ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 45 — 1997

# The Impact of Air Quality Regulation on Industrial Location

Vernon HENDERSON*

**ABSTRACT.** – This paper examines the effect of local air quality regulation on location of polluting industries in the USA. USA counties are designated annually as being in attainment or non-attainment of national air quality standards. Regulations on new and existing plants governing equipment and operating specifications are much tougher in non-attainment counties. A switch from non-attainment to attainment status greatly raises the probability of a county having a particular polluting industry.

---

## L'impact des règlements sur la qualité de l'air sur la localisation industrielle

**RÉSUMÉ.** – Dans cet article, nous essayons de déterminer quel est l'impact des règlements propres à chaque région concernant la qualité de l'air, sur la localisation des industries polluantes aux Etats-Unis. Chaque année, les contrées sont désignées selon qu'elles atteignent ou non les critères standards de bonne qualité de l'air. La réglementation est beaucoup plus sévère dans les contrées n'ayant pas rempli ces critères. Lorsqu'une région améliore son statut, la probabilité qu'une industrie polluante s'installe augmente fortement.

* V. HENDERSON: Brown University.
Support for this work by the National Science Foundation is gratefully acknowledged.

# 1 Presentation

This paper examines the impact of local air quality regulation on industrial location in the USA. I focus on 3-digit manufacturing industries which are the major emitters of volatile organic compounds (VOC). VOC emissions, along with nitrogen and atmospheric conditions such as high temperatures, are a key ingredient in the formation of ground level ozone. Exposure to ground level ozone damages human tissues and cells, affecting lung function in healthy children and adults, as well as more particularly for those with impaired respiratory systems. Regulations in the USA seek to reduce VOC and nitrogen emissions by industrial plants and motor vehicles, to reduce ozone concentrations.

There are national and local applications of regulations in the USA. National regulations affect the choice and design of new operating equipment in plants and automobiles and the choice of fuels, solvents and coolants, as examples. Local regulation involves the local implementation of national objectives and guidelines. Each county in the USA is designated each July in the Code of Federal Regulations (title 40, part 81, subsection c) as being in attainment or non-attainment of national air quality standards (NAAQS's), in different air quality dimensions including ozone. For counties not in attainment, each state is periodically required to submit plans (SIP's), indicating how the county will be brought into attainment by a particular date. Absent timely attainment, federal funding in various categories may be at risk.

For counties not in attainment of federal regulations, new plants to the county may be subject to more stringent regulations governing equipment specifications, than they would face in attainment counties. Existing plants in non-attainment areas face requirements to reduce source emissions; and new plants may be required to purchase offsets (emission rights) from existing plants (ATKINSON and TIETENBERG [1987], ROUMASSET and SMITH [1990]). All plants in non-attainment counties are more likely to be closely monitored and subject to greater enforcement efforts (see DEILY and GRAY [1991] and RUSSEL [1990]). In summary, plants in counties not in attainment of air quality standards face greater enforcement and emission reduction activity.

In making location decisions for their plants, firms evaluate county attributes, including the extent and cost of emission reduction programs. Existing plants undertaking capital expansion may want to stay where they are, even in non-attainment areas, to avoid new source performance standards that are imposed on new plants (including relocators). For new firms or new plants of existing plants the situation is different. New plants will find that counties in attainment are attractive places to locate because typically they face weaker requirements on emission reduction activity.

Nationally all this implies that there is a tendency towards a "graying process". Over time, plants will want to leave, or, at least, new plants will not want to locate in non-attainment areas. That will lead to improvements in air quality in those areas, over and above improvements due to regulation of remaining plants. But that implies that new plants will move to attainment areas, potentially decreasing air quality in those areas. While the Clean Air Act and its amendments recognize this possibility and in theory have guidelines to prevent significant deterioration of air quality in attainment areas, still in practice the tendency is there. In fact, some states and counties may be willing to be a refuge for polluting plants which generate local employment and investment, providing local air quality doesn't deteriorate so much that the county runs the risk of violating national standards. These responses also mean that the regulatory process is distorting location decisions of firms and their plants.

In this paper we examine the impact of regulation on the location decisions of five 3-digit industries which are major VOC emitters, between 1978 and 1987. These are industrial organic chemicals (SIC 286), miscellaneous plastics (SIC 307), plastic materials and synthetics (SIC 282), blast furnace and primary steel (SIC 331), and petroleum refining (SIC 291). The extent of applied regulation will be indexed by a county's attainment status in various air quality dimensions, with a focus on ozone. The issue concerns how plants in these industries alter certain behaviors, *apart* from emission reduction, in response to greater effective air quality regulation. Most of the work is focused on showing that plants will try to move their plants into attainment areas and out of non-attainment areas. But other adaptations such as a reduction in plant sizes are also examined. Firms may reduce plant sizes, as they spread out their polluting activity across several locations. This not only reduces emissions at any location so as to generally attract less attention; but in particular may move the plant out of the category of class A polluters which are the focus of EPA attention.

