# Effects of environmental regulation on foreign and domestic plant births: is there a home field advantage?

John A. List [a,b], W. Warren McHone [c], Daniel L. Millimet [d,*]

[a] *The University of Maryland, Department of Agricultural and Resource Economics, 2200 Symons Hall, College Park, MD 20742-5535, USA*
[b] *NBER, Cambridge, MA, USA*
[c] *University of Central Florida, College of Business Administration, Department of Economics, Orlando, FL 32816-1400, USA*
[d] *Southern Methodist University, Department of Economics, Box 0496, Dallas, TX 75275-0496, USA*

## Abstract

Whether environmental regulations alter capital flows remains a hotly debated issue. This paper uses county-level data to examine the location decisions of domestic and foreign firms in a single empirical model and tests for asymmetries by firm origin in the degree to which capital flows are influenced by environmental standards. We find that while domestic firms are influenced by environmental regulations, foreign firms are not. Since the benefits of foreign investment are well-documented—foreign plants typically provide more jobs and increase local wages by more than domestic plants—this result suggests a double-dividend is available: foreign plants provide an economic stimulus and are not unduly influenced by environmental protections.
© 2004 Elsevier Inc. All rights reserved.

---

* Corresponding author.
  *E-mail address:* millimet@mail.smu.edu (D.L. Millimet).
  *URL:* http://faculty.smu.edu/millimet.

## 1. Introduction

The real costs associated with environmental regulations have been the subject of widespread debate since the beginning of the environmental movement in the early 1970s. A well-documented component of these costs is the billions of dollars spent annually by manufacturing plants to comply with environmental regulations.[1] Opponents of stricter environmental standards contend that more stringent environmental regulations impose additional economic costs by distorting the spatial pattern of economic development— inducing some regions to be at a competitive disadvantage when vying for new plants and jobs. Although this concern is not universally accepted, some recent studies suggest that regions with relatively stringent air pollution control attract fewer domestic plants (see Jeppessen et al. [24] for a review). Similarly, a few recent analyses find moderate evidence that environmental regulations distort location patterns of foreign direct investment (FDI) inflows into the US (e.g., Keller and Levinson [25], Fredriksson et al. [13]), as well as FDI outflows from the US (Cole and Elliot [5]).

Despite these recent studies, to date no empirical analysis has been conducted that allows one to compare the sensitivity of domestic versus foreign plant location decisions to variation in local environmental stringency. Several regulatory features in the US represent potential reasons why one could observe heterogeneous effects. First, new plants entering 'dirty' counties must offset any increase in emissions by a reduction in emissions from another source within the same air basin. Since foreign plants might be expected to have fewer within-plant offset possibilities, their start-up costs are potentially much larger. Second, US environmental standards and regulatory prescriptions differ substantially from those in foreign countries, forcing foreign plants to undertake substantial re-engineering of their production processes to incorporate the pollution abatement capital equipment that is mandated (e.g., Jaffe et al. [23], King and Shaver [26]). Second, regulatory oversight may favor domestic plants relative to foreign entrants as the result of internal political pressure or lack of information. Vernon [39] suggests that because regulators, business leaders, and community leaders tend to have less information about foreign plants, they are encouraged to investigate, audit, and prosecute foreign-owned facilities more frequently. King and Shaver [26] extend this argument to the environmental arena. They argue that foreign-owned facilities will have more difficulty complying with environmental regulations because of their lack of familiarity with local laws, technical requirements, and business conditions. Furthermore, King and Shaver [26] argue that foreign-owned plants will be held to a higher environmental regulatory standard for all the reasons outlined in Vernon [39].

On the other hand, given that foreign plants not only create additional jobs, but also increase overall local wages by approximately 3–7 times that of domestic plants (e.g., Howenstine and Zeile [22], Figlio and Blonigen [11]), local policymakers may be more sympathetic to new foreign capital. Finally, uncertainty regarding future stringency levels may inhibit aggressive foreign investment.[2] Furthermore, foreign-owned firms may simply

---

[1] In the 1993 *Pollution Abatement Costs and Expenditures Survey*, the US Census Bureau reports that pollution abatement expenditures amounted to almost $30 billion per year in the US alone.

[2] This argument is in the spirit of the asymmetric information model of Gordon and Bovenberg [14].

be at an informational disadvantage relative to domestic firms and, therefore, unwittingly locate in areas with more stringent environmental regulation. Finally, as documented in Figueiredo et al. [12], firms locating away from their 'home' residence (in Portugal) tend to congregate in major urban centers to offset informational disadvantages, suggesting that the location choice of foreign-owned firms may be less sensitive to variation in environmental costs.[3] The authors conclude [12, p. 358]:

> Agglomeration, especially urbanization economies, can apparently help offset the information costs and other disadvantages associated with non-home investment decisions. Localization economies influence both home and non-home investors, but have more impact on non-home investors... Agglomeration compensates for the lack of local knowledge about production and market conditions, leading firms to cluster in urban areas even as they expand geographically. This represents a case of informational cascading, where firms flock into cities imitating prior location decisions to economize on search costs.

These factors suggests that it is an open issue whether there is a 'home field advantage' for domestic manufacturing producers in the US. Recent empirical work has provided some indication of that the factors outlined above do contribute to heterogeneous responses by foreign and domestic plants' to environmental regulations. For example, in their meta-analysis, Jeppessen et al. [24] report some evidence suggesting that foreign plants investing in the US are more influenced by environmental regulations than their domestic counterparts. However, these results should be interpreted as merely suggestive, as the comparison is made across studies that include different empirical models, control variables, time periods, and sampling populations. In order to determine whether foreign plants are systematically influenced by environmental regulations, and simultaneously obtain information on the magnitudes of these effects in comparison to domestic entrants, it is necessary to use a comprehensive data set that includes observations on both foreign and domestic plants.

To accomplish this goal, we use a unique county-level database for New York State (NYS) from 1980 to 1990 to integrate the domestic and foreign new plant location decision into a single empirical framework. Using both a semi-nonparametric difference-in-differences propensity score matching estimator and a standard fixed effects Poisson framework, we are able to examine whether, and to what extent, local air quality regulations influence locational decisions among foreign- and domestic-owned plants without making comparisons across data sets and estimation frameworks.

The empirical results fall into two categories. First, consonant with List et al. [29], location decisions of new domestic pollution intensive plants are adversely affected by stringent air pollution regulations. According to the parametric Poisson estimator, the expected county-level flow of dirty domestic plants decreases by 44–61% when local air quality regulation becomes very stringent. However, the difference-in-differences

---

[3] Similar evidence is also provided in Dean et al. [7], who find that FDI into China originating from OECD or non-Chinese countries tends to locate in provinces with *higher* pollution levies, and Crozet et al. [6], who document the importance of agglomeration externalities in the location choice of FDI within France.

propensity score matching estimator, which removes potential time-varying, county-specific unobservables correlated with both the flow of new pollution intensive plants and local environmental standards, yields much larger estimates: stringently regulated counties forego 0.5 to 1.9 dirty domestic plants per annum, a loss of between 150 and 600% for the average county.

Second, in all parametric Poisson specifications, and every difference-in-differences propensity score matching specification except one, we fail to find a statistically significant impact of local environmental stringency on the location of foreign-owned pollution intensive plants. This finding is striking in three respects. First, it is contrary to our intuition that suggests that a 'home field advantage' may exist for domestic plants. Second, since foreign plants not only provide more jobs, but also increase overall local wages by approximately 3–7 times that of domestic plants (e.g., Howenstine and Zeile [22], Figlio and Blonigen [11]), these results suggest that a double-dividend is available to recruiting foreign plants: they provide an economic stimulus and are relatively uninfluenced by stricter environmental protections. Finally, to the extent that the results herein hold in other contexts and with other data sets, these findings suggest that empirical studies seeking to validate the 'pollution haven hypothesis' may be more likely to find support when analyzing intra-national flows of domestic-owned capital, rather than international flows of foreign-owned capital.

The remainder of the paper is organized as follows. Section 2 contains an overview of US air quality regulations and describes the data. Section 3 presents the empirical models. Section 4 describes the results. Section 5 provides concluding remarks.

