The Importance of State and Plant Characteristics in Determining the Environmental Compliance Costs of Chemical Manufacturing Plants: Evidence from the PACE Survey, 1979 - 1990

Ellen S. Post
University of Maryland
Department of Agricultural and Resource Economics
October, 1994

Abstract

The common belief that some states are more stringent than others in their efforts to protect environmental quality from industrial pollution is accompanied by the persistent concern that industry will avoid those environmentally stringent states. These concerns are based on the assumption that interstate differences in environmental regulatory stringency are large enough to affect firms' location decisions. This study links the U.S. Census Bureau's Pollution Abatement Costs and Expenditures (PACE) data on a plant-by-plant basis with the data in the Census Bureau's Longitudinal Research Database (LRD) to examine the determinants of environmental compliance cost for chemical manufacturing plants (SIC code 28) in the United States from 1979 through 1990.

A model of the firm's profit maximization problem in the presence of environmental regulations is developed, from which a correspondence between a vector of environmental regulatory constraints and the firm's real-valued environmental compliance cost is derived. The econometric analysis, based on this correspondence, offers some insight into the relative importance of state and plant characteristics in determining a plant's environmental compliance cost. The age, size, and industry of the plant are all consistently significant predictors of its compliance cost. In contrast, the state effect is weak and not entirely consistent over time. While there are clearly differences among states in environmental regulations, these differences translate into *barely statistically discernable* differences in environmental compliance costs for the chemical plants in those states during the study period. California's reputation as a leader in environmental awareness, in particular, is not reflected in significantly greater compliance costs reported by chemical plants in that state as compared with comparable plants in other states. The results suggest, further, that (1) although newer plants may face more stringent regulations, their more pollution-efficient technologies more than compensate, and (2) environmental regulatory agencies have leaned more heavily on larger plants than smaller ones. The weak state effect within the United States does not necessarily imply a similarly weak national effect in international comparisons. It is suggested, moreover, that if there is an economically significant effect of environmental regulations on the firm location decision within the United States, it may be in the form of transaction cost differentials rather than compliance cost differentials.

I would like to thank Lars Olson, K.E. McConnell, Marc Nerlove, Nancy Bockstael, Robert Chambers, and Robert Schwab for helpful comments. I am also grateful to the staff at the Center for Economic Studies, U.S. Census Bureau, for their support in working with the PACE and LRD data, and to the U.S. EPA for financial support in obtaining access to the data.

DISCLAIMER

The Importance of State and Plant Characteristics in Determining the Environmental Compliance Costs of Chemical Manufacturing Plants: Evidence from the PACE Survey, 1979 - 1990

1. *Introduction*

It is commonly believed that some states are relatively stringent and other states relatively lax in their efforts to protect environmental quality from industrial pollution, and that industry will avoid states with tougher environmental regulations, and, conversely, be attracted to states with laxer regulations. California in particular has gained a reputation for environmental regulatory stringency, and with it the concern that industry will go elsewhere. With the passage of the North American Free Trade Agreement (NAFTA), this concern is mirrored by a larger fear that the United States as a whole will lose industry, and the jobs associated with it, to other countries with less “stringent” environmental regulations.

The notion that interstate (or international) differences in environmental regulations will affect firm location decisions is based on the assumption that there are indeed differences across states (or nations) in the degree of *environmental regulatory stringency.* Reputations for stringency, however, seem to result from specific, highly publicized legislation, such as Proposition 65 in California, rather than from any comprehensive assessment of the environmental regulatory functioning of one state relative to others. Indeed, the complexity of environmental regulatory agencies and legislation within states would make such comprehensive assessments a daunting task. States characteristically have several environmental regulatory agencies which confront a plant with an array of regulations.

The Pollution Abatement Costs and Expenditures (PACE) dataset provides an invaluable opportunity to examine the impact of a *collection* of diverse regulations on the environmental compliance cost of the profit maximizing firm.[1] The PACE survey, conducted annually by the United States Census Bureau, requests that plants report the actual costs they incurred in complying with the environmental regulations with which they were faced in a given year. It is, moreover, the only source of such plant-level information on the pollution abatement (compliance) costs of manufacturing plants in the United States.

This study links the PACE data on a plant-by-plant basis with the data in the U.S. Census Bureau's Longitudinal Research Database (LRD)[2] to examine the determinants of environmental compliance cost for chemical manufacturing plants (SIC code 28) in the United States from 1979 through 1990. It is only with plant-level data that the effects of state environmental regulations and regulatory agencies can be disentangled from the effects of characteristics of the plant itself in determining the compliance cost incurred by a plant. The linked PACE-LRD data therefore provide a unique opportunity to answer several questions of interest. Do older plants incur greater or smaller environmental compliance costs than newer plants? Do environmental regulatory agencies target larger plants in particular? Do states really matter? That is, would the same plant located in different states be expected to incur significantly different environmental compliance costs?

While it is clear that the same plant would be confronted by distinctly different sets of environmental regulations in different states, it is less clear that those differences would translate into significant differences in the environmental compliance costs of the plant. It is differences in environmental compliance costs, moreover, that would be expected to affect the firm's location decision. If state A requires greater environmental compliance costs than state B, then, all else equal, the profit maximizing firm can be expected to locate its plant in state B. What would be expected to matter is not the regulations themselves but how much they will cost the plant.

[1] "Firm" and "plant" are used interchangeably to refer to a single physical manufacturing establishment.
[2] The Census Bureau's Longitudinal Research Database (LRD) contains information about other plant characteristics, some of which are likely to affect compliance cost. The LRD is described in McGuckin and Pascoe, 1988.

A model of the firm's profit maximization problem in the presence of environmental regulations is developed in Section 2. From this model a correspondence between a vector of regulatory constraints imposed on the firm by environmental regulations and the firm's real-valued environmental compliance cost is derived. This forms the basis for the econometric model used to analyze the determinants of environmental compliance cost.

Section 3 reviews those plant-level studies that have examined the effect of environmental regulations on the firm location decision, focusing in particular on how each study attempts to measure the "stringency" of the regulations of a local jurisdiction. All such studies implicitly assume that, however it is defined, "environmental regulatory stringency" does in fact differ among states (or other local jurisdictions considered). The model developed in Section 2 is used to help identify some of the problems in previous attempts to characterize environmental regulatory stringency.

The PACE survey and data are described in Section 4, and the study sample, derived from the PACE data, is described in Section 5. The econometric model, the analyses, and the hypotheses tested to answer the questions posed above are discussed in Section 6.

The analyses show all plant characteristics considered (age, size, and industry) to be significant and consistent predictors of environmental compliance cost. The results suggest, further, that (1) although newer plants may face more stringent regulations, their more pollution-efficient technologies more than compensate, and (2) environmental regulatory agencies targeted larger plants disproportionately during the study period. In contrast to plant characteristics, the state effect is weak. Interstate differences in environmental regulations and reputations for stringency translate into barely statistically dicernable differences in environmental compliance costs. These results therefore have implications for those studies that attempt to discern the effect of environmental regulations on the firm location decision. A fuller presentation of the results and a discussion are presented in Sections 7 and 8, respectively.

2. *A Model of the Firm 's Profit Maximization Problem in the Presence of Environmental Regulations*

2.1 *The Standard Production Problem of the Profit-Maximizing Firm*

In the absence of environmental regulatory constraints, the price-taking firm's profit maximization problem may be written as

$$\max_{x,y}\{py - wx : (x, y) \in T\} \tag{2.1}$$

where $p \in R^m_+$ is the vector of output prices,

$w \in R^n_+$ is the vector of input prices,

$y \in R^m_+$ is the vector of outputs,

$x \in R^n_+$ is the vector of inputs, and

$$T = \{(x, y) \in R^{n+m}_+ : x \text{ can produce } y\}$$

is the production possibilities set, characterizing the firm's technology[3]. T is assumed to be a nonempty, closed, convex set that is bounded from above for every x. It is also assumed that there is free disposability of x (i.e., $(x, y) \in T$ and $x' \geq x \Rightarrow (x', y) \in T$), free disposability of y (i.e., $(x, y) \in T$ and $y' \leq y \Rightarrow (x, y') \in T$), and weak essentiality (i.e., $(x,0) \in T$, but $(0,y) \notin T$ for $y > 0$). In this standard model, the profit function depends only on prices, p and w.

In the short-run, there are often fixed inputs (e.g., capital), yielding a restricted profit function. Let $x \in R^n_+$ now denote the vector of variable inputs, $k \in R^K_+$ denote the vector of fixed inputs and $\bar{k} \in R^K_+$ denote the vector of values at which the elements of k are fixed. The firm's short-run problem now becomes:

$$\max_{x,y}\{py - wx : (x, k, y) \in T; k = \bar{k}\}. \tag{2.2}$$

[3] "Production possibilities set" and "technology" will be used interchangeably.

