



The Case of the Missing Pollution Haven Hypothesis*

DANIEL L. MILLIMET

Southern Methodist University

Department of Economics, Box 0496, Dallas, TX 75275-0496,

E-mail: millimet@mail.smu.edu

JOHN A. LIST

University of Maryland and NBER

2200 Symons Hall, College Park, MD 20742-5535,

E-mail: jlist@arec.umd.edu

Abstract

One particularly vexing puzzle for economists and policymakers over the past several decades concerns the empirical significance of the theoretically predicted pollution haven hypothesis. While neo-classical theory and conventional wisdom both surmise that local economies will suffer deleterious effects from stricter environmental regulations, empirical studies have largely failed to validate such claims. This study utilizes the method of matching to show that the impact of stricter regulation is *heterogeneous* spatially, varying systematically based on location-specific attributes. Previous studies that assume a homogenous response may therefore inadvertently mask the overall impact of more stringent regulations by pooling unaffected and affected regions.

Key words: pollution haven hypothesis, environmental regulation, treatment effects, propensity score, matching

JEL Classification: Q25, Q28, H00, L51, R1

1. Introduction

The “pollution haven” hypothesis (PHH) refers to the conjecture that mobile capital, particularly polluting capital, will (re)locate to regions with weak environmental standards. If valid, stricter environmental protection gives rise to a myriad of

* The authors wish to thank to Michael Crew, Anthony Heyes, an anonymous referee, Werner Antweiler, Randy Becker, Wayne Gray, Shelby Gerking, Michael Greenstone, Vern Henderson, and Arik Levinson. Participants at various university seminars and conferences also lent useful insights to this line of research.

unintended consequences, from distorting trade patterns to conferring a competitive advantage on environmentally lax areas, in turn undermining the incentives for local policymakers to adopt such protections. Concern over such outcomes is reflected in U.S. Congressional testimony (House Report 1979) and was a major impetus behind the creation of the U.S. Environmental Protection Agency (EPA) in 1968 and the passage of subsequent legislation establishing national standards for air and water quality. The recent creation of a supranational authority to oversee European environmental control, as well as debates over whether international trade agreements should include provisions on domestic environmental policy, reflect similar logic (Ederington 2001).

Such governmental actions lend credence to the conventional wisdom, which holds that environmental regulations impose significant costs in terms of slower productivity growth, diminished international competitiveness of firms, and escalating trade deficits. While the underlying logic for such arguments is intuitive, empirical estimates in the received literature, which range from positive and significant to negative and significant, lead most scholars to conclude that the relationship between environmental regulations and competitiveness is weak at best (see, e.g., Jaffe et al. 1995).¹ This puzzle has intrigued academics and policymakers for decades.

This study offers a resolution to this enigma. While empirical efforts to date typically have relied on standard regression or matching techniques to estimate the (conditional) *average effect* of stricter regulations to test the PHH, we develop an alternative approach that allows the impact of environmental stringency to vary spatially.² Accordingly, we begin with the notion that the comparative static effect of increased environmental stringency is not homogenous across space, but rather depends crucially on observable location-specific attributes. The spirit of our advance is couched in a difference-in-differences semi-non-parametric propensity score matching estimator.

We implement our approach using panel data on new plant location patterns and the exit patterns of existing plants across counties for New York State from 1980 to 1990. The empirical results are insightful. First, we find that the impact of stricter environmental standards varies tremendously across counties. Second, much of the variation can be explained via observable county-level attributes. For example, counties with high levels of unemployment tend to lose fewer existing firms and forego fewer new plants from the imposition of stricter environmental

1 See Jeppessen et al. (2002) for a more recent literature survey of empirical work on environmental regulation and capital location, and Mulatu et al. (2002) for a recent survey of environmental regulation and international competitiveness.

2 Indeed, in our own previous work we have estimated the average treatment effect. For example, Fredriksson et al. (2003) estimate the average effect of stricter U.S. state environmental regulation on inbound foreign direct investment via two-staged least squares. And, List et al. (2003a, b) utilize a difference-in-difference propensity score matching technique to estimate the *average* treatment effect on the treated (i.e., the impact of counties being labeled out of attainment of federal ozone requirements under the U.S. Clean Air Act).

standards (consistent with firms trading off higher abatement costs with abundant, cheap labor). This finding suggests that, perhaps contrary to conventional wisdom, the best time to undertake greater environmental protection is during economic downturns. Moreover, counties with greater manufacturing employment (yielding positive agglomeration externalities) and a more skilled labor force tend to also incur lower costs from increasing environmental standards.

These findings bring to the forefront one potential reason why the multitude of empirical studies testing the PHH offer such conflicting conclusions. Even within a relatively homogeneous environment, such as New York State, the impact of stricter environmental regulation runs the gamut, from positive to negative. If pollution haven effects exhibit such variation within a single state, we suspect that empirical studies aggregating across U.S. states, and especially multiple countries, mask the “true” effects by pooling relatively diverse regions. Much like finding a needle in a haystack, we conclude that successfully finding evidence of the PHH depends crucially on where one looks.

The remainder of this study is crafted as follows. Section 2 describes our data and empirical methods. Section 3 discusses the empirical results and section 4 concludes.

2. Data and Empirical Strategy

2.1. Data

Air quality regulation in the U.S. has evolved into an amalgam of local and federal authority—with the federal government responsible for setting ambient air quality standards via the 1977 Clean Air Act Amendments (CAAA). The CAAA set standards on five criteria air pollutants (sulfur dioxides, carbon monoxides, ozone, nitrogen oxides, and total suspended particulates), and directed federal authorities to label every county, beginning in 1978, as either in or out of attainment of the applicable federal standard for each pollutant, with out of attainment counties compelled to follow stricter environmental standards. Henderson (1996, 790) notes:

“For counties not in attainment, new manufacturing firms to the county will be subject to more stringent regulations governing equipment specifications. Existing firms in nonattainment areas face stricter requirements to reduce source emissions and new firms may be required to purchase offsets (emission rights) from existing firms. All firms in nonattainment counties are more likely to be closely monitored and subject to greater enforcement efforts. For ozone, in addition, auto regulations may be tougher... In summary, being out of attainment introduces a set of overall regulatory activities designed to reduce emissions, which are not faced to the same extent by counties in attainment.”

In this study we follow Henderson (1996), Becker and Henderson (2000), and others and focus on the extent that county-level ozone attainment status influences plant-level behavior.³

Building on List et al. (2003a, b), the outcomes of interest are obtained from county-level plant location data available in the comprehensive Industrial Migration File (IMF), maintained from 1980 to 1990 by New York State (NYS). This rich source of data contains information on new plant location patterns (the number of new plant births by county in a given year), as well as the exit patterns among existing plants (the number of plants lost by a county in a given year). We make use of both types of data to examine the hypotheses of interest.

