NCEE

NATIONAL CENTER FOR ENVIRONMENTAL ECONOMICS

Do Environmental Regulations Disproportionately Affect Small Businesses? Evidence from the Pollution Abatement Costs and Expenditure Survey

Randy A. Becker, Carl Pasurka Jr., and Ronald J. Shadbegian

Working Paper Series

**Do Environmental Regulations Disproportionately Affect Small
Businesses? Evidence from the Pollution Abatement Costs
and Expenditures Survey**

**Randy A. Becker, Carl Pasurka Jr.,
and Ronald J. Shadbegian**

Do Environmental Regulations Disproportionately Affect Small Businesses? Evidence from the Pollution Abatement Costs and Expenditures Survey

Randy A. Becker[a]
Carl Pasurka Jr.[b]
Ronald J. Shadbegian[b]

[a] Center for Economic Studies
U.S. Census Bureau

[b] National Center for Environmental Economics
U.S. Environmental Protection Agency

August 2013

Abstract

This paper examines whether the impact of environmental regulations differs by the size of the business. We consider the net effect of statutory, enforcement, and compliance asymmetries by estimating the relationship between plant size and pollution abatement expenditures, using establishment-level data on U.S. manufacturers from the Census Bureau's Pollution Abatement Costs and Expenditures (PACE) survey and from its Annual Survey of Manufactures and Census of Manufactures. We model establishments' pollution abatement operating costs (PAOC) per unit of economic activity as a function of establishment size, industry, state, and year. Our results show that PAOC intensity increases with establishment and firm size.

I. INTRODUCTION

After the passage of the 1970 Clean Air Act Amendments and 1972 Clean Water Act Amendments, the United States realized significant improvements in both air and water quality. These improvements were due in large part to increasing stringency of environmental regulations that led to steady declines in emissions from industrial sources. According to the Pollution Abatement Costs and Expenditures (PACE) survey, the U.S. manufacturing sector spent nearly $21 billion dollars on pollution abatement operating costs (PAOC) to comply with environmental regulations in 2005 (U.S. Census Bureau 2008). Policymakers are especially interested in how these regulations affect the economy, and there is a growing literature on the effects of environmental regulations on various plant-level outcomes, including productivity (Berman and Bui 2001a; Gray and Shadbegian 2002, 2003, 2006; Shadbegian and Gray 2005; Becker 2011b; Greenstone, List, and Syverson 2012), investment (Gray and Shadbegian 1998), location (Becker and Henderson 2000; List *et al.* 2003) and employment (Berman and Bui 2001b; Morgenstern, Pizer, and Shih 2002; and Greenstone 2002).

There is also significant interest in whether the impact of environmental regulations differs by the size of the facility. Differences in pollution abatement spending by establishment size might be expected because of statutory, enforcement, and/or compliance asymmetries (Dean, Brown, and Stango 2000; Becker 2005). Statutory asymmetries exist when regulations explicitly impose less stringent requirements on certain types of businesses, such as small businesses.[1] Enforcement asymmetries result when regulators choose to target certain

---

[1] For example, many regulations distinguish between "major" and "minor" pollution sources, based on emissions thresholds or sometimes production levels. The Regulatory Flexibility Act (RFA) of 1980 places requirements on the U.S. Environmental Protection Agency (EPA) and other federal agencies to thoroughly consider the economic impacts rules and regulations will have on small businesses. For the purposes of EPA's analyses and assessments, the definition of "small" comes from the Small Business Act, which authorizes the Small Business Administration to develop industry-specific definitions, generally based on number of employees or average annual receipts (see www.sba.gov/size). The Small Business Regulatory Enforcement Fairness Act (SBREFA) of 1996 enacted a range

establishments over others — a decision that may hinge on establishment or firm size (Gray and Deily 1996). Finally, compliance asymmetries arise when regulatory compliance involves significant fixed costs. For example, some pollution abatement is quite capital intensive, which would result in higher costs *per unit of output* for smaller establishments.

Here, we consider the net effect of these three asymmetries, by estimating the relationship between plant size and pollution abatement expenditures. In particular, we use establishment-level data from the U.S. Census Bureau's Pollution Abatement Costs and Expenditures (PACE) surveys of 1974-1982, 1984-1986, 1988-1994, 1999, and 2005. The PACE survey is the most comprehensive source of pollution abatement costs and expenditures data related to environmental protection in the U.S. manufacturing sector. These PACE data are merged to data from the Annual Survey of Manufactures (ASM) and Census of Manufacturers (CM), which include employment, value of shipments, geography, and four-digit Standard Industrial Classification (SIC) industry (or, in recent years, six-digit North American Industry Classification System (NAICS) industry). In this paper, we focus on establishments' PAOC intensity – that is, their pollution abatement operating costs per unit of economic activity – which we model as a function of establishment size, industry, state, and year. Our results show that PAOC intensity increases with establishment size. We also find that larger *firms* spend more per unit of output than do smaller firms.

The rest of the paper is organized as follows. In Section II, we survey previous studies on the effects of environmental regulations by business size. Section III describes the data and

---

of provisions, including several amendments to the RFA. In brief, SBREFA requires EPA to solicit input from a Small Business Advocacy Review Panel and to consider flexible regulatory options that minimize adverse economic impacts on small entities. It also added a provision that allows small entities adversely affected by a final rule to challenge the agency's compliance with the RFA's requirements in court. (For more information, see U.S Environmental Protection Agency 2006). Thus, on just these grounds, we might expect to find that small businesses are not disproportionately affected by environmental regulations.

the methodology. In Section IV, results are presented, and Section V offers some concluding remarks, including directions for future research.

II. LITERATURE REVIEW

Evans (1986) notes that whether or not regulations differentially affect plants or firms by size is important for three reasons. First, if economies of scale exist in regulatory compliance then, other things equal, optimal regulatory policy could require imposing less stringent regulations on smaller plants (i.e., allowing statutory and/or enforcement asymmetries). Brock and Evans (1985) argue that this differential treatment (known as tiering) can be Pareto-superior compared to untiered regulations. Second, society may (and seems to) value small business for non-economic reasons and therefore may want to ensure they are not disproportionately affected by regulation. Finally, Evans argues that the incidence of regulatory compliance costs across plant and/or firm size may indicate why certain industry groups back alternative regulatory policies.

There have been a number of published papers which examine the relationship between environmental regulations and establishment size in the U.S. manufacturing sector. Most of this previous literature uses aggregate (e.g., industry-level) data, which is less than ideal since effects are identified through the comparison of averages across industries, time periods, and/or states, rather than a comparison across actual establishments. Pashigian (1984), using data on 319 four-digit SIC industries between 1972 and 1977 (a period covering the implementation of many new federal environmental regulations), finds that compliance with environmental regulations caused an increase in plant size and a reduction in the number of plants per industry. From these results, Pashigian concludes that environmental regulations place a greater burden on smaller plants as

3

small plants find it more difficult to compete with larger plants and survive. Evans (1986) is critical of Pashigian's findings for several reasons, including the use of industry-level data. In contrast, Evans obtained data from the U.S. Census Bureau on the average pollution abatement operating costs per employee for three categories of plants: 0-99, 100-499, and 500 or more employees, for single- and multi-plant firms, from 1978-1981. In general, Evans finds evidence of diseconomies of scale in pollution abatement, with 375 of the 403 industries exhibiting diseconomies of scale, 214 of which are statistically significant. Although Evan uses more detailed data than Pashigian, they are still fairly aggregate. Dean *et al.* (2000), using data on 170 select four-digit SIC industries, find that industries with high pollution abatement costs had fewer small business formations (defined as those with less than 100 employees). Meanwhile, Millimet (2003) finds that the optimal plant size is larger in high pollution abatement intensive industries (at the two-digit SIC level) in states with more stringent environmental regulation.

Of the two papers in this literature that use establishment-level, neither focuses chiefly on the relationship between pollution abatement expenditures and establishment size. The focus of Hartman *et al.* (1997) is to develop estimates of air pollution abatement costs for seven air pollutants for various four-digit SIC industries from 1979-1985. These authors find some evidence of declining average and marginal costs of abatement with increased volume of abated pollutants, indicating economies of scale in air pollution abatement. Meanwhile, Becker (2005) also focuses just on air pollution abatement and finds mixed evidence of a disproportionate impact of abatement costs on small business. More specifically, he finds that air pollution abatement expenditures under the Clean Air Act were at times disproportionately higher for larger establishments, and sometimes the reverse, depending on the criteria air pollutant.

