinary least squares generally biases coefficient estimates toward zero because variation in the observed dependent variable understates the effects of the regressors on the true dependent variable (Chay and Powell 2001).

<sup>8</sup> We also estimated the model with OLS, truncating the data at the 99<sup>th</sup> percentile, and the results with and without fixed effects are very similar to the Tobit results.

determined by population; for example, the most polluted quarter of the city is the set of most polluted block groups that contains one quarter of the population of the urbanized area (note that the block groups are not necessarily contiguous). While we lose information in moving from a continuous variable to a dichotomous variable, the results of the dichotomous model readily lend themselves to interpretation. We apply the linear probability model (LPM), because despite well-known drawbacks, namely induced heteroskedasticity and non-conforming probabilities, the LPM delivers consistent estimates even when fixed effects are included.<sup>9</sup>

Table 1 also reports descriptive statistics for the explanatory variables. Because the means reported here are by block group unweighted by population, the percentages may not reflect overall national averages for urban areas. The average block group in this study has about 10 percent Hispanic residents, 16 percent non-Hispanic black residents, 3 percent residents of non-Hispanic Asian heritage, and 70 percent non-Hispanic white residents. Comparing the median and mean values for racial characteristics again shows significant right skew. For example, the median value for percent Hispanic residents is only 2.3 percent, indicating that a relatively small number of block groups have proportions of Hispanic residents much higher than the mean of 10 percent, whereas half of the block groups have less than 2.3 percent Hispanic residents.

The average median household income of block groups is about \$34,050, with a minimum of \$5,000 and a Census-topcoded maximum of \$150,000, while the median of median household incomes is slightly lower. The average urban block group has a population density of 3,177 persons per square kilometer, with a median value of 1,802. Some block groups, for example in New York City, have extremely high population density, ranging up to more than 300,000 persons per square kilometer. On average, 25 percent of the residents over age 18 have no high school diploma, and 18.5 percent of the households in each block group are classified as having a language other than English as their primary language.<sup>10</sup> Forty percent of the households in an average block group receive some income from interest, dividends, or net income from renting property. On average seven percent of housing units in a block group are vacant, and 61 percent of the occupied housing units in a block group are occupied by the owners. The mean area among block groups is 2.59 square kilometers, with a median of 0.52 square kilometers. The average block group contains 1,196 persons.

To explore the relationship between race and ethnicity and the pollution score of block groups we present the data for Milwaukee, Wisconsin in Figure 2. Milwaukee is chosen because relative to other urbanized areas, it has substantial black and Hispanic representation in the population, its

---

<sup>9</sup> With fixed effects, the standard logit model would give inconsistent estimates. Conditional logit (Chamberlain 1984, 1274-1278) can give consistent estimates, but consistency depends upon an assumption of independent and identically distributed residuals; furthermore the model becomes computationally untenable when the number of within-group units is high, and some cities have 2,000 or more block groups.

<sup>10</sup> The mean value for percent of households with a primary language other than English seems somewhat high. The 1990 Census assigned household primary language to all members of a household based on the language spoken at home by the householder or by other household members. As a result, some households classified as having a primary language other than English may contain household members whose primary language is English (U.S. Department of Commerce 1992).



average level of pollution is high but not extremely high, and its within-urbanized-area pollution gradient for both blacks and Hispanics is also high but not extremely high. In the four panels of Figure 2, we view, by block group around the Milwaukee urbanized area, the distribution of RSEI scores, the percent black; the percent Hispanic, and the population density in people per square kilometer. Visual inspection suggests that the most polluted block groups of Milwaukee fall into two categories, block groups predominantly inhabited by blacks or Hispanics and block groups with relatively low population density.

In Figures 3, 4, and 5, we present scatterplots that depict racial, ethnic, and income disparities between the more and less polluted half of each of the 393 urbanized areas (i.e., the block groups containing the more pollution-exposed 50% of the population and block groups containing the less pollution-exposed 50% of the population). In Figure 3, we plot the percent Hispanic in the more polluted half of each city on the vertical axis against percent Hispanic in the less polluted half of each city on the horizontal axis. The 45-degree line shows equal representation of Hispanics; cities that appear above the line have disproportionate representation of Hispanics in the more polluted half of the city, while cities that fall below the line have disproportionate representation of Hispanics in the less polluted half of the city. The size of each circle represents the pollution disparity between the more and less polluted halves of the city; a large circle indicates a sharp disparity between the more and less polluted halves of the city. Figure 4 presents the same analysis for blacks. Visual inspection strongly suggests that the more polluted halves of most cities are disproportionately Hispanic or black or both. The results reported below for the LPM affirm this result using multivariate analysis.

In Figure 5, we present a scatterplot of all 393 urbanized areas with the population-weighted average median household income in the less polluted half of the city on the horizontal axis and the average median income in the more polluted half of the city on the vertical axis. A city exactly on the 45-degree line has equal average median household income in both the more polluted half and the less polluted half. As in Figures 3 and 4, the size of the circle corresponds to the magnitude of the disparity between pollution levels in the more and less polluted halves of the city. The preponderance of the point-cloud of urbanized areas falls below the line, indicating that the more polluted half of most cities is substantially poorer than is the less polluted half.

## Regression Results

In this section, we first report results based on the Tobit model without and with fixed effects. We then report results based on the dichotomous LPM model. Table 2 reports the coefficient estimates and standard errors for the model specifications estimated with Tobit.

### *Tobit Model Without Area Fixed Effects*

The first two columns report the results of national level estimation without urbanized-area effects. In Column [1], where the model includes only the race and ethnicity variables, the coefficient on percent of blacks is positive, indicating that block groups with higher proportions of blacks tend to have higher Indicators scores, while the coefficients on percent Hispanic and

Asian are both negative, meaning that block groups with more of these ethnic groups tend to have lower Indicators scores. When we add income in Column [2], the coefficient on percent black falls rather sharply but remains positive and materially and statistically significant. This result implies that the positive relationship observed between the percent of black residents and the Indicators score observed in Column [1] is in part due to the negative correlation between percent black residents and household income. Likewise, the magnitude of the negative relationship for Hispanics increases when income controls are added, which suggests that Hispanic exposure is higher because of lower average Hispanic income. However, the magnitude of the negative relationship between percent Asian and pollution diminishes after income is included, meaning that Asians are more exposed than their incomes alone would suggest.

