NCEE
NATIONAL CENTER FOR
ENVIRONMENTAL ECONOMICS

Effects of Socio-Economic and Input-Related Factors on Polluting Plants' Location Decisions

Ann Wolverton

Working Paper Series

Working Paper # 08-08
August, 2008

U.S. Environmental Protection Agency
National Center for Environmental Economics
1200 Pennsylvania Avenue, NW (MC 1809)
Washington, DC 20460
http://www.epa.gov/economics

Effects of Socio-Economic and Input-Related Factors on Polluting Plants' Location Decisions

Ann Wolverton

NCEE Working Paper Series

Working Paper # 08-08
August, 2008

EFFECTS OF SOCIO-ECONOMIC AND INPUT-RELATED FACTORS ON POLLUTING PLANTS' LOCATION DECISIONS

August 15, 2008

By

Ann Wolverton

National Center for Environmental Economics
U.S. Environmental Protection Agency
1200 Pennsylvania Ave., NW
MC1809T
Washington, DC 20460

Abstract

Many environmental justice studies argue that firms choose to locate waste sites or polluting plants disproportionately in minority or poor communities. However, it is not uncommon for these studies to match site or plant location to contemporaneous socioeconomic characteristics instead of to characteristics at the time of siting. While this may provide important information on disproportionate impacts currently faced by these communities, it does not describe the relationship at the time of siting. Also, variables that are important to a plant's location decision – i.e., production and transportation costs - are often not included. Without controlling for such variables, it is difficult to evaluate the relative importance of socioeconomic characteristics in a firm's initial location decision. This paper examines the role of community socioeconomic characteristics at the time of siting in the location decisions of manufacturing plants while controlling for other location-relevant factors such as input costs.

When plant location is matched to current socioeconomic characteristics results are consistent with what the environmental justice literature predicts: Race is significant and positively related to plant location, while income is significant and negatively related to plant location. When plant location is matched to socioeconomic characteristics at the time of siting, empirical results suggest that race is no longer significant, though income is still significant and negatively related to plant location. Poverty rates are sometimes significant but act as a deterrent to plant location. Variables traditionally considered in the firm location literature - such as land and labor costs, the quality of labor, and distance to rail - are significant. The presence of pre-existing TRI plants in a neighborhood and average plant size are also significant.

Key Words: environmental justice, firm location.
Subject Area Classifications: Environmental Policy, Cost of Pollution Control, Distributional Effects

1. Introduction[1]

In the economics literature, a firm is assumed to evaluate potential locations for a new plant based on the principle of profit maximization. In doing so, the firm takes into account many location-specific attributes related to production and transportation costs that may affect potential profits in each of these locations. More recently, the environmental justice literature has claimed that firms discriminate against particular groups of people in siting polluting activities – in other words, that they have traded profits in favor of discrimination against poor or minority communities.

There are two major shortcomings of these studies in the context of the existing firm location literature. First, it is not uncommon for environmental justice studies to match site or plant location to contemporaneous socioeconomic characteristics instead of to characteristics at the time of siting. This relationship is important for establishing whether particular populations currently face disproportionate impacts but does not speak to whether this relationship existed at the time of siting. Second, variables that are important to a plant's location decision – i.e., production and transportation costs - are often not included in these studies. Without controlling for such variables, it is difficult to evaluate the relative importance of socioeconomic characteristics in a firm's initial location decision. This paper attempts to wed these two disparate literatures into a consistent framework by examining the role of community socioeconomic characteristics at the time of siting in the location decisions of manufacturing plants while controlling for other location-relevant factors such as input costs.

Concern over how siting decisions relate to the socioeconomic characteristics of the surrounding community stems from two studies. The GAO (1983) examined four hazardous waste landfill sites, pointing out that these sites are surrounded by predominantly poor and

[1] For their helpful comments and suggestions, I thank Don Fullerton, Dan Slesnick, Peter Wilcoxen, Sarah West, Ron Shadbegian, Arik Levinson, Matthew Kahn, Randy Becker, Vernon Henderson, and participants of the NBER Summer Institute in Environmental Economics.

minority communities. Likewise, the United Church of Christ (1987) studied U.S. commercial treatment, storage, and disposal facilities for hazardous waste (TSDFs) and found evidence of a disproportionate location pattern.[2] Goldman and Fitton (1994) reexamine and update data analyzed in the United Church of Christ study: the authors conclude that "the disproportionate impacts first identified and documented in the 1987 report…have grown more severe." These studies, and others that followed, have used the correlation between waste facility location and poor and minority communities to argue that a disturbing trend of siting discrimination exists.

Other studies that match site location to contemporaneous socioeconomic characteristics to explore the issue of environmental equity find mixed results.[3] Zimmerman (1993) finds that a greater percent of minorities live near inactive hazardous waste sites that appear on the National Priority List, but that the population living below the poverty rate does not differ significantly from the national average. Boer et. al (1997) find a significant and positive relationship between the location of hazardous waste TSDFs and race in Los Angeles, California. Sadd et.al. (1999) find little evidence of a correlation between the location of TRI plants and race in Los Angeles at the census tract level, but find evidence of a positive and significant relationship for tracts within one mile radius of the plant. Income is significant for both sets of regressions. Both Anderton et. al (1994), and Davidson and Anderton (2000) find only limited evidence of disproportionate numbers of hazardous waste facilities in minority or poor neighborhoods.

Because these studies match contemporaneous socioeconomic data to facility location, it is only possible to say that sites are disproportionately located in neighborhoods in which the current population is poor or minority. Such relationships are undoubtedly important in determining whether certain populations face disproportionate health impacts. However, they do

[2] The GAO study examine four offsite landfills in the U.S. and their surrounding populations at a fairly aggregate level. It finds that in three cases the majority of the nearby community is African American in 1980. The UCC study examines 415 U.S. hazardous waste sites at the zip code level. Using a difference of proportions test, race, household income, and property value are found to be significant in 1980.

not shed light on whether the relationship between race, income, and site location existed at the time of siting, or developed over time. To determine whether a firm considered differences in race or income across neighborhoods in its location decision, the characteristics of the neighborhood at the time the siting decision was made are the most relevant.

There are a handful of studies that have examined the relationship between neighborhood characteristics and facility location decisions at the time of siting. However, most of these studies control for few cost-relevant location variables, instead concentrating on socioeconomic factors. Been (1994) reexamines site location data used by the GAO (1983) and Bullard (1983) studies,[4] matched to socioeconomic characteristics at the time of siting. Even without additional explanatory variables, Been obtains much more mixed evidence of environmental inequity at the time of siting than the original studies. The GAO data still demonstrate a case for disproportionate siting in minority communities, while the Bullard data indicate that at least part of the relationship developed after siting. Been and Gupta (1997), using a larger data set, also obtain mixed evidence for the hypothesis that race played a role at the time of siting for active commercial hazardous waste TSDFs in the U.S. While waste disposal sites are correlated with certain 1990 socioeconomic characteristics such as race and income, neither percent poor nor percent African-American in a neighborhood are significant factors at the time of siting. The percent Hispanic remains significant at the time of siting.[5] Pastor, Sadd, and Hipp (2001) examine the location of TSDF sites in Los Angeles County and find greater evidence of disproportionate siting in established Latino and African American communities than minority move-in after the TSDF establishment. Baden and Coursey (2002) examine the location of

---

[3] A wide variety of analytical techniques are used in these studies, ranging from an examination of simple correlation coefficients and differences in means across samples to binary logit regressions.

[4] Bullard (1983) examined several landfill and incinerator sites in the Houston area and found that the vast majority of these sites are located in African-American neighborhoods.

[5] One serious problem with the study is the limited sample size of the original data sets. The GAO study only has a sample size of four, while the Bullard study (after Been's adjustments) has a sample size of ten. As Been points out, however, her research does effectively illustrate gaps in the literature.