Since the mid-1990s, there has been a series of otherwise unpublished Small Business Administration (SBA) working papers, examining the impact of environmental regulations – and federal regulations in general – on small businesses. The most recent of these studies (Crain and Crain 2010) finds that that the cost of complying with environmental regulations, on a per employee basis, is almost five times higher for small businesses than for large businesses. Harrington (2006) argues that "their claim of economies of scale in pollution abatement is utterly without foundation." In particular, he notes that Crain and Hopkins (2001) "cherry-pick" coefficients, focusing on average firm size as measured by the number of employees while ignoring value of sales (another firm size variables and one with a strong positive correlation with employment) which has the opposite sign.[2]

An important issue faced by previous studies has been the lack of official government statistics *by business size*. The Census Bureau did briefly publish such statistics as part of the PACE survey. In particular, for three years, 1976-1978, separate estimates of pollution abatement operating costs and pollution abatement capital expenditure were produced for each two-digit SIC manufacturing subsector, and for the manufacturing sector as a whole, for each of six establishment employment size classes (see U.S. Census Bureau, various years). Using value of shipments (i.e., output) statistics on these same size classes from the 1977 Census of Manufactures, we compute aggregate PAOC intensity and plot the results in Figure 1. We find that PAOC intensity increased with establishment size, peaking with the 500-999 employees category before declining a bit for the largest size category (1000+ employees).[3] Results for the 19 major manufacturing groups (two-digit SIC industries) do not necessarily tell the same

---

[2] Crain and Hopkins (2001), Crain (2005), and Crain and Crain (2010) all contain the same regression and the same regression results and interpret them the same way.

[3] The increase in PAOC intensity in the 1-49 employee category after 1976 may, in part, be an artifact. In 1977 and 1978, establishments with fewer than 20 employees were not sampled in the PACE survey. Therefore, the PAOC intensity in these years is for plants that had 20-49 employees.

story.[4,5] Nevertheless, in 48 of the 57 cases (19 groups x 3 years), PAOC intensity for the 1000+ employee category exceeded that of the 1-49 employee category, with peak PAOC intensity occurring at [frequency]: 1-49 [1], 50-99 [2], 100-249 [7], 250-499 [5], 500-999 [11], and 1000+ [31]. This certainly does not suggest a disproportionate impact on small businesses — if anything, the opposite. These data are obviously limited in that they are for just three years, from over 35 years ago, and at the two-digit SIC level.[6]

III. DATA AND METHODOLOGY

In this paper, we use the establishment-level data from the Pollution Abatement Costs and Expenditures (PACE) surveys of 1974-1982, 1984-1986, 1988-1994, 1999, and 2005.[7] Here we focus on total pollution abatement operating costs (PAOC), which includes salaries & wages, parts & materials, fuel & electricity, capital depreciation, contract work, equipment leasing, and other operating costs associated with a plant's abatement of its air and water pollution as well as its solid waste.[8] These data are merged to data on these same establishments from the contemporaneous Annual Survey of Manufactures (ASM) or Census of Manufactures (CM). While the ASM and CM contain a host of variables, here we simply use establishment employment, value of shipments, four-digit SIC industry (six-digit NAICS industry, in later

---

[4] SIC 23 (apparel and other textile products) is not among these, having been excluded from every year of the PACE survey except 1999.

[5] In some cases, employment-size classes were combined to avoid disclosing information about individual companies. In these cases, PAOC was prorated across the affected cells using 1977 value of shipments, resulting in identical PAOC intensity across these cells.

[6] Aggregate data, even at the two-digit SIC level, can obscure and distort the true relationship between PAOC intensity and establishment size, particularly if four-digit SIC industries are heterogeneous in their PAOC intensity profiles and their establishment size distributions.

[7] The PACE survey was also conducted for reference years 1973 and 1983, but the establishment-level data for those two years are unavailable. A survey for reference year 1987 was not conducted.

[8] These survey data, as well as those from the Annual Survey of Manufactures and the Census of Manufactures, are confidential, collected and protected under Title 13 and Title 26 of the U.S. Code. Restricted access to these data can be arranged through the U.S. Census Bureau's Center for Economic Studies. See www.census.gov/ces/ for details.

years), and location. We also have information to allow us to determine multi-unit firm status and establishment age/cohort. We accept any and all Census Bureau corrections and imputations (or the use of alternate, administrative data) in the event of item non-response. After restricting the sample to cases that had linkable PACE and ASM/CM records in a given year, and after eliminating non-respondents[9], establishments with signs of production inactivity, plants in Alaska and Hawaii, and those with missing or incomplete data on critical items (e.g., industry code), there are 321,526 establishment-years of observations for estimating the relationship between establishment size and environmental expenditure.

In this paper, we focus on establishments' PAOC intensity – that is, their pollution abatement operating costs per unit of economic activity. Here we employ establishment output in the denominator.[10] PAOC intensity has long been used as a proxy for the intensity of environmental regulation faced by plants and industries. For examples, see Pashigian (1984), Jorgenson and Wilcoxen (1990), List and Co (2000), Keller and Levinson (2002), and Becker (2011a).

Our main focus in this paper is not in explaining PAOC intensity, per se, but rather in estimating the relationship between PAOC intensity and establishment size — whatever the underlying explanation(s). Consequently, we begin by modeling PAOC intensity in one of the simplest ways possible: as a function of establishment size, four-digit SIC industry, state, and year.[11] To measure establishment size, we use establishment employment in linear or natural log

---

[9] Code to identify cases that are thought to be survey respondents and tabulated cases is available from the authors.
[10] Using output has an established precedent in this literature. An alternative that is sometimes used is establishment employment — and PAOC per employee might be said to encapsulate a regulator's implicit choice between environmental protection and jobs. Here, we also present results from a robustness check that uses employment rather than output in the denominator.
[11] Here we use 1987 SIC codes. Establishments from 1988-1994 are already classified on this basis. In contrast, establishments from 1974-1986 are classified using 1972 SIC codes. Approximately 84% of those have a 1972 SIC code that maps uniquely into a single 1987 SIC code. Another 11% are assigned a 1987 SIC code, from the set of possible codes for their respective 1972 SIC code, using their 1987 SIC code in the 1987 CM (requiring these plants

form or to construct a set of six (commonly used) establishment size categories: 1-49, 50-99,

100-249, 250-499, 500-999, and 1000 or more employees. This last variant of our estimating

equation is

$$
\begin{aligned}
\log(PAOC_{it}/VS_{it}) = \quad & \lambda_0 + \sum \lambda_n \cdot SIC_{nit} + \sum \lambda_s \cdot STATE_{sit} + \sum \lambda_t \cdot YEAR_{it} + \\
& \lambda_1 \cdot EMP50\_99_{it} + \lambda_2 \cdot EMP100\_249_{it} + \lambda_3 \cdot EMP250\_499_{it} + \\
& \lambda_4 \cdot EMP500\_999_{it} + \lambda_5 \cdot EMP1000plus_{it} + e_{it}
\end{aligned}
$$

where, for plant i at time t, *PAOC* is pollution abatement operating costs, *VS* is output (value of

shipments), *SIC* is a set of dummy variables indicating the plant's four-digit SIC industry,

STATE is a set of dummy variables indicating the plant's state, *YEAR* is a set of year dummies,

and *EMP50_99*, *EMP100_249*, *EMP250_499*, *EMP500_999*, and *EMP1000plus* is a series of

establishment size indicator variables based on employment. The omitted size category is the

very smallest establishments, with 1-49 employees.[12]

One very important characteristic of these data is that some 22 percent of establishments

have "zero" PAOC.[13] Moreover, the incidence of zero PAOC appears related to our variable of

interest, with incidence (from smallest to largest plants) of 51, 31, 21, 14, 9, and 5 percent,

---

to survive to 1987 and to be in a 1987 SIC industry that is in the feasible set). The remaining 5% are assigned a 1987 SIC code, from the set of possible codes for their respective 1972 SIC code, using either product-level information or through random assignment (where the probability of a certain 1987 SIC code being selected is proportional to the share of the 1972 SIC industry's value of shipments that were reclassified to that 1987 SIC industry in 1987). Meanwhile, establishments from 1999 and 2005 are classified using NAICS codes. Approximately 57% of those have a NAICS code that maps uniquely into a single 1987 SIC code. Another 37% are assigned a 1987 SIC code, from the set of possible codes for their respective NAICS code, using their 1987 SIC code in the 1997 CM or 1998-2001 ASM (requiring these plants to have been in existence in 1997 or in those ASM samples and to have been in a 1987 SIC industry that is in the feasible set). The remaining 6% are assigned a 1987 SIC code, from the set of possible codes for their respective NAICS code, through random assignment (where the probability of a certain 1987 SIC code being selected is proportional to the share of the 1987 SIC industry's value of shipments in that NAICS industry in 1997).