2.2 *Profit Maximization in the Presence of Environmental Regulations*

Environmental regulations can take several forms, all of which are straightforward to incorporate into the firm's profit maximization problem. First, however, pollutants must be added. Although the generation of pollution is a normal part of the production process, it has not been necessary up to this point to specify the pollutants generated as part of the firm's profit maximization problem because in the standard model they in no way affected profit. In the presence of environmental regulations, however, they may.

Now, in addition to fixed inputs $k \in \mathbf{R}_+^K$, variable inputs $x \in \mathbf{R}_+^n$ and output $y \in \mathbf{R}_+^m$, there is a vector of pollutants, $q \in \mathbf{R}_+^Q$, generated in the production of y by (x,k). The vector of pollutants can be considered either an output of production or an input to production. It will be convenient, however, to think of pollutants as inputs to production.

Environmental regulations can take one or several of the following forms:

(1) Limits on the amounts of pollutants that can be emitted (i.e., standards) - that is, $q \leq \bar{q}$, where $\bar{q}_k = \infty$ if there is no limit on the kth pollutant.

(2) Pollution abatement and monitoring input requirements -- i.e., requirements that a particular type of input be used -- that is, $\underline{x} \leq x$. For example, denote a particular type of mandated pollution abatement equipment as x_j. Then $\underline{x}_j > 0$.

(3) Limits on (polluting) inputs -- i.e., $x \leq \bar{x}$, where $\bar{x}_j = \infty$ if there is no limit on the jth input.

(4) A Pigouvian tax on pollutants. Let e denote the vector of such effluent taxes, where $e_j = 0$ if there is no tax on the jth pollutant.

(5) Requirements for new, non-production outputs, such as written output from self-monitoring of emissions and written verification of the compliance status of the plant. Let $\tilde{y}$ denote the vector of required minimum values of these non-production outputs. If no such outputs are required, then $\tilde{y} = 0$.

The firm's short-run profit maximization problem is now

$$\max_{x,q,y}\{py - wx - eq:\ (x,k,q,y,\tilde{y}) \in T;\ k = \bar{k};\ q \le \bar{q};\ \underline{x} \le x \le \bar{x}\}\ . \qquad (2.3)$$

The indirect objective function corresponding to this problem is denoted as

$$\pi(p, w, \bar{k}, \tilde{y}, \bar{q}, \underline{x}, \bar{x}, e, T) \equiv \max_{x,q,y}\{py - wx - eq:\ (x,k,q,y,\tilde{y}) \in T;\ k = \bar{k};\ q \le \bar{q};\ \underline{x} \le x \le \bar{x}\}, \qquad (2.4)$$

the firm's maximum achievable profit in the presence of the vector of environmental regulatory. constraints, $R = (\tilde{y}, \bar{q}, \underline{x}, \bar{x}, e)$. The maximum profit achievable in the absence of environmental regulations may be denoted as $\pi(p, w, \bar{k}, 0, \infty, 0, \infty, 0, T)$ or $\pi^{o}(p, w, \bar{k}, T)$.

The environmental compliance cost of the profit maximizing plant is the loss in profit that results from having to comply with the set of environmental regulations with which it is faced. That is, the compliance cost of a plant with technology T and fixed inputs $\bar{k}$ may be defined as.

$$cc = \pi(p, w, \bar{k}, 0, \infty, 0, \infty, 0, T) - \pi(p, w, \bar{k}, \tilde{y}, \bar{q}, \underline{x}, \bar{x}, e, T) \qquad (2.5)$$

$$\Rightarrow cc = cc(p, w, \bar{k}, \tilde{y}, \bar{q}, \underline{x}, \bar{x}, e, T) = cc(p, w, \bar{k}, R, T)\,. \qquad (2.6)$$

Compliance cost is thus a function of all of the exogenous factors in the firm's profit maximization problem.[4] A set of environmental regulations imposing the vector of constraints,

[4] The technology of the firm, T, is technically a choice variable as well -- that is, the firm has chosen its technology and always has the option to change it. Only with extreme changes in exogenous factors, however,

R, on the plant will induce larger or smaller compliance costs, for example, as the price of inputs necessary to comply with those regulations (e.g., labor to run abatement equipment) increases or decreases.

Finally, the vector of regulatory constraints, R, may itself be a function of certain characteristics of the plant being regulated. If, for example, environmental laws have "grandfather" clauses which either exempt older plants from certain requirements or allow them to meet less stringent requirements, then R would be a function of the age of the plant being regulated. New Source Performance Standards for air pollutants are an example of this.

Similarly, environmental agencies may target larger plants for more rigorous monitoring and enforcement of environmental regulations or may leave small plants alone, effectively setting the values of R for such plants to $(0,\infty,0,\infty,0)$.

Let z denote the vector of relevant plant characteristics (e.g., z_1 = the age of the plant and z_2 = the size of the plant. If environmental regulatory agencies target different industries differently, another element of z might be the industry the plant is in). Then $R = (\tilde{y}(z), \bar{q}(z), \underline{x}(z), \bar{x}(z), e(z)) \equiv R(z)$.

We now locate the plant in a particular state. We assume for now that both prices and environmental regulations, given z, are state-specific. Prices and environmental regulatory constraints in the jth state will be denoted by a superscript j. A profit maximizing plant with technology T, fixed inputs $\bar{k}$, and characteristics z located in the jth state solves the following problem:

$$\max_{x,q,y}\{p^j y - w^j x - e^j(z)q : \ (x,k,q,y,\tilde{y}^j(z)) \in T; \ k = \bar{k}; \quad q \le \bar{q}^j(z); \ \underline{x}^j(z) \le x \le \bar{x}^j(z)\}. \tag{2.7}$$

would changing technologies be profit maximizing. Although isolated cases of this have been reported in response to environmental regulations, (see, for example, Leonard, 1984, p. 92), T may generally be taken to be exogenous.

It achieves profit $\pi(p^j, w^j, \bar{k}, R^j(z), T)$ and incurs environmental compliance cost $cc(p^j, w^j, \bar{k}, R^j(z), T)$.

The relationship between environmental regulations in a state and the environmental compliance cost of a plant located in that state is thus seen to be a *correspondence.* The same set of environmental regulations may result in many different values of environmental compliance cost for a plant, depending on factor and output prices in the state, the level at which the plant's fixed factors are fixed, its technology, and those plant characteristics that determine which vector of regulatory constraints, out of a whole set of possible vectors, will be imposed on the plant.

Implicit in this model is the assumption that the environmental regulatory constraints implied by a state's environmental regulations are identical to the vector of regulatory constraints, R actually imposed on the plant. This is true only if the regulations are actually enforced and complied with. There is evidence, however, that this is not entirely the case (Russell, et al., 1986). The implications of a discrepancy between the vector of implied regulatory constraints and the vector of actual constraints are discussed below.

The above model provides the basis for an econometric model to examine the relative importance of state environmental regulations and regulatory agencies versus other factors in determining the environmental compliance costs of chemical manufacturing plants. It will also help identify some of the problems in previous attempts to characterize the environmental regulatory stringency of states or other local jurisdictions.

3. *A Review of Plant-Level Studies of the Effect of Environmental Regulations on the Firm Location Decision*

Three plant-level studies have examined the effect of environmental regulations on the location decisions of manufacturing firms. All have discussed the difficulty of characterizing environmental regulatory stringency. Noting that "there is no obvious way to measure state environmental regulations . . .," Bartik (1988) uses four different measures, two for water pollution and two for air pollution. Two of these are state (government) expenditures on air and water pollution control, respectively, per manufacturing employee in the state. Bartik argues that these variables are likely to reflect the probability of a polluter facing inspection and enforcement actions in the state. They may, however, simply reflect the extent to which more pollution-intensive plants exist in the state. Suppose, for example, states A and B have identical environmental laws and identical regulatory agencies. The plants in state A, however, use highly pollution-intensive technology T_1 while the plants in state B use the minimally polluting technology T_2 . To achieve the same level of environmental quality, state A must spend more on monitoring and enforcement actions than state B. If plants from state A were to move to state B, their compliance costs would not change, *because the two states would treat them the same.* The pollution control expenditures of state B, however, would increase. State pollution control expenditures may reflect plant characteristics rather than the stringency of regulatory agencies.

