

EFFECTS OF ENVIRONMENTAL REGULATIONS ON MANUFACTURING PLANT BIRTHS: EVIDENCE FROM A PROPENSITY SCORE MATCHING ESTIMATOR

John A. List, Daniel L. Millimet, Per G. Fredriksson, and W. Warren McHone*

Abstract—This study examines the effects of air quality regulation on economic activity. Anecdotal evidence and some recent empirical studies suggest that an inverse relationship exists between the stringency of environmental regulations and new plant formations. Using a unique county-level data set for New York State from 1980 to 1990, we revisit this conjecture using a semiparametric method based on propensity score matching. Our empirical estimates suggest that pollution-intensive plants are responding to environmental regulations; more importantly, we find that traditional parametric methods used in previous studies may dramatically understate the impact of more stringent regulations.

I. Introduction

WHETHER air quality regulation influences the formation of capital merits serious consideration. In the United States, the spirit of the Clean Air Act was to clean up source emissions rather than influence the industrial makeup of certain regions. In this sense, a finding that federal air quality regulation has influenced capital flows is against the fundamental intentions of the Clean Air Act. Indeed, an important factor leading to the creation of the Environmental Protection Agency in 1968, which resulted in a dominating federal presence in environmental policy, was that local discretion over environmental regulation may induce a “race to the bottom” to attract mobile capital.¹ Although anecdotal evidence from the popular press certainly suggests that capital responds to heterogeneous environmental standards, empirical estimates in the received literature, which range from positive and significant to negative and significant, lead most to conclude that the relationship is weak at best (see, for example, Jaffe et al., 1995).

Recently, however, several influential studies have provided strong evidence indicating that federal air quality regulations do indeed affect capital flows in pollution-intensive manufacturing plants (Henderson, 1996; Greenstone, 1998; Becker and Henderson, 2000). Yet, each study maintains the assumption that stringency of local environmental regulations, determined by county-level attainment status of the primary federal standard for various pollutant types, is strictly exogenous, lending the studies a setting

akin to a “natural experiment.”² Because intuition suggests that, for example, the location of new polluting plants in areas currently in attainment leads to higher pollution levels, and subsequently more stringent regulation if the attainment status threshold is bypassed, the assumption that attainment status is strictly exogenous is not trivial. Although quite convenient, given the county air quality distributions presented in Henderson (1996, figures 1 and 2), one could reasonably conclude that a fair number of counties are on the cusp of the nonattainment threshold. Thus, a few new pollution-intensive plants in some counties today could shift their regulatory status to out-of-attainment next period.³

The goal of this study is to examine the effects of air quality regulation on new-plant formation using a semiparametric propensity score matching method. Whereas matching methods are applicable primarily to problems of selection on observables, we employ a difference-in-difference matching estimator, similar to the efforts in Heckman, Ichimura, and Todd (1997) and Smith and Todd (2000), to control for the presence of unobservables that under normal circumstances may lead to biased estimates. As a point of reference, we compare empirical estimates from the matching method with various parametric specifications. In particular, we relax the exogeneity assumption imposed in the earlier analyses of the effect of attainment status on plant location by estimating a two-step fixed-effects Poisson model, similar in spirit to Mullahy (1997) and Windmeijer and Silva (1997). We also estimate parametric models on the data subsamples obtained from the matching algorithm to investigate differences between the various estimates.

Using a unique county-level panel data set on the location decisions of manufacturing plants from 1980 to 1990 in New York State, we obtain several interesting results. First, using parametric methods on the full data sample, we find little evidence that attainment status has a systematic effect on the location decisions of plants, regardless of the treatment of attainment status as exogenous or endogenous. Second, empirical estimates from various propensity score matching algorithms provide strong evidence that environmental regulations have statistically and economically significant effects on pollution-intensive plant formation, even after differencing out unobservables not captured by the propensity score method. This result suggests that the

Received for publication December 19, 2000. Revision accepted for publication September 20, 2002.

* University of Maryland, Southern Methodist University, Southern Methodist University, and University of Central Florida, respectively.

We thank Robert Pindyck and two anonymous reviewers for very thoughtful comments. Michael Greenstone, Shelby Gerking, and Scott Taylor also provided useful remarks. Thanks to seminar participants at the University of Maryland, the University of Wyoming, the University of Arizona, and the “International Dimension of Environmental Regulations” conference in Kerkrade, The Netherlands.

¹ The term “race to the bottom” refers to a situation where localities continually try to undercut one another in their pursuit of firms and jobs, resulting in progressively weaker environmental regulations.

² For a discussion, and use, of “natural” and “field” experiments see, for example, List (2001).

³ The importance of the endogeneity issue is highlighted in Henderson (1996), who painstakingly attempts to ensure exogeneity in his regression models (see especially pp. 803–805).

effects of environmental regulation on new-plant formation may be considerably larger than previously reported, and the difference is *not* due to the treatment of attainment status as strictly exogenous. Moreover, our empirical estimates strongly indicate that current U.S. environmental air quality regulations may be leading to a “graying” process whereby counties historically free of pollution become havens for polluters. Finally, as the parametric models treating attainment status as either exogenous or endogenous, estimated on the matched subsamples (obtained from the matching method), approach the matching estimates (but remain smaller), we conclude that sample composition only plays a partial role in explaining the estimation differences.