Superfund sites in Chicago and find that sites were disproportionately located in poor neighborhoods in the 1960s but not in the 1990s. However, they find little evidence for disproportionate exposure of African Americans either currently or at the time of siting. Jenkins, Maguire, and Morgan (2004) study compensation to communities in exchange for hosting municipal solid landfills. Controlling for tipping fees paid from the landfill to the community, they find that socioeconomic characteristics such as income and race do not matter at the city level but do appear to matter at the county level.[6]

In this paper, I match the plant to the neighborhood at the time of siting to examine whether community socioeconomic characteristics are related to facility location decisions. I examine the location decisions of manufacturing plants that report to the Toxic Release Inventory (TRI) in Texas after 1975. These plants are typically larger polluters of toxic substances and therefore pose a potential health hazard to populations located nearby. Texas seems particularly relevant as a focus for several reasons. First, during the period of study it ranks second among U.S. states in both land area and population. Second, according to the United Church of Christ study, Texas had the second highest number of blacks and Hispanics living near uncontrolled hazardous waste sites in 1980.

This study differs from other environmental justice studies of plant location in several important ways. First, I use a conditional logit model to represent the choice of a particular

---

[6] Lambert and Boerner (1995) examine TSDFs, landfills, and inactive hazardous waste sites in the St. Louis, Missouri area at the time of establishment in the context of changing socioeconomic dynamics. They do not find large initial differences in the percent of poor and minority residents between neighborhoods with and without waste sites. However, it appears that housing values grew less rapidly in neighborhoods with waste sites and that minority populations moved into these neighborhoods at a faster rate. Hersh (1995) conducts a historical analysis of the change in racial and industrial dynamics in the Pittsbugh area for firms reporting to the Toxic Release Inventory (TRI). He finds that, in general, industries and blue-collar neighborhoods located near each other for job-related reasons. Also, he notes that both white and rich residents took flight to cleaner parts of the city after firms located in a particular neighborhood, and that there was an eventual movement of minorities into more polluted areas. Krieg (1995) examines Superfund sites in and near Boston. He finds that race is associated with the number of waste sites in areas with a long history of industrial activity and that class is more closely associated with the number of waste sites in areas with more recent industrial activity.

location from a set of many neighborhoods. A firm is allowed to select a location for its plant from the actual location chosen and a given number of randomly selected alternatives drawn from the full choice set. Pastor (2001), Davidson and Anderton (2000), Been and Gupta (1999), Boer et al. (1997), and Anderton et al. (1994) all used a binary response model to examine plant or waste site location. Allowing for multiple location alternatives seems appropriate, since firms typically choose from a spectrum of competitive locations when deciding where to site a plant.

Second, I include a number of variables that are often left out of environmental justice studies but are included in other studies of firm location. Traditional location theory identifies differences in input costs as key determinants of firm location decisions. Most studies of new plant location include measures of labor costs, land costs, transportation costs, energy costs; and/or level of taxation.[7] A few environmental justice studies include variables to proxy for land and labor costs but rarely include other variables associated with firm location. Kriesel et al. (1996), while focusing on the incidence of emissions rather than plant location, is one exception. Along with land and labor costs, they include proximity to an interstate highway and find that the inclusion of these factors renders race and poverty insignificant. While several industry location variables are unavailable for any years prior to 1990 and therefore not included in this study, I do include measures of the costs of land and labor, the quality of labor, the degree of urbanization, average plant size, and distance to rail-based transportation in this study.

I also attempt to control for the effect of pre-existing TRI sites on plant location decisions. These pre-existing sites may indicate to the firm how much it can pollute if it also locates there, may proxy for factors such as zoning restrictions, or may serve as an indication of agglomeration economies. None of the previous environmental justice studies control for the possible effect of the existence older hazardous waste sites in the neighborhood.

[7] See Carlton (1983), Bartik (1985), Lee and Wayslenko (1987), McConnell and Schwab (1990), Finney (1994), and Levinson (1996).

Finally, I consider the possible effect of the collective action of a neighborhood on a plant's location decision. Neighborhoods that are less likely or able to express opposition through political or community action are posited to require less potential compensation to offset the effects of polluting activities. Hamilton (1995), using voter participation data, finds that firms take into account the likelihood of a neighborhood to engage in collective action when deciding whether to expand waste facilities. Likewise, Arora and Cason (1998) substitute variables representing a population's stake in the neighborhood for voter participation and find several of these variables to be significant. It stands to reason that such considerations also would be relevant to a firm's initial decision of where to locate a new plant.

When the locations of TRI plants in Texas are matched only to current socioeconomic characteristics results are broadly consistent with those of many previous studies and with what the environmental justice literature predicts. Race is significant and positively related to plant location, while income is significant and negatively related to plant location. Poverty is either insignificant or of the opposite sign of what the environmental justice literature would predict.

When plant location is matched to socioeconomic characteristics at the time of siting, empirical results suggest that, contrary to results cited in the environmental justice literature, race is not significantly related to plant location. Income remains significant and negatively related to plant location. Poverty rates, another variable of concern in the environmental justice literature, are sometimes significant but continue to act as a deterrent to plant location.

The presence of pre-existing TRI plants in a neighborhood and average plant size, variables left out of previous environmental justice studies, are both significant at the one percent level. Other variables traditionally considered in the firm location literature but often omitted from environmental justice studies - such as land and labor costs, the quality of labor, and distance to rail - are also significant. Most of the variables associated with collective action by the community are not significant. The percent of voters that participate in national elections is significant. However, plants appear to be attracted to these communities rather than deterred by them as Hamilton's hypothesis would predict.

2. Firm Location Theory

A firm evaluates potential locations for a new plant based on the principle of profit maximization. In doing so, the firm takes into account many location-specific attributes that may affect potential profits in each of these locations. Two types of costs that vary with location, identified in the firm location literature, are production costs and transportation costs.[8] Production costs include costs related to relatively immobile inputs such as land, labor, and housing, and costs related to operation such as taxes, public utility fees, and environmental regulations. Transportation costs include freight rates, distance to input markets, and distance to output markets. It is also important to consider any offsetting location benefits from agglomeration economies such as a shared infrastructure or labor pool.

Hamilton (1995) offers three additional theories for why a plant may locate in a poor or minority neighborhood. The first theory of plant location stems from Coase (1960): a plant is established where residents' valuation of environmental quality, and therefore the potential compensation by the firm to the neighborhood residents, is lowest. Since local willingness to pay for environmental quality is positively correlated with income, firms will tend to locate plants in poorer neighborhoods to minimize the costs of compensation.[9]

The second theory for why plants may locate in poor or minority neighborhoods is that firms locate polluting plants where the likelihood of a community engaging in collective activities is relatively low. In this case, a firm owes less to the community in the form of compensation not because the neighborhood values the externality any less than other

[8] See Harrington and Warf (1995) and Beckman and Thisse (1986).

communities, but because the transaction costs of collective action are high. A less-politically-powerful community may have a limited ability to participate in the political or legal processes through which opposition is voiced and compensation is gained.

Hamilton's final theory is that firms - or rather their managers or owners – gain utility from discriminating against a particular demographic group by locating a heavily polluting plant in that community. Since it is easier and therefore less costly to discriminate in neighborhoods with a substantial minority population, plants tend to locate in these neighborhoods.