[12] Since 1976, with the exception of 1999, establishments thought to have fewer than 20 employees were excluded from the PACE survey sample. According to early PACE publications, early surveys "indicated that establishments with less than 20 employees contributed only about 2 percent to the pollution abatement estimates while constituting more than 10 percent of the sample size." (U.S. Census Bureau, various years) In the 1977 CM, establishments with fewer than 20 employees accounted for 67.5% of the universe of manufacturing plants. Becker and Shadbegian (2005) find that, in 1999, establishments with fewer than 20 employees accounted for 3.0% of pollution abatement expenditure in the manufacturing sector, while accounting for 5.1% of value added.

[13] This includes cases of missing PAOC which, for survey respondents, is equivalent to zero PAOC for tabulation purposes.

respectively. The prevalence of zeros suggests the use of models other than ordinary least squares (OLS). Tobit specifications are often employed when there is large cluster of zeros. In results available upon request, we consider censored normal regression – a generalization of the standard Tobit that allows the censoring point to vary by observation.[14] There are at least two issues with this approach, however. First, Tobit models assume that the same process determines both the *probability* of a nonzero outcome and the *value* of a nonzero outcome. In our case, one can imagine a scenario in which establishment size has not only different but even opposite effects on the probability of nonzero PAOC and the intensity of PAOC expenditure. For example, one can imagine that smaller establishments are less likely to have nonzero PAOC — because they are small, they are perhaps rarely (if ever) inspected and therefore can forgo pollution abatement without much risk of discovery (enforcement asymmetries). However, for small establishments that *are* inspected/regulated, or otherwise decide to engage in pollution abatement, their PAOC intensity may well be higher than their larger counterparts in the industry because of various fixed and lumpy costs (compliance asymmetries).[15] A second issue is that the Tobit's coefficients measure the effects on the *latent* variable found in the standard model. If zero PAOC were due to true data censoring (i.e., akin to missing true values) then the estimated coefficients are exactly the effects of interest. Here, however, zeros can be deemed true values, and our interest is not necessarily in effects on the latent variable – a construct – but in the effects on observed values.

---

[14] PAOC intensity, as measured by *PAOC/VS* in linear form, is not normally distributed. However, the natural log of *PAOC/VS* has a skewness of -0.311 and kurtosis of 2.87, which are very near the reference values for symmetry and normality. We therefore estimate a lognormal Tobit model. For observations with zero *PAOC*, we set the dependent variable equal to $ln(0.5/VS_i)$, which is a natural censoring point: Since establishments are asked to report their expenditures in *thousands* of dollars, with rounding, a response of zero reflects expenditures of less than \$500. Because the denominator (*VS*) can obviously vary by observation, the censoring point also varies by observation. Hence the use of censored normal regression, estimated using the cnreg command in Stata.

[15] For example, the cost of capital equipment (depreciation), the salary of an environmental engineer, the annual cost of maintaining pollution abatement equipment, and the fees paid to an environmental consulting firm may all be spread over fewer units of output.

In this paper, we estimate PAOC intensity using two different methodologies: the two-part model (TPM) and the Heckman two-step selection model (Heckit). Both models are more flexible than the Tobit model and are considered to be much better suited to estimating corner solutions (see Wooldridge 2002 and Cameron and Trivedi 2010). Indeed, both models are commonly used in the health economics literature, where zero expenditures (on health care, on cigarettes, etc.) are not uncommon.[16] In both of these models, the probability of a nonzero outcome is estimated via a probit model, as in

$$\Pr(PAOC_{it} > 0) = \quad \lambda_0 + \sum \lambda_n \cdot SIC_{nit} + \sum \lambda_s \cdot STATE_{sit} + \sum \lambda_t \cdot YEAR_{it} +$$
$$\lambda_1 \cdot EMP50\_99_{it} + \lambda_2 \cdot EMP100\_249_{it} + \lambda_3 \cdot EMP250\_499_{it} +$$
$$\lambda_4 \cdot EMP500\_999_{it} + \lambda_5 \cdot EMP1000plus_{it} + e_{it}$$

,

and the second part estimates the (previous) intensity equation *on observations that are nonzero* (with the inclusion of the inverse Mills ratio, in the case of the Heckit). Among other things, these approaches relax the rather strong assumption of the Tobit model that the same process determines both the probability of nonzero PAOC and the intensity of PAOC expenditure. Instead, these approaches allow the two mechanisms to be distinct, and an explanatory variable (such as, establishment size) can have different – and even opposite – impacts on the outcomes of interest.

In the case of the TPM, these two equations are assumed to be completely independent and can be estimated separately – by probit and OLS, respectively. The Heckit, on the other hand, allows for dependence between the two parts of the model – that something akin to

---

[16] There are also precedents for using the Heckit model to examine the pollution abatement expenditures of manufacturing establishments in the United Kingdom (Collins and Harris 2002, 2005) and Ireland (Haller and Murphy 2012). In contrast to what we do here, these studies estimate the elasticity of pollution abatement expenditure (not intensity) with respect to a single, continuous measure of business size. More often than not, the estimated elasticity is less than one.

selection bias impacts the intensity equation (second part).[17] In both the TPM and Heckit, we use the same regressors in both equations. That is, the Heckit is estimated without exclusion restrictions — mainly because no good, defensible candidates exist among the available variables. Therefore, for better or worse, we rely on the nonlinearity of the functional form to achieve model identification. Meanwhile, in the TPM, the use of the same regressors in both equations makes no difference at all.

There has been significant debate – principally in the health economics literature – whether the TPM or Heckit is more appropriate in this sort of context (see Jones 2000 and references therein, plus Dow and Norton 2003). Some of the debate centers on statistical issues, while part of it focuses on theoretical matters – in particular, whether we are interested in *actual* outcomes or *potential* outcomes, and what our beliefs are regarding zero values. On the one hand, the TPM reflects actual, observed outcomes, and assumes all zeros are true and valid. On the other hand, the Heckit reflects impacts on *potential* outcomes. Here, the PAOC intensity equation is interpreted as though we observe *nonzero* PAOC intensity data for *all* plants in the sample (including the non-spenders). The Heckit's results adjust for the fact that establishments with observed nonzero PAOC are (or may be) different than those with zero PAOC. The model assumes that cases with zero PAOC expenditure have *latent* nonzero PAOC intensity that was somehow unrealized. In essence, and in contrast to the TPM, the Heckit assumes that *none* of the zeros are valid.

The Heckit is most appropriate in a case like women's wages. In this classic example, the challenge is to measure the effect of, say, education on wages. The issue is that wages are missing for women who do not work, and those who do not work are not like those who do. The Heckit treats the inherent selection bias, and the estimated effects are those on actual <u>and</u>

---

[17] This is estimated using the heckman command in Stata, with the twostep option.

potential wages (were an individual to work). At first blush, our case seems different, since we do not have "missing" PAOC data in the truest sense. But perhaps we do. Say, for example, establishments are sometimes prone to report zero PAOC expenditures though untrue. Perhaps the Census Bureau is most likely to follow up with (verify or impute) cases mostly likely to have large PAOC or a large impact on specific aggregate statistics, while smaller cases (of zero PAOC) are ignored because, in the aggregate, they make relatively little difference. If this is true then the Heckit model may make sense. In this context, the *latent* nonzero PAOC intensity that went unrealized is the nonzero PAOC intensity if the Census Bureau exerted more effort in collecting, verifying, and imputing data for these cases.

The TPM is arguably most appropriate in a case like cigarettes smoked per day. Madden (2008) suggests that in a case such as this "the participation decision dominates the consumption decision. This implies that zero consumption does not arise from a standard corner solution but instead represents a separate discrete choice." He goes on to argue that "there is unlikely to be a latent positive expected consumption.… It follows that we are interested in the marginal effects of covariates on actual as opposed to potential smoking in which case the two-part model seems more appropriate." It seems the analogous situation in our case would be industries and/or establishments that truly produce no pollution or waste, and therefore have no PAOC expenditures.