## 2. US air quality regulation and the data

### 2.1. US air quality regulation

Regulations limiting emissions of hazardous air pollutants in the US date back to the mid-nineteenth century, as many municipalities issued smoke ordinances to combat rising ambient concentration levels. The decentralized nature of air pollution control continued until the early 1960s. Disappointed with the outcomes associated with decentralized control of the environment, federal authorities began to take a much more active role in environmental regulation with passage of the National Environmental Policy Act and the first Clean Air Act Amendments (CAAA) in 1970. As environmental authority was being shifted to the federal government, technological and political changes were making industry more geographically mobile, potentially intensifying and promoting inter-jurisdictional competition (Tannenwald [38]). Coupling these temporal changes in plant mobility with the confusion of how the states should implement the new regulations of the CAAA of 1970, the federal government passed important provisioning rules via the 1977 Clean Air Act Amendments.

The 1977 Clean Air Act Amendments stipulated that starting in 1978 every county in the US was to be designated annually as being in-attainment or out-of-attainment (non-attainment) of national air quality standards for each of five criteria air pollutants—carbon monoxide, sulfur dioxide, total suspended particulates, ozone, and nitrogen oxide (other

pollutants, such as particulate matters, have subsequently been added to the list). For counties not in-attainment of federal standards, each state is required to submit periodic comprehensive plans that will bring the county into attainment status in the near future (commonly termed State Implementation Plans (SIP)). If standards are not met in due time, states run the risk of construction bans, rigorous federal intervention, and losing federal monies that help to fund state-level public goods and services.

Complying with environmental regulations in non-attainment counties can impose extraordinary costs on new plants. Plants entering a county labeled out-of-attainment are subject to a standard of 'Lowest Achievable Emission Rate' (LAER) on equipment, *without* consideration of cost. These compliance expenditures potentially run into the millions of dollars and represent a significant start-up cost for entering plants. New plants entering attainment areas, on the other hand, face a more lax regulatory standard known as 'Prevention of Significant Deterioration' (PSD). PSD entails permitting and the installation of the 'Best Available Control Technology' (BACT) for new plants that have the potential to emit more than 100 tons of a criteria pollutant in a year. The BACT is negotiated on a case-by-case basis and the economic burden on the plant is taken into consideration in arriving at a final solution. New plants that are deemed to be low-level polluters are not subject to any specific technology standards in attainment areas.

As discussed more fully below, the annual air quality status of a particular county provides a natural measure to evaluate the impact of variation in local environmental regulation.[4] In addition, attainment status represents (to our knowledge) the only county-level environmental regulatory measure available in a panel data setting. Henderson [19], Greenstone [15], and List et al. [29], among others, have utilized attainment status to measure environmental stringency.

## 2.2. The data

The data on new plants, delineated by domestic or foreign ownership, come from the Industrial Migration File (IMF), maintained until 1990 by the NYS Department of Economic Development (DED), and encompass the period 1980–1990. The goal of the IMF was to monitor all annual changes in manufacturing activity in NYS at the county-level. The IMF reports observations of individual plant openings and classifies plants by standard industrial classification (SIC) code. The information in the IMF file was assembled from a variety of sources, including regional offices of the NYS Department of Commerce, local chambers of commerce, the NYS Department of Labor, newspapers, and private reporting sources. The DED regional offices verified all reported projects before including them in the file. This method of data collection has both advantages and disadvantages.

A major advantage is that case-specific information is not suppressed (Michalke [33]). Thus, there are opportunities for improving precision and data analysis not heretofore available even from the Longitudinal Research Database (McGuckin [32]). Another

---

[4] In this sense, this empirical approach is akin to a natural experiment (see, e.g., List [27], for use of a natural, or field experiment, in a much different setting).

distinct advantage is that the data are available at the four-digit SIC level. This conveniently allows an analysis of both pollution and non-pollution intensive plants. On the negative side, since no legislative statute exists in New York that requires plants to furnish information to the state, DED makes no claim that the IMF is comprehensive. In addition, the IMF excludes some plant activities involving either small investment activity (less than $100,000) and/or modest changes in employment (less than 25 employees). Nevertheless, comparisons with Census of Manufacturers data suggest that IMF coverage is extremely broad for all but the smallest size classes.

Following previous studies, we focus on measuring the effects of county-level ozone attainment status on plant locations in pollution intensive sectors most likely to be affected by county-level ozone attainment status.[5] Given that attainment status is determined by county-level air quality readings, we follow Greenstone [15] and classify industrial sectors as 'ozone pollution intensive' if they emit at least 6% of the total industrial sector's emissions of nitrogen oxide or volatile organic compounds, the primary chemical precursors to ozone. Plants labeled pollution intensive are in SIC codes: 2611-31 (Pulp and Paper Mills), 2711-89 (Printing and Binding), 2812-19 (Industrial Inorganic Chemicals), 2861-69 (Industrial Organic Chemicals), 2911 (Petroleum Refining), 30 (Rubber and Misc. Plastics), 32 (Stone, Clay, and Glass), 3312-3 (Steel and Electrometallurgical Products), 3321-5 (Iron and Steel Foundries), 34 (Fabricated Metal Products), and 371 (Motor Vehicles and Equipment). All remaining sectors are labeled non-pollution intensive.

After classifying plant modifications by pollution intensity, the data are aggregated to the county-level. In total, we observe nearly 900 new plant openings: 220 (513) pollution intensive (non-pollution intensive) domestic plant openings and 30 (119) pollution intensive (non-pollution intensive) foreign plant openings across the 62 counties in NYS over the sample period. We then combine this measure with each county's ozone attainment designation, and other county-level attributes (discussed below).[6] Over the sample period, slightly more than 25% of the county observations are in non-attainment.

Table 1 provides a summary of the distribution of plant births by pollution intensity and ownership. The total number of new plant openings shows a rather sharp downward trend over the decade. Referring to the first panel in Table 1, we observe that the average number of plant births per year was 92 during the first half of the decade, whereas the 1986–1990 period witnessed a decline to 67 new plants per year. Breaking down the plants by foreign and domestic ownership, we find that this decrease was due largely to reductions in domestic plant births. Annual births of foreign-owned plants were stable throughout the

---

[5] As discussed in Section 2.1, the 1977 Clean Air Act Amendments set standards on five criteria air pollutants, but ozone has attracted the most regulatory attention due to the limited progress that has been made to reduce concentration levels. Thus, we follow Henderson [19] and List et al. [29] and focus on county attainment status of ozone.

[6] Although attainment status can range from in-attainment of the primary standard to out-of-attainment, with partial standards in between, ozone designation has typically been polar in nature; that is, a county is either in- or out-of-attainment. For a county to be labeled out-of-attainment, its second highest daily air quality reading must exceed 0.12 parts per million. Of the 62 NYS counties, 26 (three) are in-attainment (out-of-attainment) throughout the sample period, while 33 counties experience both regimes. Of these 33 counties, 29 are out-of-attainment for a consecutive block of time (more than two years).

Table 1
Count of plant openings in New York State by year, industry type, and foreign or domestic ownership

| | Plant births in all industries | | | |
|---|---|---|---|---|
| Year | Domestic | Foreign | Total | % Foreign |
| 1980 | 94 | 6 | 100 | 6 |
| 1981 | 59 | 14 | 73 | 19 |
| 1982 | 82 | 11 | 93 | 12 |
| 1983 | 74 | 17 | 91 | 19 |
| 1984 | 86 | 21 | 107 | 20 |
| 1985 | 73 | 12 | 85 | 14 |
| 1986 | 61 | 15 | 76 | 20 |
| 1987 | 61 | 18 | 79 | 23 |
| 1988 | 52 | 10 | 62 | 16 |
| 1989 | 51 | 14 | 65 | 22 |
| 1990 | 40 | 11 | 51 | 22 |
| *Total* | 733 | 149 | 882 | 17.4 |
| | Plant births in non-pollution intensive industries | | | |
| Year | Domestic | Foreign | NPI total | % Foreign |
| 1980 | 68 | 4 | 72 | 6 |
| 1981 | 42 | 13 | 55 | 24 |
| 1982 | 58 | 9 | 67 | 13 |
| 1983 | 56 | 13 | 69 | 19 |
| 1984 | 61 | 13 | 74 | 18 |
| 1985 | 56 | 11 | 67 | 16 |
| 1986 | 45 | 14 | 59 | 24 |
| 1987 | 35 | 14 | 49 | 29 |
| 1988 | 33 | 8 | 41 | 20 |
| 1989 | 37 | 10 | 47 | 21 |
| 1990 | 22 | 10 | 32 | 31 |
| *Total* | 513 | 119 | 632 | 18.8 |
| | Plant births in pollution intensive industries | | | |
| Year | Domestic | Foreign | PI total | % Foreign |
| 1980 | 26 | 2 | 28 | 7 |
| 1981 | 17 | 1 | 18 | 6 |
| 1982 | 24 | 2 | 26 | 8 |
| 1983 | 18 | 4 | 22 | 18 |
| 1984 | 25 | 8 | 33 | 24 |
| 1985 | 17 | 1 | 18 | 6 |
| 1986 | 16 | 1 | 17 | 6 |
| 1987 | 26 | 4 | 30 | 13 |
| 1988 | 19 | 2 | 21 | 10 |
| 1989 | 14 | 4 | 18 | 22 |
| 1990 | 18 | 1 | 19 | 5 |
| *Total* | 220 | 30 | 250 | 12.0 |

decade. Consequently, as illustrated in the last column of this frame, foreign-owned plants represent an increasing share of new plant births in the last half of the decade for NYS.