II. Data and Empirical Methodology

A. The Data

The plant location data come from the comprehensive Industrial Migration File (IMF) that was maintained until 1990 by the New York State (NYS) Department of Economic Development (DED). The intent of the IMF was to monitor all gains and losses in manufacturing activity in NYS by county on an annual basis. The data units in the IMF are case observations of individual plant openings, closings, expansions, and contractions. 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 their inclusion in the data set.

A major advantage of these data is that case-specific information is not suppressed (Michalke, 1986). Thus, there are opportunities for improving precision and data analysis not heretofore available even from the Longitudinal Research Database (McGuckin, 1990). Furthermore, data are available annually—many recent studies on manufacturing activity have extrapolated information from the Census of Manufacturers database, which is only available in five-year intervals. Although Census data are extremely comprehensive, this limitation is unfortunate, as many plants start up and die within a five-year window. For example, our computations yield hazard rates amongst new plants in the 25%–30% range, consonant with the published literature.⁴

Our data are not without cost, however. One shortcoming is that since no statute exists requiring 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. A second limitation is that the data are only available for NYS. However, we view the results as indicative of the general process underlying capital location decisions throughout the United States.

The IMF classifies plant activity by Standard Industrial Classification (SIC) code. Following previous studies, we focus on plants 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 (1998) and classify sectors based on their emission levels. Using information from EPA’s Sector Notebook Project, we label 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 area in SIC codes 2611–2631, 2711–2789, 2812–2819, 2861–2869, 2911, 2930, 2932, 3312–3313, 3321–3325, 3334, and 3371.

After classifying new plants, the data are aggregated to the county level. In total, we observe location decisions of 280 pollution-intensive plants across the 62 counties in NYS over the sample period. We then combine this measure with each county’s ozone attainment designation, as well as other county-level attributes (discussed below).⁶ Over the sample period, slightly more than 25% of the county observations are out of attainment. Summary statistics are provided in table A1 in the appendix.

B. The Empirical Models

Propensity Score Matching Method: A method of assessing the effect of environmental regulations on new-plant location patterns heretofore not utilized is the method of propensity score matching developed in Rosenbaum and Rubin (1983). It is extensively used by statisticians, but economic applications have been sparse until recently. A few notable examples include Heckman et al. (1997), Dehejia and Wahba (1999, 2002), and Smith and Todd (2000).

The fundamental problem in identifying treatment effects is one of incomplete information. Though the econometrician observes whether the treatment occurs and the outcome conditional on treatment assignment, the counterfactual is not observed. Let y_{i1} denote the outcome of observation i if

⁵ The 1977 Clean Air Act Amendments set standards on five criteria air pollutants: sulfur dioxide (SO₂), carbon monoxide (CO), ozone (O₃), nitrogen oxides (NO_x), and total suspended particulates (TSP). Since ozone has attracted the most regulatory attention due to the limited progress that has been made in reducing concentration levels, we follow Henderson (1996) 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 (3) are in (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).

⁴ Without a grasp of the underlying spatial distribution of such births, one cannot predict, a priori, the magnitude or sign of the bias (if any).

the treatment occurs (given by $T_i = 1$), and y_{i0} denote the outcome if the treatment does not occur ($T_i = 0$). If both states of the world were observable, the average treatment effect, τ , 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 advocated in Rosenbaum and Rubin (1983) is to find a vector of covariates, Z , such that

$$y_1, y_0 \perp T|Z, \quad \text{pr}(T = 1|Z) \in (0, 1), \quad (1)$$

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

$$\begin{aligned} E[y_0|T = 1, Z] &= E[y_0|T = 0, Z] = E[y_0|Z], \\ \text{pr}(T = 1|Z) &\in (0, 1), \end{aligned} \quad (1')$$

is required. To implement the matching technique, the treatment group is defined as the set of counties labeled out-of-attainment in a given year. For condition (1') to hold, the conditioning set Z should be multidimensional. Consequently, finding observations with identical values for all covariates in Z may be untenable. Rosenbaum and Rubin (1983) prove, however, that conditioning on $p(Z)$ is equivalent to conditioning on Z , where $p(Z) = \text{pr}(T = 1|Z)$ is the propensity score. $p(Z)$ is estimated via logit.

Upon estimation of the propensity score, a matching algorithm is defined in order to estimate the missing counterfactual, y_{0i} , for each treated observation i . The simplest algorithm is nearest-neighbor matching, whereby each treated observation is paired with the control observation whose propensity score is closest in absolute value (Dehejia and Wahba, 2002).⁷ Unmatched controls are discarded. The matching method, therefore, identifies a restricted control group that better approximates the treatment group in terms of pretreatment attributes. The treatment effect on the treated (TT) is given by

$$\begin{aligned} \tau_{\text{TT}} &= E[y_1|T = 1, p(Z)] - E[y_0|T = 0, p(Z)] \\ &= E[y_1 - y_0|p(Z)]. \end{aligned} \quad (2)$$

We amend the nearest-neighbor algorithm in three directions. First, though a match exists for each out-of-attainment county, the propensity scores may still be quite different. Because the unbiasedness of the matching estimator relies on the propensity scores being identical, pairs with scores significantly different are excluded. This is known as *caliper matching* (Cochran and Rubin, 1973). Defining “signif-

icantly different” is arbitrary, however. As noted in Dehejia and Wahba (2002), relaxing the definition permits more pairs to be retained (increasing efficiency), but at the expense of introducing greater bias. Consequently, we present estimates for two cutoff values.