We believe the truth is likely a blend of these two scenarios: some cases of zero PAOC are a reporting issue, while others are true zeros. We present results from the TPM and Heckit as the outer bounds on the true relationship.[18] Narrowing our analysis to just one model – one that

---

[18] Another argument in the literature is that the preferred model depends on the purpose of the analysis, and whether the interest is in *actual* or *potential* expenditure, with the Heckit being more appropriate in the latter instance. In the case of health care expenditures, it has been argued that even persons with true zero expenditures have a positive *latent* expenditure. For example, a person may have high cholesterol that is yet unknown and goes untreated.

reflects the true relationship – would require us to exert priors on which observations (or which combinations of industry, size class, state, county, year, etc.) are true zeros versus not. We do not know of an appropriate, scientifically defensible way of doing that. As results presented in the next section will demonstrate, the basic story is the same under both approaches.

IV. RESULTS

Table I presents the results using the TPM. All three specifications suggest that the *probability* of nonzero PAOC increases – and increases monotonically – with establishment size. We saw this in the raw data as well. Meanwhile, the impact of establishment size on the intensity of PAOC expenditure (conditional on having nonzero intensity) depends on how establishment size is specified. The impact is positive, negative, and U-shaped, in columns (1), (2), and (3), respectively. Given that establishment employment is log-normally distributed, columns (2) and (3) are the preferred formulations. We will look at the combined effect (of establishment size) and evaluate magnitudes below. Meanwhile, an important point can be made here. We see instances where the impact of establishment size on the probability of nonzero PAOC and on PAOC intensity are indeed different and even opposite. This suggests that the Tobit model is not appropriate here, with its strict assumption that the same process determines both the *probability* of a nonzero outcome and the *value* of the nonzero outcome.[19]

Depending on the exact nature of the policy evaluation, one might be interested in the impacts of some factor assuming that these zero expenditure cases were in fact "off the sidelines" – for example, in a world with greater availability of free medical screenings. The corollary in our case may be that zero PAOC is in fact true but is due to a lack of regulatory attention, including facility inspections. If there were more inspectors and regulators, exerting more effort in inspecting and regulating the establishment, then the latent nonzero PAOC intensity might in fact be realized. We'll note that the decision by the establishment/regulator to have nonzero PAOC expenditures at an establishment may well be linked to the known or perceived severity of pollution generated at the establishment, and hence PAOC intensity. Here, we simply note this argument for perhaps preferring one approach over the other without taking a stance on whether the reader's interest is in actual or potential PAOC intensity.

[19] A comparison of the log likelihoods confirms that the TPMs indeed fit the data better than the censored normal regressions. Nevertheless, despite the issues with Tobit specifications noted above, the results using censored

13

Table II contains the Heckit results. We see, as we did with the TPM, that the probability

of nonzero PAOC increases – and increases monotonically – with establishment size. Indeed the

results are identical. What is important are the results from the second stage, which can be

interpreted in isolation, as if we observe nonzero PAOC intensity for all plants in the sample

(including non-spenders). Here we find that the impact of establishment size on the intensity of

PAOC expenditure is positive in all specifications (though the coefficient for plants with 50-99

employees does not attain statistical significance here).[20,21] We'll note that the coefficient on

lambda (the inverse Mills ratio) is statistically significant, which suggests there is indeed sample

selectivity: some (unobserved or deliberately excluded) determinant of the probability of

nonzero expenditure is also influencing PAOC intensity.

To better compare results, we compute the marginal effects of the establishment size

dummy variables. That is, for each model, using the final column of results from Tables I and II,

we estimate the incremental effect on the level of PAOC intensity of being in the establishment

size class of interest versus the omitted size category (i.e., establishments with 1-49

employees).[22] Retransforming PAOC intensity from natural logs to levels requires special

attention and here we employ a common technique, which relies on the assumption that the error

---

normal regressions are in keeping with the results from the two preferred approaches: in all three specifications, establishment employment has a positive and monotonically increasing impact on PAOC intensity.

[20] Here and throughout, robust standard errors, clustered by establishment, are computed. In the Heckit case, standard errors are estimated via bootstrapping, using 400 replications — a number recommended by Cameron and Trivedi (2010) .

[21] That the estimates converge to a value of rho that is exactly 1 suggests there are issues with the estimates in column (2).

[22] In the TPM, predicted PAOC intensity is the product of the predicted probability of observing nonzero PAOC and the predicted PAOC intensity. In turn, the incremental effect is the sum of the effect attributable to the change in probability plus the effect attributable to the change in intensity, evaluated for each observation assuming the omitted size category. These individual incremental effects are then averaged across all observations in the sample. We thank Edward Norton for making available Stata code for computing incremental effects in two-part models (iHEA2005.do) at his UNC-Chapel Hill website. In the Heckit, predicted PAOC intensity is again a function of the estimates from the two parts. (For details, see Dow and Norton 2003.) Here, predicted PAOC intensity is computed for each observation assuming the omitted size category, and subtracted from the predicted PAOC intensity assuming the size category of interest, and this difference is averaged across all observations in the sample.

terms are normally distributed with constant variance. Alternatively, for the TPM, we also employ the nonparametric "smearing" estimator of Duan (1983), which relies on weaker assumptions about the error term.[23]

Table III contains these estimates of the average incremental effects by establishment size category. Effects are all reported in dollars of *additional* PAOC per $1000 of establishment output (value of shipments), relative to the omitted size category (i.e., establishments with 1-49 employees).[24] Standard errors (estimated via bootstrapping, using 400 replications on samples clustered by establishment) are reported in parentheses. We see that, for *all* the alternatives in this table, the very highest PAOC intensity occurs in the very largest establishments (with 1000+ employees). Meanwhile, the lowest PAOC intensity occurs in the second smallest size group (i.e., those with 50-99 employees) in the TPM and in the very smallest establishments (with 1-49 employees) in the Heckit.[25] Figure 2 plots the values from the first three columns of Table III. We see that the "normal" versus Duan retransformation makes some difference here — both curves are a similar shape, but they cross and diverge as establishment size increases.

Next, we examine the possibility that outliers are driving our results. In the last three columns of Table III, we present the corresponding results using a sample that removes observations with large (and implausibly high) values of PAOC intensity. For the purposes here, we treat observations that had PAOC expenditures that were more than 15% of establishment output, or 628 observations (0.195% of the sample), as implausibly high and drop them from our sample.[26] The results in the last three columns of Table III suggest that the previous results are

---

[23] See Cameron and Trivedi (2010) for a discussion of the retransformation problem and for appropriate Stata code.
[24] Since nominal dollars are divided by nominal dollars, results are in constant dollars.
[25] We believe that this initial decline is due to a disproportionate number of outliers among the very smallest establishments (with 1-49 employees) driving up their average PAOC intensity. This is not unexpected if PAOC (the numerator) is relatively noisy and output (the denominator) is relatively small.
[26] The dropped establishments tend to be smaller, as foreshadowed by the previous footnote. 51% of dropped cases are among the very smallest establishments (with 1-49 employees).

somewhat influenced by outliers. Relative to the full sample, the incremental effects here are usually higher (more positive, less negative) though not always. Setting aside the magnitude of the effects, the story is essentially the same. In the Heckit, PAOC intensity increases monotonically with establishment size. Controlling for industry, state, and year, the very largest establishments (with 1000+ employees) had $5.49 more PAOC per $1000 of output than the establishments with 1-49 employees. All these estimates are statistically different from zero and from each other. In the TPM, PAOC intensity essentially increases monotonically as well. Here, establishments with 1000+ employees had about $1.91 more PAOC per $1000 of output than the establishments with 1-49 employees. Establishments with 50-99 employees had the lowest PAOC intensity, according to the TPM, but by a mere 5 cents per $1000 of output, and this effect is not statistically different from zero. With the outliers removed, there is very little difference between the "normal" and Duan retransformations. In Figure 3, which plots the values from the last three columns of Table III, these two curves are virtually indistinguishable from one another. Figure 3 quite clearly reveals a "cone" bounded above by the Heckit results and below by the TPM results. As we argue in the previous section, the true relationship lies somewhere between these two extremes. Wherever the true line lies within this cone, one thing is absolutely clear: Small plants do *not* appear to have PAOC intensity that is multiples of that of large plants. Quite the opposite, larger plants spent more on pollution abatement per dollar of output than their smaller counterparts.

We now consider a few robustness checks and extensions. First, we examine the issue of outliers. Above, we presented results for the full sample and for the sample that excludes observations that had PAOC expenditures that was more than 15% of establishment output (0.195% of observations). In Figures 4 and 5, we re-graph these results along with results using

cutoffs of 100%, 25%, and 5%, which drop 0.02%, 0.10%, and 1.55% of the sample, respectively. This final scenario is fairly severe – no doubt removing some legitimate cases. Figure 4 contains TPM results, and Figure 5 contains the Heckit results. In the TPM, as more outliers are dropped, the curve shifts up – from a checkmark-shape to something that monotonically increases. These curves are all within 25 to 52 cents of one another. In the Heckit model, the curve shifts up as well, as more outliers are dropped, except in the most extreme case (PAOC of more than 5% of output) which bends below the other curves for the largest three establishment size categories. Here, the curves are all within 38 to 54 cents of one another. The fundamental story still holds, regardless of the outlier cutoff chosen, or whether outliers are dropped at all.