3. A Model of Plant Location

Levinson (1996) models new plant births in a particular location using a reduced-form latent profit function. A firm has an unobserved profit function for each possible location that is a function of location-specific variables such as factor prices, fixed inputs (land, labor) and the stringency of environmental regulation.[10]

One can add to the list of location-specific variables considered by the firm, the cost of discrimination in the form of foregone profits and the cost of required compensation. Adapting Levinson's model, a firm *i* that is considering the location of a new plant has an unobserved profit function for each feasible neighborhood *j* :

$$\pi_{ij}^{*} = g\left[f_i\left(p_j, x_j, d_j, c_j\right), s_{ij}\right] \tag{1}$$

where p_j is neighborhood-specific input prices, x_j is neighborhood-specific fixed inputs, d_j is foregone profits due to discrimination, c_j is the level of required compensation, and s_{ij} is other firm or plant-specific factors that may vary by neighborhood.

---

[9] Compensation can be thought of as both monetary and in-kind (e.g. free access to certain services, the building of a community park) forms of renumeration given by the firm to the community to offset the perceived risks of an increase in pollution due to the location of a new plant in the area.

The firm will choose to locate a plant in the neighborhood that yields the highest potential profit. An increase in the cost of a location – due to an increase in input prices, the cost of discrimination, or the level of compensation required - implies a decrease in profits. An increase in the availability of inputs implies an increase in profits.

Without price rigidities or other market failures in the input market, an increase in the availability of inputs will also result in a decrease in their price. Thus, input availability can be eliminated as a separate variable in the latent profit function. In addition, compensation is a function of two components: the value placed on environmental amenities in the neighborhood e_j, and the propensity of the neighborhood to engage in collective action a_j. Equation (1) can be rewritten as follows

$$\pi_{ij}^* = g\left[f_i\left(p_j, d_j, e_j, a_j\right), s_{ij}\right]. \tag{2}$$

4. The Econometric Model

A number of environmental justice studies use binary response models to analyze plant location decisions.[11] Most of these studies compare neighborhoods with plants or waste sites to neighborhoods without a plant or waste site. Multiple choice models seem more appropriate for the analysis of plant location, since firms choose from a spectrum of competitive locations when deciding where to site a new plant, including those that may already host a plant.

I use a conditional logit model that allows for analysis of the location decision relative to choice and plant-specific attributes. Assume that firm *i* faces *J* possible plant location alternatives and that these *J* choices are independently and identically distributed. The firm will choose location *j* when its profits are maximized in that particular location compared to all other

---

[10] Because I examine the location decisions of plants in one state, environmental regulation is not included as a variable. While stringency and enforceability may still differ substantially within a state, little data are available at such a disaggregated level.
[11] See Anderton et al (1994), Been and Gupta (1997), and Boer et al (1997).

possible choices. In other words, $\pi_{ij} \geq \pi_{ik} \quad \forall k \neq j$ where π_{ij} represents firm *i*'s profits if it decides to locate in *j* and π_{ik} represents firm *i*'s profits if it decides to locate in k. It is possible to write firm *i*'s profits as follows.

$$\pi_{ij} = \beta' z_{ij} + e_{ij} \qquad (3)$$

where z_{ij} is defined as a set of observed characteristics specific to location *j* and plant *i*.[12] Specifically, $z_j = (p_j, d_j, e_j, a_j, s_{ij})$. Assume that the error term e_{ij} has a Weibull distribution. If the firm's underlying production function is assumed to be Cobb-Douglas, then profits will be log-linear.[13]

Conditional on the decision to open a new plant, the probability that firm *i* will choose particular location k can be written as:

$$Pr(ik) = \frac{e^{\beta' z_{ik}}}{\sum_{j=1}^{J} e^{\beta' z_{ij}}} \quad . \qquad (4)$$

4.1 Defining Alternative Plant Locations

When the full choice set of feasible location alternatives is small, firms may choose among all possible location alternatives in an empirical model.[14] However, when the number of location alternatives is large relative to the number of observations, they must be restricted to allow for the practicable study of plant location. The data set used in this study has approximately 360 plants and over 2,300 possible locations from which each plant may choose.[15]

---

[12] In this particular case, the firm and plant are virtually interchangeable entities. Over 85 percent of parent firms in the data set have only one plant in Texas. Ten percent of the parent firms have two plants in Texas.
[13] See Bartik (1985), and Levinson (1996).
[14] See Levinson (1996), and Finney (1994).
[15] These 2,300 census tracts are a subset of possible locations in Texas. Feasible alternatives are limited to census tracts that are defined in both the 1980 and 1990 census. Census tracts were not defined for many rural counties in Texas in 1980 and are therefore excluded from the choice set.

I define the reduced choice set for the plant as the selected location and a given number of randomly selected alternatives from the full choice set.[16] This technique yields consistent estimates and has the added advantage that the likelihood function is identical to that used for estimating a conditional logit with the full choice set (McFadden 1978).

4.2 Independence of Irrelevant Alternatives

A conditional logit model requires the assumption of independence of irrelevant alternatives (IIA), which does not allow for correlation across subsets of choices. However, if profits are correlated across regions, then the IIA assumption is inappropriate. Just as states in a particular region in the country may have correlated disturbance terms due to regional market trends, so may census tracts in a particular city or part of the state. While it appears that correlation between states in a region is often a concern (Bartik 1985, Levinson 1996), McConnell and Schwab (1990) find little evidence of correlation between counties in four broad regions. Carlton (1983) notes that at the SMSA level, geographic regions are quite distant from each other and therefore not likely to be correlated

One method of relaxing the IIA assumption is to estimate a nested logit (McFadden 1978). A firm is assumed to choose to locate in a particular region or city according to criteria that determine feasibility. The firm then selects a particular location within that region or city as determined by profit maximization. In this latter stage, it is often assumed that the firm is choosing a particular location from a set of geographically adjacent neighborhoods, although this need not be the case. It is in this way that the probability of firm i choosing to locate in neighborhood j is adjusted for the correlation between locations in the same city or region.

A simpler but less elegant method of relaxing the IIA assumption is to include regional dummy variables in the conditional logit estimation for those characteristics likely to be

[16] See McConnell and Schwab (1990), and Friedman, Gerlowski, and Silberman (1992).

correlated across geographic units (Bartik 1985). To examine the validity of the IIA assumption for census tracts in Texas, four sets of regional dummies are explored in the context of the regressions: (1) location in a county along the Mexican border, (2) location in a county adjacent to the Gulf of Mexico, (3) location in the Houston or Dallas area, and (4) location along a border with a neighboring state.

4.3 Defining the Neighborhood

How a researcher defines the neighborhood may affect the empirical results of a study. Broad neighborhood definitions may hide important underlying trends in plant behavior, while narrow definitions may exclude areas that should be included in the neighborhood.[17] Early work related to environmental justice uses fairly broad neighborhood proxies in the form of counties or zip codes (GAO, UCC). Later studies have defined the relevant neighborhood based on concentric circles surrounding a site or at the census tract level.[18] The census tract allows for a fine degree of analysis, tends to be consistently defined over time, is generally a comparable unit of analysis,[19] and is usually defined by the community itself to reflect its own view of the neighborhood (Been and Gupta 1997). For the purpose of this study, I define the neighborhood as the census tract.

5. Data

---

[17] See Anderton et al (1994).

[18] The use of circles is meant to approximate the range of distances at which a resident is concerned about the location of a hazardous waste facility. Glickman and Hersh (1995) prefer concentric circles to a census-based neighborhood definition for several reasons. First, a census-based definition ignores how households or different socioeconomic groups are distributed within the neighborhood. Second, a census-based definition often reflects topographical features that may exclude a portion of those who, although separated by some physical feature, receive a large portion of the negative externalities from the site or plant. One reason for not using the concentric circle technique is the arbitrary choice of a radius: the circles drawn are unlikely to reflect community-defined borders between neighborhoods.

[19] A census tract usually consists of 2,500 to 8,000 people, with an average of 4,000 residents. The spatial size of a tract, however, may vary depending on population density.