A comparison of the patterns of new plant births in the non-pollution intensive (the second panel) and pollution intensive (the third panel) subgroups reveals that much of the

Table 2
Description of variables

| Variable | Mean | In-attainment Mean | Out-of-attainment Mean | Definition and source |
|---|---|---|---|---|
| Domestic new plants | | | | Actual count of new domestic plants from 1980 to 1990 labeled as having production activities that are pollution intensive. Industrial Migration File. New York State Department of Economic Development (DED). |
| *Pollution intensive* | 0.32 (0.79) | 0.24[†] (0.55) | 0.57[†] (1.20) | |
| *Non-pollution intensive* | 0.75 (1.82) | 0.45[†] (0.91) | 1.62[†] (3.09) | Domestic plants labeled as having production activities that are non-pollution intensive. |
| Foreign new plants | | | | Actual count of new foreign plants from 1980 to 1990 labeled as having production activities that are pollution intensive. Industrial Migration File. New York State Department of Economic Development (DED). |
| *Pollution intensive* | 0.04 (0.24) | 0.04 (0.24) | 0.06 (0.26) | |
| *Non-pollution intensive* | 0.17 (0.54) | 0.16 (0.53) | 0.20 (0.57) | Foreign plants labeled as having production activities that are non-pollution intensive. |
| Attainment status | 0.26 (0.44) | – | – | Intensity of county-level pollution regulations. Dichotomous variable = 1 if county is out-of-attainment of federal standards for ozone, 0 otherwise. Federal register Title 40 CFR Part 81.305. |
| Employment (100,000s) | 1.50 (3.57) | 0.08[†] (1.45) | 3.40[†] (6.22) | Total employment in manufacturing. *County Business Patterns.* |
| Wage (1000s) | 17.01 (4.1.2) | 17.34[†] (3.84) | 16.08[†] (4.74) | Total annual manufacturing payroll divided by the number of employees by county, adjusted for inflation. *County Business Patterns.* |
| Plants | 11.71 (5.39) | 12.23[†] (5.49) | 10.22[†] (4.84) | Total number of manufacturing plants. *County Business Patterns.* |
| Population (100,000s) | 2.87 (4.78) | 1.82[†] (3.40) | 5.87[†] (6.60) | County population. *Current Population Reports.* US Bureau of Census. |
| Property tax (1000s) | 0.56 (0.24) | 0.57 (0.27) | 0.54 (0.17) | Real property tax collected per capita. *Census of Governments.* |
| Per capita income (1000s) | 13.50 (4.76) | 13.58 (4.30) | 13.24 (5.88) | Real per capita income. *County Business Patterns.* |
| Highway expenditures | 0.17 (0.09) | 0.18[†] (0.10) | 0.15[†] (0.08) | Total highway expenditures. *County Business Patterns.* |

*Notes*. Data are for the 62 New York counties from 1980 to 1990. $N = 682$ (176 out-of-attainment). Standard deviations are in parentheses.

[†] Indicates difference in means across in- and out-of-attainment samples is statistically significant at $p < 0.05$ level using a two-sided $t$-test.

decline in total plant births noted between the first and second half of the decade was attributable to a sharp downward trend in non-pollution intensive industries. The average number of annual new plant births in these industries fell from 67 during the first half of the decade to 46 over the 1986–1990 period. By contrast, the average number of plant births

in pollution intensive industries only dropped from 24 to 21 per year between the first and second half of the decade.

The percentage of foreign-owned births in non-pollution intensive industries increased from an average of 16 to 25% between the first and last half of the decade. By contrast, the corresponding average for foreign-owned births in pollution intensive industries remained at 11%. Thus, by decade's end, a higher concentration of foreign-owned plants was in the non-pollution intensive sectors (relative to both the total foreign-owned plants and domestic-owned plants).

On the surface, the temporal composition of new plants in each of the broadly defined industry groups in Table 1 implies that environmental regulations could alter foreign factor flows in a different manner than domestic factor flows. Yet, the question that remains open to interpretation is whether the underlying pattern of new plant ownership remains after one controls for other important locational factors. To consider this question, we turn to the econometric framework. Descriptive statistics and definitions of all variables to be used are presented in Table 2.

Before proceeding to a discussion of the empirical models and results, we should note that we have used these data in several other studies. For example, List and McHone [28], List et al. [29], and Millimet and List [34] examine the influence of attainment status on plant births. Relatedly, List et al. [30] explore the effects of air quality regulation on the destination choice of relocating plants, while List et al. [31] compare and contrast the effects of environmental regulations on plant-level modification and closure rates to examine issues associated with that aspect of new source review. While there are similarities across these studies—empirical methodologies, regressors, etc.—the outcomes of interest in these studies and the current study are different. In this paper, we compare new plant births across country of origin to examine if foreign and domestic plants respond differently to environmental regulations.

## 3. The empirical models

### 3.1. Propensity score matching

To assess the impact of the attainment status on plant birth rates, we utilize the method of propensity score matching, developed in Rosenbaum and Rubin [36] and applied to the analysis of plant location in List et al. [29]. The method, while extensively used by statisticians, is fairly new to many economists. A few notable examples include Blundell et al. [2], Dehejia and Wahba [8,9], and Heckman et al. [18]. Blundell and Costa-Dias [1, p. 4] provide an excellent introduction to the matching method, concluding "matching methods have been extensively refined in the recent evaluation literature and are now a valuable part of the evaluation toolbox."

The fundamental problem in identifying the effect of a binary variable, such as attainment status, is one of incomplete information. While the econometrician observes whether the 'treatment' (in this case, non-attainment of the federal ozone standard) occurs and the outcome conditional on treatment assignment, the counterfactual is not observed. Let $y_{i1}$ denote the outcome of observation $i$ with the treatment ($T_i = 1$); $y_{i0}$ denotes

the outcome without the treatment ($T_i = 0$). If both states of the world were observable, the average treatment effect, $\tau$, would equal $\bar{y}_1 - \bar{y}_0$, where the former (latter) average represents the mean outcome for the treatment (control) group. However, given that only $y_1$ or $y_0$ is observed for each observation, unless assignment into the treatment group is random, generally $\tau \neq \bar{y}_1 - \bar{y}_0$.

The solution Rosenbaum and Rubin [36] advocate is to find a vector of covariates, $Z$, such that

$$y_1, y_0 \perp T \mid Z, \quad \mathrm{pr}(T = 1 \mid Z) \in (0, 1), \tag{1}$$

where $\perp$ denotes independence. If one is interested in estimating only the average treatment effect, the weaker condition

$$E[y_0 \mid T = 1, Z] = E[y_0 \mid T = 0, Z] = E[y_0 \mid Z], \quad \mathrm{pr}(T = 1 \mid Z) \in (0, 1) \tag{1'}$$

is required. For condition $(1')$ to hold, $Z$ should be multi-dimensional. Consequently, finding observations with identical values for all covariates in $Z$ may be untenable. However, Rosenbaum and Rubin [36] prove that conditioning on $p(Z)$ is equivalent to conditioning on $Z$, where $p(Z) = \mathrm{pr}(T = 1 \mid Z)$ is the propensity score. $p(Z)$ is estimated via logit.