To focus the analysis on plants most likely to be affected by environmental costs, we classify plants as pollution- or non-pollution-intensive. To obtain such a classification, we consider the industrial sector to which each firm belongs. Following Greenstone (2002) and List et al. (2003a, b), sectors that emit at least 6% of the total industrial sector's emissions of nitrogen oxide or volatile organic compounds, the primary chemical precursors to ozone, are considered pollution-intensive. Under this criteria, plants in SIC codes 2611–31, 2711–89, 2812–19, 2861–69, 2911, 30, 32, 3312–3, 3321–5, 34, and 371 are considered pollution-intensive. In total, we observe location decisions for 278 (716) new pollution-intensive (non-pollution-intensive) plants across the 62 counties in NYS over the sample period 1980–1990; we observe 561 (1708) pollution-intensive (non-pollution-intensive) plant exits due to either closure, bankruptcy, full or partial relocation out of state, or relocation within NYS. Of the new pollution-intensive (non-pollution-intensive) plant births observed, 56% (50%) were located in attainment counties; 39% (34%) of pollution-intensive (non-pollution-intensive) plant exits occurred in attainment counties. Summary statistics are provided in table 1.

2.2. Empirical Methodology

When the policy measure representing environmental regulation is binary—such as in attainment and out of attainment of the federal ozone standard—the potentially heterogeneous effect of policy stringency may be modelled in a straightforward manner. The empirical task reduces to one of identifying the distribution of the treatment effect. The fundamental problem in identifying treatment effects is one of incomplete information. While one observes treatment status and the outcome conditional on treatment status, the counterfactual is unobserved. Let y_{i1} denote the outcome of observation i if the treatment occurs (given by $T_i = 1$);

3 Our choice of pollutants is timely, as EPA recently strengthened its ozone health standard in response to a compelling body of scientific information suggesting that the existing standard was not stringent enough. This action triggered a legal requirement for EPA to designate, by July 2000, areas as in or out of attainment for the revised standard. EPA missed this deadline and the American Lung Association and eight national and local environmental groups notified EPA of their intent to sue. To avoid protracted litigation, EPA reached an agreement with the plaintiffs that requires EPA to issue attainment status designations by April 30, 2004.

Table 1. Description of Variables				
Variable	Mean	In Attainment Mean	Out of Attainment Mean	Definition and Source
New Pollution-Intensive Plant Births	0.41 (0.89)	0.31 (0.64)	0.70 (1.32)	Actual count of new 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).
New Non-Pollution-Intensive Plant Births	1.05 (2.09)	0.71 (1.25)	2.02 (3.36)	Actual count of new plants from 1980 to 1990 labeled as having production activities that are non-pollution-intensive. Industrial Migration File. New York State Department of Economic Development (DED).
Pollution-Intensive Plant Exits	0.82 (1.88)	0.43 (0.91)	1.15 (2.27)	Actual count of plant exits from 1980 to 1990 labeled as having production activities that are pollution-intensive. Industrial Migration File. New York State Department of Economic Development (DED).
Non-Pollution-Intensive Plant Exits	2.50 (6.66)	1.95 (3.11)	6.39 (11.72)	Actual count of plant exits from 1980 to 1990 labeled as having production activities that are non-pollution-intensive. Industrial Migration File. New York State Department of Economic Development (DED).
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.

Table 1. (continued)				
Variable	Mean	In Attainment Mean	Out of Attainment Mean	Definition and Source
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. <i>County Business Patterns</i> .
Firms	11.71 (5.39)	12.23 (5.49)	10.22 (4.84)	Total number of manufacturing firms. <i>County Business Patterns</i> .
Population (100,000s)	2.87 (4.78)	1.82 (3.40)	5.87 (6.60)	County population. <i>Current Population Reports</i> . U.S. Bureau of Census.
Property Tax (1000s)	0.56 (0.24)	0.57 (0.27)	0.54 (0.17)	Real property tax collected per capita. <i>Census of Governments</i> .
Per Capita Income (1000s)	13.50 (4.76)	13.58 (4.30)	13.24 (5.88)	Real per capita income. <i>County Business Patterns</i> .
Highway Expenditures	0.17 (0.09)	0.18 (0.10)	0.15 (0.08)	Total highway expenditures per capita. <i>State and Local Govt. Finances</i> .
Unemployment	8.13 (2.47)	8.05 (2.43)	8.37 (2.51)	Civilian labor force unemployment rate. <i>Bureau of Labor Statistics</i>
<i>Notes.</i> 1. Data are for the 62 New York counties from 1980 to 1990. $N = 682$ (176 out of attainment). 2. Standard deviations are in parentheses.				

y_{i0} denotes the outcome in the absence of the treatment ($T_i = 0$). If both states of the world are observable, then the treatment effect for observation i is given by $\tau_i = y_{i1} - y_{i0}$. As the treatment effect varies by observation, it is a random variable with a distribution (Wooldridge 2002).

Given that only y_{i1} or y_{i0} is observed for each observation, however, τ_i is not observed. Thus, to analyze the distribution of the treatment effect, τ_i must be estimated for all i , which requires estimating the missing counterfactual for each observation. To accomplish this task, Rosenbaum and Rubin (1983) advocate finding a vector of covariates, Z , such that

$$y_1, y_0 \perp T | Z, \quad \Pr(T = 1 | Z) \in (0, 1), \quad (1)$$

where \perp denotes independence. Because outcomes are independent of treatment assignment conditional on Z given (1), the outcomes of control (treated) observations with identical Z constitute an unbiased estimate of the missing counterfactual for each treated (control) observation.

For condition (1) to hold, the conditioning set Z should be multi-dimensional. Consequently, finding observations with differing treatment assignment, but identical values for all covariates in Z , may be untenable. To circumvent this dimensionality problem, Rosenbaum and Rubin (1983) prove that conditioning on $p(Z)$ is equivalent to conditioning on Z , where $p(Z) = \Pr(T=1|Z)$ is the propensity score. In the actual implementation, $p(Z)$ is estimated via logit.