I believe these hypotheses are straightforward. However, they are not generally accepted in the literature. McConnell and Schwab [1990], as well as Bartik [1988], find no regulatory impact on firm location decisions, considering factors such as the extent of local non-attainment of air quality standards and other measures of regulatory activity. I believe these findings occur because studies typically use cross-sectional data (see Duffy-Deno [1992] as an exception, looking at SMSA aggregate employment patterns). Cross-sectionally, high levels of industry activity do occur predominantly in non-attainment areas. Otherwise these areas would not be in non-attainment. By use of panel data I can control for this basic positive correlation, and show, for example, how a change in status affects location decisions. For example a change from non-attainment to attainment status under my hypothesis should subsequently induce new polluters into the county.

To understand that this is a real problem it is helpful to look at the raw data. The time period I look at covers 1978 to 1987 and the spatial coverage are the 742 urban counties (as of 1990) in the USA. These counties capture most industrial activity in the USA. In Table 1a I show the initial locational pattern of plants in each industry where in 1978 most plants in each industry are located in non-attainment areas. In Table 1b, I show the growth rate in the number of plants in "clean" counties which are in attainment in both 1978 and 1987 (column 1) and compare it with

Table 1a

***1978 Stock of Plants***

| | SIC | Percent of All Firms Located in Non-Attainment Counties |
|---|---|---|
| Plastic Materials & Synthetics | 282 | 91 |
| Industrial Organic Chemicals | 286 | 89 |
| Miscellaneous Plastics | 307 | 87 |
| Steel | 331 | 92 |
| Petroleum Refining | 291 | 92 |

% of all urban counties with 1978 non-attainment status = 60%.

Table 1b

***Total Percentage Change in Number of Plants 1978-1987 in Counties by Ozone Attainment Status***

| SIC | Attainment in both in '78 and '87 | Non-Attainment in both '78 and '87 | Non-Attainment in '78 Attainment in '87 |
|---|---|---|---|
| Plast. Mat'l./Synthetics (282) | 67 | 14 | 21 |
| Ind. Organic Chemicals (286) | 19 | 8 | 36 |
| Misc. Plastics (307) | 69 | 20 | 38 |
| Primary Steel (331) | 28 | 2 | 6 |
| Petroleum Refining (291) | 15 | 6 | −5 |
| employ growth in counties | 32 | 34 | |

the growth rate in the number of plants in "dirty" counties which are in non-attainment in both 1978 and 1987. Despite similar overall employment growth rates in these counties in those ten years, for all five industries the growth rate of number of plants is much higher in the attainment counties. In column 3, we also show the growth rates for counties which changed status – going from non-attainment to attainment – which accounts for 85% of status changes in our sample for ozone. In four out of five cases, improved counties also have higher growth rates than non-attainment counties.

This evidence is very suggestive of the relocation hypothesis asserted in the paper. However as already suggested, other things could be going on. For example, despite similar overall employment growth rates, non-attainment counties have much higher employment levels. Our VOC emitters could simply be decentralizing, to less dense urban locations. To sort this out requires an econometric specification.

In section 2 of the paper I specify an econometric model of industrial location, discuss estimation issues and present the data. In section 3 the results are presented and analyzed. Section 4 examines other types of evidence on the effect of regulation on industrial location.

# 2 An Econometric Model of Industrial Location

There is a substantial empirical literature on firm location decisions (see CARLTON [1983] and HERZOG and SCHLOTTMAN [1991] for a review). To answer the question particular to this paper, I adapt a specific econometric model for cross-sectional data presented in HENDERSON [1994a] to the current context which will use panel data.

In Henderson, plants are found in a particular geographic area if some criterion function is satisfied. In particular, if one or more plants in an industry can earn per plant profits $\Pi_i = \beta Z_i + \mu_i$ such that $\Pi_i$ exceeds a critical level of profits required by entrepreneurs to operate in that industry in that county, $\bar{\Pi}_i$, then one or more plants in the industry will be found in the county. That is, plants in that industry are found in the county if $\Pi_i > \bar{\Pi}_i$. $Z_i$ are arguments affecting per plant profits such as wages, local market scale and characteristics of the local industrial base which reflect the degree of positive or negative externalities in production. In principle, $Z_i$ could contain historical (lagged) as well as current measures to allow for dynamic externalities.