After estimating the propensity score, a matching algorithm is required to estimate the missing counterfactual, $y_{0i}$, for each treated observation $i$. The simplest algorithm is (single) nearest-neighbor matching, whereby each treated observation is paired with the control observation whose propensity score is closest in absolute value (Dehejia and Wahba [8]).[7] As unmatched controls are discarded, the matching method identifies a sub-sample of controls that more closely resembles the treatment group in terms of pre-treatment attributes. The treatment effect on the treated (TT) is given by

$$\tau_{\mathrm{TT}} = E\big[y_1 \mid T = 1, p(Z)\big] - E\big[y_0 \mid T = 0, p(Z)\big] = E\big[y_1 - y_0 \mid p(Z)\big]. \tag{2}$$

During the actual estimation, the nearest-neighbor algorithm is amended along three dimensions. First, while a match exists for each out-of-attainment county, the propensity scores may still differ considerably. Because the unbiasedness of the matching estimator relies on the propensity scores being identical, pairs with scores significantly different are excluded. Known as caliper matching, we present estimates for several cut-off values, or calipers (Cochran and Rubin [4]).

Second, to fully utilize our panel data, we amend the caliper matching method by restricting the pool of potential controls to which a given treated observation may be paired. Specifically, we match each treated observation twice: first, restricting matched pairs to be from the same year and region of the state (see Fig. 1);[8] second, restricting matched pairs to be the same county from different years. By matching within-year/within-region, or within-county, we explicitly remove any time-, region-, or county-specific unobservables not already controlled for by the propensity score. This is the matching method's analogy

---

[7] Typically, nearest-neighbor matching is performed *with replacement*, implying that a given control observation may be matched with multiple treatment observations. Dehejia and Wahba [9] verify that matching with replacement fares at least as well as matching without replacement, and possibly better.

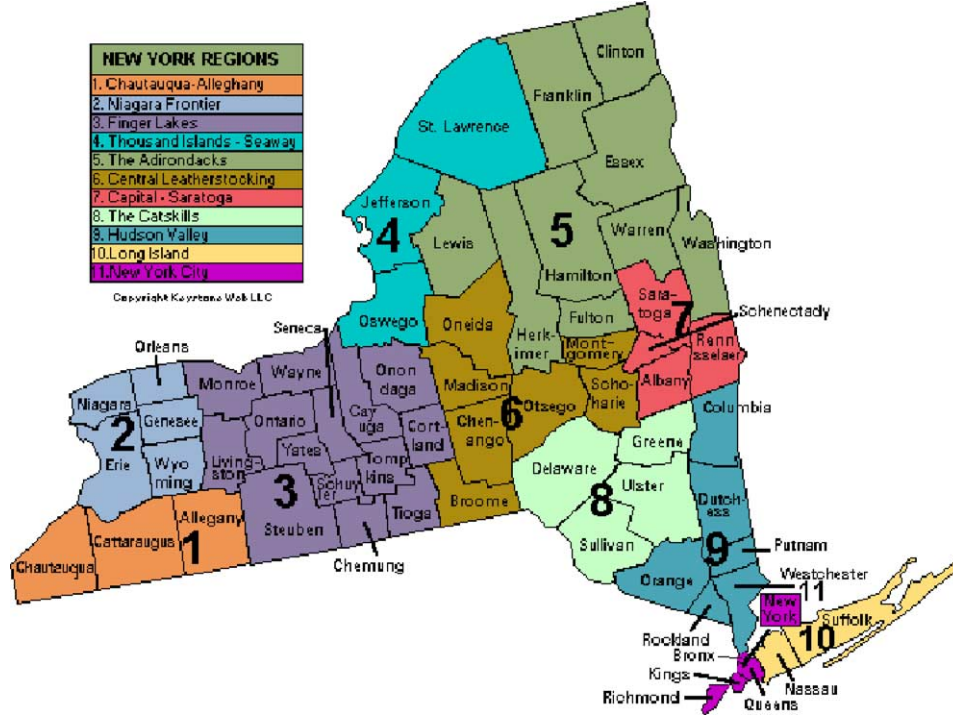[8] Eleven regions are used; see List et al. [29].

Fig. 1. Regional breakdown, New York State.

to fixed effects, and is similar in spirit to Smith and Todd's [37] claim that matches used to identify the effect of employment programs should be from the same local labor market. Thus, the estimator in (2) becomes

$$\tau_{\text{TT},tr} = E\big[y_1 \mid T = 1, p(Z), t, r\big] - E\big[y_0 \mid T = 0, p(Z), t, r\big]$$
$$= E\big[y_1 - y_0 \mid p(Z), t, r\big], \tag{3a}$$
$$\tau_{\text{TT},i} = E\big[y_1 \mid T = 1, p(Z), i\big] - E\big[y_0 \mid T = 0, p(Z), i\big]$$
$$= E\big[y_1 - y_0 \mid p(Z), i\big], \tag{3b}$$

where $t$ indexes year, $r$ indexes region, and $i$ indexes county.

Finally, as the usual matching estimators are applicable only to problems of 'selection on observables,' we amend the matching estimators in (3a), (3b) to explicitly remove certain unobservables by employing a difference-in-differences (DID) matching estimator. Since we have a count of both pollution intensive *and* non-pollution intensive plant openings for each county-year observation, and the behavior of 'clean' plants should not be affected by attainment status, any difference in the openings of 'clean' plants across the matched treatment and control groups is assumed to reflect unobservable county-specific qualities that serve to influence plant location. Thus, in the spirit of similar estimators used in Eichler and Lechner [10], Heckman et al. [18], and List et al. [29], we define the DID counterpart to (3a), (3b) as

$$\tau_{\text{DID},tr} = \tau_{\text{TT},tr} - \tau'_{\text{TT},tr}, \tag{4a}$$

$$\tau_{\text{DID},i} = \tau_{\text{TT},i} - \tau'_{\text{TT},i}, \tag{4b}$$

where $\tau'_{\text{TT},tr}$ ($\tau'_{\text{TT},i}$) is the mean difference in the count of 'clean' plant openings across the matched treatment and control groups. As the DID estimator only requires

$$E[y_0 - y'_0 \mid T = 1, Z] = E[y_0 - y'_0 \mid T = 0, Z] = E[y_0 - y'_0 \mid Z],$$
$$\text{pr}(T = 1 \mid Z) \in (0, 1) \tag{1''}$$

for identification, where $y'_0$ is the count of 'clean' plant openings, Smith and Todd [37] conclude that DID matching estimators are more robust.

We perform the matching exercises twice, estimating (4a) and (4b) once using counts of domestic-owned plant openings and once using counts of foreign-owned plant openings. Upon completing the matching estimation, balancing and specification tests are conducted. Balancing refers to the fact that after conditioning on the propensity score, the distribution of the conditioning variables, $Z$, should not differ across the treatment and control group in the matched sub-sample. Thus, after matching, we also conduct a joint test—using Hotelling's $T$-squared—to test the null hypothesis of no difference in the vectors of the means of the $Z$'s across the matched treatment and control group. The specification test—proposed in Ham et al. [16]—requires testing for mean differences in the lagged outcome across the matched treatment and control groups.[9]

### 3.2. Fixed effects Poisson

For comparison to estimates derived from (4a) and (4b), we also estimate several parametric specifications derived from the partial equilibrium framework in Henderson [17]. In making this comparison, it is important to note the major differences between the matching and parametric Poisson estimators. First, matching estimators entail fewer distributional assumptions. Second, matching estimators allow one to use additional endogenous variables that are difficult to incorporate into standard parametric count models: lagged values of the dependent variable and current values of non-pollution intensive plant openings. If lagged plant openings are correlated with both current plant openings (through either positive or negative agglomeration externalities) and attainment status (due to an increase in local emissions), consistent estimates in the parametric framework would require a valid instrument for attainment status, one that is uncorrelated with lagged plant openings. Third, matching allows for nonparametric interactions between the covariates in $Z$ in the determination of the outcome of interest (Blundell et al. [2], Bratberg et al. [3]). Fourth, matching estimators utilize only a sub-sample of the control

---

[9] In the present context, this test serves two purposes. First, if the lagged outcome differs across the treatment and control groups, then uncontrolled unobservables may be present and bias the estimated treatment effect. Second, since lagged plant openings may affect current plant births due to agglomeration externalities, for example, differences in lagged openings may have a direct effect on the outcome as well, further biasing the estimator (Henderson [20,21]). Moreover, for foreign-owned plants, lagged foreign-owned plant openings may be particularly important if they signal a county's 'openness' to foreign plants, or if agglomeration externalities are more important to foreign-owned plants due to the sharing of information on the US business environment.

group that is most 'similar' to the treatment group, whereas parametric models utilize all available observations. Finally, the matching estimators employed herein yield an estimate of the *treatment effect on the treated* (TT), whereas the parametric approach estimates the expected treatment effect for an observation chosen at random.