Upon estimating the propensity score, a matching algorithm is defined to estimate the missing counterfactual for each observation. The simplest such algorithm is single nearest-neighbor matching, whereby each treated (control) observation is paired with the control (treated) observation whose propensity score is closest in absolute value (Dehejia and Wahba 2002).⁴ After forming each matched pair, the estimated treatment effect for observation i is given by

$$\hat{\tau}_i = y_{1i} - \hat{y}_{0i} = y_{1i} - y_{0i'} \quad \text{if } T_i = 1, \quad (2a)$$

$$\hat{\tau}_i = \hat{y}_{1i} - y_{0i} = \hat{y}_{1i'} - y_{0i} \quad \text{if } T_i = 0, \quad (2b)$$

where i' indexes the match for observation i .

Prior to analyzing the distribution of the estimated treatment effect, we amend the nearest-neighbor algorithm in three directions. First, while a match exists for each observation, the propensity scores may be quite different across matched pairs. Because the unbiasedness of the matching estimator relies on the propensity scores being identical, pairs with scores too different are excluded, at the expense of some efficiency. Known as caliper matching, the choice of caliper is *ad hoc* (Cochran and Rubin 1973). We therefore utilize two values: 0.50 and 0.75.⁵

Second, we restrict the pool of potential matches for each observation to be from the same year and region (see figure 1). By matching within year and within region, we remove all time-specific and region-specific unobservables not explicitly controlled for by the propensity score. This is the matching method's version of a fixed effects panel data regression model, and is analogous to the claims made in Smith and Todd (forthcoming): matches used to identify the effect of employment programs should be from the same local labor market.

Finally, we employ a difference-in-differences matching estimator given that condition (1) may not hold in practice if we only observe a subset of the relevant conditioning variables, Z_0 , where $Z_0 \subset Z$. In other words, if unobservables exist which are correlated with both treatment assignment and the outcome of interest and are not removed by matching within-year and within-region, the Conditional Independence Assumption given by (1) will fail (the typical matching estimator is known

4 Typically, nearest-neighbor matching is performed *with replacement*, implying that a given control observation may be matched with multiple treatment observations. Dehejia and Wahba (2002) verify that matching with replacement fares at least as well as matching without replacement, and possibly better.

5 In previous work (List et al. 2003a, b), we utilize much smaller calipers. In the present exercise, such small calipers yield sample sizes which are too small in the final-stage regressions.

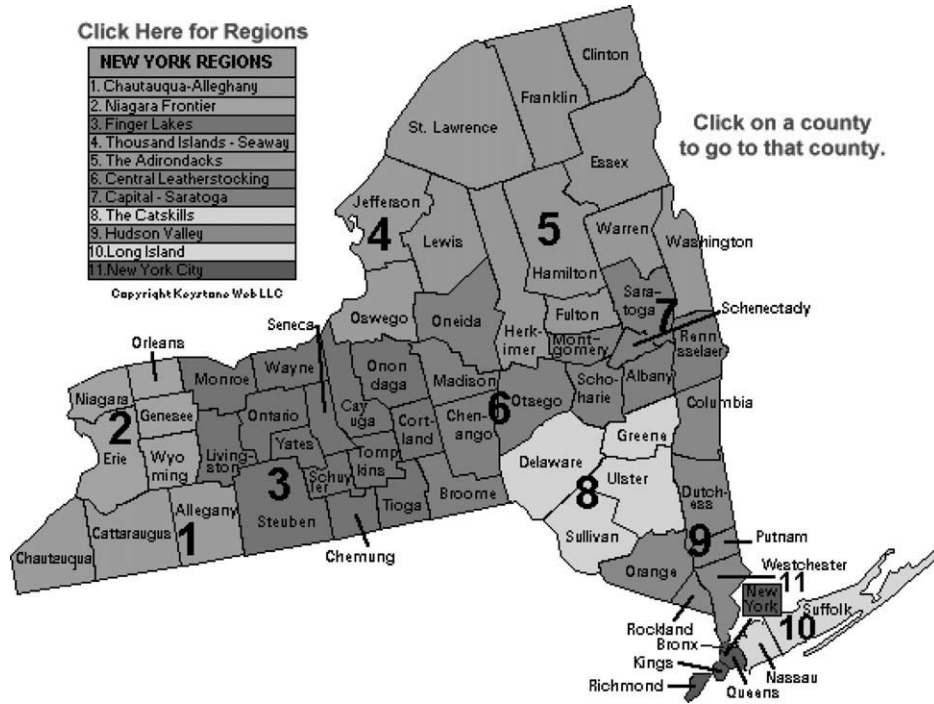


Figure 1. Regional Breakdown, New York State.

as a “selection on observables” estimator). Formally, (2a) and (2b) may re-written as

$$\hat{\tau}_i = y_{1i} - \hat{y}_{0i} = y_{1i} - y_{0i'} - (y_{0i} - y_{0i'}) = \tau_i + \underbrace{(y_{0i} - y_{0i'})}_{\text{BIAS}} \quad \text{if } T_i = 1, \quad (3a)$$

$$\hat{\tau}_i = \hat{y}_{1i} - y_{0i} = y_{1i'} - y_{0i} - (y_{1i} - y_{1i'}) = \tau_i + \underbrace{(y_{1i'} - y_{1i})}_{\text{BIAS}} \quad \text{if } T_i = 0, \quad (3b)$$

where the bias term represents the difference between the estimated and true missing counterfactual, and captures the difference in outcomes across matched pairs even if treatment assignment is identical. If the bias can be estimated, an improved estimator of the treatment effect is given by

$$\hat{\tau}_{i,DD} = \hat{\tau}_i - \hat{B} = (y_{1i} - y_{0i'}) - \hat{B} \quad \text{if } T_i = 1, \quad (4a)$$

$$\hat{\tau}_{i,DD} = \hat{\tau}_i - \hat{B} = (y_{1i'} - y_{0i}) - \hat{B} \quad \text{if } T_i = 0, \quad (4b)$$

where \hat{B} is an estimate of the bias.

Since we have counts of plants that vary in their level of pollution intensity, and the behavior of “clean” plants should not be affected by attainment status, any difference in the behavior of “clean” plants across matched pairs should reflect

unobservable, county-specific qualities that influence plant decisions but are not removed by conditioning on the propensity score, year, and region. Thus, differences in the behaviour of “clean” plants across matched pairs are used to estimate the bias in (4a), (4b). Consequently, the estimators in (4a) and (4b) become

$$\begin{aligned}\hat{\tau}_{i,DD} &= (y_{1i}^p - \hat{y}_{0i}^p) - (y_{1i}^n - \hat{y}_{0i}^n) = (y_{1i}^p - y_{0i'}^p) - (y_{1i}^n - y_{0i'}^n) \quad \text{if } T_i = 1, \\ &= \hat{\tau}_i - (y_{1i}^n - y_{0i'}^n)\end{aligned}\quad (5a)$$

$$\begin{aligned}\hat{\tau}_{i,DD} &= (\hat{y}_{1i}^p - y_{0i}^p) - (\hat{y}_{1i}^n - y_{0i}^n) = (y_{1i'}^p - y_{0i}^p) - (y_{1i'}^n - y_{0i}^n) \quad \text{if } T_i = 0, \\ &= \hat{\tau}_i - (y_{1i'}^n - y_{0i}^n)\end{aligned}\quad (5b)$$

where the p and n superscripts refer to pollution-intensive and non-pollution-intensive values of the outcome of interest, respectively.