Higher median income is strongly associated with lower pollution scores. The quadratic and linear terms together imply a concave relation with the minimum at 92,000 dollars, which means that virtually every block group lies in the domain where increasing income is associated with decreasing pollution exposure.

#### *Tobit Model With Area Fixed Effects*

The results of the fixed-effects models that correspond to the models discussed above, shown in Columns [3] and [4], reveal a striking difference between demographic correlations between cities and correlations within cities. Within urbanized areas, the strong positive relationship between percent black and pollution score persists. But when base pollution levels are permitted to vary among cities, there is a positive relationship between the percentage of Hispanics in a block group and the Indicators score. Taken together, results from the models with fixed effects (Columns [3] and [4]) and the corresponding models without fixed effects (Columns [1] and [2]) suggest that Hispanics live in cleaner cities, but that within the cities where they live, they tend to live “on the wrong side of the environmental tracks” – that is, in more polluted block groups.

The coefficient on the percent of blacks remains positive and significant in the fixed-effects model. Taken together, the results of the models without and with fixed effects indicate that blacks live both in more polluted cities in the U.S. and also in the more polluted block groups of the cities in which they live. The magnitude of the intra-city effect revealed by the fixed-effect models is stronger for Hispanics than for blacks in specifications [3] and [4]. In all the Tobit specifications, Asians are found to live in relatively less polluted neighborhoods.

The income coefficient remains stable with the addition of the 393 urbanized area fixed effects. The quadratic and linear terms imply that the minimum occurs at around 80,000 dollars per year; more than 95 percent of block groups are in the decreasing part of the function.

Visual inspection of a national RSEI-score map and regression analysis available from the authors suggest that the racial and ethnic between-city results are largely due to the concentration of Hispanics and Asians in the West and the Southwest, which generally have lower concentrations of heavy industry and TRI emitters. This analysis also finds that the inter-city

effect for blacks – blacks tend to live in significantly more polluted cities than do whites – is due to the concentration of blacks in the Rust Belt cities of the Northeast and Midwest.

To elaborate the social and demographic characteristics that drive the results, we included an extended list of covariates (Column [5]). The coefficient on the percent of population who voted in the 1992 national election is positive.<sup>11</sup> The coefficient on population density is negative which implies that within urbanized areas, less densely populated block groups have higher Indicators scores. This finding suggests support for the argument that polluting facilities, and the associated pollution burdens, should tend to be located in sparsely populated areas because planners and public health officials will work to make that happen, but is also consistent with the proposition that sparsely populated areas may have less political influence, holding constant other factors (e.g., income). The fraction of housing units that are vacant is also associated with increased pollution; high vacancy is not simply a correlate of low density ( $r = -0.0062$ ) but may reflect neighborhood disempowerment or distress. Block groups with a greater proportion of adults without high school diplomas tend to be more polluted, as do block groups with a greater proportion of households identified as having a primary language other than English. Block groups with a greater proportion of households receiving capital income tend to have lower pollution scores. The inclusion of the full set of covariates reduces the coefficient on percent black, although the value still remains positive and highly significant. The inclusion of the full set of covariates causes the sign on percent Hispanic to become negative, which suggests that the Hispanic effect is largely explained by other social and demographic characteristics, including education, unearned income, and, especially, limited English-language facility. (The correlation of percent Hispanic and percent non-English speaking is 0.85.)

Despite the high fraction of block groups with pollution scores near zero, which could bias OLS results, simple OLS regression yielded results very similar to those of Tobit, albeit with slight attenuation in the coefficients. The imposition of the upper-censoring, however, had substantial effects on the results. If the highest values of the Indicators score, which are two orders of magnitude above the 99th percentile and three orders of magnitude above the mean, are included, they drive the results and generate substantially different regression coefficients.<sup>12</sup>

#### *LPM Results*

In Table 3, we report the probability that a block group is in the most polluted fraction of its urbanized area as a function of its economic and demographic characteristics. The results are based on a fixed-effects linear probability model, and thus are most comparable to Columns [3] through [5] of Table 2. In the first three columns, we examine the probability that a block group

---

<sup>11</sup> Because percentage of population voting is measured at the county level, identification comes from urbanized areas with block groups in more than one county. Fifty-one percent of the 393 urbanized areas have block groups in more than one county, and the average urbanized area has nearly 30 percent of its block groups in counties other than the county with the plurality of block groups.

<sup>12</sup> The actual censoring value applied makes little difference. We tested censoring at centiles 90, 95, 99, and 99.5 and found very similar Tobit results for all upper-censoring points. The block groups with the very highest scores, which have median incomes close to the overall median and which are more white than is the average block group, exert substantial leverage.

is among the *more polluted half* of its city. The estimated coefficients in the linear probability model can be interpreted as percentage point changes. For example, in Column [1] we find that a block group that is 100 percent Hispanic has a probability 51 percentage points greater of being in the more polluted half of the city compared to an otherwise identical block group that is 100 percent white non-Hispanic. The results indicate that block groups that are a higher proportion Hispanic, black, or Asian are all more likely to be in the more polluted half of the urbanized area, although the Asian effect disappears when controls for income are added. We also find that over most of the observed range of incomes, income is negatively correlated with the probability of being in the more polluted half of the city, and the minimum of the income-pollution gradient occurs at a neighborhood income of 99,000 dollars. At the median, a 10,000 dollar increase in income is associated with a 7 percentage point decrease in the probability of being in the more polluted half of the city.

As in the within-city Tobit model, voter participation is positively associated with exposure to pollution. Block groups with a higher fraction of adults who did not graduate from high school are substantially more likely to be in the more exposed half. Again as in the Tobit model, capital income is associated with less pollution exposure. The results for one variable do not correspond to the Tobit results: a higher share of non-English speaking persons is associated with a lower probability of inclusion in the more polluted half of the city.