The critical level of profits required by entrepreneurs, $\bar{\Pi}_i$, can be treated as a random variable, or as a functional relationship affected by, say, local costs-of-living (for the entrepreneur and her family) and other local amenities. HENDERSON [1994a] combines this discrete choice criterion, with a non-linear continuous choice equation depicting the number of plants in the county for counties with plants and with a relationship allowing for equilibrating error. The functional relationships allow estimation of a structural form model, of the profit function, the entrepreneur's supply function and the complex error structure.

The objective here is more modest and makes use of the panel nature of the data. For 1977-1987, from County Business Patterns, I know the number of plants in each county in the major 3-digit VOC polluting industries, as well as characteristics of the local industrial situation such as wages, metro area employment, and industrial composition for any year, and for prior years back to 1977. At the 3-digit level, typically more than half of the 742 urban counties in the U.S. report zero plants in all years and typically less than 25% have positive plants in all years, so about 25% of counties experience entry and exit of the industry over the ten year period. By exit and entry, I mean respectively whether the industry leaves the county entirely in a given year and whether the industry goes from no plants to positive plants in the industry in a given year.

In this paper I focus primarily on the issue of exit and entry for reasons I will discuss below. To model exit and entry, I specify a criterion function where plants are located in county $i$ in time period $t$ if

$$\Pi_{it} = \tilde{\beta} \tilde{X}_{it} + \tilde{u}_i + e_{it} > \bar{\Pi}_{it} \tag{1}$$

$\Pi_{it}$ is the per firm profits for a firm in the industry of concern if it locates in county $i$ in period $t$, as a function of county time variant characteristics

$\tilde{X}_{it}$ and time invariant unmeasured characteristics, $\tilde{u}_i$, which is treated as a fixed/random effect. $e_{it}$ is a contemporaneous error term. $\bar{\Pi}_{it}$ is the profits required by entrepreneurs in time $t$. If $\bar{\Pi}_{it}$ is treated as a contemporaneous random variable, we combine it with $e_{it}$ so that $\beta X_{it}$ represents a structural part of the profit function. Otherwise we can specify $\bar{\Pi}_{it} = \alpha \tilde{Z}_{it} + \delta_{it}$, where $\tilde{Z}_{it}$ might depict local amenities (for the entrepreneur's family) and combine the LHS and RHS of the inequality in (1) to obtain a reduced form specification. In both cases the criterion for whether county $i$ has plants in the industry located there in time period $t$ is that $\overset{*}{y}_{it} \equiv \beta X_{it} + u_i + \varepsilon_{it} > 0$ with an indicator function for the presence of the industry in the county

$$
\text{(2)} \qquad \begin{aligned} y_{it} &= 1, \quad \text{if} \quad \overset{*}{y}_{it} = \beta X_{it} + u_i + \varepsilon_{it} > 0 \\ &= 0, \quad \text{otherwise.} \end{aligned}
$$

The formulation in (1) can be adjusted to model the continuous choice of number of plants in the county, conditional upon the industry being there. To do that requires making the number of plants (due to localized external economies of industry scale) an argument in $\Pi_{it}$, equating $\Pi_{it}$ and $\bar{\Pi}_{it}$, and then solving for the number of plants. Later we will footnote results on estimating an equation for the number of plants, conditional on the county having the industry. Those GMM estimates will suffer from selectivity bias (since the estimation is not joint with (1)) and the problem that the number of plants itself is not continuous number but is an integer typically taking values under four.

Returning to the discrete choice problem of whether the county has the industry or not, from equation (2) we have the $\text{Prob}\,(y_{it} = 1) = \text{Prob}\,(\varepsilon_{it} > -\beta X_{it} - u_i) = 1 - F\,(-\beta X_{it} - u_i)$. For a logistic distribution

$$
\text{(3)} \qquad \text{Prob}\,(y_{it} = 1) = \frac{\exp\,(\beta X_{it} + u_i)}{1 + \exp\,(\beta X_{it} + u_i)}.
$$

If we treat the $u_i$ fixed effects in equation (3) as nuisance parameters to be estimated, estimates of $\beta$ are not consistent in short panels (although for $T$ approaching even ten the bias may be quite small (HECKMAN, [1981]). Also note that for counties where the industry never appears or always appears, the $u_i$'s are not identified (are $-\infty$ and $+\infty$ respectively). CHAMBERLAIN'S [1980] fixed effect logit is a resolution to these problems since it eliminates the fixed effects in estimation (although they then become non-recoverable). For counties where $0 < \sum_{y_{it}}^{T} < T$, Chamberlain looks at the joint probability of $Y_i$, the time vector of $y_{it}$'s for a particular county $i$ [1]. He then conditions on the minimum sufficient statistic for $u_i$, $\sum y_{it}$; and maximizes a likelihood