After obtaining the observation-specific treatment effects in (5a), (5b), we examine the variation in $\hat{\tau}_{i,DD}$ to determine if it is systematically related to observable attributes, Z_0 , or the estimated propensity score, $p(Z_0)$. In this effort, we estimate several empirical specifications—separately on the sub-sample of treated and control observations—within a simple parametric framework where the estimated treatment effect constitutes the dependent variable. Specifically, in the final stage, we estimate specifications of the form

$$\hat{\tau}_{ik,DD} = \begin{cases} \alpha_{k0} + Z_{0ik}\beta_{k0} + \varepsilon_{i0} & \text{if } T_i = k, \quad k = 0, 1, \\ \alpha_{k1} + \beta_{k1}p(Z_{0ik}) + \varepsilon_{i1} \end{cases} \quad (6)$$

where β reflects the effect of the covariates on treatment effect heterogeneity as long as the covariates are independent of the difference between the actual bias in (3a), (3b) and the bias correction employed in (5a), (5b).

To see this final point, note that (5a), (5b) may be re-written as

$$\begin{aligned}\hat{\tau}_{i,DD} &= \hat{\tau}_i - (y_{1i}^n - y_{0i'}^n) \\ &= \tau_i + \underbrace{(y_{0i} - y_{0i'})}_{\text{BIAS}} - \underbrace{(y_{1i}^n - y_{0i'}^n)}_{\hat{B}} \quad \text{if } T_i = 1, \\ &= \tau_i + v_i\end{aligned}\quad (7a)$$

$$\begin{aligned}\hat{\tau}_{i,DD} &= \hat{\tau}_i - (y_{1i'}^n - y_{0i}^n) \\ &= \tau_i + \underbrace{(y_{1i'} - y_{1i})}_{\text{BIAS}} - \underbrace{(y_{1i'}^n - y_{0i}^n)}_{\hat{B}} \quad \text{if } T_i = 0. \\ &= \tau_i + v_i\end{aligned}\quad (7b)$$

Since the regressand in (6) is $\tau_i + v_i$, we implicitly maintain that v_i is uncorrelated with Z_0 and $p(Z_0)$; otherwise, the coefficients, β , will represent the combined effects of the covariates on τ_i and v_i . In addition, if $v_i \sim N(0, \sigma_v^2)$, then (6) reduces to the classical measurement error model, where measurement error in the dependent variable increases the standard errors, but does not alter the properties of the coefficient estimates.

Variables contained in Z_0 include the following county-level attributes: manufacturing employment, average manufacturing wage, population, per capita income,

property taxes, unemployment rate, and highway expenditures. This conditioning set is more exhaustive than those typically used in similar parametric analyses, and reflects variables that are commonly thought to affect plant location decisions. Thus, identification of the distribution of the treatment effects requires that there be no county-specific unobservables that are correlated with attainment status and differentially affect pollution-intensive versus non-pollution-intensive plants. As noted in Black and Smith (2004), this assumption is untestable. However, as noted in List et al. (2003), it is weaker than the identifying assumptions utilized in the past (e.g., Henderson (1996) assumes ozone attainment status is strictly exogenous conditional on a subset of the attributes in Z_0).

Prior to continuing, it is worth highlighting the benefits of the two-step matching approach relative to its parametric analog. The corresponding parametric model would be a dynamic fixed effects count data model.⁶ In addition, since matching implicitly relaxes the assumption of linearity of the controls, the parametric model would need to include higher order and interaction terms (Black and Smith 2004). Finally, the parametric model would need to incorporate data on the behavior of “clean” plants, as well as interactions between the difference-in-difference treatment effect estimate and the covariates examined in the second-stage analysis. While such a model is feasible in principal (see, e.g., Blundell et al. 2002), the computational advantages of matching are self-evident.

3. Empirical Results

3.1. Baseline Results

Prior to obtaining the primary results, we estimate the first-stage logit, where the dependent variable equals one if the county is out of attainment in the given year. The controls include the variables listed in Z in section 2, the square of each, a complete set of 21 interaction terms, and time dummies. The results, which are available upon request, indicate that the level of manufacturing employment, population, per capita income, property taxes, and many of the interaction terms are associated with the probability of being out of attainment.⁷ Furthermore, the time dummies are highly significant, and suggest a declining probability of being out of attainment over the time period analyzed.

In terms of the matching exercise, there are 506 (176) in (out of) attainment observations. Of the 506 (176) control (treated) counties, 398 (100) have no potential match, as the entire region is in (out of) attainment in a given year. As a

6 As shown in List et al. (2003), matching in these data also controls for lagged values of the outcome of interest.

7 Interesting to note is the fact that the probability of non-attainment follows an inverted-U pattern with respect to per capita income, consonant with the environmental Kuznets curve literature (see, e.g., Grossman and Krueger (1995) and more recently Millimet et al. (2003)).

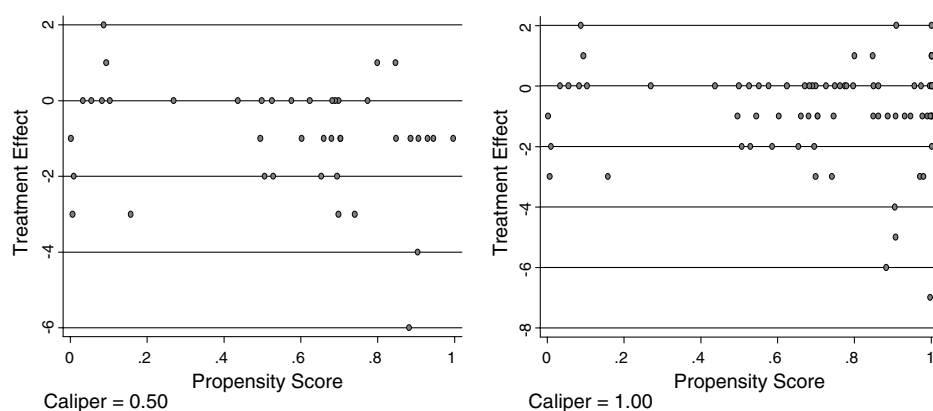


Figure 2. Distribution of Treatment Effects for Matched Treatment Observations: Pollution-Intensive Plant Births (Note: Estimated observation-specific treatment effect on the treated, obtained using equation (5a)).

result, we are left with at most 108 (76) possible matches from the control (treatment) group.