When we turn to the *most polluted tenth* of cities, the explanatory power of the model drops, but most of the results persist (Columns [4] through [6] of Table 3). In Column [4], which shows results for a model that includes only the race and ethnicity variables, we find that Hispanics and blacks are substantially more likely than non-Hispanic whites to live in the most polluted tenth of cities, and Asians are less likely to live in the most polluted tenth. When median household income is included in the regression, the race and ethnicity effects decline but remain positive and significant. When we include the full set of neighborhood covariates, the signs on the coefficients for percent Hispanic and percent black actually reverse, although the size of the coefficients is small. This result suggests that for inclusion in the most polluted portions of the city, other social indicators, e.g., education and income, explain the correlations between pollution and Hispanic and black. The income-pollution gradient remains strongly negative for determining the probability that a block group is among the most polluted 10% of block groups in the city. Higher neighborhood incomes imply a steadily declining probability of membership in the most polluted tenth. We find negative associations between inclusion in the most polluted tenth and both population density and capital income; we find a positive relationship between presence in the most polluted tenth and both vacant units and adults with less than high school education.<sup>13</sup>

The models explain a small proportion of the variation in the Indicators score. Without fixed effects, the pseudo-R-squared values for the Tobit regressions reported in Table 2 are extremely low (below 0.01). When we estimated the same models using OLS, the R-squared values in the models without fixed effects reached only 0.12 for the models with the full set of explanatory

---

<sup>13</sup> We examined, but do not present here, the correlates of the worst quarter and of centiles 75 through 90, and found results that were generally intermediate between those for the worst tenth and the worse half.

variables. Even when we estimated the models with 393 UA fixed effects, the pseudo-R-squared of the Tobits climbed only to around 0.06. While R-squared is a poor measure of fit for the linear probability model, the within-city R-squared for the fixed-effect linear probability models ranged from 0.01 for the sparsest specification of the regression for the most polluted tenth to 0.09 for the richest specification of the regression for the more polluted half.

Relatively few studies of the demographic correlates of pollution report the percentage of variation explained; however, of those that do, most report a higher percentage explained than is explained by our model. We experimented with adding additional explanatory variables to the model specification, including variables included by authors whose models had substantially greater R-squared, and found that the additional variables did not raise the R-squared of our model substantially. Aggregated analyses are likely to have upward biased R-squared because of the ecological fallacy. Our highly disaggregated analysis reduces its effect. Despite a low proportion of explained variation, we find that there are statistically significant correlations between demographic variables and pollution, as evidenced by the high t-statistics in the multivariate results. Clearly, future research on the correlates of exposure has much rich territory to cover.

## **Discussion and Conclusions**

Previous environmental justice research has generally failed to address the demographics of pollution in terms of toxicity and exposure, focusing instead on proximity to pollution sources or on the mass of pollutants released. Many previous studies also used relatively large units of spatial resolution. And most national studies control for regional variation, if at all, only by estimating correlations separately for a small number of regions in the country (e.g., Arora and Cason 1999) or separately for densely and sparsely populated regions (e.g., Bouwes, Hassur, and Shapiro 2001).

This paper addresses these issues by using data from the EPA's Risk-Screening Environmental Indicators model. Our results indicate that in the urban U.S. as a whole, block groups with more blacks have higher levels of exposure to toxic pollution from facilities that report to the TRI, while block groups with more Hispanics and Asians/Pacific Islanders have lower levels. When we control for differences among urban areas, however, we find that within cities, Hispanics, as well as blacks, tend to live in more polluted neighborhoods. In national-level comparisons, this disparity is offset for Hispanics by the fact that they tend to live in cities with relatively low levels of industrial toxics. Blacks, by contrast, tend to live in more polluted cities as well as in the more polluted neighborhoods within cities.

There are several important caveats regarding the data underlying this analysis. First, our dependent variable represents only a subset of environmental pollutants that people face. Pollution from mobile sources, such as automobiles, and from small point sources, such as dry cleaners, are not included. The data also do not include exposure to pollutants in other media, such as water pollution, or other non-residential exposure such as in the workplace. Second, the measure of toxicity does not include possible synergistic or multiplicative effects of different

pollutants. Third, the dependent variable is based on annual averages of atmospheric dispersion patterns.

Several caveats are also in order regarding our methodology. A cross-sectional analysis cannot elucidate causal relationships between demographics and pollution. The correlations between the demographic variables and Indicators scores could be caused by a number of underlying factors. For instance, blacks or Hispanics may have lower levels of wealth than whites on average, even when their incomes are the same, which would constrain housing choices. Blacks or Hispanics also may tend to have less access to information about the health effects of pollution. Or there may be institutional racism in housing or credit markets, or in the processes by which industrial facilities are sited. This study cannot ascertain which of these or other processes underlie the results. We note, however, that the strong effect of race and ethnicity, controlling for income, suggest that voluntary move-in spurred by low income is not a plausible explanation for these differences in exposure.

This study offers several key insights to inform future research and policy:

- Our results suggest that different minority groups should be included separately in econometric analysis rather than lumped together as “all nonwhite” or “all minority” residents.
- National environmental justice studies should account for variations in base levels of pollution, in order to avoid collapsing variation within cities and variation among cities into a single coefficient.
- Results may depend on the unit of analysis used. For analysing *exposure*, though not necessarily point sources, we believe that the unit of analysis should be as small as possible because analysis at larger levels can obscure heterogeneity.