To analyze plant location in the context of environmental equity, I use the Toxic Release Inventory (TRI), the U.S. Census of Population and Housing, the U.S. Census of Manufactures, the County and City Data Books, and several directories of manufacturers for the state of Texas.[20] .Each TRI plant in Texas is matched to the census tract in which it is located.[21] Any plant that appears in the TRI at least once between 1988 and 1993 is eligible for inclusion.[22] A plant is dropped from the data set if the street address in the TRI cannot be matched to a census tract, if the plant's establishment date is unavailable,[23] or if the plant was established prior to 1976.[24] Plants established between 1976 and 1985 are matched to the 1980 census tract and socioeconomic characteristics from the U.S. Census of Population and Housing. Plants established between 1986 and 1993 are matched to the 1990 census tract and accompanying socioeconomic data.[25] Two steps are taken to ensure that plants select a location from the same alternative choice set across census years. First, only counties for which tract definitions are assigned in both 1980 and 1990 are included in the study. Second, tract definitions that have changed from one census year to the next are aggregated to a common tract definition. The Census of Manufactures is also used to add average wage and average plant size in the same industry at the county and MSA-level to the data set. Information from the County and City Data Book is utilized to add voter participation data.

---

[20] The establishment date for each plant is collected from the Bureau of Business Research Directory of Texas Manufacturers: Volume I (1990-1993), the Harris Texas Manufacturers Directory (1995), the Texas High Technology Directory (1995), and the Texas Manufacturers Register (1994). It is important to note that the main source of plant establishment dates, the Bureau of Business Research Directory of Texas Manufactures, stopped collecting this information after 1993.

[21] Due to possible unreliability of the data in the first reporting year, 1987 TRI data are not used.

[22] Plants that use more than 10,000 pounds or manufacture more than 25,000 pounds of the 329 listed toxic chemicals are required to report how much of each chemical is released into the air, land, or water. Many plants submit more than one report, depending on the number of chemicals they emit that are above the reporting standard. One limitation is that the TRI excludes a number of important toxic chemicals and important non-industrial sources of pollution, such as dry cleaning establishments.

[23] I was unable to find an establishment date for 221 of 1,675 plants listed in the TRI from 1988 to 1993.

[24] Plants sited prior to 1976 are not included in the analysis due to unavailability of many of the variables in the 1970 U.S. Census. About 990 TRI plants for which I have establishment dates were sited before 1976. While older plants' location decisions are not analyzed, these plants are taken into account as a possible explanatory variable for the location of plants of more recent vintage.

5.1 Input-Related Costs

The cost of inputs to production - such as land, labor and transportation- is important in a firm's evaluation of possible plant locations. To capture differences in the cost of land, I use the average property value of owner-occupied housing in a neighborhood, $PROPERTY_j$.[26][27] It is assumed that housing values represent a more general trend associated with the price per acre, regardless of whether it is dedicated to residential or industrial use. Given a firm's desire to minimize cost when establishing a plant in a particular location, it is expected that higher property values decrease the likelihood of locating in that neighborhood.[28]

Labor costs in a given location are related to the wage a plant pays to its workers, the ease with which the plant is able to hire workers, and the qualifications of those workers. All else equal, a firm will prefer a location that provides access to a large yet inexpensive pool of available workers who match the hiring needs of the plant. The average wage of a production worker in manufacturing at the county level, $WAGE_j$, is used as a proxy for the market wage in a particular neighborhood. The quality of workers available in a neighborhood is measured by the percent with at least a high school diploma, $HIGHSCH_j$.[29]

To capture transportation costs, I include a measure of the average distance of a given neighborhood from the nearest railroad, $RAIL_j$, which was derived by overlaying census tract boundary information with geographic information system data that identify the location of

---

[25] Results are not sensitive to one or two year shifts in the set of years matched to each Census.
[26] Both property values and household income are adjusted to 1980 dollars. The consumer price index for the southern region of the United States is used to make this adjustment.
[27] Use of property values in the same regression as income is potentially problematic since they are highly correlated. I also explore the use of the percent of housing in a neighborhood that was built prior to 1970 as a proxy for property values. This measure is not expected to be a perfect substitute for property values since it only captures housing stock age, but it allows me to explore the robustness of the results.
[28] Property values may be positively related to plant location if higher land values also proxy for increased quality or usefulness of land, and if quality is an important consideration in the location decision.

major railways throughout Texas.[30] It is generally expected that the further away the facility from rail transportation, the higher its shipping costs.[31]

Arora and Cason (1998) also include percent of population employed in manufacturing, $MANUF_{ij}$. In their analysis, the percent of workers in manufacturing is meant to reflect the way in which workers evaluate the trade-off between jobs and environment, which in turn may depend on whether their income is directly generated from the polluting industry. However, from the firm's perspective, percent employed in manufacturing may function as an indicator of how many workers matching the hiring needs of a new plant are located in the area.

A firm may consider how many pre-existing polluting plants there are in a particular neighborhood as a criterion in its initial location decision. The number of polluting plants may indicate the existence of agglomeration economies. The firm would have an incentive to locate a plant in a neighborhood with other polluting plants to take advantage of spillover effects that reduce productivity costs. The number of polluting plants may also serve as an indicator of the amount a plant can pollute if it locates in that neighborhood (for instance, due to zoning, or people's tolerance for relatively high levels of pollution). To account for the role that such factors play in the location decision, the number of pre-existing TRI facilities in the same census tract is included as a variable, $OLDSITE_j$. It is expected that the higher the number of pre-existing TRI plants in a neighborhood, the more likely a plant will locate in that neighborhood.

To take advantage of localization economies, a plant may locate near other plants in the same industry. The larger the industry, the lower the production costs to a locating plant. The average number of workers per plant at the two-digit SIC level in a MSA, referred to as

[29] Percent with high school degrees is used instead of college educated, because percent with college degrees is highly correlated with income while percent with high school degrees is not.
[30] I also explore a variable measuring the average distance to a major highway. It was not significant in any of the regressions nor changed the sign or significance of other variables.
[31] I also examined whether environmental regulatory costs are relevant in this context. In alternate regressions not reported here, I included a measure of whether a plant was located in a county out of attainment for criteria air pollutants. This variable was not significant for any of the regressions.

$SCALE_{ij}$, is used to control for industry size.[32] It is expected that the larger the average plant size in the area, the more likely a plant will locate in that neighborhood.

Finally, a variable measuring whether a tract is in an urban area, $URBAN_j$, is also included in the analysis. Urban areas offer access to large labor pools, a potentially large output market, better infrastructure, and easy access to public services. However, they also tend to have higher taxes, more traffic, and more crime.[33] If the benefits of locating in an urban area are greater than the costs, then urban areas may attract more plants. If however the additional costs of locating in an urban area dominate the benefits, the urban area may attract fewer plants.

5.2 Compensation Costs

The potential compensation a firm pays to a neighborhood depends on willingness to pay for environmental quality and the neighborhood's propensity for collective action. Residents' willingness to pay for environmental amenities is most closely associated with income levels, $INCOME_j$. If environmental quality is a normal good, then higher income will lead to greater demand for clean air and a higher level of compensation for plants to dirty the air. Firms will seek to avoid this cost by locating in neighborhoods with lower incomes. The percent of households living below the poverty line, $POVERTY_j$, is also included as a variable.[34] If a firm compensates each member of the neighborhood, then the more people living in a neighborhood,

[32] Where a plant locates also depends on plant-specific characteristics: what type of product it manufactures, desired plant scale, type of technology used, input mix used in production. To include firm-specific factors, one would have to interact them with *all* possible location choices. Unless the data set is very large or the choice set is small, there are not enough degrees of freedom to accommodate such a procedure.
[33] Ideally, I would include a measure of the level of taxes faced by plants in each neighborhood. However, such data are unavailable for the 1980s.
[34] The percent in poverty is highly correlated with income and percent non-white so I explore an alternate measure of poverty, the percent without a phone in their home. This measure has the benefit of being fairly highly correlated with poverty (67 percent) but being far less correlated with the income and race variables.

the more costly to the firm and the less likely it will locate a plant. To measure this effect, the number of people affected, POP_j, and the housing vacancy rate, $VACANCY_j$, are included.