After matching the 76 treated observations with a viable match, the average difference in propensity scores across the pairs is 0.45. Retaining only those pairs whose propensity scores differ by less than 0.50 (0.75) restricts the sample to 42 (54) matched pairs. Similarly, after matching the 108 control observations with a feasible match, the average difference in propensity scores is 0.31. Discarding those pairs with propensity scores that differ by more than 0.50 (0.75) limits the sample to 69 (93) matched pairs.

To assess the variation in treatment effects, figures 2–5 plot estimates of τ_i against the propensity score. Figures 2 and 4 plot the treatment effects for the treated observations (out of attainment counties) when new plant births and plant exits are the outcomes of interest. Figures 3 and 5 plot the analogous results for the matched controls. Additionally, each figure contains two plots, one containing treatment effect estimates obtained only from matched pairs within the 0.50 caliper, and the second utilizing all matched pairs (an implicit caliper of unity).

All four figures indicate sizeable variation in the treatment effects across the observations utilized—variation lost when focusing solely on the *average* treatment effect on the treated (the usual parameter reported in matching studies; hereafter denoted as ATT) and the *average* treatment effect on the untreated (i.e., the expected cost of non-attainment on the control group; hereafter denoted as ATU). Specifically, using the 42 matched pairs retained in the 0.50 sample, the ATT is -0.95 (0.24) using plant births (exits) as the outcome of interest. Thus, the average out of attainment county forgoes nearly one new plant and loses roughly one-quarter of an existing plant each year due to its treatment designation. However, the cost of being out of attainment ranges from a *loss* of six to a *gain* of two new

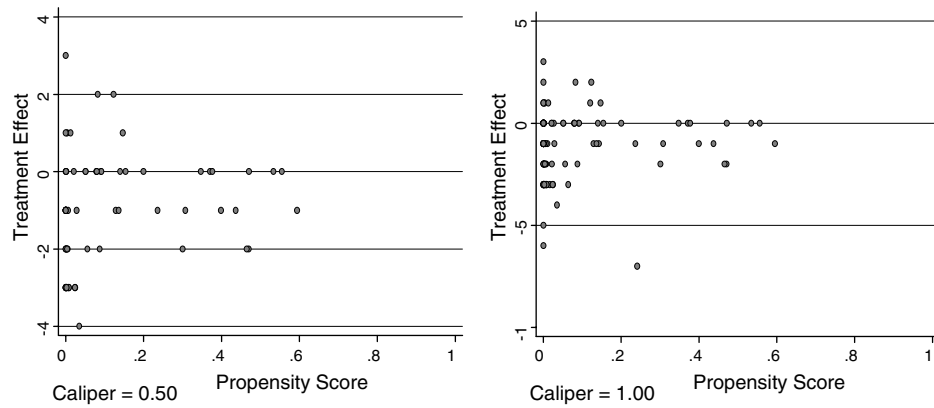


Figure 3. Distribution of Treatment Effects for Matched Control Observations: Pollution-Intensive Plant Births (Note: Estimated observation-specific treatment effect on the untreated, obtained using equation (5b)).

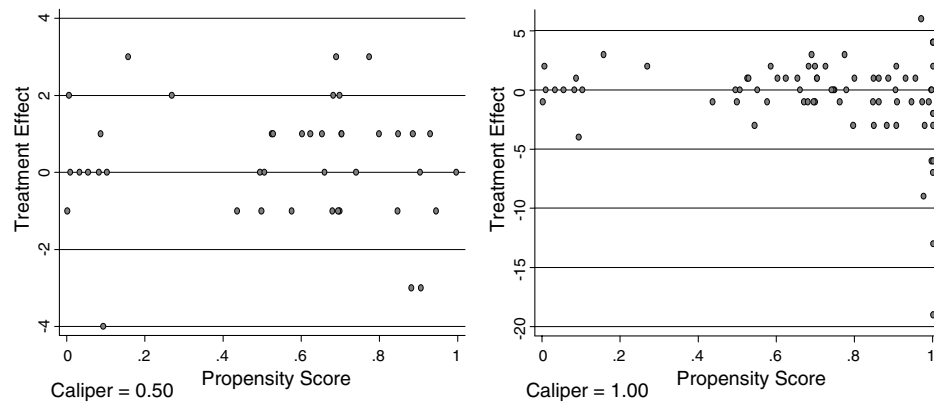


Figure 4. Distribution of Treatment Effects for Matched Treatment Observations: Pollution-Intensive Plant Exits (Note: See figure 2).

plant births, and from a *loss* of four to a *gain* of three plant exits.⁸ The ATT is -0.80 (-0.62) using all 76 possible matched pairs, with a range of -7 to $+2$ (-19 to $+6$).

For the control group, using the 69 matched pairs retained in the 0.50 sample, the ATU is -1.00 (0.57) using plant births (exits) as the outcome of interest. Thus,

⁸ One should be careful to realize that a treatment effect of, say, -4 plant exits implies that being out of attainment “saved” a county from four plant exits that otherwise would have occurred had the county been in attainment. Thus, saying that the treatment effect for exits varies from -4 to $+3$ means that being out of attainment *saved* some counties from losing up to four plants, and caused other counties to *lose* up to three plants.

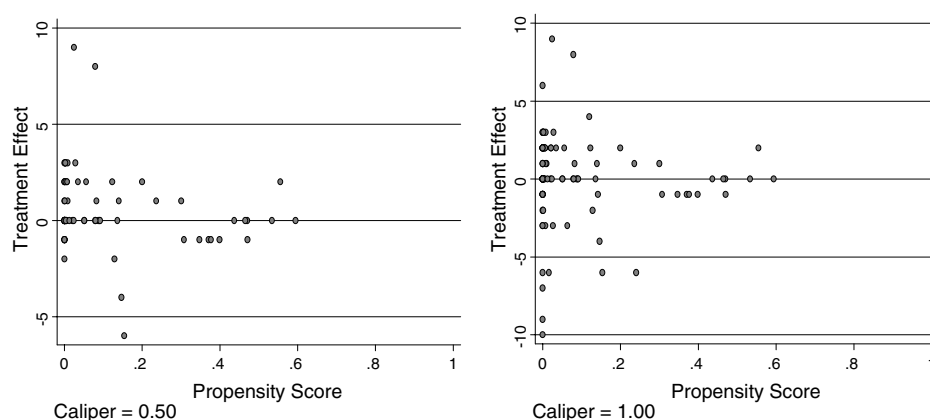


Figure 5. Distribution of Treatment Effects for Matched Control Observations: Pollution-Intensive Plant Exits (Note: See figure 2).

the average county that is in attainment would forego one new plant and would lose roughly one-half of an existing plant each year if its treatment designation was changed to non-attainment. As before, however, the predicted impact of being out of attainment varies, ranging from a *loss* of four to a *gain* of three new plant births, and from a *loss* of six to a *gain* of nine plant exits. The ATU is -1.04 (0.14) using all 108 possible matched pairs, with a range of -7 to $+3$ (-10 to $+9$).