We have identified only two previous national studies that use an indicator of air pollution that accounts both for the relative toxicity of different pollutants and for their atmospheric dispersion. Their results are only partially comparable due to differences in methodology; however, with respect to the coefficients on race and ethnicity variables, the results here are generally consistent. Brooks and Sethi (1997) report a positive correlation between the percent of blacks in a zip code and their index of pollution. Bouwes, Hassur, and Shapiro (2001) report a positive correlation between their index of pollution and both the percent of blacks and the percent of Hispanics.<sup>14</sup>

What are the conclusions for environmental justice policy? The results here echo those of the many other researchers who have found that pollution burdens do fall disproportionately on blacks in the U.S. – and on Hispanics and Asians within regions. This result further suggests that

---

<sup>14</sup> Because the Bouwes, Hassur, and Shapiro (2001) study also uses the RSEI model to generate a dependent variable, the opposite sign on the Hispanic variable may seem surprising. However, three key choices in methods may account for the difference. Their measure of pollution was weighted by population, so that more populated areas have higher values of pollution even if the unweighted pollution level is the same; they use only observations for which the value of the Indicators score is greater than zero; and the unit of their analysis is the square-kilometer cell rather than the Census block group.

regional and state policy makers should examine and seek to redress disparate exposure. The results for blacks imply that environmental justice should remain a priority for national as well as regional environmental policy. In addition, the results here highlight the value of the EPA's Risk-Screening Environmental Indicators data for environmental justice analyses. The fine geographic resolution and well-developed dependent variable make these data exceptionally well-suited for analyses of the demographics of pollution.

**Table 1. Descriptive statistics (N=136,362 block groups in urbanized areas)**

<b>Variable</b>	<b>Mean</b>	<b>Median</b>	<b>Std. Dev.</b>	<b>Minimum</b>	<b>Maximum</b>
Indicators score	716.28	139.3	10369.79	0.00	1734694
Hispanic (percent)	9.88	2.28	18.59	0.00	100.00
Black (percent)	16.03	2.16	28.67	0.00	100.00
Asian/Pacific Islander (percent)	2.98	0.24	6.65	0.00	100.00
Native American (percent)	0.46	0.00	1.65	0.00	94.44
Other race (percent)	0.12	0.00	0.74	0.00	57.14
Median household income (\$000)	34.05	31.11	18.30	5.00	150.00
Population density (persons/km <sup>2</sup> )	3177.4	1801.9	5826.6	0.11	397555.6
Population voting in 1992 national election (percent)	41.02	41.68	6.89	7.33	57.96
Households with non-English primary language (percent)	18.50	12.39	18.80	0.00	100.00
Adults without high school diploma (percent)	24.97	21.22	17.32	0.00	100.00
Households with interest, dividend, or net rental income (percent)	40.43	41.05	20.74	0.00	100.00
Vacant housing units (percent)	7.13	5.05	7.77	0.00	94.41
Owner-occupied housing units (percent)	61.49	66.85	27.88	0.00	100.00
Block group area (km <sup>2</sup> )	2.59	0.52	22.00	0.001	3202.21
Block group population	1196.09	978.00	980.04	1	35682



**Table 2. Results for Tobit estimation, with and without area fixed effects**

	[1]	[2]	[3]	[4]	[5]
Area fixed effects	No	No	393 UAs	393 UAs	393 UAs
Hispanic (percent)	-2.25** (0.09)	-4.68** (0.09)	3.07** (0.08)	0.96** (0.08)	-0.55** (0.14)
Black (percent)	2.87** (0.06)	0.24** (0.06)	2.32** (0.04)	0.81** (0.048)	0.24** (0.06)
Asian/Pacific Islander (percent)	-10.64** (0.25)	-8.26** (0.24)	-0.123 (0.188)	-0.91** (0.19)	-0.99** (0.22)
Median Household Income (\$000)		-18.4** (0.3)		-10.2** (0.2)	-6.03** (0.27)
Square of Median household income		0.0997** (0.0023)		0.061** (0.002)	0.038** (0.002)
Population density (persons/km <sup>2</sup> )					-0.00116** (0.00024)
Population voting in 1992 national election (percent)					0.37 (0.32)
Households with non-English primary language (percent)					0.231# (0.139)
Adults without high school diploma (percent)					3.16** (0.11)
Households with interest, dividend, or net rental income (percent)					-0.64** (0.10)
Vacant housing units (percent)					2.34** (0.17)
Owner-occupied housing units (percent)					0.022 (0.061)
R-squared (percent)	0.29%	0.73%	5.52%	5.71%	5.77%

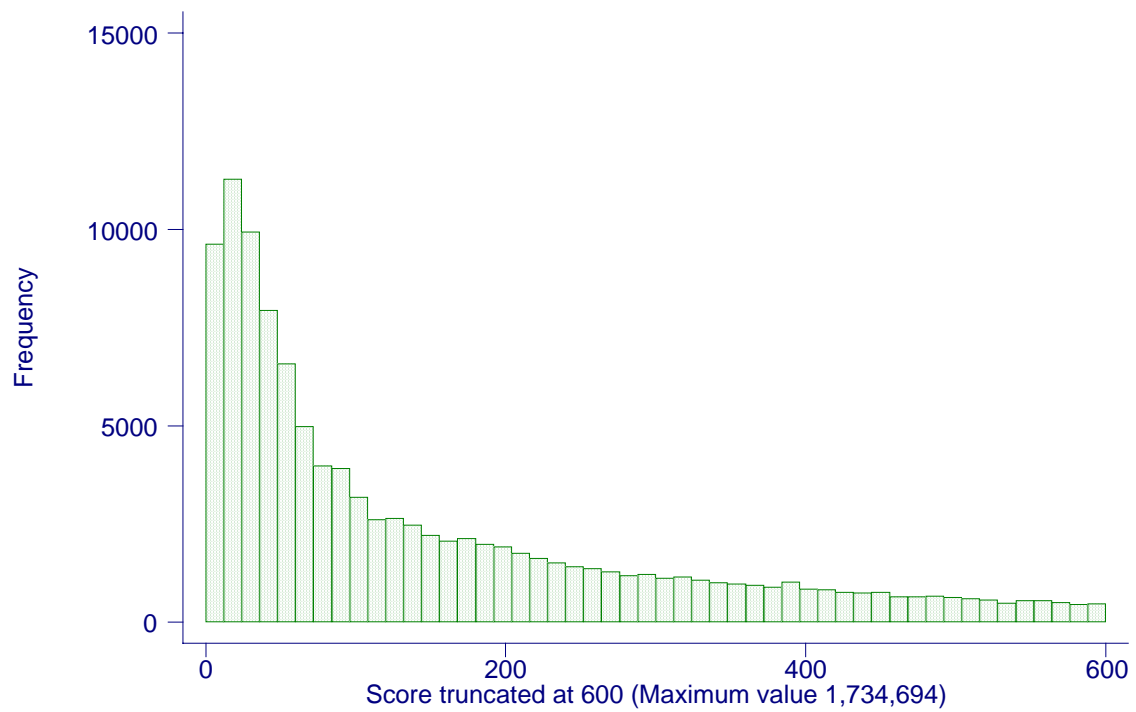
NOTES: Sample size: 136,362 block groups. Dependent variable is RSEI score: 1,376 observations left-censored at or below 1; 6,209 observations right-censored at or above 2,300. All variables in percentages range from 0 to 100. Constant included in columns [1] and [2]. Standard errors are in parentheses. # Significant at  $p < .10$ ; \* Significant at  $p < .05$ ; and \*\* Significant at  $p < .01$ .