The second two variables Bartik considers are based on comparisons of state-specific compliance costs per dollar of product shipped with the national average, specific to each two-digit SIC code. Again, these variables may reflect differences in the distributions of plant characteristics in different states rather than behavioral differences in environmental regulatory agencies. As Bartik notes, for example, "this measure does not control for the mix of new versus existing plants in a state. Because new plants face stiffer environmental regulation than existing plants, states that attract more new plants will have higher average compliance costs." Suppose, as above, that states A and B are identical, except that all the plants in state A are age z_A and all the plants in state B are age z_B , where $z_A < z_B$. Then, if both states have the same "grandfather" clauses, although $R^A(z) = R^B(z) = R(z)$, for all z,

$cc(p, w, \bar{k}, R(z_A), T) > cc(p, w, \bar{k}, R(z_B), T)$.[5] Differences in compliance cost, in this case, reflect differences in plant age rather than differences in the stringency of state regulatory agencies.

McConnell and Schwab (1990) also consider several indirect measures of stringency. They focus on the motor vehicle assembly industry (SIC 3711), for which volatile organic compounds (VOC's) are a major pollutant. Counties are the units of location considered. Because VOC's are an important source of ozone, one indirect measure used is an indicator of whether or not the county is in attainment for ozone. Another measure involves the degree to which the county is out of attainment, if it is not in attainment. In addition, several state-level measures are considered, such as the state-specific pollution abatement operating costs in industry 37 per dollar of shipments in that industry in the state. The state-level measures have the same problems as those in Bartik (1988), namely, they may not be reflecting what they are intended to reflect. The county-level measures of attainment, on the other hand, do not reflect possible differences across counties in the degree to which regulations are actually enforced. They would be inadequate measures, moreover, for any industry, such as the chemical industry, that emits a wide array of pollutants rather than a single dominant one.

Levinson (1992) considers three indirect measures of what he refers to as the "price of waste disposal services" in a state. Similar to some of the measures used by Bartik (1988) and McConnell and Schwab (1990), Levinson (1992) uses as one indirect measure a state's gross pollution abatement operating costs (in 1982) per production worker. Similar to a second measure used by Bartik (1988), Levinson also considers state expenditures on air quality programs per manufacturing plant. The shortcomings of both of these indirect measures have already been described. Finally, Levinson considers the "Green Index" of a state (Hall and Kerr, 1991), an index based on the number of pollution-regulating laws a state has adopted out of a possible twenty-one laws. However, the relationship between the number of laws adopted'

[5]Following Bartik, this example ignores the effect of age on the pollution-intensity of a plant. Older plants tend to have older vintage technologies, which are likely to be more polluting. Whether newer plants have higher or lower average compliance costs than older plants depends on which of two opposing age-related effects is dominant. This is tested in the analyses described in Sections 6, 7, and 8.

and any reasonable definition of the actual stringency of a state may be tenuous at best. The adoption of a law does not necessarily imply the enforcement of the law. The enforcement of environmental regulations is very difficult, and there is evidence that the compliance of industrial plants with such laws is far from complete (Russell, et al., 1986). Even if a law is enforced, moreover, the degree of enforcement may vary considerably from one state to another. Finally, even if all adopted laws were enforced and enforced to equal degrees in all states, the laws themselves are likely to differ substantially in the degree to which they actually impose compliance costs on polluting plants. The mere number of laws adopted in a state is therefore a very weak indicator of the environmental compliance cost that would be incurred by a plant in that state.

A reasonable definition of the environmental regulatory stringency of a state should reflect the impact of environmental regulations, and the agencies that administer those regulations, on the compliance costs of the plants in that state. As shown above, however, a plant's compliance cost is affected by several factors, only some of which reflect the behavior of environmental regulatory agencies. To disentangle the effect of a state's environmental regulatory behavior from other factors affecting the plant's compliance cost (e.g., the pollution-intensity of its technology), *plant-level* data are needed.

4. *The PACE Data*

The PACE questionnaire has been sent every year since 1973[6] to a sample of manufacturing plants in the United States. The sample is a probability sample designed to overrepresent large plants, which contribute a disproportionately large proportion of total industry activity, and underrepresent small plants. In the 1987 Census of Manufactures (CM)[7],

[6] 1987 was an exception in which PACE data were not collected. Moreover, the plant-level PACE data now exist only from 1979 onward. The PACE data prior to 1979 exist only in the form of aggregate published statistics.

[7] The Census of Manufactures is intended to be a complete enumeration, rather than a sample of plants.

for example, only 22 percent of all plants in the chemical industry had total value of shipments (tvs) of at least $10,000,000, but this 22 percent of plants constituted 91 percent of total industry tvs. In contrast, in the 1988 PACE sample, drawn from the 1987 CM, 87 percent of all chemical plants had tvs of at least $10,000,000. Similarly, in the 1989 PACE sample, drawn from the same CM, 83 percent of all chemical plants had tvs of at least $10,000,000. The comparisons are similar for the 1982 and 1977 CMs and the corresponding PACE samples drawn from each. The PACE survey, then, is designed to collect most of its information about the approximately one fifth of all plants that are responsible for the bulk of industry activity. Information on the smaller plants is therefore relatively sparse.

Because plants are selected into the PACE sample with different probabilities, depending on the size of the plant, a sample weight has been calculated for each plant in the PACE dataset. This weight is the inverse of the probability of the plant's having been selected into the sample. The largest plants, selected with certainty, therefore have weights of 1.0. The smallest plants can have weights as large as 20, 30, or 40.

The PACE survey requests that plants report the actual costs they incurred in complying with environmental regulations in a given year. In particular, establishments are asked to report their (1) capital expenditures for abatement, (2) operating costs for abatement[8], (3) payments to government for pollution removal, and (4) costs recovered through abatement activities. A sample PACE survey form is shown in Appendix A.

Because the PACE survey is conducted every year on a sample of manufacturing plants and these samples are different but not exclusive of each other, a given plant may appear in the PACE dataset in more than one year. For example, in a subset of the PACE dataset consisting of 2002 plants in the chemical industry (SIC code 28) for 'which data on abatement operating costs were non-missing and non-imputed, 700 (35 percent) appear in the PACE dataset in only

---

[8] i.e., the increase in production costs attributable to pollution abatement requirements. Abatement operating costs, categorized by kind of cost, in the PACE survey are (1) depreciation, (2) labor costs, (3) costs of materials and supplies, and (4) sevices, equipment leasing, and other costs.

one year during the period from 1979 through 1990 (excluding 1987), 283 (14 percent) appear in two years, 219 (11 percent) appear in three years, and so on. Only 63 plants (3 percent) appear in all eleven years. The PACE data are therefore incomplete (unbalanced) panel data. Many plants are observed more than once, but most are not observed in every year of the survey.

Finally, the PACE data are survey data reported by individuals (e.g., plant managers). In many cases, information was not reported. Moreover, even when abatement cost data are not missing, abatement costs are sometimes reported as zero. The frequency of reported zeros is large enough that they should not be ignored. Abatement costs (e.g., capital expenditures for abatement or operating costs or a combination of the two) therefore have the characteristics of a truncated or censored variable, and analyses in which these costs are the dependent variable should take this into account.

5. *The Study Sample*

Any chemical plant included in the PACE survey in a year during the period 1979-1990 was a potential observation in the study sample.[9] Plant-years were eliminated from the sample, however, on the basis of several criteria. In particular, an observation was eliminated if (1) its PACE data could not be linked with its LRD data, (2) it had missing data for any of the variables in the model, (3) at least one of its variables or its PACE weight was determined to be spurious or a gross outlier[10], (4) it was designated as an administrative record[11], or (5) its total value of shipments Was less than $10,000,000.[12] Finally, to ensure that estimates of state-

[9] An observation is a plant-year - i.e., a plant in a given year. The same plant in two different years constitutes two (correlated) observations.

[10] This is discussed more fully in the dissertation from which this paper is derived.

[11] Some plants in the Annual Survey of Manufactures (ASM), from which the PACE samples are drawn, are required to answer only a small subset of the survey questions. Answers to the remaining questions are imputed by the Census Bureau. These plants, generally small, are designated as *administrative records*. This exclusion criterion is, however, subsumed by the fifth exclusion criterion.

[12] Total value of shipments is in 1982-constant dollars.

specific coefficients be based on at least five observations, any plant-year not previously eliminated was eliminated if (6) it was from a state with fewer than 5 observations in the study sample in that year.