Proceeding, in equilibrium, the count of new (pollution intensive) plants in county $i$ of origin $j$ ($j$ = domestic, foreign) at time $t$, $Y_{ijt}$ ($Y \in \{0, 1, 2, \ldots\}$), is given by

$$Y_{ijt} = \exp(X_{ijt}\beta)\eta_{ijt} + \varepsilon_{ijt}, \tag{5}$$

where $X_{ijt}$ is a vector of county attributes, including attainment status, and time dummies, $\eta_{ijt}$ captures all unobservable, time-varying attributes of county $i$ specific to plant openings from origin $j$, and $\varepsilon_{ijt}$ is an error term satisfying $E[\varepsilon \mid X, \eta] = 0$. Given the inclusion of a constant in $X_{ijt}$, $E[\eta] = 1$ can be assumed without loss in generality (Mullahy [35]).

Equation (5) is estimated using the fixed effects (FE) Poisson model of Hausman et al. [17] (hereafter HHG). In this model each unit of observation is conditioned on total plant openings over the sample period to account for county-specific unobservables that do not vary over time. which accounts for time invariant, county-specific unobservables by conditioning on total plant openings for each unit of observation over the sample period. Resulting estimates of $\beta$ are consistent if $E[\eta \mid X] = E[\eta]$; in other words, regressors in $X$ may be correlated with *time-invariant*, county-specific unobservables, but not *time-specific*, county-specific unobservables.[10] Even if this assumption holds, standard errors from the Poisson model are sensitive to over- or under-dispersion in the data. As such, we also estimate fixed effects negative binomial models, which relax the restriction of equality of the conditional mean and variance functions. Note, in the Poisson framework coefficient estimates have an elasticity interpretation when regressors are in natural logarithmic form and are semi-elasticities when the regressors enter linearly.

In the actual estimation, there are 62 NYS counties, eleven time periods, and two counts of plant openings: foreign- and domestic-owned pollution intensive openings. Thus, we have 124 observations for each period, or 1364 total observations. Factors included in $X$ follow the received literature (e.g., List et al. [29]) and are: attainment status (one for non-attainment, zero for in-attainment), manufacturing employment, manufacturing wages, population, and property taxes. In addition, we also include a dummy variable equal to one if the dependent variable is a count of foreign-owned plant openings (zero otherwise), as well as the foreign dummy interacted with all the remaining covariates. Finally, we include time dummies, where the time dummies are allowed to differ for domestic- and foreign-owned plant openings. Allowing for time dummies specific to foreign entrants controls for temporal shocks that are specific to foreign investors in all counties, such as policies enacted by NYS officials toward foreign investors or other macroeconomic factors that affect foreign plants. These macroeconomic variables include exchange rate fluctuations, interest rates, and any other factors that foreign investors must consider when making investment decisions.

The fixed effects that are removed by conditioning each observation on total births over the entire sample period are not included in the preceding list of controls. These fixed

---

[10] HHG estimates are also consistent in the presence of correlation between variables in $X$ and time-specific unobservables that are constant across counties if $X$ includes period-specific dummies.

effects may either be defined at the county level, or at the county-ownership (i.e., foreign or domestic) level. Allowing for county fixed effects specific to foreign entrants controls for county-specific, time invariant unobservables that only impact foreign investors in a particular county, such as ethnocentric attitudes, or the presence of a large immigrant population. Thus, in the least restrictive model, the time dummies and county fixed effects are allowed to differentially affect plant openings by plant origin.

## 4. Empirical results

### 4.1. Propensity score matching results

Table 3 presents the first-stage logit estimates used to form the propensity score. The specification includes higher order terms and interactions of the various variables to facilitate the balancing of the covariates across the matched treatment and control groups (Dehejia and Wahba [9]).

The first matching algorithm pairs each out-of-attainment county with the in-attainment county from the same year and region with the nearest propensity score. Although there are 176 treatment (non-attainment) observations, only 76 have a possible match.[11] Of the 76 matched pairs, we retain those with propensity scores that differ by less than 1, 5, 10, and 20%. Under these cut-offs, 8, 16, 20, and 27 matched pairs are obtained. Using the same cut-offs but matching each treatment county to itself at a different point in time yields 9, 11, 18, and 25 matches, respectively.[12] The number of matches are scarce in this case due to the fact that few counties exist that were labeled out-of-attainment in one year and in-attainment in another year that meet the caliper constraints.

Table 4 presents the estimated mean differences between the matched treatment and control groups for the primary variables of interest, along with $p$-values associated with the null hypothesis that the means are equal, for each of the eight matched sub-samples. Values can be read as follows: the "0.50" entry in the third row of the first column suggests that treated counties had, on average, 0.50 more clean domestic plant births than control counties had when we set the caliper to be 0.01 and only examine within-region/within-year pairs. In addition, results of the Hotelling's $T$-squared tests for differences in the vectors of means of the various control variables are also displayed. Examination of the results yields four key findings. First, both the within-region/within-year and within-county matching algorithm balance the vector of controls (at the $p < 0.10$ level) in every specification. This includes the variables specifically controlled for in the first-stage logit, as well as property taxes and highway expenditures. Thus, all matching algorithms satisfy

---

[11] All counties in five regions are in non-attainment in various years. Thus, there is no control observation for these counties to serve as a match.

[12] There are at most 143 possible matches using the within-county matching algorithm as three counties are in non-attainment all eleven years of the sample, and therefore have no possible match. The number of actual matches utilized, however, differs little from the within-year, within-region matching due to the fact that few counties exist that were labeled out-of-attainment in one year and in-attainment in another year that have propensity scores that differ by less than the respective calipers.

Table 3
First-stage logit estimates of the determinants of attainment status

| Independent variable | Coefficient (SE) | |
|---|:---:|---:|
| Man. wages | 3.63E−03 | (2.55E−03) |
| Man. wages$^2$ | −2.23E−07 | (1.41E−07) |
| Man. wages$^3$ | 4.27E−12 | (2.74E−12) |
| Man. plants | 1.40[*] | (0.58) |
| Man. plants$^2$ | −0.09[*] | (0.05) |
| Man. plants$^3$ | 1.84E−03[*] | (1.04E−03) |
| Population | −1.85E−06 | (6.28E−06) |
| Population$^2$ | 7.37E−12 | (6.12E−12) |
| Population$^3$ | −3.14E−18[*] | (1.82E−18) |
| Per capita income | 4.73E−03[*] | (1.25E−03) |
| Per capita income$^2$ | −1.86E−07[*] | (9.64E−08) |
| Per capita income$^3$ | 2.63E−12[*] | (1.40E−12) |
| Man. wages[*] Man. plants | −9.57E−06 | (3.20E−05) |
| Man. wages[*] Population | 1.08E−09[*] | (4.53E−10) |
| Man. wages[*] Per capita income | −1.61E−08 | (6.61E−08) |
| Man. plants[*] Population | −8.61E−07[*] | (3.54E−07) |
| Man. plants[*] Per capita income | 1.67E−05 | (3.04E−05) |
| Population[*] Per capita income | −8.88E−10[*] | (4.10E−10) |
| Time effects | Yes | |
| Log-likelihood | −145.8 | |
| Pseudo $R^2$ | 0.63 | |
| N | 682 | |

*Notes*:

(1) Dependent variable is equal to one if county is out-of-attainment of federal ozone standards during the year, zero otherwise.

(2) Standard errors are in parentheses beside the coefficient estimates and are adjusted for clustering within counties.

(3) Time effects are jointly significant at the $p < 0.01$ level.

[*] Significant at the $p < 0.10$ level using a two-sided alternative.

the balancing test. Second, all the matching algorithms also pass the specification test proposed in Ham et al. [16] at the $p < 0.05$ level as lagged pollution intensive plant openings (both domestic- and foreign-owned) are balanced as well.