**Table 3. Results for LPM estimation**

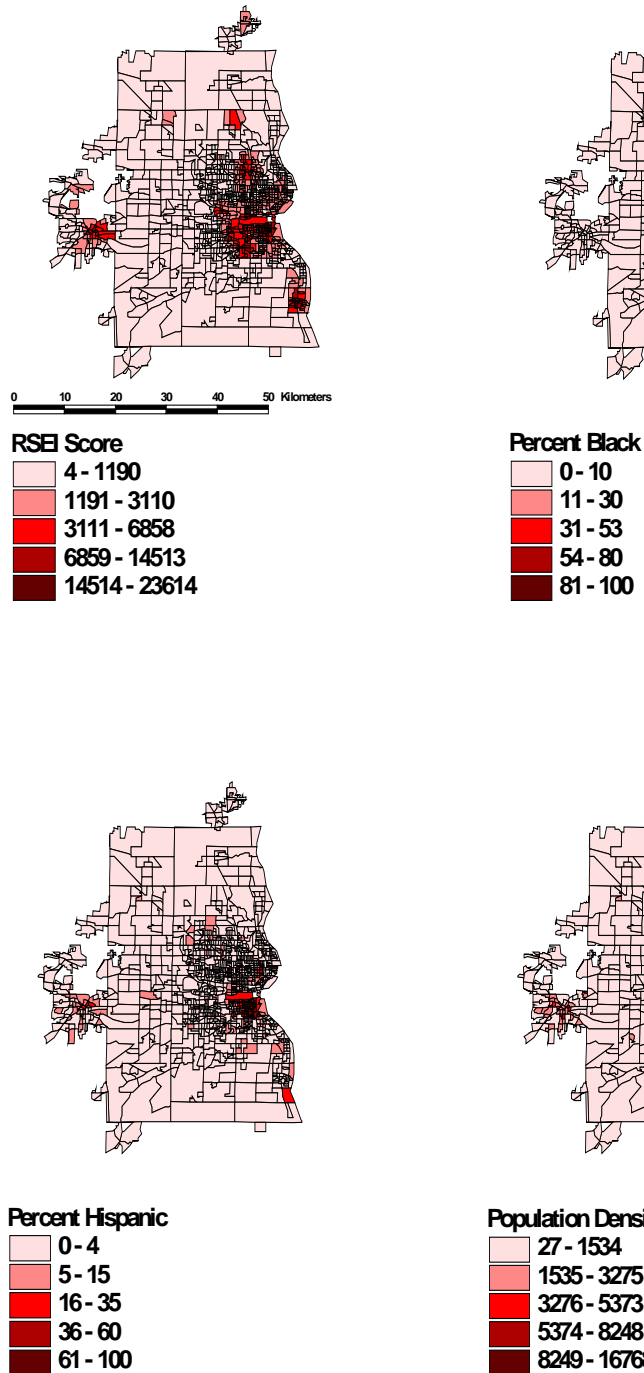
	[1]	[2]	[3]	[4]	[5]	[6]
Most polluted...of UA	Half	Half	Half	Tenth	Tenth	Tenth
Hispanic (percent)	0.517** (0.009)	0.228** (0.010)	0.137** (0.017)	0.194** (0.006)	0.094** (0.006)	-0.008 (0.011)
Black (percent)	0.337** (0.005)	0.138** (0.006)	0.087** (0.007)	0.083** (0.003)	0.014** (0.004)	-0.033** (0.004)
Asian/Pacific Islander (percent)	0.103** (0.022)	0.000 (0.022)	0.040 (0.026)	-0.037** (0.015)	-0.073** (0.015)	-0.050** (0.017)
Median Household Income (\$000)		-0.0121** (0.0002)	-0.0082** (0.0003)		-0.0043** (0.0001)	-0.0011** (0.0002)
Square of Median household income		0.000061** (0.000002)	0.000040** (0.000002)		0.0000222** (0.0000013)	0.0000052** (0.0000014)
Population density (persons/km <sup>2</sup> )			0.000005** (.0000003)			-0.000003** (0.0000002)
Population voting in 1992 national election (percent)			0.326** (0.038)			-0.015 (0.025)
Households with non-English primary language (percent)			-0.096** (0.017)			-0.00006 (0.011)
Adults without high school diploma (percent)			0.335** (0.014)			0.245** (0.009)
Households with interest, dividend, or net rental income (percent)			-0.028* (0.012)			-0.082* (0.008)
Vacant housing units (percent)			0.05** (0.02)			0.080** (0.013)
Owner-occupied housing units (percent)			-0.0026 (0.0072)			-0.0043 (0.0047)
R-squared (within)	5.3%	8.6%	9.2%	1.2%	2.2%	3.3%

NOTES: Sample size: 136,362 block groups. Dependent variable is dichotomous with 1 indicating inclusion in the more polluted segment of the city. All variables in percentages range from 0 to 100. Coefficients express the percentage point change in probability of exposure per unit change in the explanatory variable. Standard errors are in parentheses. # Significant at  $p < .10$ ; \* Significant at  $p < .05$ ; and \*\* Significant at  $p < .01$ .