The fifth exclusion criterion restricts the analysis to consider only the relatively larger plants. Because the PACE sampling design so heavily favors larger plants, the relatively small number of small plants that enter the PACE sample have large weights. In some cases, PACE weights are as large as 20, 30, or 40. Generalizing inferences to the population of *all* plants in the chemical industry would require relying on the relatively sparse, but heavily weighted data representing the approximately four fifths of all chemical plants that are responsible for only about ten percent of industry tvs. Information from large plants, moreover, is believed to be generally more reliable and less variable than that from small plants.[13]

If analyses use unweighted data, inferences cannot be generalized to any population beyond the particular plants in the sample. That solution seems very limiting. Using all the plants in a weighted analysis, on the other hand, has the serious drawback that results may be heavily influenced by a small number of the least reliable observations, putting the reliability of the results themselves in jeopardy.

The sample was therefore limited to only those plant-years with tvs of at least $10,000,000. This eliminates only seven percent of the study sample, but restricts the population to which inferences can be generalized to the approximately twenty percent of the industry's largest plants responsible for about 90 percent of industry activity.[14]

---

[13]The greater variability of data from small plants vs. larger plants is borne out in analyses which model the variance of the error term as a function of log(tvs) (see Sections 6 and 7).

[14] There is an additional advantage to limiting the sample to the larger plants. The PACE weights were calculated so that it is possible to aggregate up to *national* figures for each 4-digit industry. A plant of size x in 4-digit industry y with PACE weight z, for example, represents z plants in 4-digit industry y in the United States. A plant of size x in industry y with PACE weight z in state s (e.g., Michigan) represents z similar plants *in state s* only if the size and industry distributions in state s are the same as in the United States as a whole. Because this is potentially a problem for the estimation of state-specific coefficients in PACE-weighted analyses, both weighted and unweighted analyses were run for the three years with the largest PACE weights in the sample to determine the extent to which state-specific coefficient estimates were actually affected by the weighting. In all three years examined, differences in coefficient estimates between weighted and unweighted

The study sample consists of 8,031 observations made over the course of eleven years on 1,877 plants. The number and percent of plants that appear in the sample for N years, for N=1, 2, ... 11, is shown in Table 1. The number and percent of plants that opened prior to 1970 and after 1970 is shown in Table 2.[15] The number and percent of observations (plant-years) in each of the eleven survey years is given in Table 3. Finally, the number of plants in the sample by state and year is given in Table 4. State names corresponding to the abbreviations used in this paper are given in Appendix B.

6. *The Econometric Model and Analyses*

As shown in Section 2, the plant's compliance cost depends in part on certain characteristics of the plant itself, such as its technology, T, and its fixed factors, $\bar{\mathbf{k}}$, and in part on certain state-specific factors, such as prices, $\mathbf{p}^j$ and $\mathbf{w}^j$, and the vector of regulatory constraints, $\mathbf{R}^j$. The latter may itself depend on certain characteristics of the plant, z. It will now be assumed that capital is the only fixed factor, and that the "size of the plant" means "the level at which its capital is fixed". It will also be assumed that the only elements of z are the age and size of the plant.

The two characteristics of the plant's technology that are most likely to affect its environmental compliance cost are the particular industry to which it belongs and the vintage of its technology. Although the chemical industry is, overall, one of the most pollution-intensive 2-digit SIC code industries, there appears to be substantial variation in pollution-intensity among the thirty 4-digit SIC code industries within the chemical industry. Therefore,

---

analyses were typically less than one percent. Limiting the study sample to the non-small plants eliminated most of the large PACE weights, also greatly reducing the extent of the potential problem described here.

[15]The U.S. Environmental Protection Agency was established in 1970. That year may therefore serve as a kind of benchmark of when environmental regulations might have begun to have had any impact on the location decisions of new plants.

4-digit SIC code dummies are included among the exogenous variables in the econometric model.

The vintage of the technology is likely to have a significant impact on environmental compliance cost because newer technologies tend to be more efficient and less pollution-intensive than older ones. Although information on the vintage of each plant's technology is not available, the age of each plant is, and this is likely to be highly correlated with the vintage of the plant's technology. Therefore the age of the plant is included among the exogenous variables in the model.

Finally, the size of the plant is likely to be a dominant determinant of the plant's overall level of pollution and therefore its compliance cost. The total value of shipments (tvs) is used as a proxy for plant size.[16]

Monetary values for both compliance cost and total value of shipments (tvs) are in 1982-constant dollars and are logged so that a few relatively larger values will not dominate parameter estimates.[17] As a first order approximation, the effects of 4-digit SIC code industry, age, and log(size) on log(compliance cost) are additive and separable from $f(p^j, w^j, R^j(\text{age, size}))$, some function of the remaining variables in the model. The function $f(p^j, w^j, R^j(\text{age, size}))$ is assumed to satisfy the conditions necessary to ensure that $\pi \equiv \pi^{\circ} - cc$ has the properties of a restricted profit function.[18] The basic form of the econometric model up to this point is

$$\log cc_i = \mu + \gamma_k + \theta \text{age}_i + \lambda \log \text{size}_i + f(p^j, w^j, R^j(\text{age}_i, \text{size}_i)) + \varepsilon_i, \qquad (6.1)$$

[16] Because tvs = py, and the level of output, y, is itself a function of the vector of regulatory constraints, R, it is possible that compliance cost and tvs are jointly determined. This possible endogeneity problem is discussed more fully below.

[17] A very small number of observations had tvs=0, but these were eliminated by the fifth inclusion criterion. About five percent of the plant-years in the study sample reported zero abatement operating cost, and about a third of the plant-years reported zero capital expenditures for abatement, causing these to be considered censored dependent variables, as noted above. The problem of defining a truncation point, given that the log of 0 is undefined, is discussed below.

[18] For the properties of a restricted profit function, see, for example, Chambers, (1988), pp. 279-281.

if the ith plant is in the kth 4-digit industry and the jth state. We would expect that $\theta > 0$ and $0 \leq \lambda \leq 1$.

It is assumed that $f(p^j, w^j, R^j(\text{age}, \text{size}))$ can be approximated by the following decomposition:

$$f(p^j, w^j, R^j(\text{age}_i, \text{size}_i)) \approx \alpha(p^j, w^j, \overline{R}^j) + \eta\,\text{age}_i + \xi \log \text{size}_i,$$

where $\alpha(p^j, w^j, \overline{R}^j) \equiv \alpha_j$ is the effect on $\log cc_i$ of the average vector of regulatory constraints in the jth state, $\overline{R}^j$, in conjunction with state-specific prices, and $\eta\,\text{age}_i$ and $\xi \log$ size, are the deviations from that average effect resulting from deviations of $R^j(\text{age}_i,$ size,) from $\overline{R}^j$ due to the ith plant's particular age and size, respectively. If "grandfather" clauses are the predominant way in which regulations differentiate among plants of different ages, then $\eta < 0$. If environmental regulatory agencies treat plants of all sizes the same, then $\xi = 0$; if such agencies target the larger plants in particular, then $\xi > 0$.

The econometric model is now

$$\log cc_i = \mu + \alpha_j + \gamma_k + \delta\,\text{age}_i + \beta \log \text{size}_i + \varepsilon_i, \tag{6.2}$$

where $\delta = (\theta + \eta)$ and $\beta = (\lambda + \xi)$.

The age of the plant is seen to potentially affect its compliance cost in two opposite directions. The relative strengths of those effects can be examined by a hypothesis test on δ . Similarly, the size of the plant potentially affects its compliance cost in two ways. The value of ξ reflects the degree to which regulatory constraints are a function of plant size. The value of λ reflects the degree to which there are economies of scale in compliance cost in this industry, if regulatory constraints are not a function of plant size. Given prior beliefs about λ , inferences about ξ can be made based on a hypothesis test on β .

The average effect of state-specific environmental regulations on the compliance cost of a plant will be captured by state dummies[19]. The coefficient of the jth state dummy, α_j , is

[19] If a plant chose to locate in a particular state in part because of the environmental compliance cost it expected to incur in that state relative to other states, then compliance cost and state are jointly determined, and state dummies are not entirely exogenous. About two thirds of the plants in the study sample opened prior to 1970, however, that is, before environmental regulations were likely to be much of a consideration (see Table

correctly interpreted as "the effect of being located in the jth state," rather than "the effect of the environmental regulations of the jth state." Because $\overline{\mathbf{R}}^j$ is not the only state-specific argument in $\alpha(\mathbf{p}^j, \mathbf{w}^j, \overline{\mathbf{R}}^j)$, the two interpretations are not identical. It would be extremely difficult, however, to isolate the effect of a set of environmental regulations because the $\overline{\mathbf{R}}^j$ corresponding to that set is a detailed vector rather than a real-valued scalar. For two reasons, however, "the effect of being located in the jth state," captured by α_j, may actually be of more interest. First, it is in conjunction with prices that regulatory constraints decrease the plant's profit from what it would have been in the absence of regulations, and that is what matters to the plant. The higher the prices of inputs needed to comply or outputs foregone in complying with a set of environmental regulations, the greater the decrease in profit induced by that set of regulations. It is therefore the *joint* effect of $\mathbf{p}^j$, $\mathbf{w}^j$, and $\overline{\mathbf{R}}^j$ that is important.