The variation in treatment effects noted heretofore is especially extreme when one considers that the representative county obtains only 0.4 new pollution-intensive manufacturing plants per annum, and loses only 0.8 pollution-intensive manufacturing plants per annum. This result highlights the crux of our study: because the effect of more stringent regulations is extremely heterogeneous, focusing on the (conditional) average effect yields potentially unreliable inferences.

That being said, if the variation in the treatment effect represents purely random influences, then concentration on the average treatment effect (on the treated or untreated) is warranted, as local policymakers could not forecast their specific cost of non-attainment. If the variation is systematic, however, then policymakers could more accurately weigh the costs of increased local environmental stringency if the pattern could be discerned. Moreover, future research into the empirical validity of the PHH may benefit by focusing on locations most likely to incur a high cost of stringent regulations.

To analyze the spatial variation in the treatment effects, tables 2–5 contain the empirical results from several OLS regressions using the estimated treatment effect on the treated (or treatment effect on the untreated) as the dependent variable.⁹ Tables 2 (treated) and 3 (controls) present the results for the treatment effect on new

9 In all cases, we report standard errors that are robust to heteroskedasticity and within-county correlation.

plant births; tables 4 (treated) and 5 (controls) present the results for the treatment effect on plant exits. For each sample (discussed below), three specifications are estimated. The first specification includes manufacturing employment, manufacturing wages, population, per capita income, property taxes, the unemployment rate, and highway expenditures as covariates. The next specification adds lagged pollution-intensive plant births or exits to capture agglomeration externalities. In the final specification, we condition only on the propensity score. If the variation in the cost of being out of attainment represents purely random influences, one would expect low R^2 s and all of the coefficients to be statistically insignificant.

3.1.1. New Plant Births

Columns 1–3 (4–6) in table 2 present empirical results using the estimated treatment effect on the treated in the 0.50 (0.75) sample; columns 1–3 (4–6) in table 3 present the analogous results using the estimated treatment effect on the untreated. Examination of the diagnostics indicates that our specifications explain a fair amount of the variation in the regressands, especially when analyzing the variation in the treatment effect on the treated using the first and second specifications and the 0.50 caliper (columns 1 and 2, table 2), where the R^2 is 0.39 and 0.48, respectively.

In terms of coefficient estimates, we find a positive and statistically significant (at the 90% confidence level or better) association between the treatment effect on the treated and manufacturing employment, average manufacturing wage, and the unemployment rate. Alternatively, we find a negative association between population and the estimated treatment effect on the treated in the 0.50 sample using the first specification (column 1, table 2). The positive (negative) coefficients imply that *higher (lower)* levels of these variables are associated with a *reduction* in the number of foregone plant births as a result of being labeled out of attainment. Consequently, non-attainment counties with greater manufacturing employment, higher average manufacturing wage, higher unemployment rate, and smaller population forego fewer new plant births as a result of their attainment status.

The fact that greater employment helps mitigate the cost of stricter regulation is consonant with positive agglomeration externalities offsetting the added costs due to environmental regulation (Henderson 1997). Similarly, a higher unemployment rate helps offset the costs of environmental regulation through the presence of a larger pool of available labor. Interestingly, the fact that more stringent regulation imposes a lower cost when unemployment is high suggests that economic slowdowns, not booms, present a greater opportunity to improve environmental protection if curbing the expected decline of new plants is a policy objective.

On the other hand, the fact that counties with a larger population forego more new plant births, on average, as a result of being out of attainment may be attributable to a belief on the part of firms that larger counties will remain out of attainment for a longer duration. The sample correlation between non-attainment and population is 0.37, and the average population of the three counties out of attainment in all 11 years in the sample is 1.53 million, compared with an overall sample

mean of 0.29 million. Lastly, the fact that higher wages are associated with a lower cost of non-attainment is a bit perplexing. One possible explanation is that the wage proxies for the skill level of the labor force, where having access to more skilled workers helps offset the costs of stricter environmental regulation. We will return to this in section 4.

In addition to statistical significance, coefficient magnitudes are economically meaningful. For example, given that the ATT in the 0.50 sample is -0.95 , a “typical” county (in the treatment group) foregoes one new pollution-intensive manufacturing plant each year it is labeled out of attainment. However, if the typical county increases its manufacturing employment by roughly 16,000, or 11%, *ceteris paribus*, this negative effect is entirely offset. Similarly, an increase in the unemployment rate of 2.7% would also eliminate the cost of non-attainment.

In terms of the second specification, displayed in column 2 of table 2, all of the coefficients now become statistically insignificant, including the coefficient on lagged new plant births. However, the coefficients are similar to those in column 1, the R^2 increases, and the coefficients are jointly significant ($F = 10.19$, $p = 0.00$), suggesting problems with multicollinearity.

Columns 4 and 5 of table 2 display the results from the first two specifications using the 0.75 sample. While the signs and magnitudes of many of the coefficients are similar to those found in columns 1 and 2, the only statistically significant coefficient is that on the unemployment rate in column 4. Thus, while the offsetting effect of a higher unemployment rate remains, the other previous findings are not robust to increasing the caliper.

Examination of the results from the first two specifications using the estimated treatment effect on the untreated as the dependent variable (columns 1, 2, 4, and 5 in table 3) indicates that while the association between unemployment and the estimated treatment effect is statistically significant at conventional levels in three of the four models, no other coefficient is significantly different from zero. As in table 2, the coefficient is positive, suggesting that firms rationally trade off higher regulatory costs and the availability of labor. However, for the “typical” county that is presently in attainment, the unemployment rate would have to be approximately 15% (compared with a mean unemployment rate of 8.1%) to completely offset the expected negative effect of becoming out of attainment.