**Figure 1. Histogram of Indicators Scores by Block Group**



**Figure 2. Pollution, Race and Ethnicity, and Population Density  
Census Block Groups of Milwaukee, Wisconsin**



Notes: Population density is reported in persons per square kilometer; the average density for the Milwaukee Urbanized Area is 3180 persons per square kilometer.

Figure 3. Percent Hispanic population in more and less polluted halves of cities

Size of circle represents pollution disparity between halves

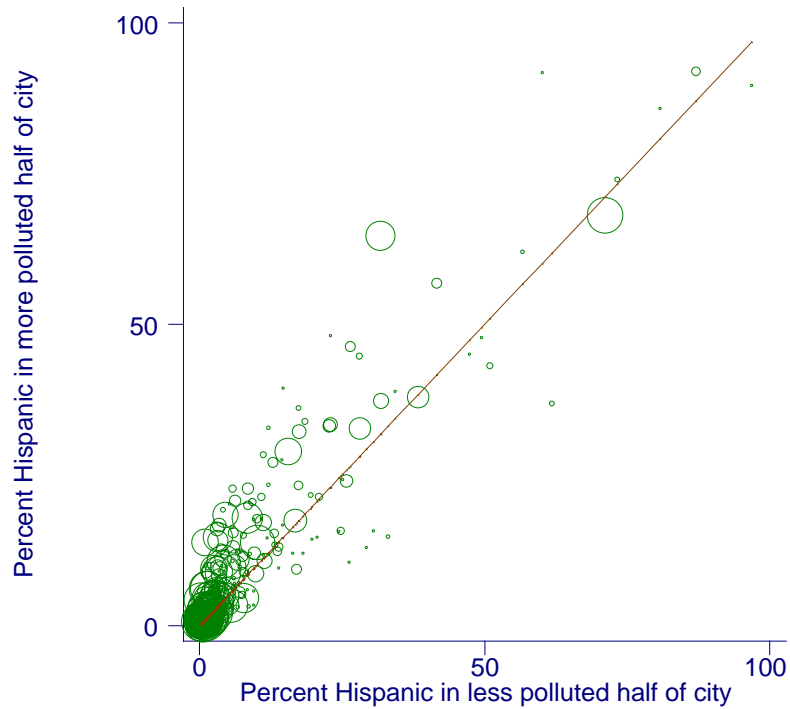


Figure 4. Percent of black population in more and less polluted halves of cities

Size of circle represents pollution disparity between halves

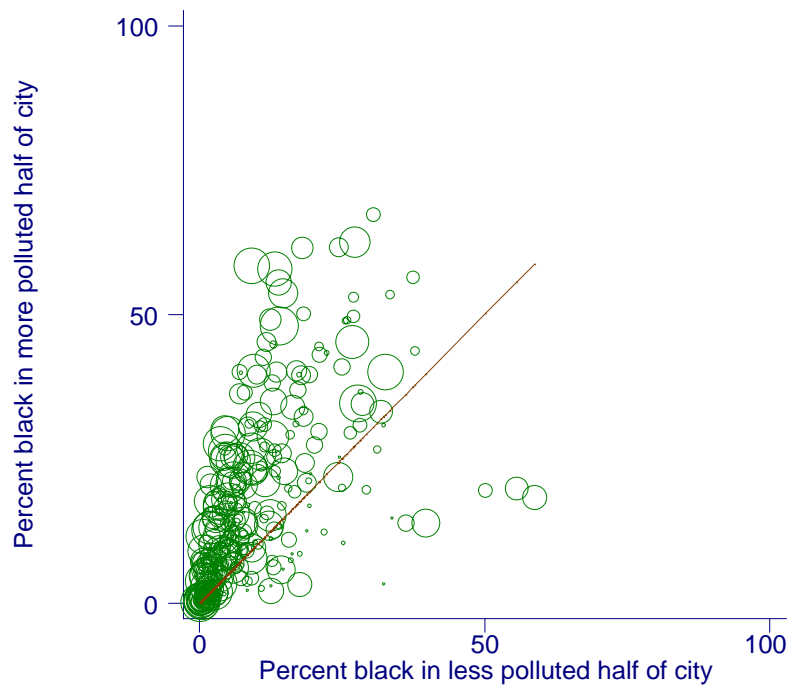
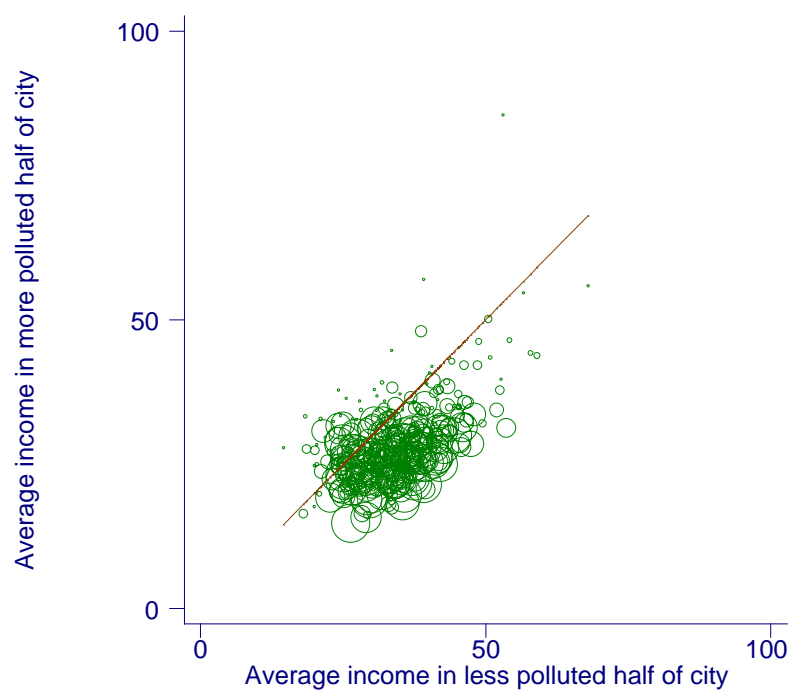