Third, all eight specifications yield negative point estimates for the impact of non-attainment on domestic-owned plant openings, with three being statistically significant at the $p < 0.05$ level, and a fourth being statistically significant at the $p < 0.10$ level. The magnitudes of the coefficients are also economically significant, implying a loss of between 0.5 and 1.9 domestic-owned plant openings for each year that a county is out-of-attainment. Since the average county receives 0.32 new (pollution intensive)

Table 4
Propensity score estimates of attainment status effect

| Independent variable | Matching algorithm | | | | | | | |
| | Within region and year | | | | Within county | | | |
| | Maximum difference | | | | Maximum difference | | | |
| | (0.01) | (0.05) | (0.10) | (0.20) | (0.01) | (0.05) | (0.10) | (0.20) |
|---|---|---|---|---|---|---|---|---|
| Propensity score | 0.00 | 0.00 | 0.01 | 0.04 | 0.00 | 0.01 | 0.04 | 0.06 |
| | ($p = 0.98$) | ($p = 0.98$) | ($p = 0.90$) | ($p = 0.62$) | ($p = 1.00$) | ($p = 0.97$) | ($p = 0.75$) | ($p = 0.60$) |
| Dirty domestic plant births ($\tau_{TT,\bullet}$) | 0.38 | −0.13 | −0.20 | −0.22 | −1.22 | −0.91 | −0.67 | 0.20 |
| | ($p = 0.06$) | ($p = 0.65$) | ($p = 0.39$) | ($p = 0.21$) | ($p = 0.07$) | ($p = 0.11$) | ($p = 0.08$) | ($p = 0.72$) |
| Clean domestic plant births | 0.50 | 0.19 | 0.30 | 0.37 | 0.67 | 0.36 | −0.22 | 1.52 |
| | ($p = 0.39$) | ($p = 0.57$) | ($p = 0.27$) | ($p = 0.10$) | ($p = 0.26$) | ($p = 0.47$) | ($p = 0.53$) | ($p = 0.29$) |
| Net domestic plant births ($\tau_{DID,\bullet}$) | −0.13 | −0.31 | −0.50 | −0.59 | −1.89 | −1.27 | −0.44 | −1.32 |
| | ($p = 0.82$) | ($p = 0.34$) | ($p = 0.08$) | ($p = 0.01$) | ($p = 0.00$) | ($p = 0.03$) | ($p = 0.27$) | ($p = 0.21$) |
| Lagged dirty domestic plant births (1 year) | 0.29 | 0.21 | 0.24 | 0.13 | 0.38 | 0.49 | 0.31 | 0.91 |
| | ($p = 0.15$) | ($p = 0.15$) | ($p = 0.11$) | ($p = 0.27$) | ($p = 0.48$) | ($p = 0.28$) | ($p = 0.29$) | ($p = 0.12$) |
| Dirty foreign plant births ($\tau_{TT,\bullet}$) | 0.00 | −0.06 | −0.05 | −0.04 | −0.11 | −0.09 | −0.11 | 0.00 |
| | ($p = 1.00$) | ($p = 0.70$) | ($p = 0.70$) | ($p = 0.70$) | ($p = 0.72$) | ($p = 0.73$) | ($p = 0.52$) | ($p = 1.00$) |
| Clean foreign plant births | 0.75 | 0.25 | 0.20 | 0.15 | −0.44 | −0.36 | −0.17 | −0.04 |
| | ($p = 0.09$) | ($p = 0.40$) | (0.41) | ($p = 0.42$) | ($p = 0.51$) | ($p = 0.55$) | ($p = 0.68$) | ($p = 0.90$) |
| Net foreign plant births ($\tau_{DID,\bullet}$) | −0.75 | −0.31 | −0.25 | −0.19 | 0.33 | 0.27 | 0.06 | 0.04 |
| | ($p = 0.02$) | ($p = 0.15$) | ($p = 0.15$) | ($p = 0.16$) | ($p = 0.43$) | ($p = 0.48$) | ($p = 0.84$) | ($p = 0.85$) |
| Lagged dirty foreign plant births (1 year) | 0.43 | 0.21 | 0.18 | 0.13 | 0.50 | 0.44 | 0.27 | −0.02 |
| | ($p = 0.18$) | ($p = 0.18$) | ($p = 0.18$) | ($p = 0.18$) | ($p = 0.08$) | ($p = 0.07$) | ($p = 0.07$) | ($p = 0.90$) |
| $H_0$: Vectors of means of controls are equal | $F(7, 8) = 1.31$ | $F(7, 24) = 1.57$ | $F(7, 32) = 1.66$ | $F(7, 46) = 1.41$ | $F(7, 10) = 0.21$ | $F(7, 14) = 0.14$ | $F(7, 28) = 0.22$ | $F(7, 42) = 1.42$ |
| | ($p = 0.36$) | ($p = 0.19$) | ($p = 0.15$) | ($p = 0.22$) | ($p = 0.98$) | ($p = 0.99$) | ($p = 0.98$) | ($p = 0.22$) |

Table 4 (*Continued*)

| Independent variable | Matching algorithm | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Within region and year | | | | Within county | | | |
| | Maximum difference | | | | Maximum difference | | | |
| | (0.01) | (0.05) | (0.10) | (0.20) | (0.01) | (0.05) | (0.10) | (0.20) |
| Number of matched pairs | 8 | 16 | 20 | 27 | 9 | 11 | 18 | 25 |
| Number of unique controls | 8 | 15 | 17 | 18 | 6 | 7 | 10 | 12 |

*Notes*:

(1) Figures represent mean difference between treatment counties (out-of-attainment) and control counties (in-attainment). *P*-values in parentheses are for the tests that the mean difference across the treatment and controls groups are equal.

(2) 'Dirty' plants are those defined as pollution intensive (see text); 'clean' are non-pollution intensive plants.

(3) 'Unique controls' reports the number of control counties that are matched with at least one treatment county.

(4) Test of equality of the vectors of means across the matched treatment and control group based on Hotelling's *T*-squared test. Set of controls includes number of manufacturing plants, manufacturing wages, manufacturing employment, population, property taxes, per capita income, and highway expenditures.

domestic-owned plant openings per year, this represents a sizeable deterrent effect of local environmental standards.

Finally, while the four within-region/within-year specifications yield negative point estimates for the impact of non-attainment on foreign-owned plant openings, only one is statistically significant at the $p < 0.10$ level. The four within-county matching specifications all find a positive effect of non-attainment on foreign-owned plant births, although none of the estimates are statistically significant. Thus, while there is some evidence to suggest that foreign-owned plants also incorporate spatial variation in environmental stringency into decisions concerning plant location, it is far less robust than for domestic-owned plants.

As noted in the introduction, these findings are consistent with (at least) three possible explanations. First, foreign-owned firms face weaker enforcement of environmental standards due to the economic benefits bestowed upon the local community. Second, foreign-owned firms may simply be at an informational disadvantage relative to domestic firms and, therefore, unwittingly locate in areas with more stringent environmental regulation. Finally, the location decisions of foreign-owned firms may be less sensitive to environmental costs as other factors (such as agglomeration externalities, or geography) dictate the location decision.

## 4.2. Fixed effects Poisson results

Table 5 contains the FE Poisson estimation results.[13] Four specifications are estimated, depending on whether the county fixed effects and time dummies are allowed to have differential effects on domestic- versus foreign-owned plants. Column (4) presents the least restrictive model, allowing both to vary by plant origin. In addition, for each specification, we test for equal effects of attainment status on plant openings by origin, as well as jointly test whether the set of covariates have equal effects on domestic- and foreign-owned plant openings.

Examination of the results reveals three important conclusions. First, in all four specifications, consonant with the difference-in-differences propensity score matching estimates, we find a negative and statistically significant impact of non-attainment on domestic-owned plant openings. The point estimates indicate that the expected number of plant openings by domestic plants is reduced by 44–61% each year a county is in non-attainment. Since the average county receives 0.32 new domestic-owned plants each year, this translates into an expected loss of 0.14–0.20 new domestic-owned plants per year, of considerably smaller magnitude than the matching estimates. Second, all four specifications fail to find a statistically significant effect of non-attainment on openings of foreign-owned plants. However, we never reject the null of equal effects of non-attainment on plant openings by plant origin. Finally, we find statistically significant effects of some of the controls on plant openings. Specifically, both domestic- and foreign-owned plants are less likely to locate in counties with high manufacturing employment, high property taxes,

---

[13] Results based on fixed effects negative binomial models are virtually identical to the estimates reported and are available upon request.