1. The joint probability is

$$
\text{Prob}\,(Y_i) = \frac{\exp\left\{u_i\left(\sum_{t=1}^{T} y_{it}\right) + \beta \sum_{t=1}^{T} X_{it}\, y_{it}\right\}}{\prod_{t=1}^{T} [1 + \exp\,(\beta X_{it} + u_i)]}
$$

function, where the conditional probability for each county of the sequence $Y_i$ is

$$\text{(4)} \qquad \text{Prob}\left(Y_i \middle| \sum_{t=1}^{T} y_{it}\right) = \frac{\exp\left(\beta \sum_{t=1}^{T} X_{it}\, y_{it}\right)}{\sum_{d \in C_i} \exp\left(\beta \sum_{t=1}^{T} X_{it}\, d_t\right)}$$

where $C_i = \{(d_1, \ldots, d_T) | d_t = 0 \text{ or } 1 \text{ and } \sum d_t = \sum y_{it})\}$. For $T = 2$, the probability of the sequence (0, 1) is simply $\exp[\beta(X_{i2} - X_{i1})]/(1 + \exp\beta(X_{i2} - X_{i1}))$.

As with a regular fixed effects model we are looking to time differences in the $X_{it}$ variables to explain entry and exit from a county. Here the data on exit and entry are very noisy. Plants with either no employees or just 1 or 2 employees routinely drop in and out of some counties so that up to 20% of the sequences examined in the estimation involve an industry appearing in just for one year with minimal employment. Moreover, in any sequence of industry exit and entry, there is no way of knowing if the same plant(s) is involved each year. That is, I don't know plant or firm exit and entry, only industry exit and entry. I did clean the data up a little by imposing two restrictions. If a county records positive employment in all years but zero plants in any year, I record it as having positive plants in all years. Second, if a county has 1 plant and only 0-2 employees (mostly 0 or 1) in a year and has zero plants the years both before and after, I record it as zero plants (*i.e.* a non-entrant). While this cleaning tends to mostly sharpen results (lower standard errors in basic estimation) it has virtually no impact on coefficients. I would add that I could sharpen the results even more with further cleaning [2].

## 2.1. Data and Estimation Issues

As noted earlier, the basic economic data are from County Business Patterns 1977-1987, giving the number of plants in each 3-digit industry in each county each year. My focus is on environmental regulation, but I must control for county economic conditions. From the CBP data a variety of control variables such as wages, market scale (county and MSA total employment and manufacturing employment), and diversity indices can be calculated as reported in HENDERSON [1994b]. Our specification will be quite parsimonious given the stark nature of the discrete choice. In preliminary estimation, the only control variables with consistent impacts were market scale measures.

The key regulatory variables involve the county's attainment status in various air quality dimensions – ozone (03), sulfur oxides ($SO_X$), carbon

2. For example, I could impose the restriction that, if employment is greater than or equal to, say, 5 in a year but plants are zero, then the county is recorded as having the industry. Or if there is a long sequence of plants (but less than 10 years) with large employment (say, over 100) interrupted with a year of zero plants (and zero employment) –*i.e.* probably missing values– I could impose the restriction that the year have positive plants.

monoxide (CO), nitrogen oxides ($NO_X$), and total suspended particulates (TSP). For 1978 to 1987, attainment status may involve attainment of the primary standard, non-attainment of the primary standard, partial attainment of the primary standard, or non-attainment of a secondary (weaker) standard (but attainment of the primary standard). For ozone there is no secondary standard and "partial attainment" in our sample applies almost exclusively to Los Angeles and surrounding counties, which are the worst ozone counties in the U.S. Therefore for ozone, I construct a dichotomous variable with value 0 if the county is in attainment of the primary standard, and 1 otherwise. Given the Chamberlain logit, action in estimation will come from counties which change status during the panel time period (1978-87). For the roughly 170 counties (sequences of 0's and 1's) which are part of the entry-exit group forming the likelihood function, there are 50-60 (depending on the industry) status changes over ten years.