Figure 5. Average median household income in more and less polluted halves of cities  
Size of circle represents pollution disparity between halves



## REFERENCES

- Anderton, D. L., A. B. Anderson, P. H. Rossi, J. M. Oakes, M. R. Fraser, E. W. Weber, and E. J. Calabrese. 1994. Hazardous waste facilities: "Environmental equity" issues in metropolitan areas. *Evaluation Review* 18, no. 2: 123-140.
- Anselin, L. 1988. *Spatial econometrics: methods and models*. Boston: Kluwer Academic Publishers.
- Arora, S., and T. N. Cason. 1999. Do community characteristics influence environmental outcomes? Evidence from the Toxics Release Inventory. *Southern Economic Journal* 65, no. 4: 691-716.
- Asch, P., and J. J. Seneca. 1978. Some evidence on the distribution of air quality. *Land Economics* 54, no. 3: 278-297.
- Been, V., and F. Gupta. 1997. Coming to the nuisance or going to the barrios? A longitudinal analysis of environmental justice claims. *Ecology Law Quarterly* 24, no. 1: 1-56.
- Boer, J. T., M. Pastor, J. L. Sadd, and L. D Snyder. 1997. Is there environmental racism? The demographics of hazardous waste in Los Angeles County. *Social Science Quarterly* 78, no. 4: 793-810.
- Boerner, C., and T. Lambert. 1995. *Environmental justice in the city of St. Louis: The economics of siting industrial and waste facilities*. St. Louis: Center for the Study of American Business. Working Paper 156.
- Bouwes, N. W., and S. M. Hassur. 1997. *Toxics Release Inventory Relative Risk-Based Environmental Indicators methodology*. Washington, D.C.: U.S. EPA Office of Pollution Prevention and Toxics. Available at [http://www.epa.gov/opptintr/env\\_ind/docs/methpdf.exe](http://www.epa.gov/opptintr/env_ind/docs/methpdf.exe); accessed 24 August 2000.
- Bouwes, N. W., and S. M. Hassur. 1998. *Ground-truthing of the air pathway component of OPPT's Risk-Screening Environmental Indicators model (draft)*. Washington, D.C.: U.S. EPA Office of Pollution Prevention and Toxics. Available at [http://www.epa.gov/opptintr/env\\_ind/docs/grndpdf.exe](http://www.epa.gov/opptintr/env_ind/docs/grndpdf.exe); accessed 2 November 2001.
- Bouwes, N. W., and S. M. Hassur. 1999a. *Estimates of stack heights and exit gas velocities for TRI facilities in OPPT's Risk-Screening Environmental Indicators model*. Washington, D.C.: U.S. EPA Office of Pollution Prevention and Toxics. Available at [http://www.epa.gov/opptintr/env\\_ind/docs/stackpdf.exe](http://www.epa.gov/opptintr/env_ind/docs/stackpdf.exe); accessed 2 November 2001.

- Bouwes, N. W., and S. M. Hassur. 1999b. *User's manual for EPA's Risk-Screening Environmental Indicators model, version 1.02*. Washington, D.C.: U.S. EPA Office of Pollution Prevention and Toxics. Available from [http://www.epa.gov/opptintr/env\\_ind/docs/userpdf.exe](http://www.epa.gov/opptintr/env_ind/docs/userpdf.exe); accessed 18 November 1999.
- Bouwes, N. W., S. M. Hassur, and M. D. Shapiro. 2001. *Empowerment through risk-related information: EPA's Risk-Screening Environmental Indicators project*. Amherst, MA: Political Economy Research Institute. Working Paper DPE-01-06.
- Bowen, W. M., M. J. Salling, K. E. Haynes, and E. J. Cyran. 1995. Toward environmental justice: Spatial equity in Ohio and Cleveland. *Annals of the Association of American Geographers* 85, no. 4: 641-663.
- Boyce, J. K., A. R. Klemmer, P. H. Templet, and C. E. Willis. 1999. Power distribution, the environment, and public health: A state-level analysis. *Ecological Economics* 29: 127-140.
- Brooks, N., and R. Sethi. 1997. The distribution of pollution: Community characteristics and exposure to air toxics. *Journal of Environmental Economics and Management* 32: 233-250.
- Bryant, B., ed. 1995. *Environmental justice: Issues, policies, and solutions*. Washington, D.C.: Island Press,
- Chamberlain, G. 1984. Panel data. In *Handbook of Econometrics*, vol. 2, ed. Z. Griliches and M. D. Intriligator, 1247-1318. New York: North-Holland.
- Chay, K. Y., and M. Greenstone. 2001. Air quality, infant mortality, and the Clean Air Act of 1970. Unpublished working paper. Available at <http://emlab.berkeley.edu/users/kenchay/ftp/binresp/aerpaper.pdf>; accessed 11 January 2002.
- Chay, K. Y., and J. L. Powell. 2001. Semiparametric censored regression models. Unpublished working paper. Available at <http://elsa.berkeley.edu/~kenchay/ftp/binresp/jepfinal.pdf>; accessed 9 December 2001.
- Efron, B., and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. New York: Chapman and Hall.
- Freeman, M. A. 1972. The distribution of environmental quality. In *Environmental quality analysis*, ed. A. V. Kneese and B. T. Bower, 243-278. Baltimore: Johns Hopkins University Press.