Second, to take advantage of 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 perform the matching exercise three times: first, restricting matched pairs to be from the same year; second, restricting matched pairs to be from the same year and same region of the state (see figure A1 in the appendix); and, finally, restricting matched pairs to be the same county from a different year. By matching within year, within year and 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 to fixed effects, and is similar to the claims made in Smith and Todd (2000): matches used to identify the effect of employment programs should be from the same local labor market. Thus, the estimator in (2) becomes

$$\begin{aligned} \tau_{\text{TT},t} &= E[y_1|T = 1, p(Z), t] - E[y_0|T = 0, p(Z), t] \\ &= E[y_1 - y_0|p(Z), t], \end{aligned} \quad (3a)$$

$$\begin{aligned} \tau_{\text{TT},r} &= E[y_1|T = 1, p(Z), t, r] - E[y_0|T = 0, p(Z), t, r] \\ &= E[y_1 - y_0|p(Z), t, r], \end{aligned} \quad (3b)$$

$$\begin{aligned} \tau_{\text{TT},i} &= E[y_1|T = 1, p(Z), i] - E[y_0|T = 0, p(Z), i] \\ &= E[y_1 - y_0|p(Z), i], \end{aligned} \quad (3c)$$

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

Finally, we complement these main results by employing a difference-in-differences (DID) matching estimator. Since we have a count of both pollution-intensive and non-pollution-intensive plant formations for each county-year observation, and the location decisions of “clean” plants should not be affected by attainment status, any differences in the birth of “clean” plants across the matched treatment and control groups is assumed to reflect unobservable county-specific qualities that are attractive to new plants. Thus, in the spirit of similar estimators used in Smith and Todd (2000) and Ham, Li, and Reagan (2001), we define the DID counterpart to (3a)–(3c) as

$$\tau_{\text{DID},t} = \tau_{\text{TT},t} - \tau'_{\text{TT},t}, \quad (4a)$$

$$\tau_{\text{DID},r} = \tau_{\text{TT},r} - \tau'_{\text{TT},r}, \quad (4b)$$

$$\tau_{\text{DID},i} = \tau_{\text{TT},i} - \tau'_{\text{TT},i}, \quad (4c)$$

where $\tau'_{\text{TT},t}$ ($\tau'_{\text{TT},r}$, $\tau'_{\text{TT},i}$) is the mean difference in the birth of “clean” plants across the matched treatment and control groups. As the DID estimator only requires

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 (2002) verify that matching with replacement faces at least as well as matching without replacement, and possibly better.

$$\begin{aligned} E[y_0 - y'_0 | T = 1, Z] &= E[y_0 - y'_0 | T = 0, Z] \\ &= E[y_0 - y'_0 | Z] \Pr(T = 1 | Z) \in (0, 1) \end{aligned} \quad (1'')$$

for identification, where y'_0 is the count of new “clean” plants, Smith and Todd (2000) conclude that DID matching estimators are more robust. Below we present estimates of the estimators in (3a)–(3c) and (4a)–(4c).

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 subsample. Thus, after matching, we also test for differences in the mean of the Z 's. The specification test proposed by Ham et al. (2001) requires testing for mean differences in the lagged outcome across the matched treatment and control groups. In the present context, this test serves two purposes. First, if the lagged outcome differs across the treatment and control groups, that suggests the presence of uncontrolled unobservables that may bias the estimated treatment effect. Second, since lagged plant births may affect current plant births due to agglomeration externalities, for example, then any differences in lagged births may have a direct effect on the outcome as well, further biasing the estimator.

Parametric Approach: For comparison purposes, we also estimate several parametric models derived from the partial equilibrium framework in Henderson (1996) and Becker and Henderson (2000).⁸ In equilibrium, the count of new (pollution-intensive) plants in county i at time t , Y_{it} ($Y \in \{0, 1, 2, \dots\}$), is given by

$$Y_{it} = \exp(X_{it}\beta) \eta_{it} + \epsilon_{it} \quad (5)$$

where X_{it} is a vector of county attributes, including attainment status, η_{it} captures all unobservable, time-varying attributes of county i and may be correlated with some of the variables in X_{it} , and ϵ_{it} is an error term satisfying $E[\epsilon | X, \eta] = 0$. Given the inclusion of a constant in X_{it} , we can assume $E[\eta] = 1$ without loss in generality (Mullahy, 1997).