Second, $\alpha(\mathbf{p}^j,\mathbf{w}^j,\overline{\mathbf{R}}^j)$ will capture the effect of the jth state's environmental regulations *as enforced by the environmental regulatory agencies in that state.* If, as is likely, states differ in the degree to which they actually enforce their regulations, it is the *actual* constraints imposed by their regulations, rather than the constraints "on the books" that affect the plant's compliance cost and would therefore affect location decisions. If, for example, $\mathbf{R}' = (\tilde{\mathbf{y}} > 0, \overline{\mathbf{q}} < \infty, \underline{\mathbf{x}} > 0, \overline{\mathbf{x}} < \infty, e > 0)$ is the vector of regulatory constraints implied by the laws on the books, but these laws are not enforced and are known to be unenforced, then the vector of constraints atually facing the plant is $R = (0,\infty,0,\infty,0)$. It is the vector of actual constraints on the plant that will affect its profit maximizing decisions.

While the PACE questionnaire elicits information on outlays for pollution abatement, the plant's *total* environmental compliance cost consists not only of outlays for pollution

2). Although state dummies are treated as exogenous in the main set of analyses, the effect of the possible endogeneity of states is explored in a separate set of analyses.

abatement, but opportunity cost as well (see Section 2).[20] Even total outlays for pollution abatement cannot be fully measured for most plants in the PACE dataset, because only a small subset of plants reported all the necessary information. Total compliance cost must therefore be approximated by those variables that are relatively well reported, namely, capital expenditures for abatement or abatement operating costs (or a combination of the two). Although approximations based on the PACE data are likely to understate a plant's total compliance cost, this would present a problem only if there is some systematic bias in the degree of understatement.

For several reasons, abatement operating cost appears to be the best variable in the PACE dataset to approximate total compliance cost. In contrast to abatement capital expenditures, annual operating costs for abatement are quite stable over time. Plant-specific coefficients of variation for abatement capital expenditures were routinely larger, and often an order of magnitude larger, than the corresponding plant-specific coefficients of variation for abatement operating costs. Parameter estimates based on abatement operating cost would thus be less vulnerable to the destabilizing effect of missing years.

Abatement operating costs, moreover, appear to reflect abatement capital expenditures rather well. Among plants that were in the dataset for at least nine of the eleven years (i.e., had no more than two missing years), the correlation between plant-specific mean abatement capital expenditures and plant-specific mean abatement operating costs is 0.85.[21]

Finally, abatement operating costs appear to surpass all other abatement costs. In 1983, for example, total abatement operating cost in SIC code 28 was $2090.1 million, while total

[20] If, for example, compliance with regulations took the form of limiting output, outlays for pollution abatement would be zero but there would be a positive compliance cost equal to the decrease in profit resulting from choosing a suboptimal level of output.

[21] Because of the way plants were selected into the PACE sample, the subset of plants appearing in the PACE dataset in at least 9 out of 11 years is not representative of the general population of plants. The PACE sample is skewed towards larger plants, and this subset of the PACE sample is even more skewed in that direction. This selectivity bias problem arises any time plants are selected from the PACE sample on the basis of the number of years they appear in the sample.

abatement capital expenditures was \$395.4 million and payments to government for pollution removal was \$108.1 million.[22]

Because abatement operating costs (aoc) are likely to yield more precise, and therefore more reliable estimates, because they are the largest abatement costs, and because they appear to be reflective of capital expenditures for abatement as well, they seem to provide the most reasonable measure of environmental compliance costs.

Model (6.2) thus becomes:

$$\log(aoc_i) = \mu + \alpha_j + \gamma_k + \delta age_i + \beta \log(tvs_i) + \varepsilon_i \qquad (6.3)$$

Abatement operating cost (aoc), however, is a censored dependent variable, with over ten percent of plants reporting zero aoc in some years (see Table 5). In addition, if all years are pooled in the analysis, observations from the same plant will be correlated.

Efficient and consistent parameter estimates in models with limited dependent variables and serial correlations can, in theory, be obtained by maximum likelihood estimation. However, the likelihood function involves N truncated multivariate normal densities, the ith density corresponding to the n_i correlated observations from the ith plant, i=1, ... N. Because these density functions are truncated, maximum likelihood estimation involves evaluating multidimensional integrals. For more than 3 dimensions (i.e., $n_i > 3$), this becomes computationally infeasible. Alternative estimation procedures, which trade some efficiency for computational feasibility, are therefore necessary.

If the error is not correlated with the exogenous variables, then using a standard Tobit estimation method on the pooled data, ignoring the correlations among the errors, yields consistent, though inefficient, estimates (Robinson, 1982). The estimate of the covariance matrix of the parameter estimates, however, is not consistent. Therefore, inferences based on tests which depend on consistent estimates of the variances of parameter estimates (such as F-tests and t-tests) are not valid. Other methods of circumventing the serial correlation problem

[22] U.S. Department of Commerce, Bureau of the Census, Current Industrial Reports, Pollution Abatement Costs and Expenditures, 1983, MA-200 (83)-1.

created by pooling years introduce serious selectivity biases into the sample.[23] These biases stem from the fact that the PACE samples are probability samples.

Because the consistent and efficient method of analyzing these data across all years is computationally infeasible, while various alternative, less efficient but consistent methods of analyzing the pooled data have serious drawbacks, each of the eleven years was analyzed separately. Assuming the model is well specified, this approach yields consistent estimates and allows valid inferences to be made to the general population of "larger" chemical manufacturing plants, as defined in Section 5. Although separate-year analyses are less efficient than an analysis in which years are pooled, sample sizes in each year are still substantial (see Table 3).[24]

It is unclear whether all of the reported zero abatement operating costs are true zeros. Estimating abatement operating costs may have been difficult, and some reported zeros may simply reflect the unwillingness of plant personnel to do so. It is unlikely, in any case, that the specification of exogenous factors that predict the value of a positive abatement operating cost is the same specification that predicts whether an abatement operating cost is reported to be zero or positive. A Tobit model is thus likely to give distorted estimates of the parameters relevant to the ninety-five percent of plant-years that reported positive abatement operating costs. Therefore, the following model, due to Cragg (1971), is used instead for each of the eleven separate-year analyses:

$$Y_i^* = Z_i' \Gamma + \eta_i \tag{6.4}$$

---

[23] The creation of selectivity bias by methods intended to avoid the problem of serial correlation is described in the dissertation from which this paper is derived..

[24] If separate-year analyses suggest that the relationships in the model are in fact time-independent, a single set of parameter estimates (for all years) could be obtained from the eleven separate-year vectors of estimates by applying a method suggested by Malinvaud (1970) and Chamberlain (1984), and applied to labor data by Jakubson (1988). This method is a two-step procedure in which (1) a Tobit (or other) analysis is run on the above model, separately for each time period, t=1, ... T, and (2) a minimum distance estimator is then applied to the resulting T vectors of estimates to impose any desired restrictions on the final parameter estimates - for example, the restriction that each parameter in the model is the same across years.

$$Y_i > a^{25} \text{ whenever } Y_i^* > 0,$$
$$Y_i = a \text{ whenever } Y_i^* \leq 0$$
$$Y_i = X_i' B + \varepsilon_i, \text{ given } Y_i > a, \tag{6.5}$$

where

Y* is a latent (unobserved) variable such that whenever Y*>0, reported abatement operating cost is positive, and whenever Y* ≤ 0, reported abatement operating cost is zero,

Y is log(reported abatement operating cost),

Z is a vector of exogenous variables that affect whether a plant reports a zero or positive abatement operating cost,

X is a vector of exogenous variables that affect abatement operating cost, given that it is positive, and

η and ε are random errors.

Assuming that η_i is distributed N(0,1), (6.4) can be reformulated as

$$\text{Prob}(Y_i > a) = \Phi(Z_i' \Gamma) \tag{6.4'}$$
$$\text{Prob}(Y_i = a) = 1 - \Phi(Z_i' \Gamma),$$

where Φ is the cumulative distribution function of a standard normal random variable.