- General Accounting Office. 1983. *Siting of hazardous waste landfills and their correlation with racial and economic status of surrounding communities*. Washington, D.C.: Government Printing Office.
- Glickman, T. S., and R. Hersh. 1995. *Evaluating environmental equity: The impacts of industrial hazards on selected social groups in Allegheny county, Pennsylvania*. Washington, D.C.: Resources for the Future. Discussion paper 95-13.
- Greene, William H. 1993. *Econometric analysis*. 2<sup>nd</sup> ed. Englewood Cliffs, NJ: Prentice Hall International.
- Hamilton, J. T. 1993. Politics and social costs: estimating the impact of collective action on hazardous waste facilities. *Rand Journal of Economics* 24, no. 1: 101-125.
- Hamilton, J. T. 1995a. Pollution as news: Media and stock market reactions to the Toxics Release Inventory data. *Journal of Environmental Economics and Management* 28: 98-113.
- Hamilton, J. T. 1995b. Testing for environmental racism: Prejudice, profits, political power? *Journal of Policy Analysis and Management* 14, no. 1: 107-132.
- Hird, J. A., and M. Reese. 1998. The distribution of environmental quality: An empirical analysis. *Social Science Quarterly* 79, no. 4: 693-716.
- Johnston, J., and J. DiNardo. 1997. *Econometric methods*. 4<sup>th</sup> ed. New York: The McGraw-Hill Companies.
- Kriesel, W., T. J. Centner, and A. G. Keeler. 1996. Neighborhood exposure to toxic releases: Are there racial inequities? *Growth and Change* 27 (fall): 479-499.
- Kruvant, W. J. 1975. People, energy, and pollution. In *The American energy consumer*, ed. K. K. Newmand and D. Day, 125-167. Cambridge, MA: Ballinger.
- Massey, D. S. 1996. The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demography* 33, no. 4:395-412.
- Massey, D. S., and N. Denton. 1993. *American apartheid: segregation and the making of the underclass*. Cambridge, MA: Harvard University Press.
- Mohai, P., and B. I. Bryant. 1992. Environmental racism: Reviewing the evidence. In *Race and the incidence of environmental hazards: A time for discourse*. Boulder, Colorado: Westview Press.

- Morello-Frosch, R., M. Pastor, and J. Sadd. 2001. Environmental justice and southern California's "risky landscape": The distribution of air toxics exposures and health risks among diverse communities. *Urban Affairs Review* 36, no. 4: 551-578.
- Oakes, J. M. 1997. The location of hazardous waste facilities. Ph.D. diss., University of Massachusetts Amherst.
- Odland, J. 1988. *Spatial autocorrelation*. Scientific Geography Series. Newbury Park, CA: Sage Publications.
- Office of Pollution Prevention and Toxics. 1999. Fact sheet: Risk Screening Environmental Indicators. Available from [http://www.epa.gov/opptintr/env\\_ind/docs/fact.pdf](http://www.epa.gov/opptintr/env_ind/docs/fact.pdf); accessed 24 August 2000. Washington, D.C.: U.S. EPA Office of Pollution Prevention and Toxics.
- Palm, R. 1981. *The geography of American cities*. New York: Oxford University Press.
- Pastor, M. 2001. *Building social capital to protect natural capital: The quest for environmental justice*. Amherst, MA: Political Economy Research Institute, Working Paper DPE-01-02.
- Pastor, M., J. Sadd, and R. Morello-Frosch. 2000. Who's minding the kids? Pollution, public schools, and environmental justice in Los Angeles. Unpublished working paper.
- Pellow, D. N. 2000. Environmental inequality formation: Toward a theory of environmental injustice. *American Behavioral Scientist* 43, no. 4: 581-601.
- Pollack, P. H. III, and M. E. Vittas. 1995. Who bears the burdens of environmental pollution? Race, ethnicity and environmental equity in Florida. *Social Science Quarterly* 76, no. 2: 294-310.
- Powell, J. L. 1984. Least absolute deviations estimation for the censored regression model. *Journal of Econometrics* 25: 303-325.
- Rabin, Y. 1989. Expulsive zoning and the legacy of Euclid. In *Zoning and the American dream: Promises still to keep*, ed. C. M. Haar and J. S. Kayden, 101-121. Chicago: Planners Press and American Planning Association in association with the Lincoln Institute of Land Policy.
- Sadd, J. L., M. Pastor, J. T. Boer, and L. D. Snyder. 1999. "Every breath you take. . .": The demographics of toxic air releases in southern California. *Economic Development Quarterly* 13, no. 2: 107-123.
- United Church of Christ. Commission for Racial Justice. 1987. *Toxic wastes and race in the United States: A national report on the racial and socioeconomic*

*characteristics of communities with hazardous waste sites*. New York: United Church of Christ.

- U.S. Department of Commerce. Economics and Statistics Administration. Bureau of the Census. 1992. 1990 Census of Population and Housing: Summary Tape File 3 on CD-ROM. Washington, D.C.: U.S. Department of Commerce.
- U.S. Department of Commerce. Economics and Statistics Administration. Bureau of the Census. 1994a. County and City Data Book. Washington, D.C.: U.S. Department of Commerce. Available at <http://fisher.lib.virginia.edu/ccdb>; accessed 15 June 2001.
- U.S. Department of Commerce. Economics and Statistics Administration. Bureau of the Census. 1994b. *Geographic areas reference manual*. Available from <http://www.census.gov/geo/www/garm.html>; accessed 16 October 2001.
- U.S. Department of Commerce. Economics and Statistics Administration. Bureau of the Census. 2001. TIGER Frequently Asked Questions. Available at <http://www.census.gov/geo/www/tiger/faq.txt>; accessed 2 November 2001.
- U.S. Environmental Protection Agency. Office of Air Quality Planning and Standards. Emissions, Monitoring and Analysis Division. 1995. *User's Guide for the Industrial Source Complex (ISC3) Dispersion Models, Volume II - Description of Model Algorithms*. Research Triangle Park, NC: U.S. Environmental Protection Agency. Available at <http://www.epa.gov/scram001/userg/regmod/isc3v2.pdf>; accessed 2 November 2001.
- U.S. Environmental Protection Agency. 2000. *Summary of 1998 Toxics Release Inventory data*. Available at <http://www.epa.gov/tri/tridata/tri98/data/1998datasumm.pdf>.