Equation (5) can be estimated via the fixed-effects (FE) Poisson model of Hausman, Hall, and Griliches (1984; hereafter HHG). Resulting estimates of β will be consistent if $E[\eta | X] = E[\eta]$; in other words, regressors in X may be correlated with *time-invariant*, county-specific unobservables, but not *time-specific*, county-specific unobservables.⁹ Even if this assumption holds, standard errors from the Poisson model are sensitive to over- or underdispersion in

the data. Thus, we also estimate negative binomial models, which relax the restriction of equality of the conditional mean and variance functions.

If $E[\eta | X] \neq E[\eta]$, then another estimator is needed. Mullahy (1997) and Windmeijer and Silva (1997) discuss various instrumental variables (IVs) and two-step solutions, given a sufficient number of instruments, contained in the vector W_{it} . IV solutions are typically estimated using a generalized method of moments (GMM) framework. Two-step estimators involve, in the case of continuous endogenous regressors, either replacing endogenous variables with their predicted values from a first-stage regression (and adjusting the standard errors), or replacing X with W in (5) and backing out an estimate of β (and its covariance matrix) using a minimum-distance (MD) estimator.

If the source of endogeneity is a dichotomous treatment variable, as is the case with county-level attainment status, the two-step solution is not as straightforward. Simply replacing the attainment dummy with its predicted probability will not produce consistent estimates of β (Windmeijer and Silva, 1997). If one envisions plant location decisions as dependent upon the latent variable underlying attainment status, however, then consistent estimates of the effect of that latent variable is obtained by replacing the treatment variable in (5) with its predicted linear index. The estimating equation then becomes

$$Y_{it} = \exp(\delta T_{it}^* + X'_{it}\beta') \eta_{it} + \epsilon_{it}, \quad (5')$$

where T_{it}^* is a latent variable such that if $T_{it}^* > 0$ then the county is out of attainment ($T_{it} = 1$), and if $T_{it}^* \leq 0$ then it is in attainment ($T_{it} = 0$). Therefore δ is the parameter of interest. X'_{it} contains the remaining variables in X_{it} excluding attainment status, and β' is the corresponding parameter vector. Consistent estimates can be obtained by replacing T^* with $W\gamma$, where $T_{it}^* = W_{it}\gamma + v_{it}$. Assuming γ is unknown, it can be estimated via logit or probit (and the second-stage standard errors must be adjusted).

Modeling plant births as a function of latent attainment status is perhaps a more appealing way of organizing location decisions in the present context. If, for example, a county is currently in attainment, but current air quality is near the federally prescribed threshold, then a plant may still opt not to start up, *ceteris paribus*, to avoid potentially more stringent regulations in the future.

Prior to continuing, it is important to highlight the differences between matching estimators and the parametric Poisson estimators from the previous section. On the positive side, the matching estimator entails relatively few distributional assumptions. Moreover the matching estimator allows one to use additional endogenous variables that are difficult to incorporate into standard parametric count models: lagged values of the dependent variable and current and lagged values of “clean” manufacturing plant births. In a parametric world, one must rely on valid exclusion restrictions that are uncorrelated with these omitted

⁸ Note that we have estimated this parametric model elsewhere with these data, but they are parsed somewhat differently (see, for example, List & McHone, 2000).

⁹ 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. The breakdown comes from <http://visitnewyorkstate.net/regions>.

endogenous variables. Finally, matching estimators identify a restricted subsample of control observations that are most “similar” to the treatment group, whereas parametric models utilize all available observations. The major disadvantages are that: (i) matching estimators yield an estimate of the treatment effect on the treated (TT), not the expected treatment effect for an observation chosen at random, and (ii) every treated observation may not have a match.

III. Empirical Results

A. Propensity Score Matching Estimates

Table A2 in the appendix presents the first-stage logit estimates used to form the propensity score. The specification is similar to the first-stage equation used in the two-step HHG model (discussed below) with the inclusion of higher-order terms and interactions to facilitate the balancing of the covariates across the matched treatment and control groups (Dehejia & Wahba, 2002).

The first matching algorithm pairs each out-of-attainment county with the in-attainment county from the same year having the nearest propensity score. Of the 176 treatment (out-of-attainment) observations, we retain those with propensity scores that differ by less than 1% or 5%. Under these cutoffs, 37 and 81 matched pairs are formed. Using the same cutoffs but restricting matches to be from the same year and region of NYS (see figure A1) reduces the number of matched pairs to 8 and 16, respectively. Finally, matching each treatment county to itself at a different point in time and using the same cutoffs yields 9 and 11 matches, respectively.¹⁰

Table 1 presents estimated mean differences between the treatment and control groups, along with p -values associated with the null that the means are equal, for each of the six matched subsamples. Examination of the table yields three important insights. First, within-year and within-county matching algorithms balance the mean of all the covariates (at the $p < 0.10$ level) across the treatment and control groups using the 1% cutoff; the within-year, within-region algorithm balances the means using the 5% cutoff. This includes the variables specifically controlled for in the first-stage logit, as well as property taxes, the proportion of the population with a high school diploma, and highway expenditure. Thus, these algorithms satisfy the balancing test.