The vector of parameters in (6.4'), Γ, is estimated by a probit analysis. The vector of parameters in (6.5), B, is estimated by a truncated regression on those observations with positive abatement operating costs. In all analyses, observations are weighted by their PACE weights. Because it is believed that the precision of reported information increases with plant size, heteroskedasticity in (6.5) is modeled by making the variance of the error depend on plant size. In particular, the distributional assumption on ε_i is that ε_i is distributed N(0, σ_i^2), where

$$\sigma_i^2 = \sigma^2 \exp(\theta \log(tvs_i)).$$

The parameter θ is estimated along with the vector. B in the truncated regression analyses.

---

[25]If Y were abatement operating cost, then the truncation point, a, would be 0. However, the relationship between log(aoc) and log(tvs) appears to fit a linear model much better than that between aoc and tvs. Because log(0) is undefined, there is no clear choice for a. This is discussed below.

Because abatement operating cost is logged, the exact truncation point is undefined. To circumvent this problem and preserve the true truncated nature of abatement operating costs, all reported zero abatement operating costs are set to $1.00. Because aoc is actually reported in units of thousand-dollars, aoc is set to 0.001 ($1.00) whenever it is zero, making the truncation point -6.9.[26]

Although the parameter values in (6.4') are not forced to be the same as those in (6.5), it is assumed that X=Z. The specific models that are estimated, separately for each year, then, are

$$\Pr ob(aoc_i > 0) = \Phi(\mu' + \sum_{k=1}^{n_y-1} \alpha'_k state_k + \sum_{j=1}^{m_y-1} \gamma'_j industry_j + \delta' age_i + \beta' \log(tvs_i)) \quad (6.6)$$

$$\Pr ob(aoc_i = 0) = 1 - \Phi(\mu' + \sum_{k=1}^{n_y-1} \alpha'_k state_k + \sum_{j=1}^{m_y-1} \gamma'_j industry_j + \delta' age_i + \beta' \log(tvs_i)).$$

If $aoc_i > 0$,

$$\log(aoc_i) = \mu + \sum_{k=1}^{n_y-1} \alpha_k state_k + \sum_{j=1}^{m_y-1} \gamma_j industry_j + \delta age_i + \beta \log(tvs_i) + \varepsilon_i \quad (6.7)$$

where

ε_i is distributed $N(0, \sigma_i^2)$, of as defined above,

$state_k$ = 1 if the ith plant is in the kth state, = 0 otherwise,

$industry_j$ = 1 if the ith plant is in the jth 4-digit SIC code industry, = 0 otherwise, and

$(n_y$ - 1) is the number of states and $(m_y$ - 1) the number of 4-digit industries in the y-th year analysis.[27]

---

[26] Where to set the truncation point when limited dependent variables are logged can be a serious problem for the Tobit model, in which observations at the truncation point are included in the estimation of model parameters. It has very little effect, however, on the estimation of parameters in the Cragg model. It has no effect at all on the estimation of parameters in the probit model (6.4'). It has only a very slight effect on the estimation of parameters in the truncated regression model because observations at the truncation point are not included. To verify this, truncated regressions were run for 1979 using truncation points of -100 and -6.9. Differences in parameter estimates were typically less than one percent.

[27] The number of states and the number of 4-digit industries in analyses varied somewhat from year to year because elimination of observations that did not meet all the inclusion criteria sometimes resulted in a state and/or 4-digit industry having too few observations. Any state with fewer than 5 observations and any 4-digit industry with fewer than two observations was excluded from an analysis. The dummy for the state of Massachusetts was omitted from every analysis to avoid singularity of the X'X matrix. Similarly, the dummy

In addition, several joint hypotheses are tested. First, to test whether the relationship between (positive) abatement operating cost and plant size differs across states, the following model is also estimated for each year:

If $aoc_i > 0$,

$$\log(aoc_i) = \mu + \sum_{k=1}^{n_y-1} \alpha_k state_k + \sum_{j=1}^{m_y-1} \gamma_j industry_j + \delta age_i + \beta_k \log(tvs_i) + \varepsilon_i \qquad (6.8)$$

for the ith plant in the kth state.

Equation (6.7) is a restricted version of (6.8) in which the coefficient of log(tvs) is constrained to be the same in all states. A likelihood ratio test is then used to test the hypothesis,

$$H_0: \beta_1 = \beta_2 = \ldots = \beta_{n_y} = \beta .$$

Second, to test whether the relationship between (positive) abatement operating cost and plant age differs across states, the following model is estimated for each year:

If $aoc_i > 0$,

$$\log(aoc_i) = \mu + \sum_{k=1}^{n_y-1} \alpha_k state_k + \sum_{j=1}^{m_y-1} \gamma_j industry_j + \delta_k age_i + \beta \log(tvs_i) + \varepsilon_i \qquad (6.9)$$

for the ith plant in the kth state.

Equation (6.7) is a restricted version of (6.9) in which the coefficient of plant age is constrained to be the same in all states. A likelihood ratio test is used to test the hypothesis,

$$H_0: \delta_1 = \delta_2 = \ldots = \delta_{n_y} = \delta .$$

Finally, to test whether states overall make a difference in positive abatement operating costs, if the above two hypotheses are accepted, the following model is estimated:

If $aoc_i > 0$,

$$\log(aoc_i) = \alpha + \sum_{j=1}^{m_y-1} \gamma_j industry_j + \delta age_i + \beta \log(tvs_i) + \varepsilon_i . \qquad (6.10)$$

Equation (6.10) is a restricted version of (6.7) in which the coefficients of all state dummies are constrained to be zero. A likelihood ratio test is then used to test the hypothesis,

$$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_{n_y-1} = 0$$

---

for 4-digit SIC code industry 2821 was omitted for the same reason, making 29 the maximum possible number of industry coefficients in a model.

or, equivalently,

$$H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_{n_\gamma} = \alpha$$

-- i.e., that the state in which the plant is located does not affect its abatement operating cost, given that it is positive. The distributional assumptions for ε_i in (6.8), (6.9), and (6.10) are the same as in (6.7).

To examine the effect of the possible endogeneity of states (see footnote 19), the above analyses were also run on that subset of the plants in the study sample that opened prior to 1970, before environmental regulations were likely to have been much of an issue.

7. *Results*

The results of the probit analyses, estimating the process by which abatement operating costs are reported as zero or positive, and the truncated regression analyses, estimating the process by which the values of positive abatement operating costs are determined, are reported separately. The process by which an abatement operating cost is reported as zero or as positive was generally not successfully modelled by equation (6.6). Tests of overall significance of the model showed that in all but two years the estimated model did not fit statistically significantly better than a model with all zero coefficients. The very small numbers of zeros in several years may have made the explanatory process in those years particularly difficult. Even in years with substantial numbers of zeros, however, the model failed. It is possible that certain factors are associated with the absence of abatement operating costs, and these factors were simply not in the model. It is also possible that some, perhaps a substantial proportion, of the reported zero costs were not actual zero costs, but instead reflect an unwillingness of plant personnel to estimate this cost. There is, unfortunately, no obvious way to distinguish between this possibility and real zero abatement operating costs.

Because almost ninety-five percent of all plant-years in the sample reported positive abatement operating costs, however, the results of the truncated regression analyses are of more interest. The results provide little support for the hypotheses that the relationship between environmental regulations (and/or their enforcement) and plant size or age varies by state, at least among the relatively larger plants examined. The first joint hypothesis, that the coefficient of log(tvs) is the same for all states (model (6.7) compared with (6.8)), is rejected in only one out of eleven years.[28] The second joint hypothesis, that the coefficient of plant age is the same for all states (model (6.7) compared with (6.9)), is rejected in only one out of seven years[29].

The third joint hypothesis, that there are no state effects on abatement operating costs, given that the first two joint hypotheses are accepted (model (6.10) compared with (6.7)), is rejected in eight of the eleven years. There is thus evidence, if somewhat weak, that states do significantly affect the abatement operating costs of the larger chemical manufacturing plants. No pattern over time is evident in the rejection (vs. acceptance) of this hypothesis.

Given the overall evidence that neither the coefficient of log(tvs) nor the coefficient of plant age is state-specific, and the evidence that the coefficients of state dummies are not all the same, the model represented by equation (6.7) appears to be the best choice. Tables 6a and 6b give the estimated coefficients from this model, separately for each of the eleven years.

Overall, model (6.7) appears to fit the data quite well. Although no goodness-of-fit statistic is provided for the truncated regressions, R^2 values are given for the Ordinary Least Squares used to provide the initial values for the iterative procedure necessary to maximize the log likelihood function for the truncated regression model. These R^2 values range from 0.679, in 1979, to 0.786 in 1985.[30]

---

[28] Rejection of null hypotheses was based on the 95% level of confidence.