Next, we examine the impact of defining PAOC intensity using establishment employment in the denominator rather than output (value of shipments). As in Figure 3, we exclude those 628 observations with PAOC expenditures that were more than 15% of establishment output. The results are depicted in Figure 6. In the Heckit, as before, PAOC intensity increases monotonically with establishment size. Controlling for industry, state, and year, the very largest establishments (with 1000+ employees) had $959 more PAOC per employee (in 1979 dollars) than the establishments with 1-49 employees. In the TPM, as before, establishments with 50-99 employees had the lowest PAOC intensity. Beyond that, PAOC intensity increases, and establishments with 1000+ employees had about $225 more PAOC per employee than the establishments with 1-49 employees. Again, the true relationship lies somewhere between these two extremes and again suggests larger plants have more intense pollution abatement spending than their smaller counterparts.

Next, we recognize that the public discussion is often about *firm* size, not *establishment*

size. With the data at hand, however, we cannot compute firm-level PAOC intensity since some of a firm's establishments may not be in a given year's PACE survey.[27] An alternative would be to compute establishment-level PAOC intensity, as we do here, but regress it on (and plot it against) firm rather than establishment size categories.[28] Here, we instead use information on whether the establishment belongs to a multi-unit firm (i.e., there are other establishments in the firm) or is a single-unit firm (i.e., the observed establishment *is the entire firm*). We run our previous TPM and Heckit models on these two different samples.[29] The results are plotted in Figure 7. We see, unambiguously, establishments in multi-unit firms spend more than otherwise similar single-unit establishment/firms. It is well-known that multi-unit firms are larger than single-unit firms on average. Here, it is less ambiguous: For single-unit establishments, firm and establishment sizes are one and the same. For multi-unit establishments, firm size will be at least as great as the size category the establishment is in. Therefore, not only can we say that, controlling for year and industry, larger *establishments* spend more on PAOC per unit of output than their smaller counterparts, but larger *firms* also spend more than their smaller counterparts, with the largest establishments in the larger firms spending the most of all.[30,31]

Next, we explore the components of PAOC. In particular, the PACE survey collects PAOC by media (i.e., air, water, and solid waste) and by type of cost. Regarding the latter,

---

[27] And a firm's non-manufacturing establishments are not even eligible to receive the survey.

[28] With some effort, an establishment's firm size can be constructing using the Census Bureau's business register (previously known as the Standard Statistical Establishment List), which is a database of virtually all private, non-farm business entities (establishments and companies) in the United States with paid employees. Containing establishment-level information on employment, payroll, industry, location, and firm affiliation, the business register serves as the sampling frame for nearly all of the Census Bureau's business surveys. An alternative to this would be to use the various Censuses of Manufactures to calculate the size of the firm's total manufacturing employment, ignoring any and all activity outside the manufacturing sector. However, this is only available once every five years.

[29] Establishments belonging to multi-unit firms account for about 81% of the total sample here. 84% of multi-unit establishments have non-zero PAOC, compared to 54% among single-unit establishments.

[30] And this does not account for environmental expenditures that may occur in a firm's corporate headquarters, unless those headquarters are co-located with a production facility in the PACE survey and the facility included those corporate expenditures.

[31] Evans (1986) also found this to be true, in most manufacturing subsectors.

because of changes in the survey instrument over time, we are only able to construct three

consistently defined cost categories: labor, depreciation, and all other costs (which include

materials, supplies, fuel & electricity, services & contract work, equipment leasing, and

miscellaneous expenses).[32] We run our previous TPM and Heckit models using the intensity of

expenditure on these separate components of PAOC. We find that PAOC per unit of output

increases with establishment size for all three media: air, water, and solid waste. That is, an

increase in all three media categories underlie the increase seen above in the total. We find the

same by type of cost: expenditures (per unit of output) on labor, depreciation, and "other" all

increase with establishment size. While all observable components of PAOC contribute to the

results seen above using total PAOC, we note that, conditional on having nonzero PAOC, larger

establishments devote a higher proportion of their PAOC to air and water (versus solid waste)

and labor and depreciation (versus other costs), suggesting that these components contribute

especially to the size-expenditure relationship.

To this point, we have used the same set of explanatory variables in our estimating

equations. We now consider alternate sets of controls variables. In the first column of Tables IV

and V, we present the regression results from the TPM and Heckit model, respectively, that

underlie the estimated incremental effects presented in the last three columns of Table III. That

is, column (1) in both of these tables contains the coefficients from our baseline empirical

specification, on the sample that removes 628 outliers (i.e., observations that had PAOC

expenditures that was more than 15% of establishment output). From this specification, we

subtract the state dummy variables in column (2). The resulting establishment size coefficients

will then also capture any state-level variation that existed in regulatory stringency, enforcement

intensity, and environmental compliance costs. We see that excluding state dummies has

---

[32] The 1999 survey does not breakdown total PAOC into these subcategories at all.

relatively little impact on the establishment size coefficients. In Table VI, we present the corresponding estimates of the average incremental effects. We see that in the Heckit, the upward slope is increased very slightly (by about 2%), relative to our baseline specification, and in the TPM, the estimates are either identical or differ by only a penny. Whatever state-level variation existed is neither obscuring nor driving the positive relationship seen between PAOC intensity and establishment size.

Finally, in column (3) of Tables IV and V, we add other establishment-level controls to the baseline specification — ones known to be related to establishment size. One is a dummy variable indicated whether the establishment belongs to a multi-unit firm, which is a factor we explored above. The other is a measure of establishment age.[33] Here we employ a series of cohort dummy variables indicating a plant's first observed appearance in a Census of Manufactures, which is conducted every five years, in years ending in 2 and 7, with the exception of 1963, which is also the earliest CM with available microdata. The expected effects here are ambiguous. Establishments built in the distant past no doubt had less efficient, more polluting capital equipment than newer cohorts of plants, which may raise their pollution abatement operating costs per unit of output. However, environmental regulations sometimes 'grandfather' older, existing establishments from the most stringent requirements. But that grandfathered status is usually surrendered when there is a significant expansion or modernization of operations. So, a given cohort of establishments will no doubt have a mix of establishments with different vintages of production and abatement equipment and with different regulatory requirements – a mix that changes over time – leaving ambiguous the net impact of cohort on PAOC intensity.

---

[33] There is evidence in the job creation literature that firm age actually underlies employment growth often attributed to firm size (Haltiwanger *et al.* 2013).

We see in Tables IV and V that establishments in multi-unit firms have a higher probability of nonzero PAOC as well as a greater intensity of PAOC expenditure. Becker and Henderson (2000) and Becker (2005) have previously found that multi-unit status impacts the timing and intensity of regulation and expenditure. With more at stake in terms of reputation and assets, such firms may be more likely to engage in pollution abatement, all else being equal. Meanwhile, the cohort effects (not reported here, in the interest of space) reveal that the older cohorts tend to have the highest PAOC intensity.

Of particular interest here is whether the inclusion of these additional establishment-level controls impacts the estimates of the average incremental effects. In the final column of Table VI, we see that estimates are in fact different. In the Heckit, the estimated effects are all 20–48% lower than in the baseline specification, and in the TPM, the estimates are all lower as well. Clearly, some of the (positive) effect of establishment size on PAOC intensity seen in the basic specification is instead attributable to corporate structure and establishment age. Nevertheless, even after controlling for these other establishment characteristics, the positive relationship between PAOC intensity and establishment size is still apparent. In the Heckit, PAOC intensity increases monotonically with establishment size, and the very largest establishments (with 1000+ employees) had $4.41 more PAOC per $1000 of output than the establishments with 1-49 employees (compared with $5.49 in the baseline specification). In the TPM, establishments with 50-99 employees had the lowest PAOC intensity, with PAOC intensity increasing monotonically after that point, just as in the baseline specification. Here, establishments with 1000+ employees had $1.34 more PAOC per $1000 of output than the establishments with 1-49 employees (compared to $1.93 in the baseline specification). The basic story is consistent throughout our analyses.