Arora and Cason (1998) include variables that affect a population's "stake" in the neighborhood as well as their desire to free ride. Neighborhood residents that have a higher stake will tend to engage in collective action more often, while those with a lesser stake in the well-being of the neighborhood will attempt to free ride on the participation of other residents. A firm wishing to minimize the potential costs of compensation will locate in neighborhoods that have fewer of the more active types of residents. Included as collective action variables are the average number of children per household, $CHILD_j$; the percent over the age of 65, $AGE65_j$; and the percentage of households that lived in the neighborhood for at least five years, $NOMOVE_j$. Neighborhoods that have a higher percent of households with children, older residents, or longer-term residents are expected the engage in higher levels of collective action and should therefore be negatively related to a firm's propensity to locate a plant in that neighborhood. Finally, I also include $VOTE_j$, the percent of residents in a county that voted in the previous Presidential election, as a measure of a neighborhood's propensity to engage in collective action (see also Hamilton 1995, Zimmerman 1993, and Arora and Cason 1998).[35]

5.3 Discrimination Costs

Two variables are closely associated with the theory that when firms want to discriminate against a particular population, they will find it easier and therefore less costly to do so in neighborhoods in which a greater number of that group live. Percent nonwhite in a census tract, $NONWHT_j$, is used to represent the racial composition of a neighborhood. Ethnic or

[35] Voter participation data for all counties in Texas are available for a number of state elections from 1986 to the present. However, data of this type for prior to 1986 are only available for some counties.

immigrant populations may also be subject to discrimination.[36] Percent foreign-born, $FOREIGN_j$, captures the segment of the population least likely to be assimilated into mainstream American culture and arguably most subject to discriminatory behavior on the part of the firm. The more minorities or foreign-born residents, the more likely it is for a firm to discriminate through the simple act of locating a polluting plant in that neighborhood.

6. Previous Studies

This section briefly presents results for a set of regressions that match all Texas plants reporting to TRI regardless of establishment date to 1990 characteristics. This exercise is directly comparable to data matched in previous studies to contemporaneous socioeconomic characteristics. I run regressions that use variables similar to those presented in three published studies.[37] In each case, a logit regression is used where a value of one signifies the presence of a TRI plant of any vintage, and zero signifies the absence of a TRI plant. The results are useful for two reasons. First, these regressions allow for the examination of the pattern of plant location and neighborhoods to ascertain whether certain populations are disproportionately affected. Second, it allows for an assessment of the consistency of this data set with others that have found evidence of environmental inequity

[36] Because Hispanics are included in both percent nonwhite and percent white in the US Census, using percent Hispanic directly in the regression may be both confusing and redundant. However, in Texas the percent foreign-born is strongly correlated with percent Hispanic. I also constructed variables to measure racial concentration, how much a census tract was above or below the county average. These variables were never significant and did not change the results for other variables.

[37] Goldman and Fitton (1994), Davidson and Anderton (2000), and Kriesel et. al (1996). Note that Goldman and Fitton only examine mean values and correlation coefficients. Davdison and Anderton calculate odds ratios. Also, while the set of variables used for each of these regressions is similar to those used by the authors, they are not identical. For instance, Kriesel et al use a measure of proximity to an interstate highway as an independent variable, which is not included here.

Table 1: Logit Regressions Using All TRI Plants and 1990 Socioeconomic Characteristics

Variable	Goldman and Fitton	Davidson and Anderton	Kriesel et al
	Coefficient Estimates		
CONSTANT	5.78 ***	-2.63 ***	-6.14 **
NONWHT (percent nonwhite residents)	0.49 **	0.59 **	0.44 *
FOREIGN (percent foreign residents)	1.44 **	1.29 **	1.09 *
POVERTY (percent persons living in poverty)	-2.19 ***	0.25	0.45
LNINCOME (log of median household income)	-0.73 ***		
URBAN (percent persons living in urbanized area)	-0.35 **		
HIGHSCH (percent with high school degrees)		1.67 **	1.60 **
LNVALUE (log of average housing value – owner occupied)		-0.05 **	-0.05 **
MANUF (percent employed in manufacturing)		7.20 ***	7.06 ***
UNEMPL (percent of residents over age 16 that are unemployed)		-1.96	
WAGE (log of average county-level wage)			0.23
LNPOP (log of total population)			0.15 *

* indicates significance at the 10 % level, and ** indicates significance at the 5 % level, and *** indicates significance at the 1 % level. Numbers reported are coefficient estimates and not marginal effects.

For all logit regressions in Table 1, percent nonwhite and foreign are significant and positively related to plant location, confirming the potential for current disproportionate impacts. Percent in poverty is not significant, except in the Goldman and Fitton regression, and is opposite in sign from what the environmental justice literature predicts. Median income is only included in one regression but is also significant and negatively related to plant location. In other words, richer neighborhoods are less likely to attract a TRI plant. In the regressions in which it is included, property value is also negatively related to plant location. All three sets of results are consistent with allegations of environmental inequity and are broadly consistent with the studies they attempt to mimic. The one exception worth noting is that race is not a significant variable in regressions reported by Kreisel et. al, while it is significant here.

7. Summary Statistics

When TRI plant location is matched to socioeconomic characteristics at the time of siting, the data set consists of 361 TRI plants established in Texas after 1975. Approximately 70 percent of the plants were established between 1976 and 1985 (266 plants), while the remaining plants were established between 1986 and 1993 (95 plants).[38] Across both time periods, approximately 75% of the plants are in five two-digit industries: chemicals and allied products (SIC 28), rubber and miscellaneous plastics (SIC 30), fabricated metals (SIC 34), electronic and other electrical equipment (SIC 36), and industrial/commercial machinery and computer equipment (SIC 35). Ten industries account for about 90% of the plants in both time periods.

Several characteristics differ between tracts with a TRI plant established between 1976 and 1985 and tracts without a TRI plant established in this time period (see Table 2). Tracts in which a plant locates tend to have lower property values and poverty rates, and a greater number of people living in them.[39] They also tend to be less urban, have twice as many pre-existing TRI sites, and have larger plants in the same industry. Counter intuitively, tracts in which a plant locates also tend to have a greater percentage of longer-term residents living in them. Unlike studies that have matched site location to contemporaneous demographic variables, percent nonwhite is slightly lower in tracts in which a plant locates at the time of siting. Likewise, there are fewer foreign-born residents in neighborhoods with a TRI plant at the time of siting. Average income and percent who voted are not substantially different across the two sub-samples.

[38] Recall that the business directories relied on for plant establishment dates stopped collecting this information in 1993. For this reason, we are unable to expand the dataset further out in time. It is unknown how this biases the data for the later set of years.