[29] The test of whether the coefficient of plant age is state-specific could not be performed in four of the eleven years because of multicollinearity problems in the unrestricted model.

[30] Because the number of observations with aoc=0 is less than 15% in each year, the coefficient estimates from the truncated regressions would not be expected to be enormously different from those generated by OLS (although the initial OLS regressions do not adjust for heteroskedasticity). The goodness-of-fit of the OLS models should therefore give a rough idea of the goodness-of-fit of the truncated regressions.

Differences in abatement operating cost among 4-digit SIC code industries within the chemical industry are often substantial, statistically significant, and consistent over time. Some industries, such as Pharmeceutical Preparations (SIC code 2834), Soap and Other Detergents (SIC code 2841), Toilet Preparations (SIC code 2844), and Paints and Allied Products (SIC code 2851) always have statistically significant negative coefficients. Others, such as Industrial Inorganic Chemicals, N.E.C.[31] (SIC code 2819), Cyclic Crudes and Intermediates (SIC code 2865), and Industrial Organic Chemicals, N.E.C. (SIC code 2869) always have positive coefficients which are almost always significant. Unless environmental regulatory agencies are targeting certain industries, these differences presumably reflect real differences in the pollution-intensity of industries within the chemical industry.

The coefficient of log(tvs) is the elasticity of abatement operating cost with respect to the total value of shipments (that is, approximately, the level of output) of the plant-year. This parameter is always significant at the 99% level and is not significantly different from 1 in any of the years except 1979, when it is significantly greater than 1 (see footnote 2, Table 6a). The implications of this result are discussed in Section 8.

The coefficient of age is always positive and statistically significant at the 95% level. It is significant at the 99% level in ten of the eleven years. The year-specific coefficients of age show only small variation, ranging from 0.015 in 1988 to 0.036 in 1985, with an average of 0.024.

In contrast to all other coefficients in the model, the coefficients of state dummies are virtually never statistically significantly different from zero.[32] This result, however, requires careful interpretation. To avoid singularity of the X'X matrix in the estimation of model parameters, one state must be omitted from the model. Massachusetts is arbitrarily chosen. This is equivalent to constraining α_{MA} to be 0. The lack of statistical significance of the.

---

[31]N.E.C. means "not elsewhere classified".

[32]Out of 308 coefficients of state dummies, only 3 were "significant" at the 95% level, which is less than would be expected by random chance.

coefficients of state dummies thus means that the effect of each of the other states is not statistically distinguishable from that of Massachusetts. (A different choice of omitted state dummy might have resulted in some statistically significant coefficients.) The question of whether states overall affect a plant's abatement operating cost is more relevant and is appropriately answered by testing the third joint hypothesis, discussed above. The rejection of this hypothesis in eight of the eleven years is unlikely to have happened by random chance alone. The evidence thus suggests that, considering all states, the location of a plant does affect its abatement operating cost, even though any binary comparison between a particular state and Massachusetts shows no discernable difference.[33] The fact that virtually no individual coefficients are statistically significantly different from zero, however, suggests that whatever state effect exists, it is relatively weak.[34]

The results of the analyses on the subset of plants that opened prior to 1970 were not appreciably different from the results described above. In particular, model (6.7) was still the best choice model, and there was weak evidence of a state effect in that model.[35]

8. *Discussion*

The results presented in Section 7 offer some insight into how environmental regulations, interacted with plant characteristics to affect the compliance costs of chemical manufacturing plants during the period 1979-1990. Several interesting results emerged.

First, the results suggest that environmental regulatory agencies did in fact target larger plants disproportionately. It is likely that chemical manufacturing plants enjoy economies of

---

[33] The choice between Michigan and Texas, for example, may have shown statistically discernably different abatement operating costs.

[34] A comparable joint hypothesis for industry effects, for instance, would undoubtedly have been rejected in all eleven years.

[35] The third joint hypothesis was rejected in seven out of eleven years. The set of years in which this hypothesis was rejected, however, was not. entirely contained within the set of years in which it was rejected using the larger sample. The results of the analyses on the subset of plants that opened prior to 1970 are presented more fully in the dissertation from which this paper is derived.

scale in compliance cost. If plants of all sizes were treated the same by environmental regulatory agencies ($\xi = 0$), then the elasticity of abatement operating cost with respect to tvs (β) would be expected to be significantly less than 1, reflecting these economies of scale. The fact that this elasticity is not significantly different from 1 in ten of the eleven years and is significantly greater than 1 in the remaining year, suggests that environmental regulatory agencies leaned more heavily on the larger plants than on the smaller ones during the study period. Recalling that the sample, and the population of plants to which inferences may be generalized, is limited to those “larger” plant-years with tvs of at least $10,000,000, it is likely that this result would be strengthened further had plants of all sizes been included.

The fact that the coefficients of plant age are consistently positive and significant suggests that any effect of “grandfathering” of older plants is outweighed by the greater pollution-intensity of older plants. Newer plants may face more stringent regulations, on average, but their more pollution-efficient technologies appear to more than make up for it.

Finally, the results on possible state effects warrant further discussion. First, it is noteworthy that, although there appears to be some effect of state on a plant’s abatement operating cost, it is the weakest of all the effects tested. Both plant age and log(tvs) are statistically significant in every year, usually at a high level of confidence. Although an overall 4-digit industry effect is not tested, it is clear from the consistently highly significant coefficients of industry dummies that an overall industry effect would also be significant in every year. In contrast, although an overall state effect tests significant in eight years, it does not show up as significant in three years. Even in those years in which an overall state effect is statistically significant, moreover, almost no individual state coefficients are significant, suggesting that the state effect is a weak effect.

Second, the results of these analyses do not support the image of California as substantially more stringent than most other states. Although California has gained the reputation as a leader in environmental awareness, this reputation did not translate into significantly greater abatement operating costs reported by chemical plants in that state as compared with comparable chemical plants in other states during the study period.[36]

There are several possible reasons why the state effect is not more pronounced and why states that might have been expected to stand out do not. First, the PACE dataset is likely to be somewhat biased towards sameness across states. Sufficiently stringent regulations in a state might have caused some plants in that state to close or to move out of the state. Because individual plants are not followed over time in the PACE survey, the closing of a plant or the relocation of a plant out of a state would not be detectable in the study sample, nor would the decision of a firm to close its plant in one state and open a different plant in another state or abroad because of differences in environmental regulatory stringency.

Similarly, sufficiently stringent regulations in a state might have induced changes in technology in some plants in that state (see footnote 4), with a corresponding decrease in profit. While abatement operating cost is likely to generally understate a plant's total environmental compliance cost,[37] as defined in Section 2, the understatement may be greater in those cases involving technological change.

Second, it is possible that the largest proportion of compliance cost comes from initial compliance, that is, the acquisition and installation of abatement and monitoring capital, and that much of this was done prior to 1979. Plants in the study sample with higher abatement capital expenditures, however, generally also had higher abatement operating cost (see the discussion of the measurement of compliance cost in Section 6). Therefore, unless this relationship changed over time, interstate differences in abatement capital expenditures before 1979 should be reflected in interstate differences in abatement operating costs after 1979.

Third, there may be a substantial gap between the vector of constraints implied by environmental regulations "on the books" and the vector of regulatory constraints that a

---

[36] It is possible that the abatement costs reported by the plants in California were downward biased, but there is no obvious reason why this should be the case.

[37] If costs recovered through abatement activities are substantial, aoc may actually *overstate*-total compliance cost. In most cases, however, aoc is more likely to understate compliance cost.

chemical manufacturing plant actually faces. In addition, interstate differences among vectors of implied regulatory constraints may be substantially greater than interstate differences among vectors of actual regulatory constraints. A study of monitoring and enforcement of compliance among plants by state environmental agencies (Russell et al., 1986) revealed a pattern of infrequent investigation of plants by state agency officials and even more infrequent enforcement actions. The study found, in particular, that (1) self-monitoring by plants is very widely relied upon, (2) auditing of self-monitoring sources is infrequent, (3) enforcement usually relies heavily on "voluntary compliance," (4) there is only a small probability that a Notice of Violation, issued to an audited plant found to be in violation of the terms of its permit, will lead to an actual monetary penalty, and (5) penalties that are imposed are generally small, making it questionable whether such penalties alone provide the incentive necesssary for future compliance. Studies by Harrington (1981), McInnes and Anderson (1981), and the U.S. General Accounting Office (1983) (cited by Russell et al. (1986)) suggest, moreover, that the degree of noncompliance with the terms of air and water permits has been considerable. In the GAO study, for example, 21 percent of industrial wastewater dischargers examined were found to be in "significant noncompliance during the period."