21

V. CONCLUDING REMARKS

Results in this paper show that spending on pollution abatement operating costs per unit of output increases with establishment size. In the Heckman two-step selection (Heckit) model, controlling for industry, state, and year, the very largest establishments (with 1000+ employees) had $5.49 more PAOC per $1000 of output than the establishments with 1-49 employees. In the two-part model (TPM), establishments with 1000+ employees had about $1.92 more PAOC per $1000 of output than establishments with 1-49 employees. We argue that the true relationship lies somewhere between these two extremes, approaching the TPM results as the proportion of zeros that are "true" rises. Our results apply to firm size as well as establishment size. These results are contrary to the claims by some; we certainly do not see environmental compliance costs for small businesses that are multiples of those for large (Crain and Crain 2010). Our findings suggest that there are statutory and/or enforcement asymmetries that favor smaller establishments and that outweigh any compliance asymmetries.

In future work, we hope to explore these results further. Here, we presented results for the entire manufacturing sector. However, pollution abatement expenditures are highly concentrated. Four manufacturing subsectors (i.e., two-digit SIC / three-digit NAICS) dominate in sheer expenditure as well as in PAOC intensity. In particular, paper, chemicals, petroleum/coal products, and primary metals together accounted for 63% of total PAOC in the manufacturing sector in both 1974 and 2005 – the first and last year studied here. This paper's analyses could be performed separately for each of these major manufacturing subsectors.

Analyses on market entry and exit may also be revealing. While we find no cost disadvantage for *existing* small businesses, regulations sometimes impose more stringent

requirements on *new* facilities. If compliance with regulations is capital intensive, and it often is, compliance asymmetries may deter the entry of new, small establishments, particularly if they are not offset by regulatory exemptions. This possibility should be kept in mind when interpreting our results. That said, much of the public discussion is about impacts on existing businesses, not potential businesses. Meanwhile, provisions that exempt existing facilities from newer, more stringent regulations make the effects on exit more ambiguous and/or less immediate.

For better or worse, public discussion also tends to focus on the size dimension alone. Here we demonstrate that corporate structure (i.e., single- versus multi-unit firm status) and establishment age are both related to establishment size and to environmental compliance costs. Nevertheless, even though estimated effects are somewhat lower with the inclusion of additional control variables, especially these additional plant-level controls, we consistently find that spending on pollution abatement operating costs per unit of output increases with establishment size.

REFERENCES

Becker, Randy A. "Air Pollution Abatement Costs under the Clean Air Act: Evidence from the PACE Survey," *Journal of Environmental Economics and Management*, 50(1), July 2005, 144-169.

Becker, Randy A. "On Spatial Heterogeneity in Environmental Compliance Costs," *Land Economics*, 87(1), February 2011a, 28-44.

Becker, Randy A. "Local Environmental Regulation and Plant-level Productivity," *Ecological Economics*, 70(12), October 2011b, 2516-2522.

Becker, Randy A. and Vernon Henderson. "Effects of Air Quality Regulations on Polluting Industries," *Journal of Political Economy*, 108(2), April 2000, 379-421.

Becker, Randy A. and Ronald J. Shadbegian. "A Change of PACE: Comparing the 1994 and 1999 Pollution Abatement Costs and Expenditures Surveys," *Journal of Economic and Social Measurement*, 30(1), 2005, 63-95.

Berman, Eli, and Linda T. Bui. "Environmental Regulation and Labor Demand: Evidence from the South Coast Air Basin," *Journal of Public Economics*, 79(2), February 2001a, 265-295.

Berman, Eli, and Linda T. Bui. "Environmental Regulation and Productivity: Evidence from Oil Refineries," *Review of Economics and Statistics*, 83(3), August 2001b, 498-510.

Brock, William A. and David S. Evans, "The Economics of Regulatory Tiering," *RAND Journal of Economics*, 16(3), Autumn 1985, 398-409.

Cameron, A. Colin, and Pravin K. Trivedi. *Microeconometrics Using Stata (Revised Edition)*. College Station, TX: Stata Press, 2010.

Collins, Alan, and Richard I. D. Harris. "Does Plant Ownership Affect the Level of Pollution Abatement Expenditure?" *Land Economics*, 78(2), May 2002, 171-189.

Collins, Alan, and Richard I. D. Harris. "The Impact of Foreign Ownership and Efficiency on Pollution Abatement Expenditure by Chemical Plants: Some U.K. Evidence," *Scottish Journal of Political Economy*, 52(5), November 2005, 747-768.

Crain, W. Mark. *The Impact of Regulatory Costs on Small Firms*. Washington, DC: U.S. Small Business Administration, Office of Advocacy, September 2005.

Crain, Nicole V. and W. Mark Crain. *The Impact of Regulatory Costs on Small Firms*. Washington, DC: U.S. Small Business Administration, Office of Advocacy, September 2010.

Crain, W. Mark, and Thomas D. Hopkins. *The Impact of Regulatory Costs on Small Firms*. Washington, DC: U.S. Small Business Administration, Office of Advocacy, October 2001.

Dean, Thomas J., Robert L. Brown, and Victor Stango. "Environmental Regulation as a Barrier to the Formation of Small Manufacturing Establishments: A Longitudinal Examination," *Journal of Environmental Economics and Management*, 40(1), July 2000, 56-75.

Dow, William H. and Edward C. Norton. "Choosing Between and Interpreting the Heckit and Two-Part Models for Corner Solutions," *Health Services & Outcomes Research Methodology*, 4(1), March 2003, 5-18.

Duan, Naihua. "Smearing Estimate: A Nonparametric Retransformation Method," *Journal of American Statistical Association*, 78(383), September 1983, 605-610.

Evans, David S. "The Differential Effect of Regulation Across Plant Size: Comment on Pashigian," *Journal of Law and Economics*, 29(1), April 1986, 187-200.

Gray, Wayne B. and Mary E. Deily. "Compliance and Enforcement: Air Pollution Regulation in the U.S. Steel Industry," *Journal of Environmental Economics and Management*, 31(1), July 1996, 96-111.

Gray, Wayne B. and Ronald J. Shadbegian. "Environmental Regulation, Investment Timing, and Technology Choice," *Journal of Industrial Economics*, 46(2), June 1998, 235-256.

Gray, Wayne B. and Ronald J. Shadbegian. "Pollution Abatement Costs, Regulation, and Plant-Level Productivity," in *Economic Costs and Consequences of Environmental Regulation*, Wayne B. Gray (ed), Ashgate Publishing, 2002.

Gray, Wayne B. and Ronald J. Shadbegian. "Plant Vintage, Technology, and Environmental Regulation," *Journal of Environmental Economics and Management*, 46(3), November 2003, 384-402.

Greenstone, Michael, John A. List, and Chad Syverson. "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing," *National Bureau of Economic Research Working Paper*, 18392, September 2012.

Haller, Stefanie A. and Liam Murphy. "Corporate Expenditure on Environmental Protection," *Environmental and Resource Economics*, 51(2), February 2012, 277-296.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. "Who Creates Jobs: Small vs. Large vs. Young," *Review of Economics and Statistics*, forthcoming (May 2013).

Harrington, Winston. "Grading Estimates of the Benefits and Costs of Federal Regulation: A Review of Reviews," *RFF Discussion Paper*, 06-30, September 2006.

Hartman, Raymond S., David Wheeler, and Manjula Singh. "The Cost of Air Pollution Abatement," *Applied Economics*, 29(6), 1997, 759-774.

Heyes, Anthony. "Is Environmental Regulation Bad for Competition? A Survey," *Journal of Regulatory Economics*, 36(1), August 2009, 1-28.

Jones, Andrew M. "Health Econometrics," in *Handbook of Health Economics, Volume 1*, A.J. Cuyer and J.P. Newhouse (eds.), Oxford, UK: North Holland, 2000.

Jorgenson, Dale W. and Peter J. Wilcoxen. "Environmental Regulation and U.S. Economic Growth," *RAND Journal of Economics*, 21(2), Summer 1990, 314-340.

Keller, Wolfgang, and Arik Levinson. "Pollution Abatement Costs and Foreign Direct Investment Inflows to U.S. States," *Review of Economics and Statistics*, 84(4), November 2002, 691-703.

List, John A. and Catherine Y. Co. "The Effects of Environmental Regulations on Foreign Direct Investment," *Journal of Environmental Economics and Management*, 40(1), July 2000, 1-20.

List, John A., Daniel L. Millimet, Per G. Fredriksson, and W. Warren McHone. "Effects of Environmental Regulations on Manufacturing Plant Births: Evidence from a Propensity Score Matching Estimator," *Review of Economics and Statistics*, 85(4), November 2003, 944-952.