While I am focusing on industries responsible for VOC emissions, these industries may pollute in other air quality dimensions and may be affected by the attainment status in those dimensions. These dimensions, unlike ozone, both experience very few changes over the ten years and such changes typically involve a switch in or out of partial attainment or secondary non-attainment status. For these dimensions ($SO_X$, CO, TSP), I construct an index (omitting $NO_X$ as having no status changes). For each dimension for each county a zero is given if the county is in attainment of the primary standard, a 1/2 if the county is in either partial attainment of the primary standard or non-attainment of the secondary standard (only), and a one if the county is in non-attainment of the primary standard. I then sum these numbers across the three air quality dimensions to get an index of county compliance in these other air quality dimensions, representing the force of regulation in other dimensions. A county has an index between 0 ("cleanest" and hence least regulated) and 3 ("dirtiest" and most regulated).

The major econometric issue in the estimation of the model, where the elements of the L.L.F. are based on equation (4), concerns correlation between the remaining part of the error term, the contemporaneous component $\varepsilon_{it}$, and the $X_{it}$ explanatory variables. For a fixed effects model (with short $T$), the $\varepsilon_{it}$ must be independent of not just the current $X_{it}$, but all $X_{is}$, $S = 1, \ldots, T$ (even more so for the Chamberlain logit, since the likelihood event for each county is the sequence $Y_i$). From other work (HENDERSON [1994c]), entry or exit of a polluting industry to a county will significantly affect current ozone readings. That, in turn, may subsequently change county attainment status. So a contemporaneous industry specific county shock today may be correlated with future attainment status. Ordinary logit results will be biased in a predictable fashion. A higher $\varepsilon_{is}$ is associated with both a higher chance of entry and a higher chance of non-attainment. We expect the non-attainment coefficient to be negative, but this simultaneity problem is likely to bias it towards being positive. Thus it is desirable to instrument for the ozone status variable. This is done by using predicted values from a first stage logit (see below) for ozone attainment status, in the second stage logit estimation.

For other variables such as wages and county or metro area employment levels, informal analyses indicated no problems of endogeneity. The suggestion then is that the $\varepsilon_{it}$ is an industry specific shock at the county level. Any effects on the industry of general county wide shocks only come indirectly through, say, the effect on total county employment which is a measure of local demand for plant output in the industry. For example, for the panel of counties which always have plants from 1978 to 1987, Hausman tests based on GMM estimates of continuous equations (see footnote 4 below) indicated for all five industries, that treating the $X_{it}$ as all strictly exogenous to all years produces the same estimates as treating the $X_{it}$ as only predetermined (*i.e.*, $X_{is}$ for $s < t$ only are unaffected by $\varepsilon_{it}$). I also conducted SMITH and BLUNDELL [1986] type experiments where residuals from first stage regressions of the $X_{it}$ on a list of instruments are added to the logit equation. The coefficients of such residuals should provide a measure of potential correlation between the $X_{is}$ and the $\varepsilon_{is}$. For the variables I use in the estimation below no correlation was found. So the only $X_{it}$ I treat as endogenous are attainment status variables.

For the first stage estimation, logit equations for ozone attainment status are estimated separately for each year, using as explanatory variables all exogenous variables for all years (cf. PITT and ROSENZWEIG, [1990]). The list of instruments includes county land area, a coastal location dummy, 1970 county population as time invariant variables. For time variant variables, state fuel prices, metro area manufacturing employment (in all other than the own industry) and county and metro area total employment (in all other than the own industry) figures for 1977-1987 are used in each year's equation as instruments. The first stage results are good for every year and over time.

## 2.2. Implementation

Ordinary and 2-stage Chamberlain logits are estimated for each industry. As noted earlier, the data on industry exit and entry to counties is very noisy and the Chamberlain logit (discrete choice on effectively first differenced variables) is stark and demanding of the data. The sample sizes for sequences (number of counties where $0 < \sum y_{it} < 10$) averages about 170 per industry. I was unable to capture a sophisticated lag structure as in HENDERSON [1994b] for the $X_{it}$ or even to represent typical $X_{it}$ such as wages and diversity indices –such measures produced inconsistent, insignificant and sometimes perverse outcomes. The basic control variable is a scale measure of local demand for products, where either metro area or county total employment (in all industries except the own 3-digit one) produced similar results. Thus except for the scale measure, a time dummy and environmental regulatory variables, all other effects are represented in the fixed/random component of the error term (which is effectively differenced out) or the contemporaneous $\varepsilon_{it}$.

Controlling for time, the year across counties, may be important. For example if the number of plants is growing nationally over time and ozone attainment status is improving over time, we don't want positive time effects attributed to ozone status changes. A similar comment applies to business cycle issues. I have a limited number of sequences and entering 10 years of time dummies for some industries swamps the results. So I was a little more discriminating in use of time dummies. For only two of the industries is the number of plants growing over the 10 year period –miscellaneous plastics (SIC 307) and plastic materials and synthetics (SIC 282). In addition for business cycles (expansion), 1983 appeared to be the critical year. For all industries a time dummy for 1983 is in every estimating equation. For SIC 282 and 307, I will also report the effects of including a full set of time dummies.