If the degree of noncompliance by plants is proportional to the severity of the regulatory constraints they face, then noncompliance would tend to smooth out differences in the compliance costs of similar plants in different states, thereby weakening a state effect. Apparently substantial differences in environmental laws or in the reputations of states for "greenness" may not translate into correspondingly substantial differences in the compliance costs of the plants in those states if noncompliance is a significant factor.

The translation of environmental regulations into actual constraints on the plant is further complicated by the high degree of discretion accorded to state and local regulatory agencies in the administration of environmental laws. As Selmi and Manaster (1989) observe, "The difficulty -- even impossibility -- of having legislation enacted with sufficient detail to eliminate the need for the exercise of bureacratic discretion and technical judgment is certainly evident in the environmental field." Environmental laws at times appear to have been written with this in mind.

Finally, the notion of a state-specific vector of environmental regulatory constraints, even for plants of the same size and age, is an oversimplification. In addition to federal and state regulations, plants are often required to comply with local ordinances. When these impose more severe constraints than federal or state laws would, they alter the vector of regulatory constraints facing the plant. In Indiana, for example, discharge to sewage treatment plants is subject to local sewage ordinances. During the 1980's, Indiana had no state-wide limits for discharge to sewage treatment plants. Instead, 45 local sewage authorities set local limits, which may have varied substantially.[38] In California, the regulation of air pollution is administered by several separate air districts. The particular air district in which the plant is located determines that part of its regulatory constraint vector related to air pollution.

The geographic configuration of plants within a state or local jurisdiction can also affect the regulatory constraints placed on those plants. If there is an upper limit to the acceptable concentration of a pollutant in a body of water, for example, then the per-plant effluent limit for that pollutant may decrease as the number of plants discharging into that body of water increases. A high level of within-state variablility due to local ordinances and local conditions may make any state effects particularly difficult to detect.

The results of the analyses may also have been affected by two possible sources of endogeneity. If plants are making location decisions as well as production decisions, then states could be endogenous. The endogeneity of location, after all, underlies all studies of the effects of environmental regulations on the firm location decision. About two-thirds of the plants in the study sample opened prior to 1970 (Table 2), however, and these plants are unlikely to have considered environmental regulations in their location decisions. Any endogeneity of state would be related to the relatively newer one-third of the plants in the sample. A comparison of

---

[38] Personal communication with Philip Preston, Pretreatment Coordinator, Permits Section, Operations Branch, Indiana Department of Environmental Management, Indianapolis, IN.

the analyses on the entire sample with the analyses on the subset of plants that opened prior to 1970, however, showed little difference.

The other possible source of endogeneity is tvs, used as a proxy for plant size $(\bar{k})$. Although $\bar{k}$ is exogenous, tvs may be partly endogenous, because it closely approximates output level (tvs=py), a choice variable that 'is a function of R and may therefore be jointly determined with compliance cost. In the short-run, however, the dependence of y on $\bar{k}$ is likely to far exceed its dependence on R. That is, changes in tvs reflecting output decisions related to R are likely to be only small variations in a much larger pattern determined by capital constraints. The degree of endogeneity is therefore likely to be quite small. This possible endogeneity of tvs, however, will be investigated in future analyses.

A major difficulty in assessing the effect of environmental regulations on where firms locate their plants is that each state has a *set* of environmental regulations, and often *several* regulatory agencies that impact the plant. Each set of regulations corresponds to a vector of implied constraints on the plant's profit maximization problem, which may differ from the *actual* constraints faced by the plant. It has been argued here that the environmental regulatory stringency of a state should reflect the impact, on average, of environmental regulations, and the agencies that administer those regulations, on the environmental compliance cost of plants in that state. The coefficients of the state dummies in the analyses presented here, then, offer a reasonable way of comparing the relative stringencies of the states examined during the study period.

The results of these analyses therefore have implications for studies, such as those reviewed in Section 3, that attempt to discern the effect of environmental regulations on the firm location decision. Based on the pollution abatement costs actually reported by chemical manufacturing plants in the PACE surveys from 1979 through 1990, interstate differences in stringency appear to have been barely statistically discemable. It is not surprising, then, that it has been difficult to discern any effects of these differences on the location decisions of plants within the United States. There are, however, two caveats to this conclusion.

First, the weakness of the state effect within the United States does not necessarily imply a similarly weak effect of national-level environmental regulations in international comparisons. The fact that states mattered collectively in explaining compliance cost in eight of the eleven years examined suggests that, if international differences in environmental regulations are substantially greater than interstate differences within the United States, such international differences could indeed translate into significant differences in compliance costs in different countries.

Second, compliance costs, as defined above, are not the only costs a plant could incur in dealing with environmental regulatory agencies. It is possible that the more important impact of environmental regulations comes in the form of transaction costs incurred in the process of having to obtain all the permits necessary to be allowed to begin production. This was the dominant response of business executives involved in location decisions. When asked whether environmental regulations affected the decision of where to locate their plants, business executives generally responded that compliance costs were not a concern, but that the process of satisfying all the requirements to obtain the necessary permits -- in particular, the opportunity cost of waiting for the permitting process to be completed -- could be (Stafford, 1985). The results of these plant-level analyses suggest that, if there is an economically significant effect of environmental regulations on the firm location'decision within the United States, it may be in the form of transaction cost differentials rather than compliance cost differentials. The investigation of this hypothesis would be an area of useful further research.

References

Bartik, T.J., 1988. "The Effects of Environmental Regulation on Business Location in the United States," *Growth and Change,* Summer 1988, pp. 22-44.

Chamberlain, 1984. "Panel Data." Chapter 22 of *Handbook of Econometrics,* vol. 2, Z. Griliches and M. Intriligator, eds., Amsterdam: North Holland Publishing Co.

Chambers, R.G., 1988. *Applied Production Analysis: A Dual Approach,* Cambridge University Press.

Cragg, J., 1971. "Some Statistical Models for Limited Dependent Variables With Application to the Demand for Durable Goods," *Econometrica,* 39: 829-844.

Hall, B. and Kerr, M.L., 1991. *1991-1992 Green Index: A State-By-State Guide to The Nation's Environmental Health,* Island Press, Washington, D.C.

Harrington, W., 1981. *The Regulatory Approach to Air Quality Management*, Resources for the Future, Washington, D.C.

Jakubson, 1988. "The Sensitivity of Labor-Supply Parameter Estimates to Unobserved Individual Effects: Fixed- and Random-Effects Estimates in a Nonlinear Model Using Panel Data," *Journal of Labor Economics,* 6(3): 302-329.

Leonard, H.J., 1984. *Are Environmental Regulations Driving U.S. Industry Overseas?,* The Conservation Foundation, Washington, D.C.

Levinson, A., 1992. "Environmental Regulations and Manufacturers' Location Choices: Evidence from the Census of Manufactures," Unpubl. paper.

Malinvaud, E., 1970. *Statistical Methods of Econometrics*, Amsterdam: North Holland Publishing Co.

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

McGuckin, R.H. and Pascoe, G.A. Jr., 1988. "The Longitudinal Research Database: Status and Research Possibilities," *Survey of Current Business,* November 1988, pp. 30-37.

McInnes, R.G., and Anderson, P.H., 1981. "Characterization of Air Pollution Control Equipment Operation and Maintenance Problems," Report, U.S. Environmental Protection Agency, Washington, D.C.

Robinson, P.M., 1982. "On the Asymptotic Properties of Estimators of Models Containing Limited Dependent Variables," *Econometrica,* 50(1): 27-41.

Russell, C.S., Harrington, W., and Vaughan, W., 1986. *Enforcing Pollution Control Laws,* Resources for the Future, Washington, D.C.

Selmi, D.P., and Manaster, K.A., 1989. *State Environmental Law,* Clark, Boardman, and Callaghan, Deerfield, II.

Stafford, H.A., 1985. "Environmental Protection and Industrial Location," *Annals of the Association of American Geographers* 75(2): 227-240.

U.S. Department of Commerce, 1981. *U.S. Industrial Outlook 1981.*

U.S. Department of Commerce, Bureau of the Census, 1983. *Current Industrial Reports, Pollution Abatement Costs and Expenditures, 1983,* MA-200 (83)-1.

U.S. General Accounting Office, 1983. *Wastewater Dischargers Are Not Complying with EPA Pollution Control Permits,* U.S. GAO, Washington, D.C., GAO/RCED-84-53 (December).