Madden, David. "Sample Selection versus Two-Part Models Revisited: The Case of Female Smoking and Drinking," *Journal of Health Economics*, 27(2), August 2008, 300-307.

Millimet, Daniel L. "Environmental Abatement Costs and Establishment Size," *Contemporary Economic Policy*, 21(3), July 2003, 281-296.

Millimet, Daniel L., Santanu Roy, and Aditi Sengupta, "Environmental Regulations and Economic Activity: Influence on Market Structure," *Annual Review of Resource Economics*, 1, 2009, 99-117.

Morgenstern, Richard D., William A. Pizer, and Jhih-Shyang Shih. "Jobs versus the Environment: An Industry-Level Perspective," *Journal of Environmental Economics and Management*, 43(3), May 2002, 412-436.

Pashigian, B. Peter. "The Effect of Environmental Regulation on Optimal Plant Size and Factor Shares," *Journal of Law and Economics*, 27(1), April 1984, 1-28.

Shadbegian, Ronald J. and Wayne B. Gray. "Pollution Abatement Expenditures and Plant-Level Productivity: A Production Function Approach," *Ecological Economics*, 54(2-3), August 2005, 196-208.

U.S. Census Bureau. *Pollution Abatement Costs and Expenditures, ___*. Washington, DC: U.S. Government Printing Office, various years.

U.S. Environmental Protection Agency. *Final Guidance for EPA Rulewriters: Regulatory Flexibility Act as Amended by the Small Business Regulatory Enforcement Fairness Act*, November 2006.

Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

TABLE I
Intensity of Pollution Abatement Operating Costs:
Two-Part Model[†]

	(1)	(2)	(3)
Probability (PAOC/VS >0)			
Establishment employment / 100	+0.0894***		
	(0.0035)		
Log (establishment employment)		+0.4616***	
		(0.0044)	
50-99 employees			+0.5077***
			(0.0119)
100-249 employees			+0.8918***
			(0.0120)
250-499 employees			+1.2557***
			(0.0144)
500-999 employees			+1.5172***
			(0.0181)
1000+ employees			+1.9086***
			(0.0266)
Constant	yes	yes	yes
Industry effects (4-digit SIC)	yes	yes	yes
State effects	yes	yes	yes
Year effects	yes	yes	yes
Pseudo R-squared	0.1882	0.2364	0.2302
Log(PAOC/VS) conditional on (PAOC/VS >0)			
Establishment employment / 100	+0.0032***		
	(0.0006)		
Log (establishment employment)		−0.0239***	
		(0.0056)	
50-99 employees			−0.2751***
			(0.0185)
100-249 employees			−0.3193***
			(0.0184)
250-499 employees			−0.3043***
			(0.0196)
500-999 employees			−0.2404***
			(0.0218)
1000+ employees			−0.1353***
			(0.0255)
Constant	yes	yes	yes
Industry effects (4-digit SIC)	yes	yes	yes
State effects	yes	yes	yes
Year effects	yes	yes	yes
R-squared	0.3941	0.3939	0.3969
Number of observations	321,526	321,526	321,526

[†] Statistical significance at the 10%, 5%, and 1% level are indicated by single, double, and triple asterisks, respectively. Robust standard errors, clustered by establishment, are in parentheses.

TABLE II
Intensity of Pollution Abatement Operating Costs:
Heckman Two-Step Selection (Heckit) Model[†]

	(1)	(2)	(3)
Probability (PAOC/VS >0)			
Establishment employment / 100	+0.0894***		
	(0.0035)		
Log (establishment employment)		+0.4616***	
		(0.0044)	
50-99 employees			+0.5077***
			(0.0119)
100-249 employees			+0.8918***
			(0.0120)
250-499 employees			+1.2557***
			(0.0144)
500-999 employees			+1.5172***
			(0.0181)
1000+ employees			+1.9086***
			(0.0266)
Constant	yes	yes	yes
Industry effects (4-digit SIC)	yes	yes	yes
State effects	yes	yes	yes
Year effects	yes	yes	yes
Log(PAOC/VS) conditional on (PAOC/VS >0)			
Establishment employment / 100	+0.0051***		
	(0.0008)		
Log (establishment employment)		+0.2281***	
		(0.0106)	
50-99 employees			+0.0199
			(0.0265)
100-249 employees			+0.1556***
			(0.0344)
250-499 employees			+0.3138***
			(0.0419)
500-999 employees			+0.4574***
			(0.0469)
1000+ employees			+0.6575***
			(0.0512)
Constant	yes	yes	yes
Industry effects (4-digit SIC)	yes	yes	yes
State effects	yes	yes	yes
Year effects	yes	yes	yes
Rho	+0.2745	+1.0000	+0.8538
Sigma	+1.2634	+1.6840	+1.4182
Lambda	+0.3468***	+1.6840***	+1.2110***
	(0.0248)	(0.0328)	(0.0336)
Number of observations	321,526	321,526	321,526

[†] Statistical significance at the 10%, 5%, and 1% level are indicated by single, double, and triple asterisks, respectively. Robust standard errors, clustered by establishment, are in parentheses.

TABLE III
Average Incremental Effect by Establishment Size Category[†]
(in dollars of additional PAOC per $1000 of establishment output)

	Full sample			Sample with outliers removed		
	Two-part model		Heckit model	Two-part model		Heckit model
	Normal	Duan (1983)	Normal	Normal	Duan (1983)	Normal
50-99 employees	−0.31 (0.09)	−0.51 (0.18)	+0.18 (0.13)	−0.05 (0.08)	−0.05 (0.08)	+0.48 (0.12)
100-249 employees	+0.10 (0.09)	+0.16 (0.15)	+0.98 (0.17)	+0.35 (0.07)	+0.35 (0.07)	+1.33 (0.16)
250-499 employees	+0.56 (0.09)	+0.91 (0.21)	+2.05 (0.24)	+0.79 (0.08)	+0.78 (0.08)	+2.44 (0.23)
500-999 employees	+1.03 (0.10)	+1.68 (0.32)	+3.17 (0.31)	+1.26 (0.08)	+1.24 (0.08)	+3.60 (0.29)
1000+ employees	+1.73 (0.12)	+2.81 (0.51)	+5.03 (0.41)	+1.93 (0.11)	+1.91 (0.11)	+5.49 (0.41)
Number of observations	321,526	321,526	321,526	320,898	320,898	320,898

[†] Estimated incremental effects are based on column (3) of Table I and Table II, respectively, and on a sample that excludes outliers. Effects are relative to the omitted size category (i.e., establishments with 1-49 employees). Standard errors are in parentheses.

TABLE IV
Intensity of Pollution Abatement Operating Costs:
Two-Part Model with Alternate Control Variables[†]

	(1)	(2)	(3)
Probability (PAOC/VS >0)			
50-99 employees	+0.5134***	+0.5153***	+0.4445***
	(0.0119)	(0.0119)	(0.0121)
100-249 employees	+0.8981***	+0.9014***	+0.7578***
	(0.0121)	(0.0120)	(0.0125)
250-499 employees	+1.2627***	+1.2679***	+1.0738***
	(0.0144)	(0.0144)	(0.0152)
500-999 employees	+1.5244***	+1.5310***	+1.3120***
	(0.0181)	(0.0181)	(0.0191)
1000+ employees	+1.9165***	+1.9245***	+1.6831***
	(0.0266)	(0.0266)	(0.0274)
Multi-unit firm			+0.3732***
			(0.0107)
Cohort effects	*no*	*no*	*yes*
Constant	*yes*	*yes*	*yes*
Industry effects (4-digit SIC)	*yes*	*yes*	*yes*
State effects	*yes*	*no*	*yes*
Year effects	*yes*	*yes*	*yes*
Pseudo R-squared	0.2307	0.2293	0.2391
Log(PAOC/VS) conditional on (PAOC/VS >0)			
50-99 employees	–0.2249***	–0.2254***	–0.2350***
	(0.0179)	(0.0179)	(0.0180)
100-249 employees	–0.2619***	–0.2640***	–0.2849***
	(0.0178)	(0.0178)	(0.0184)
250-499 employees	–0.2417***	–0.2466***	–0.2759***
	(0.0191)	(0.0191)	(0.0201)
500-999 employees	–0.1741***	–0.1794***	–0.2148***
	(0.0213)	(0.0213)	(0.0224)
1000+ employees	–0.0656***	–0.0686***	–0.1150***
	(0.0251)	(0.0251)	(0.0262)
Multi-unit firm			+0.0268*
			(0.0140)
Cohort effects	*no*	*no*	*yes*
Constant	*yes*	*yes*	*yes*
Industry effects (4-digit SIC)	*yes*	*yes*	*yes*
State effects	*yes*	*no*	*yes*
Year effects	*yes*	*yes*	*yes*
R-squared	0.4001	0.3954	0.4007
Number of observations	320,898	320,898	320,898

[†] Statistical significance at the 10%, 5%, and 1% level are indicated by single, double, and triple asterisks, respectively. Robust standard errors, clustered by establishment, are in parentheses.