# 3 Results

The basic set of results is given in Table 2. Apart from a control for the year 1983 (not reported) and for county employment lagged one period, the Table reports on the effect of current ozone attainment status. The 1983 time effect is always positive and significant, as is the county employment effect. Ozone attainment status relevant for an industry in year $t$ is listed as the status recorded in July of that year, rather than lagged status. The presumption is that plants know well in advance (based on the previous year's recorded ozone readings) what this year's status will be and react to that.

In ordinary logit results, the coefficient on non-attainment status is only negative and reasonably significant for two of the industries –industrial organic chemicals and miscellaneous plastics. Recall I expect a negative coefficient, since a switch from attainment (dummy variable has value zero) to non-attainment (dummy variable has value one) should cause exit (and vice versa for entry).

The 2-stage results are much more persuasive. In all cases, as anticipated the non-attainment status coefficient becomes more negative. For all industries except petroleum refining, the coefficient is large and negative. Since fixed effects and constant terms are differenced out in the Chamberlain logit, we can't recover predicted probabilities. However we can start at an arbitrary base and evaluate, from that base, the impact of a change in attainment status.

Specifically, for each industry for the base, I take a county in non-attainment status and assign it a base probability of having the industry in year $t$ of 0.25. I then calculate the positive effect of moving to attainment status [3]. The results are reported in Table 3. The effects are very large. Probabilities rise from 0.25 to anywhere from 0.39 to 0.71. This would suggest that attainment status is a key variable in firm location decisions.

I also experimented with the index (0-3) of attainment for each county in other air quality dimensions. Except for plastic materials (SIC 282),

3. The calculation uses equation (3). If the base probability is 0.25, then $\exp(\beta X + u) = .33$, and $\beta X + u \approx 1.1$. To calculate the effect of a status improvement, we add (– 1 times) the status coefficient in Table 2 to $-1.1$ and recalculate the probability.

TABLE 2

***Chamberlain Logit Results*** [1] ***(standard errors in parentheses)***

| | Industrial Organic Chemicals (SIC 286) | | Plastic Materials and Synthetics (SIC 282) | | Petroleum Refining (SIC 291) | | Misc. Plastics (SIC 307) | | Primary Steel (SIC 331) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Ordinary | 2-Stage | Ordinary | 2-Stage | Ordinary | 2-Stage | Ordinary | 2-Stage | Ordinary | 2-Stage |
| ln (county employ. in all other industries) | 1.04** (.421) | .735* | .879** (.404) | .757* | 1.32** (.487) | 1.14** | .507* (.302) | .461 | 1.26** (.397) | 1.10** |
| County Non-attainment status: | | | | | | | | | | |
| Actual | −.430* (.253) | | .154 (.228) | | −.316 (.292) | | −.625** (.285) | | −.183 (.237) | |
| Instrumented | | −2.03** | | −.860** | | −.064 | | −1.76** | | −.650* |
| Number of sequences $0 < \sum y_{it} < 10$ | 164 | 164 | 192 | 192 | 143 | 143 | 155 | 155 | 210 | 210 |

* Significant at .05 level; but 2-stage standard errors are incorrect (and not reported).

** Significant at .10 level; but 2-stage standard errors are incorrect (and not reported).

[1] To aid in convergence, county employment figures are scaled.

this index in ordinary Chamberlain logit results has no impact. For plastic materials it has a significant negative coefficient of $-.96$. However to the extent that entry of the own industry can affect later attainment status in other air quality dimensions, there is also a simultaneity issue with this variable. Readily instrumenting for this index is a problem, since it takes seven discrete values. I did experiment using a first stage Tobit (*i.e.*, treating the variable as continuous but truncated at zero), obtaining similar shaped distributions for actual and predicted values.

The coefficients for the instrumented value of this index in the Chamberlain logit all become strongly negative, except (again) for petroleum refining. For industrial organic chemicals, plastic materials and synthetics, miscellaneous plastics, and primary metals the coefficients are respectively $-.93$, $-1.93$, $-.59$, and $-.56$. The move from the index having a middle value such as 1.5 to a county being in complete attainment in all other air quality dimensions (index $= 0$) has an effect on the probability of the county attracting an industry similar to that in Table 3.