Table 5
Parametric estimates of the determinants of county-level plant births

| Independent variable | Fixed effects Poisson (HHG) | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Non-attainment* domestic plant | −0.46* (0.25) | −0.44* (0.25) | −0.61* (0.26) | −0.58* (0.27) |
| Non-attainment* foreign plant | −0.77 (0.49) | −0.64 (0.54) | −0.14 (0.54) | 0.09 (0.65) |
| ln(employment)* domestic plant | −6.91* (2.19) | −7.00* (2.22) | −7.32* (2.28) | −7.37* (2.29) |
| ln(employment)* foreign plant | −6.54* (2.24) | −6.35* (2.24) | −6.12 (7.01) | 1.97 (10.07) |
| ln(wage)* domestic plant | −0.34 (2.44) | −0.04 (2.45) | 0.17 (2.54) | −0.64 (2.67) |
| ln(wage)* foreign plant | 1.46 (2.66) | 0.39 (2.91) | 2.04 (3.54) | 9.28 (8.65) |
| ln(population)* domestic plant | 9.43* (5.28) | 9.69* (5.28) | 9.18 (5.60) | 9.09 (5.62) |
| ln(population)* foreign plant | 9.14* (5.30) | 9.35* (5.30) | 10.54 (14.47) | 5.74 (18.82) |
| ln(prop. tax)* domestic plant | −3.59 (2.29) | −3.08 (2.31) | −3.87* (2.33) | −4.28* (2.51) |
| ln(prop. tax)* foreign plant | −7.12* (2.50) | −7.93* (2.62) | −4.31 (2.73) | 2.53 (6.32) |
| Foreign | 1.95 (10.44) | 20.17 (18.16) | – | – |
| County effects | Yes | Yes | No | No |
| County × origin effects | No | No | Yes | Yes |
| Period effects | Yes | No | Yes | No |
| Period × origin effects | No | Yes | No | Yes |
| $H_0$: Equal effects of non-attainment | $p = 0.53$ | $p = 0.73$ | $p = 0.42$ | $p = 0.34$ |
| $H_0$: Equal effects of controls | $p = 0.01$ | $p = 0.01$ | $p = 0.77$ | $p = 0.49$ |
| Log-likelihood | −367.09 | −360.66 | −321.87 | −315.79 |
| N | 1364 | 1364 | 1364 | 1364 |

*Notes*:

(1) Dependent variable is the count of annual pollution intensive manufacturing plant births from 1980 to 1990.

(2) Non-attainment equals one if county is out-of-attainment of federal ozone standard, zero otherwise. Foreign equals one if dependent variable is a count of foreign plant births, zero if domestic.

(3) Standard errors are in parentheses beneath coefficient estimates.

(4) ln indicates the natural logarithm of the variable.

\* Significant at the 10% level using a two-sided alternative.

and small populations. Interestingly, though never statistically significant at conventional levels, the estimates are consistent with the notion that foreign plants are attracted to high-wage counties, in contrast to domestic plants. Equality of the vector of coefficients on the four control variables is rejected in columns (1) and (2), but not (3) and (4).[14]

---

[14] Since equality of the response coefficients is not rejected in columns (3) and (4), we re-estimated these specifications imposing equal effects of the controls, while still allowing the non-attainment effect to vary by plant

In sum, the parametric FE Poisson and semi-nonparametric difference-in-differences propensity score matching estimators provide generally consonant results: a statistically significant, negative effect of non-attainment on domestic-owned, but not foreign-owned plant openings. However, given that the magnitude of the effect is much smaller in the parametric framework, it is worth briefly exploring the reasons for this difference. Several potential explanations exist. First, the matching method estimates the average treatment effect *on the treated*, while the Poisson framework estimates the overall average treatment effect.[15] Second, the matching method uses only a sub-sample of the original sample, discarding control observations deemed too different from the treatment group. Third, the matching method permits the nonparametric interactions between the various control variables. Finally, our parametric method controls for both county and time effects simultaneously (the matching method controls for either county effects, or region and time effects, but never both county and time effects directly), but the matching method utilizes data on non-pollution intensive plant openings to control for time-specific, county-specific unobservables, which are not controlled for in the parametric estimation.

We can explore the role played by some of these potential explanations. The first pertains to the distinction between the average treatment effect and the average treatment effect on the treated. To see if this distinction matters in practice in the current application, we can use the matching technique to estimate the average treatment effect as well. To proceed, we label 'in-attainment' as the treatment and re-do the matching by pairing each in-attainment county with an out-of-attainment county. This process provides estimates of the average treatment effect *on the untreated*. The weighted average of the average treatment effect on the treated and untreated constitutes the overall average treatment effect. Performing this exercise using the four specifications that yielded statistically significant results for domestic-owned plant openings in Table 4, we obtain the following estimates of the average treatment effect:

- within-region/within-year matching with 10% (20%) caliper $= -0.57$ ($-0.55$);
- within-county matching with 1% (5%) caliper $= -1.68$ ($-1.11$).

These estimates differ very little from those reported in Table 4.

The second potential explanation deals with the fact that the method of matching utilizes only a sub-sample of the original data, whereas the parametric framework uses all observations. In principle, one could re-estimate the parametric model using only the matched sub-samples utilized in Table 4, as in List et al. [29]. However, in the present case, the sample sizes from Table 4 are clearly inadequate. The third potential explanation pertains to the fact that the method of matching permits the nonparametric interaction of the various controls. To test the relevance of this, one could re-estimate the parametric models with a full set of interactions, as well as higher order terms, for the various control variables. We do not attempt this given the enormous demands that are already placed on

---

origin. The results are unchanged: we still do not reject equality of the non-attainment effect across domestic- and foreign-owned plants.

[15] See Wooldridge [40] for a discussion of various treatment effect parameters.

the data through the inclusion of county fixed effects and time dummies that are allowed to differ by plant origin.

The final potential culprit is the fact that, through the use of data on non-pollution intensive plants and maintenance of the assumption that 'clean' plant openings should not be affected by attainment status, the matching method controls for time-specific, county-specific unobservables while the parametric technique only controls for time invariant, county-specific unobservables and time-varying, county invariant unobservables. To analyze the ramifications of this difference, we estimate a more complex FE Poisson model, incorporating plant openings by domestic- and foreign-owned plants in non-pollution intensive sectors.

To proceed, we estimate (5) where the dependent variable is now $Y_{ijkt}$, the count of new plants in county $i$ of origin $j$ ($j =$ domestic, foreign) with pollution intensity $k$ ($k =$ pollution intensive, non-pollution intensive) at time $t$. Since each county-year pair supplies four observations, the total sample size increases to 2728. In addition, in the least restrictive model, we allow the covariates, including attainment status, the county fixed effects, and the time dummies to have differential effects for each of the four data types: domestic/non-pollution intensive, domestic/pollution intensive, foreign/non-pollution intensive, and foreign/pollution intensive. Thus, we estimate four effects of non-attainment, $\beta_{d,npi}$, $\beta_{d,pi}$, $\beta_{f,npi}$, and $\beta_{f,pi}$, respectively.

Maintaining the assumption, as in the matching method, that non-pollution intensive plant openings should not be affected by attainment status (i.e., $\beta_{d,npi} = \beta_{f,npi} = 0$), differences from zero in the estimated values of these parameters may be attributable to time-specific, county-specific unobservables not controlled in the previous Poisson model (e.g., subsidies to attract new plants or worker training programs). Consequently, we can obtain difference-in-differences FE Poisson estimates, analogous to the difference-in-differences matching estimates, from

$$\beta_{d,\mathrm{DD}} = \beta_{d,pi} - \beta_{d,npi}, \tag{6a}$$

$$\beta_{f,\mathrm{DD}} = \beta_{f,pi} - \beta_{f,npi}, \tag{6b}$$

where $\beta_{d,\mathrm{DD}}$ ($\beta_{f,\mathrm{DD}}$) is the difference-in-differences FE Poisson estimator of the effect of non-attainment on domestic- (foreign-)owned plant openings. The estimate obtained for domestic-owned plant openings is $-0.81$ ($\chi^2(1) = 6.35$, $p = 0.01$); the estimate for foreign-owned plant openings is $0.28$ ($\chi^2(1) = 0.15$, $p = 0.69$).[16] While a reduction in domestic pollution intensive plant openings of 81% is nearly 40% larger than the point estimate obtained from the least restrictive model in Table 5 (column (4)), it is still smaller in magnitude than the matching estimates provided in Table 4.[17] Consequently, a sizeable discrepancy in the magnitudes of the parametric and semi-nonparametric estimates remains, suggesting a role for the other two explanations (sample composition and non-linearity of the effects of the control variables).

---

[16] The null of equal effects of non-attainment is not rejected at conventional levels ($\chi^2(1) = 1.93$, $p = 0.16$).

[17] An 81% decrease in the expected number of new domestic plant openings translates into a loss of 0.26 plants per annum for the average county. The smallest, statistically significant matching estimate in Table 4 suggests a loss of at least 0.50 plants annually.