Second, the three algorithms passing the balancing test also pass the specification test proposed in Ham et al. (2001): lagged pollution-intensive births and lagged net births (defined as dirty births minus clean births) are balanced as well. Finally, of these three algorithms, two (within-year and within-county) yield negative, statistically significant (at the $p < 0.10$ level) estimates of the TT ($\tau_{TT,t} =$

-0.32 , $p = 0.08$; $\tau_{TT,i} = -1.33$, $p = 0.09$); the within-year, within-region estimate, though negative, is not significant ($\tau_{TT,r} = -0.19$, $p = 0.60$). However, as aforementioned, a more robust estimate is the DID estimator. Again, two of the three DID estimators (the within-year, within-region one and the within-county one) yield negative, statistically significant estimates ($\tau_{DID,r} = -0.69$, $p = 0.05$; $\tau_{DID,i} = -1.33$, $p = 0.03$); the within-year estimate is negative, but not significant ($\tau_{DID,t} = -0.35$, $p = 0.27$).

Given the robustness of the DID estimators, and recognizing that the within-year, within-region matching algorithm is more likely to yield an unbiased estimate of the treatment effect than matching within year alone—since it removes any region-specific unobservables that may affect the location decisions of pollution-intensive plants not already removed through the use of net births as the outcome measure (e.g., differences in political activism across regions)—implies an estimated “cost” of being out of attainment that is between 0.7 and 1.3 new plants per year. As the average county obtains 0.4 new (pollution-intensive) plants per year, this represents a sizable percentage loss.

B. Parametric Results

To facilitate comparison with the received literature, we begin by estimating (5), treating attainment status as exogenous. Choosing the most appropriate specification is difficult in view of the numerous specifications that have been utilized (see, for example, Jeppessen, List, & Folmer, 2002). To provide a fair comparison, we estimate four specifications. The baseline specification follows Henderson (1996) and includes county FEs, county-level attainment status, and a measure of scale (manufacturing employment) as the determinants of new-plant formations. We then add additional controls for real manufacturing wages, population, and real property taxes. Finally, we reestimate the previous two specifications including time dummies.

Before discussing coefficient estimates, we should note that the empirical results from Poisson and negative-binomial specifications are generally similar, so we focus on the HHG estimates.¹¹ Columns (1) and (2) in table 2 presents empirical results from the specifications treating attainment status as exogenous and including time effects. Empirical results suggest that attainment status is a significant determinant of new-plant formations in pollution-intensive industries, suggesting that being out of attainment reduces the flow of births by approximately 50%.¹² Since the mean number of new pollution-intensive plants is 0.4, these estimates imply that being out of attainment results in an average annual loss of nearly 0.2 new plants. Although economically significant, the point estimates are considerably

¹⁰ The number of “unique” controls used is 33 and 44 for the within-year algorithm (1% and 5% cutoff, respectively), and 8 and 15 (6 and 7) for the within-year, within-region (the within-county) algorithm.

¹¹ For brevity, we only present selected results here. All results not shown are available at <http://faculty.smu.edu/millimet/pdf/ny1results.pdf>.

¹² We should note that in the specifications omitting time effects (results not shown), attainment status coefficients are negative and significantly different from zero at the $p < 0.05$ level in both specifications.

TABLE 1.—PROPENSITY SCORE ESTIMATES OF ATTAINMENT-STATUS EFFECT

Independent Variable	Matching Algorithm					
	Within Year Max. Difference		Within Region & Year Max. Difference		Within County Max. Difference	
	(0.01)	(0.05)	(0.01)	(0.05)	(0.01)	(0.05)
Propensity score	−0.00 (0.99)	0.00 (0.97)	0.00 (0.98)	0.00 (0.98)	0.00 (1.00)	0.01 (0.97)
New dirty plants ($\tau_{TT,}$)	−0.32 (0.08)	−0.69 (0.00)	0.38 (0.25)	−0.19 (0.60)	−1.33 (0.09)	−1.18 (0.07)
New clean plants	0.03 (0.95)	−0.59 (0.08)	1.25 (0.07)	0.50 (0.36)	0.00 (1.00)	−0.18 (0.84)
Net new plants ($\tau_{DID,}$)	−0.35 (0.27)	−0.10 (0.68)	−0.88 (0.12)	−0.69 (0.05)	−1.33 (0.03)	−1.00 (0.08)
Lagged new dirty plants (1 year)	−0.07 (0.79)	−0.06 (0.70)	0.71 (0.08)	0.43 (0.10)	1.00 (0.12)	1.04 (0.05)
Lagged net new plants (1 year)	0.53 (0.31)	0.71 (0.04)	0.50 (0.41)	0.44 (0.43)	0.00 (1.00)	−0.14 (0.74)
Man. wages (\$1000s)	−0.73 (0.33)	−0.20 (0.66)	−0.06 (0.98)	−0.91 (0.44)	0.54 (0.60)	−0.01 (0.99)
Man. employment (\$1000s)	−38.86 (0.27)	−52.88 (0.07)	29.94 (0.63)	4.05 (0.93)	3.11 (0.98)	2.53 (0.98)
Man. plants	−0.72 (0.52)	−0.76 (0.32)	−0.79 (0.82)	−2.37 (0.26)	0.59 (0.70)	0.48 (0.74)
Population (1000s)	−53.91 (0.50)	−40.74 (0.57)	59.49 (0.55)	4.61 (0.96)	−0.65 (1.00)	−0.31 (1.00)
Per capita income (\$1000s)	−0.09 (0.89)	0.15 (0.72)	−0.66 (0.79)	−0.61 (0.66)	0.33 (0.84)	−0.20 (0.90)
Property tax	−31.38 (0.40)	7.85 (0.73)	−389.13 (0.06)	−186.81 (0.10)	1.22 (0.98)	1.00 (0.98)
High school graduates (%)	−1.10 (0.34)	−0.85 (0.29)	−3.61 (0.32)	−3.39 (0.12)	−1.09 (0.70)	−0.89 (0.71)
Highway expenditure	−0.01 (0.38)	0.01 (0.31)	−0.16 (0.09)	−0.07 (0.16)	−0.00 (0.97)	−0.00 (0.92)
Number of matched pairs	37	81	8	16	9	11
Number of unique controls	33	44	8	15	6	7