TABLE 3

***Effect on Probability of Industry Being in County of an Improvement in Ozone Attainment Status (starting probability of having the industry = 0.25)***

| | Ind. Organic Chemicals (286) | Misc. Plastics (307) | Primary Steel (331) | Plastic Mat. & Synthetics (282) |
|---|---|---|---|---|
| New probability with change to attainment status | 0.71 | 0.66 | 0.39 | 0.44 |

Finally, in terms of industries growing in the time period 1978-1987 which are plastics related, the insertion of a full set of time dummies for plastic materials (SIC 282) has little effect. However the effect on miscellaneous plastics is devastating. Time dummies simply swamp everything else.

# 4 Other Considerations

Table 1 suggests that, apart from the sample of exit and entry counties, for counties always having the industry ($\sum y_{it} = 10$) the number of plants in the county will respond to county attainment status. For the balanced panel of counties always having an industry, for each of the five industries, I estimated a GMM model of the number of plants in county $i$ in year $t$ [4]. Explanatory control variables included county employment, wage rates, and an Hirschman-Herfindahl index of county diversity across all 2-digit industries. Results were not sensitive as to whether these variables were treated as strictly exogenously or merely predetermined. Such estimation suffers from selectivity bias (*i.e.*, a full model could be a fixed effects Tobit, incorporating the exit and entry counties), and from a small numbers problem. Except for miscellaneous plastics (SIC 307), the number of plants in any county in the balanced panel is typically under four. This suggests the dependent variable should be properly modeled as a discrete dependent variable.

For the balanced panel of counties, in the GMM model where the number of plants is treated as a continuous dependent variable, the explanatory variable of focus is county attainment status for ground level ozone. That variable is treated as endogenous (but the estimation method doesn't explicitly recognize its dichotomous nature). For miscellaneous plastics the coefficient is negative and significant as expected, but the coefficient is small. A switch in status to non-attainment reduces the number of plants in the county by only 1.5%. For petroleum (SIC 291) and industrial organic chemicals (SIC 286) the coefficient for county attainment status is negative and significant at the .10 level. There the effect is larger. A switch in status to non-attainment reduces the number of plants in petroleum by 12.5% and in industrial organic chemicals by 6%. Effects in steel (331) and plastic materials (282) were not significant at any reasonable level.

Apart from the number of plants, environmental regulation may also affect plant sizes. This could occur for two inter-related reasons. Smaller plants with lower absolute emissions attract less regulatory attention, in the way the EPA and state regulatory agencies select "audits". Second, as a transition phenomenon, as firms move plants to attainment areas, they may continue to operate older plants in non-attainment areas, but perhaps at a reduced activity level. Data for industrial organic chemicals (SIC 286) strongly suggested this possibility. There average plant size nationally declined continuously

4. In the continuous version of the model applied to counties where $\sum y_{it} = 10$, the basic equation is $F_{it} = \beta X_{it} + u_i + \varepsilon_{it}$, where $F_{it}$ is the number of plants (in logs) in a county $i$ in time period $t$. This equation is first differenced to remove the $u_i$ so the estimating equation is $\Delta F_{it} = \beta \Delta X_{it} + \Delta\varepsilon_{it}$. This model is estimated by GMM where each year (1979-87) is treated as separate equation with cross-equation constraints on coefficients, with serial correlation accounted for (as in generalized 3-stage least squares in HAYASHI [1992]). To test for endogeneity of the $X_{it}$, the model can be estimated treating $X_{it}$ as strictly exogenous (to all $\varepsilon_{is}$) versus treating the $X_{it}$ as predetermined (only $X_{is}$ $s < t-1$, given first differencing are used as instruments for year $t$). A Hausman test for wages, MSA employment, and a diversity index could not reject the hypothesis of strict exogeneity. Moreover these coefficients seemed insensitive to specification. Only the regulatory variables were affected. The main issue with the model is selectivity bias in defining the balanced panel and the fact that except for industry 307 where the number of firms can range into the hundreds, the integer number of firms in a county is typically 1-3 (although the maximum values are typically 40-50). Treating the number of firms as a continuous variable may be then problematical. Not surprisingly the sharpest results are for industry 307.

from 1978 to 1987, starting at 185 employees per plant and ending at 137. Of course this decline could occur for other reasons.

To test empirically for the possibility that regulation affected plant sizes, I modeled the change in the log of average plant size in a county over time, for each industry. That is, I examined average plant size growth rates over the period 1978-1987, where size is measured by per plant employees. Annual growth rates of plant sizes rather than plant sizes themselves are used so any fixed/random effects are differenced out. To control for changes in technology, as explanatory variables I include a full set of time dummies for each industry. For all industries these time dummy are almost always negative. While that may represent technology changes, it could also represent regulation. Plants are down sizing over time to try to avoid regulatory scrutiny.