## 5. Conclusion

When the federal government reversed more than 100 years of state and local authority of polluters in the 1960s and 1970s, substantial costs were incurred to design and implement regulations, monitor and report regulatory compliance, and invest in research and development. These costs, and other private compliance costs, were ultimately borne by stockholders, business owners, consumers, and taxpayers. Given the magnitude of such costs, intuition dictates that plants should incorporate spatial differences in such costs into their decision calculus when making plant-level location choices. This study explores whether foreign capital and domestic capital responds asymmetrically to local environmental standards. If asymmetries do exist, it is critical for policymakers to more fully understand these behavioral differences given the well-documented economic benefits of foreign investment (even relative to domestic investment). In addition, in a normative sense the results provide insights into the political economy of regulatory federalism.

To provide initial insights into this question, we use a series of novel empirical methods to examine plant location choices in New York State from 1980 to 1990. Our empirical methods test for differential responses of foreign and domestic investment flows to environmental regulations by integrating the domestic and foreign investment decision into a single empirical model. Making use of both a semi-nonparametric difference-in-differences propensity score matching estimator and a standard fixed effects Poisson model, we find that domestic plants are responsive to environmental standards whereas foreign plants are not. While more research on this observed differential is warranted since our sample of foreign-owned pollution intensive plant openings is not large, our results indicate that there may be a 'foreign plant double-reward' from increased environmental standards: while other studies indicate that foreign-owned plants in the US tend to create more jobs and increase area wages relative to domestic-owned plants, our results imply that foreign plants are less sensitive to interjurisdictional differences in environmental regulations. This finding opens up the possibility that localities can introduce more stringent environmental regulations without risking the loss of substantial foreign capital, capital that provides the greatest local economic stimulus.

# References

[1] R. Blundell, M. Costa-Dias, Alternative approaches to evaluation in empirical microeconomics, Portuguese Economic Journal 1 (2002) 91–115.

[2] R. Blundell, L. Dearden, B. Sianesi, Evaluating the impact of education on earnings in the UK: models, methods and results from the NCDS, Journal of the Royal Statistical Society, in press.

[3] E. Bratberg, A. Grasdal, A.E. Risa, Evaluating social policy by experimental and nonexperimental methods, Scandinavian Journal of Economics 104 (2002) 147–171.

[4] W. Cochran, D. Rubin, Controlling bias in observational studies, Sankyha 35 (1973) 417–446.

[5] M. Cole, R.J.R. Elliott, FDI and the capital intensity of 'dirty' sectors: A missing piece of the pollution haven puzzle, University of Birmingham discussion paper in Economics 02-04, 2002.

[6] M. Crozet, T. Mayer, J.-L. Mucchielli, How do firms agglomerate? A study of FDI in France, Regional Science and Urban Economics 34 (2004) 27–54.

[7] J.M. Dean, M.E. Lovely, H. Wang, Foreign direct investment and pollution havens: evaluating the evidence from China, unpublished manuscript, Syracuse University, 2002.

[8] R.H. Dehejia, S. Wahba, Casual effects in nonexperimental studies: reevaluating the evaluation of training programs, Journal of the American Statistical Association 94 (1999) 1053–1062.

[9] R.H. Dehejia, S. Wahba, Propensity score matching for nonexperimental causal studies, Review of Economics and Statistics 84 (2002) 151–161.

[10] M. Eichler, M. Lechner, An evaluation of public Employment programs in the East German State of Sachsen-Anhalt, Labour Economics 9 (2002) 143–186.

[11] D. Figlio, B. Blonigen, The effects of direct investment on local communities, Journal of Urban Economics 48 (2000) 338–363.

[12] O. Figueiredo, P. Guimarães, D. Woodward, Home-field advantage: location decisions of Portuguese entrepreneurs, Journal of Urban Economics 52 (2002) 341–361.

[13] P.G. Fredriksson, J.A. List, D.L. Millimet, Corruption, environmental policy, and FDI: Theory and evidence from the United States, Journal of Public Economics 87 (2003) 1407–1430.

[14] Roger H. Gordon, Lans A. Bovenberg, Why is capital so immobile internationally? Possible explanations and implications for capital income taxation, American Economic Review 86 (1996) 1057–1075.

[15] M. Greenstone, The impacts of environmental regulations on industrial activity: evidence from the 1970 and 1977 Clean Air Act amendments and the Census of Manufactures, Journal of Political Economy 110 (2002) 1175–1219.

[16] J.C. Ham, X. Li, P.B. Reagan, Matching and selection estimates of the effect of migration on wages for young men, unpublished manuscript, Ohio State University, 2001.

[17] J. Hausman, B.H. Hall, Z. Griliches, Econometric models for count data with an application to the patents—r&d relationship, Econometrica 52 (1984) 909–938.

[18] J.J. Heckman, H. Ichimura, P.E. Todd, Matching as an econometric evaluation estimator: evidence from evaluating a job training program, Review of Economic Studies 64 (1997) 605–654.

[19] J.V. Henderson, J. Vernon, Effects of air quality regulation, American Economic Review 86 (1996) 789–813.

[20] J.V. Henderson, J. Vernon, Externalities and industrial development, Journal of Urban Economics 42 (1997) 449–470.

[21] J.V. Henderson, J. Vernon, Marshall's scale economies, Journal of Urban Economics 53 (2003) 1–28.

[22] N.G. Howenstine, W.J. Zeile, Characteristics of foreign-owned US manufacturing establishments, Survey of Current Business, 1994.

[23] A. Jaffe, S. Peterson, P. Portney, R. Stavins, Environmental regulation and the competitiveness of US manufacturing: What does the evidence tell us? Journal of Economic Literature XXXIII (1995) 132–163.

[24] T. Jeppessen, J.A. List, H. Folmer, Environmental regulations and new plant location decisions: evidence from a meta-analysis, Journal of Regional Science 42 (2002) 19–49.

[25] W. Keller, A. Levinson, Pollution abatement costs and foreign direct investment inflows to US states, Review of Economics and Statistics 84 (2002) 691–703.

[26] A. King, J.M. Shaver, Are aliens green? Assessing foreign establishments' environmental conduct in the United States, unpublished manuscript, NYU, 1999.

[27] John A. List, Preference reversals of a different kind: the more is less phenomenon, American Economic Review 92 (2002) 1498–1507.

[28] J.A. List, W.W. McHone, Measuring the effects of air quality regulations on 'dirty' firm births: Evidence from the neo- and mature- regulatory periods, Papers in Regional Science 79 (2000) 177–190.

[29] J.A. List, D.L. Millimet, P.G. Fredriksson, W.W. McHone, Effects of environmental regulations on manufacturing plant births: Evidence from a propensity score matching estimator, Review of Economics and Statistics 85 (2003) 944–952.

[30] J.A. List, D.L. Millimet, W.W. McHone, Effects of air quality regulation on the destination choice of relocating plants, Oxford Economic Papers 55 (4) (2003) 657–678.

[31] J.A. List, D.L. Millimet, W.W. McHone, The unintended disincentive in the clean air act, Advances in Economic Analysis & Policy 4 (2004), article 2, http://www.bepress.com/bejeap/advances/vol4/iss2/art2.

[32] R.H. McGuckin, Longitudinal Economic Data at the Census Bureau: A New Database Yields Fresh Insights on Some Old Issues, Center for Economic Studies, Bureau of Census, Washington, DC, 1990.

[33] J.C. Michalke, Almost Everything You Ever Wanted to Know About the New York State Department of Commerce Industrial Migration File, NYS Department of Economic Development, Division of Policy and Research, Albany, NY, 1986.

[34] D.L. Millimet, J.A. List, The case of the missing pollution haven hypothesis, Journal of Regulatory Economics, in press.

[35] J. Mullahy, Instrumental-variable estimation of count data models: Applications to models of cigarette smoking behavior, Review of Economics and Statistics 79 (1997) 586–593.

[36] P. Rosenbaum, D. Rubin, The central role of the propensity score in observational studies for causal effects, Biometrika 70 (1983) 41–55.

[37] J. Smith, P. Tood, Does matching address Lalonde's critique of nonexperimental estimators, Journal of Econometrics, in press.

[38] R. Tannenwald, State regulatory policy and economic development, in: New England Economic Review, Proceedings of a Symposium on The Effects of State and Local Public Policies on Economic Development, 1997, pp. 83–97.

[39] R. Vernon, In the Hurricane's Eye: The Troubled Prospects of Multinational Enterprises, Harvard Business School Press, Cambridge, MA, 1998.

[40] J.M. Wooldridge, Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge, MA, 2002.