Entries 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.

“Dirty” plants are those defined as pollution-intensive (see text); “clean” are all remaining manufacturing plants.

“Unique controls” reports the number of control counties that are matched with at least one treatment county.

smaller than the matching estimates presented in the previous section.

Our initial supposition as to why the HHG estimates are 3.5 to 6.5 times smaller in magnitude than the matching estimates is that the parametric estimates are biased due to the endogeneity of attainment status. To explore this hypothesis, we estimate the two-step model in (5'). Consistency of the estimates relies on finding a valid instrument for attainment status. Because attainment status depends on the lagged level of air quality, which in turn depends on lagged manufacturing activity, which in turn may affect plant startup decisions on account of positive Marshall-Arrow-Romer or negative Jacobs externalities, one is hard pressed to argue that some particular attribute influences current attainment status and not current births conditional on attainment status.

In an attempt to circumvent this problem, we exploit a natural phenomenon that has heretofore not been utilized: wind direction. Since county-level attainment status is not based on own emissions, but rather observed air quality readings, emissions from neighboring counties may influence attainment status. We therefore use the proportion of all contiguous western neighbors (the jet stream flows from west to east) that are out of attainment to identify the model.¹³ Before examining the two-step results, we note that (western) neighboring attainment status is a highly significant determinant of own attainment status ($p < 0.01$; see table A2). Thus, there is no question of bias due

¹³ For counties located on the NYS border, we obtained data on the attainment status of neighboring counties in other states to form the appropriate instrument.

TABLE 2.—PARAMETRIC ESTIMATES OF THE DETERMINANTS OF COUNTY-LEVEL PLANT LOCATION

Independent Variable	Fixed-Effects Poisson			
	HHG		Two-Step HHG	
	(1)	(2)	(3)	(4)
Nonattainment	−0.35 (0.22)	−0.50* (0.23)	−0.07 [−0.26, 0.13]	−0.08 [−0.28, 0.10]
ln(employment)	−4.63* (1.70)	−7.13* (2.11)	−4.41* [−7.35, −1.53]	−6.63* [−10.37, −3.08]
ln(wage)	—	−1.50 (2.24)	—	−0.80 [−5.09, 3.32]
ln(population)	—	10.91* (4.90)	—	9.27* [0.67, 17.21]
ln(prop. tax)	—	−3.19 (2.07)	—	−2.98 [−6.68, 0.42]
Period effects	Yes	Yes	Yes	Yes
County effects	Yes	Yes	Yes	Yes
Log likelihood	−325.0	−319.4	−326.2	−321.6
N	682	682	682	682

Dependent variable is the count of new plants annually from 1980 to 1990.

Nonattainment equals 1 if county is out of attainment of federal standards, 0 otherwise.

Models (1) and (2) treat attainment status as exogenous [estimated using HHG (1984) quasi maximum likelihood (QML)]; models (3) and (4) treat attainment status as endogenous (estimated using a two-step QML procedure). The proportion of western neighboring counties that are out of attainment is used as the instrument (see table A2).

Standard errors are in parentheses beneath coefficient estimates; 90% bootstrap confidence intervals—based on 1,000 repetitions—in brackets. * indicates significant at the 10% level using a two-sided alternative.

ln indicates the natural logarithm of the variable.

Time effects are jointly significant in models (1) and (2) at the $p < 0.15$ level.

to a weak instrument (Bound, Jaeger, & Baker, 1995). Moreover, as the coefficient is positive, this is consonant with our logic that emission spillovers affect the attainment status of eastern neighboring counties.

In terms of the actual results—displayed in columns (3) and (4) of table 2—the point estimates are −0.07 and −0.08, respectively, although the confidence intervals are extremely wide (90% confidence intervals are obtained via 1,000 bootstrap repetitions). When interpreting the magnitude of the two-step coefficients, one must note that the two-step model does not provide an estimate of the treatment effect per se, since the coefficient refers to a one-unit increase in latent attainment status. Since the predicted mean of latent attainment status ($Z_{it}\hat{\gamma}$) is −2.3, a 2.3-unit increase in latent attainment is required to move the average county from in to out of attainment. Multiplying the estimated two-step coefficients yields an estimated treatment effect of roughly −0.2, with a 90% confidence interval of approximately [−0.6, 0.3]. This implies a loss of around 0.1 new plant formations per annum from being out of attainment. These estimates remain significantly smaller than the matching estimates of between 0.7 and 1.3, and in fact are not statistically significant. This result holds in the two-step estimation of the specifications omitting time dummies as well (results not shown).