Lastly, empirical results contained in columns 3 and 6 in tables 2 and 3 are obtained from regressing the estimated treatment effects on the propensity score. The coefficient is statistically insignificant in all four cases. Nonetheless, we still believe this to be an interesting finding. We interpret the coefficient in this specification as reflecting the cost of uncertainty. In other words, as the propensity score captures the *ex ante* probability of a county being out of attainment, the coefficient in this specification reflects the cost of a higher *ex ante* probability on the cost of being out of attainment for a county that *ex post* is observed to be out of attainment (table 2) or in attainment (table 3). A positive (negative) coefficient would imply that counties with a higher probability of being out of attainment incur a lower (higher) cost if they are then found to actually be out of attainment.

The fact that we find insignificant effects suggests that a county's *ex ante* probability of being out of attainment has no bearing on its treatment effect once its attainment status is known.

3.1.2. Plant Exits

Tables 4 and 5 are identical in organization to the previous two tables, except plant exits is now the outcome of interest. As a result, a higher treatment effect signifies a greater exodus of firms from a county due to being labeled out of attainment. As before, columns 1–3 (4–6) in table 4 present empirical results using the estimated treatment effect on the treated in the 0.50 (0.75) sample; columns 1–3 (4–6) in table 5 present the analogous results using the estimated treatment effect on the untreated. Examination of the diagnostics indicates that again our specifications explain a fair amount of the variation in the treatment effects, especially when analyzing the variation in the treatment effect on the untreated using the first and second specifications (columns 1, 2, 4, and 5 in table 5), where the R^2 ranges from 0.40 to 0.55.

In terms of coefficient estimates, we find a negative and statistically significant (at the 90% confidence level or better) association between manufacturing employment and per capita income and the treatment effect on the treated, and a positive association between population and property taxes and the estimated treatment effect on the treated in the 0.50 sample using the first specification (column 1, table 4). The negative (positive) coefficients imply that *higher (lower)* levels of these variables are associated with a *reduction* in the number of plant exits as a result of being out of attainment, the converse from section 2. Consequently, non-attainment counties with greater manufacturing employment, higher per capita income, smaller population, and lower property taxes experience fewer plant exits as a result of their attainment status.

The fact that greater employment and smaller population reduce the number of exits is consonant with empirical results in table 2, where these attributes were found to also reduce the number of foregone new plant births. The positive effect of property taxes on exits due to non-attainment is intuitive, as this implies that counties with lower property taxes experience fewer exits as the lower taxes presumably help offset the costs of environmental regulation. Finally, the fact that counties with higher per capita income incur a smaller exodus of plants due to non-attainment status makes sense if income proxies for the local demand for manufacturing output, or the provision of public goods that individuals and firms both value (e.g., crime prevention).

As in section 2, coefficient magnitudes are economically meaningful. For example, given that the ATT in the 0.50 sample is 0.24, a “typical” county (in the treatment group) loses one-quarter of an existing pollution-intensive plant each year it is designated out of attainment. However, if the typical county increases its manufacturing employment by roughly 3,000, or 2%, *ceteris paribus*, this effect is entirely offset. Similarly, a decrease in per capita property taxes of

approximately \$25 per annum, or 5%, would also eliminate the cost of non-attainment.

Moving to the second specification, displayed in column 2 (table 4), all of the previous statistically significant coefficients remain significant. In addition, higher manufacturing wages are associated with more plant exits, and greater lagged plant exits is negatively correlated with current exits, due to non-attainment. Both results are somewhat surprising. First, while the positive association between manufacturing wages and the treatment effect is not surprising by itself, it is when one recalls that higher wages were found to be negatively associated with foregone new plant births in section 2. One explanation for this apparent inconsistency is that per capita income and manufacturing wages are highly correlated and, as a result, one should not view the effects of these two variables on the treatment effect in isolation.¹⁰ Whereas in table 2 only the wage coefficient is significant, in table 4 the coefficients on wages and per capita income are both significant (and of opposite sign). This suggests that higher wages are associated with a greater cost of non-attainment (fewer new plant births and more exits of existing plants) once one adequately controls for the income/human capital of the labor force.

The second result—the negative association between lagged plant exits and current plant exits due to non-attainment—is also peculiar at first glance. On the one hand, assuming positive agglomeration externalities, a greater exodus of plants in the past may reduce the incentives of current plants to remain in the face of stricter environmental regulation. However, since the negative effect is obtained holding manufacturing employment fixed, this is not a valid interpretation. Instead, a greater loss of plants in the previous period may imply greater (unobservable) concessions or promises of future concessions if existing firms do not exit.

Columns 4 and 5 in table 4 display the results from the first two specifications using the 0.75 sample. While the signs and magnitudes of many of the coefficients are similar to those found in columns 1 and 2, none of the results are statistically significant at conventional levels. As in section 2, the findings in columns 1 and 2 are not robust to increasing the caliper.

Examination of the results from the first two specifications using the estimated treatment effect on the untreated as the dependent variable (columns 1, 2, 4 and 5 in table 5) indicates that the association between unemployment, highway expenditures, and manufacturing employment and the estimated treatment effect is statistically significant at conventional levels in at least three of the four models. And, the association between manufacturing wages and population and the estimated treatment effect is statistically significant at conventional levels in two of the four models. As in tables 2 and 3, the coefficient on unemployment suggests that firms

10 In the full sample, the correlation between manufacturing wages and per capita income is 0.83.

rationally trade off higher regulatory costs and the availability of labor, as counties with higher unemployment would lose fewer existing plants if designated out of attainment. However, other statistically significant coefficients are inconsistent with results in the previous tables: manufacturing employment (population) is positively (negatively) associated with the number of exits, and greater highway expenditure is associated with more plant exits. The coefficient on lagged plant exits, on the other hand, is negative, consonant with the results reported for the treatment group in table 4.

The final set of empirical results is presented in columns 3 and 6 in tables 4 and 5, and is obtained by regressing the estimated treatment effects on the propensity score. The coefficient is negative and statistically insignificant for the control group only (columns 3 and 6, table 5). In these samples, $\hat{\beta} \approx -2.50$ implies, for example, that a county in attainment with a propensity score of 0.50 (i.e., *a priori* there is a 50% chance of being out of attainment) would lose more existing plants if it was actually designated out of attainment than a county with a propensity score of, say, 0.99 (i.e., *a priori* it is known with near certainty that the county will be out of attainment). In particular, the county with a propensity score of 0.50 will lose more than one additional existing pollution-intensive plant as a result of the stricter environmental regulation.