Table 2: Summary Statistics

Variable	Tracts Matched to 1980 Characteristics		Tracts Matched to 1990 Characteristics	
	With a TRI Plant Established Between 1976 and 1985 (N=266)	**Without a TRI Plant Established Between 1976 and 1985** (N=2,186)	**With a TRI Plant Established After 1985** (N=97)	**Without a TRI Plant Established After 1985** (N=2,300)
NONWHT (percent nonwhite)	20.17 (23.03)	23.08 (26.28)	26.14 (18.62)	29.61 (26.03)
FOREIGN (percent foreign-born)	5.11 (5.93)	6.27 (7.24)	8.90 (8.44)	9.38 (9.49)
VACANT (percent of housing units vacant)	9.12 (8.02)	9.00 (6.63)	12.17 (7.03)	12.66 (8.30)
PROPERTY (ave. housing value - owner occupied)	$32,778.42 (18,618.31)	$38,392.72 (26,703.56)	$9,529.51 (5,614.80)	$13,983.63 (12,944.52)
INCOME (ave. household income)	$9,375.14 (3,027.44)	$9,856.88 (4,792.40)	$10,342.80 (4,139.15)	$10,692.34 (6,644.34)
POVERTY (percent in poverty)	11.58 (9.91)	14.22 (12.04)	16.97 (10.74)	19.61 (14.63)
HIGHSCH (percent with high school degree or more)	30.60 (8.39)	28.42 (8.46)	27.48 (6.37)	25.25 (8.10)
CHILD (percent with children under age 18 years)	43.41 (14.08)	40.37 (14.74)	40.10 (11.77)	36.83 (12.48)
AGE_65 (percent over age 65)	7.98 (5.96)	9.53 (6.28)	8.99 (6.26)	11.00 (6.19)
NOMOVE (percent not moved in at least 5 years)	42.93 17.24	46.88 (17.32)	46.07 (11.98)	50.17 (14.47)
AVEWAGE (county-level average wage per plant)	$16,292.64 (3,620.80)	$16,728.94 (4,265.16)	$21,157.86 (5,901.44)	$21,092.05 (5,506.05)
OLDSITE (number of pre-1976 TRI sites)	2.45 (4.01)	0.21 (0.63)	4.35 (6.40)	0.37 (1.47)
URBAN (percent living in urban area)	61.83 (45.27)	78.89 (39.05)	85.57 (26.83)	86.74 (29.87)
SIZE (MSA average number of workers per plant)	84.72 (92.58)	75.60 (86.13)	84.33 (95.96)	70.06 (87.10)
POP (total population)	5,477.40 (4,727.47)	4,830.21 (3,212.35)	8,591.28 (7,215.93)	5,908.84 (5,351.19)
RAIL (average miles to nearest railroad)	1.37 (1.61)	1.42 (1.66)	1.24 (1.27)	1.42 (1.66)
VOTE (percent that voted in the county for President)	34.58 (4.63)	33.67 (4.97)	34.43 (5.06)	34.78 (5.62)

* All means and standard deviations for variables in percentage from are multiplied by 100 for easy interpretation. In the data, percents range from 0 to 1.

39 Percent urban and population are not highly correlated: the correlation coefficient is only 0.11.

Many of the differences between tracts with and without a TRI plant established between 1986 and 1993 are similar to those already discussed for plants established a decade earlier. However, unlike for neighborhoods with plants established between 1976 and 1985, there no longer appears to be any difference in the percent urban for neighborhoods with and without a TRI plant. Also, the average number of older TRI plants in these neighborhoods has increased to roughly four times the number of old sites in neighborhoods without a plant.

Thus, initial evidence suggests that higher percent nonwhite, percent foreign, and percent poor are not associated with plant location at the time of siting. However, without controlling for other important variables, one cannot definitively say whether race or poverty is an important consideration in a plant's location decision. It is possible that other variables are masking a race or income effect. To clearly establish whether this is the case, a conditional logit is estimated, which will present the effect of race and poverty while controlling for other factors.

8. Results

The results of several conditional logit estimations, when each plant has 49 alternatives to its actual choice randomly selected from the overall alternative choice set, are presented in Table 3.[40] The first three columns present regression results for equations similar to those found in other studies. In each case, the data and estimation technique utilized differ from what was used originally. Columns (1) and (2) re-estimate equations similar to those analyzed in Davidson and Anderton (2000) and Kreisel et. al (1996) but match plant locations to socioeconomic characteristics at the time of plant siting. Column (3) re-estimates the equation from Been and Gupta (1997). The last two columns present new results. Column (4) presents results when only variables representative of discrimination and compensation costs are included in the regression. Columns (5) presents results when input costs are added to the regression in (4), consistent with the established literature on firm location. Column (6) represents a robustness check on the results in column (5). While property values and poverty rates are variables that have traditionally been included in the environmental justice literature, they are often highly collinear with other variables in the regression such as income and race.

First, note that in each specification percent nonwhite and percent foreign are both insignificant.[41] These results run directly counter to those of studies that have matched all plant or waste site locations to the current socioeconomic characteristics of the neighborhood. They also differ from Been and Gupta's results: they find no evidence of a relationship between waste site location and percent African-American at the time of siting, but do find evidence of a relationship between waste site location and percent Hispanic.

Second, note that while percent poor is significant at the five percent level for most regressions, the sign is the opposite of what is predicted in the environmental justice literature. Been and Gupta find a similar result for sites established in the 1980s or 1990s but offer no reason for why this may be the case. One possible explanation may be that firms interpret the percent poor in way similar to the unemployment rate: the higher the percent poor, the worse the general state of the economy and the more unskilled and potentially unemployable workers there are. Third, the input costs added in (5) and (6) are all highly significant and render percent graduated from high school insignificant in (5).

Only one of the variables related to environmental amenity costs is significant for the regressions in Table 4: average income.[42] As predicted, income negatively affects the likelihood

---

[40] The conditional logit is also estimated with nine and 24 alternatives randomly selected from the choice set. For both cases, none of the previously significant results change sign or significance.

[41] Results for regressions that include percent African-American and percent Hispanic instead of percent non-white and percent foreign are identical in sign and significance to those presented here.

[42] Several variables are in log form due to the Cobb-Douglas production function assumption: income, wage, property value, population, and scale. If these variables are included in the regression in non-log form, there is no real change in sign or significance. All other variables – except number of pre-existing sites - are in percentage form because: (a) by using percentage, one has already adjusted the data into log-form equivalents, and (b) the log of a percentage approaches negative infinity as the percentage becomes very small (see Bartik 1985 for a more detailed discussion).

of plant location. Two variables potentially related to collective action are significant: percent that have lived in the neighborhood for at least five years, and percent who voted in the previous Presidential election. As expected, plants are less likely to locate in neighborhoods with more established households living in them. However, percent vote is positively related to plant location, running counter to the theory that firms avoid locating plants in neighborhoods with more politically involved communities. Thus, these results offer mixed support for the theory that firms locate plants in neighborhoods where compensation costs are low.

Most of the signs on the input cost coefficients are fairly intuitive. Plant location is negatively related to wages and property values.[43] As the cost of land, transportation, and workers increases, a firm is less likely to locate a plant in the neighborhood. The location of polluting plants is positively related to the number of pre-existing TRI plants in the neighborhood and is significant at the one percent level. This result is robust to the inclusion or exclusion of any of the other variables included in regressions (5) and (6). The presence of pre-existing polluting plants may serve as a signal that more pollution will not be strongly resisted by the community. Alternatively, the pre-existing plant variable may be acting as a proxy for zoning laws that limit industrial activities in some areas and encourage them in others. It is also possible that locating near other polluting plants allows a plant to take advantage of agglomeration economies from locating near other polluting manufacturing firms. The coefficient on plant size is also positive and significant at the one percent level. The larger are plants in the same industry, the more likely that a TRI plant will locate in that neighborhood. The percent of persons employed in manufacturing is also positive and significant, and may serve as an indicator to the firm of a ready and skilled labor force. If this variable is excluded from the regression, the other variable possibly representative of labor force quality – percent graduated from high school – is still not significant. Finally, firms are less likely to locate a plant in urban areas, indicating

[43] MSA-level wage is also available. Use of this variable in the regression yields similar results.

that the costs of urban areas outweigh possible benefits. This may be due to factors such as higher tax rates and more crime in urban areas.

It is also worth noting that various regional dummies are included in regressions (4), (5), and (6) to allow for the possible relaxation of IIA - proximity to the border with Mexico, proximity to the Gulf of Mexico, proximity to a neighboring state, and in the Houston or Dallas areas. While not reported in Table 3, they are not significant for any specification and do not significantly improve or change the results.