Given that the discrepancy between the parametric and matching estimates cannot be explained by a failure to treat attainment status as endogenous, we seek an alternative explanation. At least three possible explanations exist: (i)

TABLE 3.—PARAMETRIC ESTIMATES OF THE DETERMINANTS OF COUNTY-LEVEL PLANT LOCATION

Independent Variable	Fixed-Effects Poisson			
	HHG		Two-Step HHG	
	0.01 Sample (1)	0.05 Sample (2)	0.01 Sample (3)	0.05 Sample (4)
Nonattainment	−1.35* (0.74)	−0.62* (0.33)	−0.24 [−12.55, 8.24]	−0.21 [−0.49, 0.28]
ln(employment)	−2.01 (2.00)	0.31 (0.67)	−1.29 [−103.63, 54.74]	−0.58 [−3.71, 3.86]
ln(wage)	−6.13 (4.82)	2.37 (1.83)	−3.32 [−94.00, 79.87]	2.38 [−6.93, 8.91]
ln(population)	3.29 (2.33)	0.57 (0.66)	2.32 [−55.32, 125.14]	1.35 [−2.02, 5.20]
ln(prop. tax)	−1.38 (1.87)	−0.30 (0.92)	−0.61 [−137.16, 19.89]	0.37 [−3.59, 3.38]
Period effects	Yes	Yes	Yes	Yes
County effects	Yes	Yes	Yes	Yes
Log likelihood	−11.0	−54.3	−13.0	−54.2
N	74	162	74	162

Sample restricted to the matched subsamples used in columns (1) and (2) of table 1. See table 2.

the matching algorithm, by restricting the estimation to a matched subsample more similar in terms of other attributes, produces an estimator that is not subject to “outliers”; (ii) the parametric assumptions of the Poisson model are invalid; and (iii) the matching algorithm yields an estimate of the treatment effect on the treated, while the parametric approach estimates the treatment effect on a random observation from the population.

As there is no formal specification test for the Poisson model (other than testing for over- or underdispersion), we focus first on the role of sample composition. Thus, we proceed by estimating the parametric models—exogenous and two-step HHG—on the matched subsamples from the within-year matching algorithm (since the sample sizes are relatively large). Empirical results are displayed in table 3.

Estimated treatment effects are considerably larger than those reported in table 2. The new parametric estimates range from −0.6 to −1.4 on treating attainment status as exogenous (an increase of up to nearly threefold), and approximately −0.5 to −0.6 in the two-step models (a similar increase of up to threefold).¹⁴ Moreover, empirical results from the exogenous HHG models are statistically significant despite the small sample sizes; yet the two-step estimates remain imprecise. Taking the point estimates literally, however, implies an estimated loss of 0.2 to 0.6 (of ≈ 0.2) new pollution-intensive plants per annum from being out of attainment when attainment status is treated as exogenous (endogenous). Though larger, these point estimates remain smaller than the matching estimates presented in table 1. Thus, although sample composition plays a role

¹⁴ The point estimates from the two-step models are −0.24 and −0.21. Multiplying these by 2.3 yields the range −0.5 to −0.6.

in explaining the discrepancy across methods, it does not tell the entire story.¹⁵

IV. Concluding Remarks

This study examines whether plant location is influenced by environmental air quality regulations.¹⁶ Some recent innovative studies provide evidence that suggests environmental regulations are an important factor in the startup decisions of new manufacturing plants. Using the Census of Manufacturers database, Greenstone (1998), and Becker and Henderson (2000) present evidence indicating an inverse relationship exists between the stringency of air quality regulations and certain types of capital formation, implying that the spirit of the Clean Air Act has been violated.

We extend these studies by using various parametric methods and a semiparametric method based on propensity score matching, which has been shown to estimate treatment effects from nonexperimental data more precisely than standard parametric estimators. Via the matching technique we are able to take advantage of the panel nature of our data to control for time- and location-specific unobservables, as well as lagged values to the outcome of interest, in a straightforward manner. Our major findings are that: (i) pollution-intensive plants respond quite adversely to more stringent environmental regulations, and (ii) the matching method produces empirical estimates considerably larger than parametric estimates which treat attainment status as exogenous *or* endogenous within a Poisson framework. These findings suggest that the current state of the literature may dramatically understate the impact of pollution regulations.