To control for economic conditions I included the change in the log of average wages in the county (in all industries except the own industry), as calculated from the total wage compensation and total employee figures for each county in County Business Patterns. These wage coefficients are negative and significant (at a .05 level for all industries except steel where significance is at the .10 level). They suggest a 1% increase in wages leads to at .40 to 1.6% decline in average plant size, the anticipated wage effect.

The variable of focus is county attainment status. The question is how to include it in the regression on plant size growth rates. In the time interval 1978-1987, I expect plant sizes to shrink in non-attainment counties relative to plant sizes in attainment counties. Reducing regulatory attention by reducing size is more important in non-attainment counties.

For all industries, the (level) indicator variable for county ozone attainment status has a negative sign, as does the index of attainment status in other air quality dimensions. But only for miscellaneous plastics (SIC 307) are any of the attainment status variables significant at any reasonable level, and then only at a .10 level. For miscellaneous plastics, being in non-attainment reduces plant size by 2.5%. However it is not clear that modeling plant growth rates as a function of the level indicator of attainment status is appropriate. Rather we could focus on the impact of changes in status to sharpen the results, and to be consistent with the methodology in the rest of the paper. Changes in status (the first differenced attainment status variable) produced only one significant coefficient –again for miscellaneous plastics (at a 0.10 level). For miscellaneous plastics a change from attainment to non-attainment reduces average plant size by 15%. That's a fairly large effect.

# 5 Conclusion

Air quality regulation induces plants to leave non-attainment ("dirty") counties and move to attainment ("clean") counties. That may create a graying effect, where relocations help improve air quality in non-attainment areas, but worsen it in attainment areas.

Apart from affecting plant location decisions, air quality regulation may cause firms to reduce plant sizes (so as to attract less EPA attention).

# • References


ATKINSON, S., TIETENBERG T. (1987). – "Economic Implications of Emissions Trading Rules for Local and Regional Pollutants", *Canadian Journal of Economics*, 20, pp. 370-386.

BARTIK, T. J. (1988). – "The Effects of Environmental Regulation on Business Location in U.S.", *Growth and Change*, 19, pp. 22-44.

CARLTON, D. (1983). – "The Location and Employment of New Firms", *Review of Economics and Statistics*, 65, pp. 446-449.

CHAMBERLAIN, G. (1980). – "Analysis of Covariance with Qualitative Data", *Review of Economic Studies*, 47, pp. 225-238.

DEILY, M.E., GRAY W. B. (1991). – "Enforcement of Regulations in a Declining Industry", *Journal of Environment Economics and Management*, 21, pp. 260-274.

DUFFY-DENO, K. T. (1992). – "Pollution Abatement Expenditures and Regional Manufacturing Activity", *Journal of Regional Science*, 32, pp. 419-436.

HAYASHI, F. (1992). – "Comment", *Journal of Business of Economic Statistics*, 10, pp. 15-17.

HECKMAN, J. J. (1981). – "The Incidental Parameters Problem" in *Structural Analysis of Discrete Data with Econometric Applications*, C.F. Manski and D. McFadden (eds.), pp. 179-95. MIT Press.

HENDERSON, J. V. (1994a). – "Where Does an Industry Locate?", *Journal of Urban Economics*.

HENDERSON, J. V. (1994b). – "Externalities and Industrial Development", *NBER Working Paper #4730.*

HENDERSON, J. V. (1994c). – "The Impact of Air Quality Regulation on Ozone", *mimeo.*

HERZOG, H.W. Jr., SCHLOTTMANN A. M. (1991). – "Industrial Location and Public Policy", *University of Tennessee Press.*

MCCONNELL, V., SCHWAB R. (1990). – "Industry Location Decisions", *Land Economics*, 66, pp. 57-81.

PITT, M.M., ROSENZWEIG M. R. (1990). – "Estimating the Intrahousehold Incidence of Illness", *International Economic Review*, 31, pp. 969-989.

PORTNEY, P.R. (1990a). – "Policy Watch: Economics and the Clean Air Act", *Journal of Economic Perspectives*, 4, pp. 173-182.

ROUMASSET, J.A., SMITH K.R. (1990). – "Exposure Trading", *Journal of Environmental Economics and Management*, 18, pp. 276-291.

RUSSEL, C.S. (1990). – "Monitoring and Enforcement" in P. R. Portney *et. al.* (ed.) *Public Policy for Environmental Protection*, RFF Washington.

SMITH, R.J., BLUNDELL R.W. (1986). – "An Exogeneity Test for a Simultaneous Equation Tobit Model", *Econometrica*, 54, pp. 697-683.