I conduct a robustness check on the results in column (5) by including alternate measures of poverty and the cost of land that do not suffer from multi-collinearity problems. These results are reported in column (6). Property values and income are very highly correlated (84 percent), while poverty rates are highly correlated with both income and percent non-white (both are about 60 percent). The percent of households without phone service is of the predicted sign found in the environmental justice literature but is not significant. This is not true when percent of housing built before 1970 is substituted for property values. In this case plants are less likely to locate in older neighborhoods, which seems counter-intuitive from a property value perspective. It may be the case that this variable also reflects differences in infrastructure and services that are more likely to be state-of-the art in newer neighborhoods. Most of the key results for other variables remain unchanged – no variable changes sign; percent that have lived in the neighborhood for at least five years is no longer significant and percent with at least a high school education becomes significant.

Table 3: Conditional Logit Estimation – 50 choices

Variable	(1) Davidson and Anderton	(2) Kriesel et. al	(3) Been and Gupta	(4) Without Input Costs	(5) With Input Costs	(6) With Input Costs
NONWHT	-0.03 (0.28)	0.09 (0.28)	-0.12 (0.28)	-0.28 (0.29)	0.52 (0.33)	0.44 (0.28)
FOREIGN	0.25 (0.80)	-0.46 (0.85)	0.08 (0.83)	-0.16 (1.02)	0.46 (0.97)	1.20 (1.02)
INCOME			-0.23 *** (0.06)	-0.21 *** (0.05)	-0.16 *** (0.06)	-0.21 *** (0.06)
POVERTY	-0.56 (0.69)	-1.83 *** (0.71)	-0.50 (0.71)	- 0.95 (0.67)	-2.21 *** (0.83)	
NOPHONE						1.10 (0.85)
HIGHSCH	1.90 *** (0.67)	1.43 ** (0.68)	1.98 *** (0.63)	2.27 *** (0.67)	0.61 (0.76)	1.74 ** (0.84)
VACANT				0.67 (0.78)	0.99 (0.78)	-0.47 (1.10)
WCHILD				1.28 ** (0.47)	0.58 (0.50)	0.09 (0.60)
AGE65				-0.20 (1.51)	-2.00 * (1.49)	0.39 (1.52)
NOMOVE				-0.94 ** (0.40)	-1.05 ** (0.44)	0.12 (0.51)
POP		0.12 (0.09)	0.26 *** (0.09)	0.08 (0.09)	0.13 (0.09)	0.09 (0.09)
VOTE				3.74 *** (1.37)	3.54 *** (1.51)	2.82 * (1.51)
UNEMPL	-8.04 ** (3.90)		-10.30 ** (4.50)			
PROPERTY	-0.10 *** (0.02)	-0.10 *** (0.02)	-0.10 *** (0.02)		-0.03 (0.02)	
BUILT < 1970						-2.21 *** (0.35)
WAGE		-0.92 *** (0.24)			-1.08 *** (0.32)	-0.93 *** (0.31)
PLANT SIZE					0.36 *** (0.08)	0.36 *** (0.08)
OLDSITE					0.27 *** (0.03)	0.26 *** (0.03)
URBAN					-1.00 *** (0.14)	-0.77 *** (0.15)
MANUF	5.27 *** (0.53)	6.03 *** (0.60)	5.95 *** (0.54)		4.11 *** (0.70)	3.89 *** (0.76)
RAIL					-0.14 ** (0.04)	-0.18 *** (0.04)
Pseudo R^2	0.05	0.05	0.06	0.03	0.18	0.19
Log L	-1,350.05	-1,342.11	-1,339.59	-1,377.86	-1,171.45	-1,155.62

* indicates significance at the 10% level, ** indicates significance at the 5% level, and *** indicates significance at the 1% level. Standard errors are in the parentheses.

One measure of the possible explanatory power of these regressions is a pseudo-R^2. Note that the regressions based on prior literature generally perform poorly with an R-squared of 5 percent or less. Specifications (5) and (6), which include input costs, have a pseudo-R^2 of 0.17 and 0.18, respectively. The inclusion of the number of pre-existing TRI plants already established in the neighborhood is paramount to the performance of specifications (5) and (6): without it in the regression, the pseudo R^2 has a value of only about 0.10.[44]

A likelihood ratio test is used to examine the overall significance of each of the theories posited by Hamilton for why plants may locate in poor or minority neighborhoods. The null hypothesis in each case is that the set of variables most closely associated with each theory is not relevant to the location decision. The results indicate that discrimination costs are not relevant to the decision of where to locate a polluting plant (see Table 4). However, costs related to the compensation costs are significant – though for specification (6) they are only significant at the 10 percent level. When likelihood ratio tests are conducted for environmental disamenity costs and collective action costs separately, each set of variables is still significant in the case of specification (5), but only environmental disamenity costs are significant in the case of specification (6).[45] Variables related to input costs, which are traditionally included in the firm location literature, are the most relevant to a plant's location decision and consistently significant at the one percent level across the two specifications.[46]

---

[44] Not including the number of pre-existing sites does not change the basic finding of insignificance for the race variables or the significant, negative coefficients on income and percent in poverty.
[45] Environmental amenity costs include average income, percent poor, population, and the vacancy rate. Collective action costs include the remaining compensation variables.
[46] There is no change to the significance of the likelihood ratio tests when percent employed in manufacturing is included as a variable representative of compensation costs (job v. environment trade-off by workers) instead of as an input cost (potential qualified labor pool). There is also no change when percent with a high school education is included as a measure of quality of the labor pool – as an input cost - or as a measure of the propensity to engage in collective action – as a compensation cost variable.

Table 4: Likelihood Ratio Tests

Sets of Variables [47]	Likelihood Ratio Test for Specification (5)	Likelihood Ratio Test for Specification (6)
Discrimination Costs	2.60	3.38
Compensation Costs	51.28 ***	14.74 *
Inputs Costs	412.84 ***	440.90 ***

* indicates significance at the 10 % level, ** indicates significance at the 5 % level, and *** indicates significance at the 1 % level.

I also calculate the marginal effects for the significant variables in specification (5). These marginal effects are only valid for the actual location and the random sample of 49 alternative locations and should not be generalized to the entire sample of alternative locations. One can consider the following rather unrealistic story: A firm is asked to choose one of 50 census tracts in which to locate, selected at random from the state of Texas. Given that it will locate in one of these 50 tracts, how important are the variables included in the regression to the decision process? In most cases, the marginal effects are very small: for instance, evaluated at the mean, a one percent increase in the percent who voted results in a change in the probability of plant location of 0.4 percent. Two marginal effects are worth noting: A ten employee increase in plant size increases the likelihood of plant location by 2.6 percent. Likewise, an increase in the number of pre-existing sites by one, increases the likelihood of plant location by 2.4 percent.

[47] Discrimination costs include percent nonwhite and percent foreign. Compensation costs include average income, percent poor, percent vacant, population, percent with children, percent over age 65, and percent that have not moved in the last five years. Input costs include average wage, percent high school educated, percent urban, average property value, number of pre-existing TRI plants, distance to rail, and average plant size. Results do not change if percent poor is included as an input cost variable instead of an environmental disamenity variable.

Table 5: Marginal Effects for Significant Variables in Specification (5)[48]

Variable	Own Marginal Effects for Actual Location
ln (INCOME)	-0.0004
PERCENT NOMOVE	-0.0002
PERCENT VOTE	0.004
ln (CNTYWAGE)	0.000
ln (PLANT SIZE)	0.026
ln (OLDSITE)	0.024
PERCENT URBAN	-0.0002
PERCENT IN MANUFACTURING	0.003
Ln (ARAIL)	-0.004

9. Conclusion

Many environmental justice studies argue that firms choose to locate waste sites or polluting plants disproportionately in minority or poor communities. There are two major shortcomings of these studies. First, site or plant location is matched to current socioeconomic characteristics instead of to neighborhood characteristics at the time that the siting decision was made. Second, variables that are important to a plant's location decision - the cost and quality of labor, the average size of a plant in the same industry, the existence of older polluting plants - are often not included in these studies. Without controlling for such variables, it is difficult to evaluate the importance of socioeconomic characteristics such as race and poverty to a firm's initial location decision.