3.2. Sensitivity Analysis

In addition to the robustness checks already discussed, we perform two additional sensitivity analyses. First, because the estimated treatment effect for each county under single nearest-neighbor matching is an integer (positive or negative), we re-estimate the specifications using an ordered probit model as opposed to OLS. The results, which are available upon request, do not change qualitatively. Second, we re-do the matching using a different matching algorithm: radius matching. Whereas single, nearest-neighbor matching uses only the outcome of a single observation to estimate the missing counterfactual, radius matching uses the average outcome over all neighbors within a specified caliper to estimate the missing counterfactual. Formally, (5a), (5b) become

$$\hat{\tau}_{i,DD} = (y_{1i}^p - \hat{y}_{0i}^p) - (y_{1i}^n - \hat{y}_{0i}^n) = \left(y_{1i}^p - \frac{1}{J_i} \sum_{j=1}^{J_i} y_{0j}^p \right) - \left(y_{1i}^n - \frac{1}{J_i} \sum_{j=1}^{J_i} y_{0j}^n \right) \quad \text{if } T_i = 1, \quad (8a)$$

$$\hat{\tau}_{i,DD} = (\hat{y}_{1i}^p - y_{0i}^p) - (\hat{y}_{1i}^n - y_{0i}^n) = \left(\frac{1}{J_i} \sum_{j=1}^{J_i} y_{1j}^p - y_{0i}^p \right) - \left(\frac{1}{J_i} \sum_{j=1}^{J_i} y_{1j}^n - y_{0i}^n \right) \quad \text{if } T_i = 0, \quad (8b)$$

where J_i is the number of neighbors within the specified caliper.¹¹ As before, we utilize a caliper of 0.50 and 0.75. It is well known that radius matching trades off greater bias for increased efficiency relative to single, nearest-neighbor matching.¹²

The results are displayed in columns 7–12 in tables 2–5. While the statistical significance of certain coefficients is altered relative to the previous results based on single, nearest-neighbor matching (sometimes attaining significance at conventional levels, and at times losing significance), in general the results are robust, as the signs and magnitudes vary little.

4. Concluding Remarks

The onset of a chronic U.S. trade deficit following the beginning of the modern environmental movement led many to suspect that environmental regulation played a major causal role in impairing the “competitiveness” of U.S. firms (Jaffe et al. 1995). The debate rages on today, as conventional wisdom continues to be at odds with the plethora of empirical studies that observe an insignificant link between environmental regulations and competitiveness. Using a new empirical strategy, this study sheds light on these previous ambiguous results, as well as suggests possible remedies that localities may utilize to offset the negative effects of more stringent regulations on capital inflows.

Overall, our findings suggest that the estimated costs (in terms of foregone plant births or induced plant exits) of more stringent air quality controls are quite heterogeneous spatially. Our empirical estimates suggest that while certain regions will experience little, if any, deleterious effect, others will experience considerable losses. This finding has importance in both a positive and normative sense. Future research examining how environmental regulations affect polluters should attempt to sharpen insights by allowing the effects to be heterogeneous, both spatially and across industry. We suspect that once all heterogeneity is taken into account, the empirical findings will begin to match expectations.

References

- Becker, R., and J. V. Henderson. 2000. “Effects of Air Quality Regulations on Polluting Industries.” *Journal of Political Economy* 108(2): 379–421.
- Black, D. A., and J. A. Smith. 2004. “How Robust is the Evidence on the Effects of College Quality? Evidence from Matching.” *Journal of Econometrics* 121(1–2): 99–124.

11 Note that with radius matching, the averaging used to estimate the missing counterfactual implies that the estimated treatment effects will no longer assume only integer values.

12 Radius matching involves greater bias since it utilizes information from observations with propensity scores that are further from the propensity score of the observation being matched. However, since information is being utilized from more of the sample, efficiency is improved.

- Blundell, R., R. Griffith, and F. Windmeijer. 2002. "Individual Effects and Dynamics in Count Data Models." *Journal of Econometrics* 108(1): 113–131.
- Cochran, W. G., and D. B. Rubin. 1973. "Controlling Bias in Observational Studies: Review." *Sankhyā-A* 35(4): 417–446.
- Dehejia, R. H., and S. Wahba. 2002. "Propensity Score Matching for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84(1): 151–161.
- Ederington, J. 2001. "International Coordination of Trade and Domestic Policies." *American Economic Review* 91(5): 1580–1593.
- Fredriksson, P. G., J. A. List, and D. L. Millimet. 2003. "Bureaucratic Corruption, Environmental Policy and Inbound US FDI: Theory and Evidence." *Journal of Public Economics* 87(7–8): 1407–1430.
- Greenstone, M. 2002. "The Impact of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufacturers." *Journal of Political Economy* 110(6): 1175–1219.
- Grossman, G. M., and A. B. Krueger. 1995. "Economic Growth and the Environment." *Quarterly Journal of Economics* 110(2): 353–377.
- Henderson, J. V. 1996. "Effects of Air Quality Regulation." *American Economic Review* 86(4): 789–813.
- Henderson, J. V. 1997. "Externalities and Industrial Development." *Journal of Urban Economics* 42(3): 449–470.
- Jaffe, A. B., S. R. Peterson, P. R. Portney, and R. N. Stavins. 1995. "Environmental Regulation and the Competitiveness of US Manufacturing: What Does the Evidence Tell Us?" *Journal of Economic Literature* 33(1): 132–163.
- Jeppessen, T., J. A. List, and H. Folmer. 2002. "Environmental Regulations and New Plant Location Decisions: Evidence from a Meta-Analysis." *Journal of Regional Science* 42(1): 19–49.
- List, J. A., D. L. Millimet, P. G. Fredriksson, and W. W. McHone. 2003a. "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator." *Review of Economics and Statistics* 85(4): 944–952.
- List, J. A., D. L. Millimet, and W. W. McHone. 2003b. "Effects of Air Quality Regulation on the Destination Choice of Relocating Firms." *Oxford Economic Papers* 55(4): 657–678.
- Millimet, D. L., J. A. List, and T. Stengos. 2003. "The Environmental Kuznets Curve: Real Progress or Misspecified Models?" *Review of Economics and Statistics* 85(4): 1038–1047.
- Mulatu, A., R. Florax, and C. Withagen. 2002. "Environmental Regulations and Competitiveness." *Discussion Paper #01–139/3*. Tinbergen Institute.
- Rosenbaum, P. R., and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1): 41–55.
- U.S. House of Representatives, 1979. "Legislative History of the Clean Air Act," House Report 91-1146, U.S. 6PO.
- Wooldridge, J. M., 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.