REFERENCES

- Becker, R., and J. V. Henderson, "Effects of Air Quality Regulations on Polluting Industries," *Journal of Political Economy* 108 (2000), 379–421.
- Bound, J., D. A. Jaeger, and R. M. Baker, "Problems with Instrumental Variables Estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association* 90 (1995), 443–450.
- Cochran, W., and D. Rubin, "Controlling Bias in Observational Studies," *Sankhyā* 35 (1973), 417–446.
- Dehejia, R. H., and S. Wahba, "Casual Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association* 94 (1999), 1053–1062.
- , "Propensity Score Matching for Nonexperimental Causal Studies," *Review of Economics and Statistics* 84 (2002), 151–161.
- Greenstone, M., "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* (forthcoming).
- , "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* 110:6 (2002), 1175–1219.
- Ham, J. C., X. Li, and P. B. Reagan, "Matching and Selection Estimates of the Effect of Migration on Wages for Young Men," Ohio State University mimeograph (2001).
- Hausman, J., B. H. Hall, and Z. Griliches, "Econometric Models for Count Data with an Application to the Patents/R&D Relationship," *Econometrica* 52 (1984), 909–938.
- Heckman, J. J., H. Ichimura, and P. E. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies* 64 (1997), 605–654.
- Henderson, J. V., "Effects of Air Quality Regulation," *American Economic Review* 86 (1996), 789–813.
- Jaffe, A., S. Peterson, P. Portney, and R. Stavins, "Environmental Regulation and the Competitiveness of US Manufacturing: What Does the Evidence Tell Us?" *Journal of Economic Literature*, XXXIII (1995), 132–163.
- Jeppessen, T., J. A. List, and H. Folmer, "Environmental Regulations and New Plant Location Decisions: Evidence from a Meta-analysis," *Journal of Regional Science* 42:1 (2002), 19–49.
- List, J. A., "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sports-cards," *American Economic Review* 91 (2001), 1498–1507.
- List, J., and 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* 79:2 (2000), 177–190.
- McGuckin, R. H., "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).
- Michalke, J. C., "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).
- Mullahy, J., "Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior," *Review of Economics and Statistics* 79 (1997), 586–593.
- Rosenbaum, P., and D. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70 (1983), 41–55.
- Smith, J., and P. Todd, "Does Matching Address Lalonde's Critique of Nonexperimental Estimators," *Journal of Econometrics* (forthcoming).
- Windmeijer, F. A. G., and J. M. C. Silva, "Endogeneity in Count Data Models: An Application to Demand for Health Care," *Journal of Applied Econometrics* 12 (1997), 281–294.

¹⁵ Note, however, that there is significant overlap in the confidence intervals of the estimates across the two methods.

¹⁶ This issue is also germane when computing the benefits and costs of proposed environmental regulations.

APPENDIX

TABLE A1.—DESCRIPTION OF VARIABLES

Variable	Mean	In-Attainment Mean	Out-of-Attainment Mean	Definition and Source
New pollution-intensive plants	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, NYS DED.
New non-pollution-intensive plants	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, NYS 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.
ln(employment)	10.81 (1.33)	10.55 (1.15)	11.59 (1.53)	Natural logarithm of total employment in manufacturing. <i>County Business Patterns</i> .
ln(wage)	9.71 (0.23)	9.74 (0.22)	9.65 (0.25)	Natural logarithm of total annual manufacturing payroll divided by the number of employees by county, adjusted for inflation. <i>County Business Patterns</i> .
ln(population)	11.66 (1.25)	11.39 (1.07)	12.47 (1.38)	Natural logarithm of county population. <i>Current Population Reports</i> , U.S. Bureau of Census.
ln(property tax)	6.26 (0.34)	6.27 (0.35)	6.25 (0.28)	Natural logarithm of real property tax collected per capita. <i>Census of Governments</i> .

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

TABLE A2.—FIRST-STAGE LOGIT ESTIMATES OF THE DETERMINANTS OF ATTAINMENT STATUS

Independent Variable	Coefficient (SE)			
	(1)		(2)	
Neighboring attainment status	2.85*	(0.33)	—	—
Man. employment	1.99E-06	(1.29E-06)	—	—
Property taxes	-1.85E-03*	(8.75E-04)	—	—
Man. wages	-3.95E-06	(7.08E-05)	3.63E-03	(2.55E-03)
(Man. wages) ¹			-2.23E-07	(1.41E-07)
(Man. wages) ²			4.27E-12	(2.74E-12)
Man. plants			1.40*	(0.58)
(Man. plants) ¹			-0.09*	(0.05)
(Man. plants) ²			1.84E-03*	(1.04E-03)
Population	1.62E-06*	(5.09E-07)	-1.85E-06	(6.28E-06)
Population ¹			7.37E-12	(6.12E-12)
Population ²			-3.14E-18*	(1.82E-18)
Per capita income			4.73E-03*	(1.25E-03)
(Per capita income) ¹			-1.86E-07*	(9.64E-08)
(Per capita income) ²			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		Yes	
Log likelihood	-180.7		-145.8	
Pseudo R^1	0.54		0.63	
N	682		682	

Dependent variable is equal to 1 if county is out of attainment of federal ozone standards during the year, 0 otherwise. Neighboring attainment status is the percentage of western contiguous neighbors that are out of attainment.

Time effects jointly significant at the 1% level.

¹ Standard errors are in parentheses beside the coefficient estimates and are adjusted for clustering within counties. * indicates significant at the 10% level using a two-sided alternative.

² Model (1) is used in the two-step FE Poisson estimation. Model (2) is used to generate the propensity score estimates.