I address both of these issues in this paper. First, I match actual TRI plant location to the characteristics of the neighborhood at the time of siting for plants established in Texas between 1976 and 1993. I include socioeconomic characteristics such as race, income and poverty, as

[48] In most cases, the marginal change evaluated is a change of one percent in an independent variable from the mean value. For county wage and income the change is 10 percent from the mean (prior to applying the

well as variables often omitted from the environmental justice literature such as the quality and cost of labor, plant size, transportation costs, and the existence of older polluting manufacturing plants. Using the assembled data set, I estimate a conditional logit to allow a firm to choose from several alternative locations.

When the location of TRI plants are matched to 1990 socioeconomic characteristics, race and income variables are significant and of the expected sign. These results confirm that a disproportionate number of minority and lower income populations currently live near TRI plants. However, when plant location is matched to neighborhood socioeconomic characteristics at the time of siting, race variables are not significant determinants of TRI plant location in Texas. Percent poor, another variable of concern in the environmental justice literature, is a significant factor, but rather than attracting plants to an area it acts as a deterrent.

Two of the most important variables in this study are left out of previous environmental justice studies: the presence of pre-existing TRI plants in a neighborhood and the average size of a plant in the same industry and region. The fact that plants tend to locate where other TRI plants are already established may indicate that firms use the existence of other polluting plants as an indicator of the acceptability of pollution increases in the area or of the industrial distribution imposed by zoning. Both the existence of older TRI plants and average plant size may indicate that firms are taking advantage of agglomeration economies. Variables traditionally considered in the firm location literature such as land and labor costs, plant size, and the quality of labor are also significant variables.

natural log); for oldsite, the change is one facility; and for plant size, the change is 10 employees (prior to

Bibliography

Anderton, D., A. Anderson, J. Oakes, and M. Fraser (1994). "Environmental Equity: The Demographics of Dumping." Demography. v. 31, n. 2: 229-248.

Arora, S. and T, Cason (1998). "Do Community Characteristics Influence Environmental Outcomes? Evidence from the Toxics Release Inventory." Journal of Applied Economics.

Baden, B., and D. Coursey (2002). "The Locality of Waste Sites Within the City of Chicago: A Demographic, Social, and Economic Analysis." Resource and Energy Economics. Vol. 24: 53-93.

Bartik, T. (1985). "Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States." Journal of Business and Economic Statistics. v. 3, n. 1: 14-22.

Baumol, W. and W. Oates (1988). The Theory of Environmental Policy. Second Edition. New York: Cambridge University Press.

Been, V. (1994). "Locally Undesirable Land Uses in Minority Neighborhoods: Disproportionate Siting or Market Dynamics?" The Yale Law Journal. v. 103, n. 6: 1383-1421.

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. v. 24, n. 1: 1-56.

Boer, T., M. Pastor, J. Sadd and L. Snyder (1997). "Is There Environmental Racism? The Demographics of Hazardous Waste in Los Angeles County?" Social Science Quarterly. v. 78, n. 4: 793-809.

Brooks, N. and R. Sethi (1997). "The Distribution of Pollution: Community Characteristics and Exposure to Air Toxics." Journal of Environmental Economics and Management. v. 32: 233-250.

Bullard, R. (1983). "Solid Waste Sites and the Black Houston Community." Social Inquiry. v. 53: 273.

Coase, R. (1960). "The Problem of Social Cost." The Journal of Law and Economics.

Davidson, P. and D. Anderton (2000). "Demographics of Dumping II: A National Environmental Equity Survey and the Distribution of Hazardous Materials Handlers." Demography. V. 37, n. 4: 461-466.

Finney, M. (1994). "Property Tax Effects on Intrametropolitan Firm Location: Further Evidence." Applied Economic Letters. v. 1: 29-31.

---

applying the natural log).

Friedman, J., D. Gerlowski, and J. Silberman. "What Attracts Foreign Multinational Corporations? Evidence from Branch Plant Location in the United States." Journal of Regional Science. v. 32, n. 4: 403-418.

GAO (1983). Siting of Hazardous Waste Landfills and Their Correlation with Racial and Economic Status of Surrounding Communities. GAO: Washington, DC.

Glickman, T. and R. Hersh (1995). "Evaluating Environmental Equity: The Impacts of Industrial Hazards on Selected Social Groups in Allegheny County, Pennsylvania." Discussion Paper 95-13. Resources for the Future.

Goldman, B. and L. Fitton (1994). Toxic Wastes and Race Revisited. UCC: Center for Policy Alternatives.

Gray, W. and R. Shadbegian (2002). "'Optimal' Pollution Abatement – Whose Benefits Matter, and How Much?" Working Paper.

Greene, W. (1993). "Models for Multiple Choice." Econometric Analysis. Macmillan: New York.

Hamilton, J. (1995). "Testing for Environmental Racism: Prejudice, Profits, Political Power?" Journal of Policy Analysis and Management. v. 14, n. 1: 107-132.

Hersh, R. (1995). "Race and Industrial Hazards: A Historical Geography of the Pittsburgh Region, 1900-1990." Discussion Paper 95-18. RFF.

Jenkins, R., K. Maguire, and C. Morgan (2004). "Host Community Compensation and Municipal Solid Waste Landfills." Land Economics, v. 80, n. 4: 513-528.

Krieg, E. (1995). "A Socio-Historical Interpretation of Toxic Waste Sites: The Case of Greater Boston." The American Journal of Economics and Sociology. v. 54, n. 1: 1-14.

Kriesel, W., T. Centner, and A. Keeler (1996). "Neighborhood Exposure to Toxic Releases: Are There Racial Inequalities?" Growth and Change. v. 27: 479-499.

Lambert, T. and C. Boerner (1995). Environmental Inequity: Economic Causes, Economic Solutions. Policy Study 125. Center for the Study of American Business.

Levinson, A. (1996). "Environmental Regulations and Manufacturers' Location Choices: Evidence from the Census of Manufactures." Journal of Public Economics. v. 62: 5-29.

McConnell, V. and R. Schwab (1990). "The Impact of Environmental Regulation on Industry Location Decisions: The Motor Vehicle Industry." Land Economics. v. 66, n. 1: 67-81.

McFadden, D. (1978). "Modelling the Choice of Residential Location" in Spatial Interaction Theory and Residential Location. A. Karlquist, ed. North Holland: Amsterdam.

Pastor, M., J. Sadd, and J. Hipp (2001). "Which Came First? Toxic Facilities, Minority Move-In, and Environmental Justice." Journal of Urban Affairs. Vol. 23, no. 1: 1-21.

Rinquist, E. (1997). “Equity and the Distribution of Environmental Risk: The Case of TRI Facilities.” Social Science Quarterly. v. 78, n. 4: 811-829.

Sadd, J., M. Pastor, T. Boer, and L. Snyder (1999). “’Every Breath You Take…’ The Demographics of Toxic Air Releases in Southern California.” Economic Development Quarterly. Vol. 13, no. 2: 107-123.

United Church of Christ (1987). Toxic Waste and Race in the United States. UCC: Commission for Racial Justice.

Wolverton, A. (2002). “The Demographic Distribution of Pollution: Does Neighborhood Composition Affect Plant Pollution Behavior?” Working Paper.

Zimmerman, R. (1993). "Social Equity and Environmental Risk." Risk Analysis. v. 13, n. 6: 649-666.