TABLE V
Intensity of Pollution Abatement Operating Costs:
Heckman Two-Step Selection (Heckit) Model with Alternate Control Variables[†]

|	(1)	(2)	(3)
Probability (PAOC/VS >0)			
50-99 employees	+0.5134***	+0.5153***	+0.4445***
	(0.0119)	(0.0119)	(0.0121)
100-249 employees	+0.8981***	+0.9014***	+0.7578***
	(0.0121)	(0.0120)	(0.0125)
250-499 employees	+1.2627***	+1.2679***	+1.0738***
	(0.0144)	(0.0144)	(0.0152)
500-999 employees	+1.5244***	+1.5310***	+1.3120***
	(0.0181)	(0.0181)	(0.0191)
1000+ employees	+1.9165***	+1.9245***	+1.6831***
	(0.0266)	(0.0266)	(0.0274)
Multi-unit firm			+0.3732***
			(0.0107)
Cohort effects	*no*	*no*	*yes*
Constant	*yes*	*yes*	*yes*
Industry effects (4-digit SIC)	*yes*	*yes*	*yes*
State effects	*yes*	*no*	*yes*
Year effects	*yes*	*yes*	*yes*
Log(PAOC/VS) conditional on (PAOC/VS >0)			
50-99 employees	+0.0867***	+0.0916***	+0.0399
	(0.0260)	(0.0262)	(0.0253)
100-249 employees	+0.2373***	+0.2441***	+0.1439***
	(0.0339)	(0.0342)	(0.0315)
250-499 employees	+0.4067***	+0.4137***	+0.2807***
	(0.0414)	(0.0420)	(0.0380)
500-999 employees	+0.5573***	+0.5660***	+0.4154***
	(0.0442)	(0.0450)	(0.0408)
1000+ employees	+0.7648***	+0.7776***	+0.6068***
	(0.0509)	(0.0519)	(0.0471)
Multi-unit firm			+0.2553***
			(0.0178)
Cohort effects	*no*	*no*	*yes*
Constant	*yes*	*yes*	*yes*
Industry effects (4-digit SIC)	*yes*	*yes*	*yes*
State effects	*yes*	*no*	*yes*
Year effects	*yes*	*yes*	*yes*
Rho	+0.8892	+0.8956	+0.9052
Sigma	+1.4146	+1.4242	+1.4200
Lambda	+1.2579***	+1.2755***	+1.2853***
	(0.0332)	(0.0336)	(0.0333)
Number of observations	320,898	320,898	320,898

[†] Statistical significance at the 10%, 5%, and 1% level are indicated by single, double, and triple asterisks, respectively. Robust standard errors, clustered by establishment, are in parentheses.

TABLE VI
Average Incremental Effect by Establishment Size Category[†]
(in dollars of additional PAOC per $1000 of establishment output)

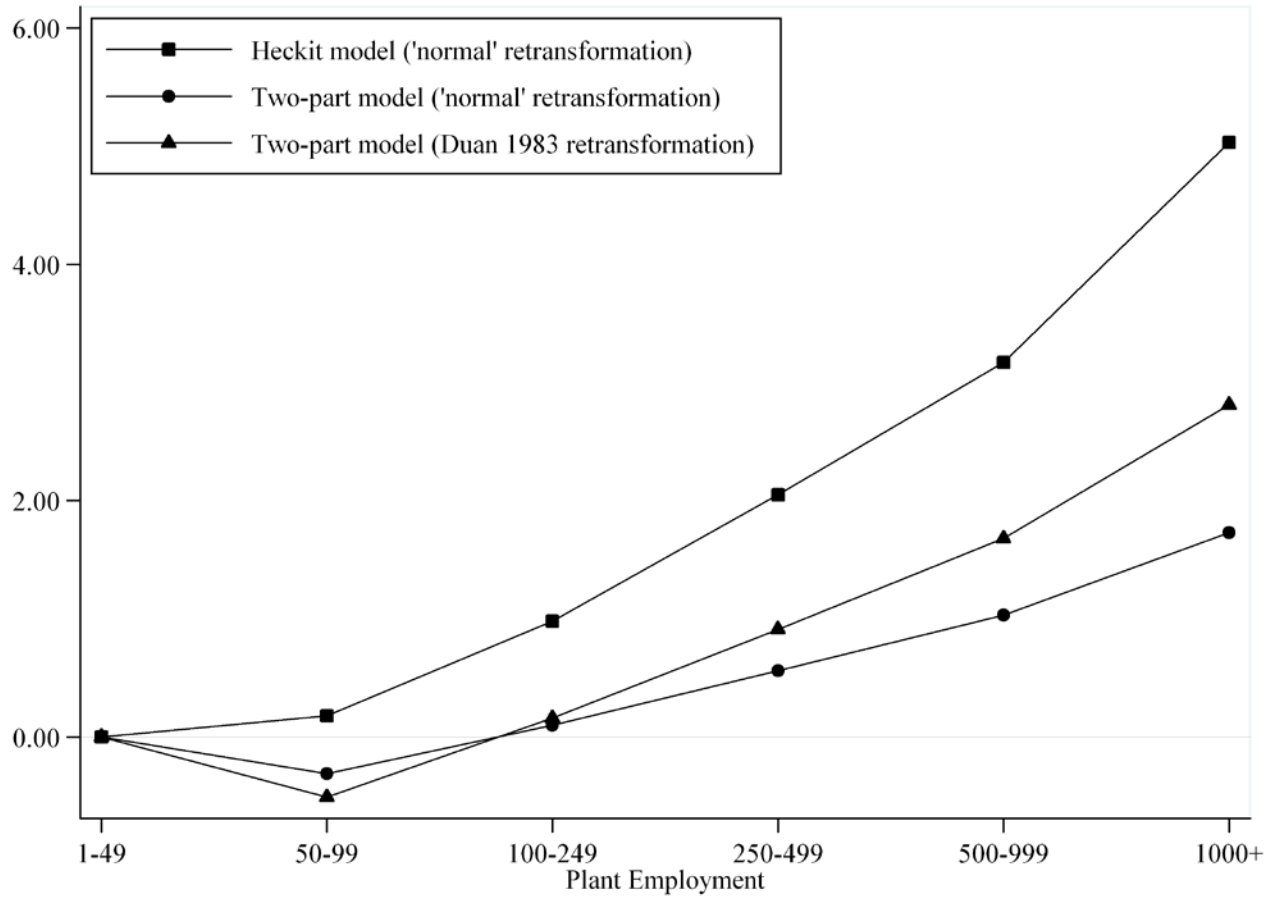|	Baseline specification		Without state effects		With additional plant controls	
Two-part model						
50-99 employees	−0.05	(0.08)	−0.04	(0.08)	−0.36	(0.09)
100-249 employees	+0.35	(0.07)	+0.36	(0.07)	−0.14	(0.09)
250-499 employees	+0.79	(0.08)	+0.79	(0.08)	+0.22	(0.09)
500-999 employees	+1.26	(0.08)	+1.25	(0.08)	+0.67	(0.10)
1000+ employees	+1.93	(0.11)	+1.93	(0.11)	+1.34	(0.12)
Heckit model						
50-99 employees	+0.48	(0.12)	+0.50	(0.12)	+0.25	(0.13)
100-249 employees	+1.33	(0.16)	+1.37	(0.17)	+0.86	(0.16)
250-499 employees	+2.44	(0.23)	+2.48	(0.24)	+1.75	(0.22)
500-999 employees	+3.60	(0.29)	+3.65	(0.30)	+2.75	(0.27)
1000+ employees	+5.49	(0.41)	+5.61	(0.42)	+4.41	(0.37)
Number of observations	320,898		320,898		320,898	

[†] Estimated incremental effects are based on columns (1), (2), and (3), respectively, of Table IV and Table V, respectively. Effects are relative to the omitted size category (i.e., establishments with 1-49 employees). Standard errors are in parentheses.

FIGURE 1
PAOC Intensity in the Manufacturing Sector by Establishment Size Category
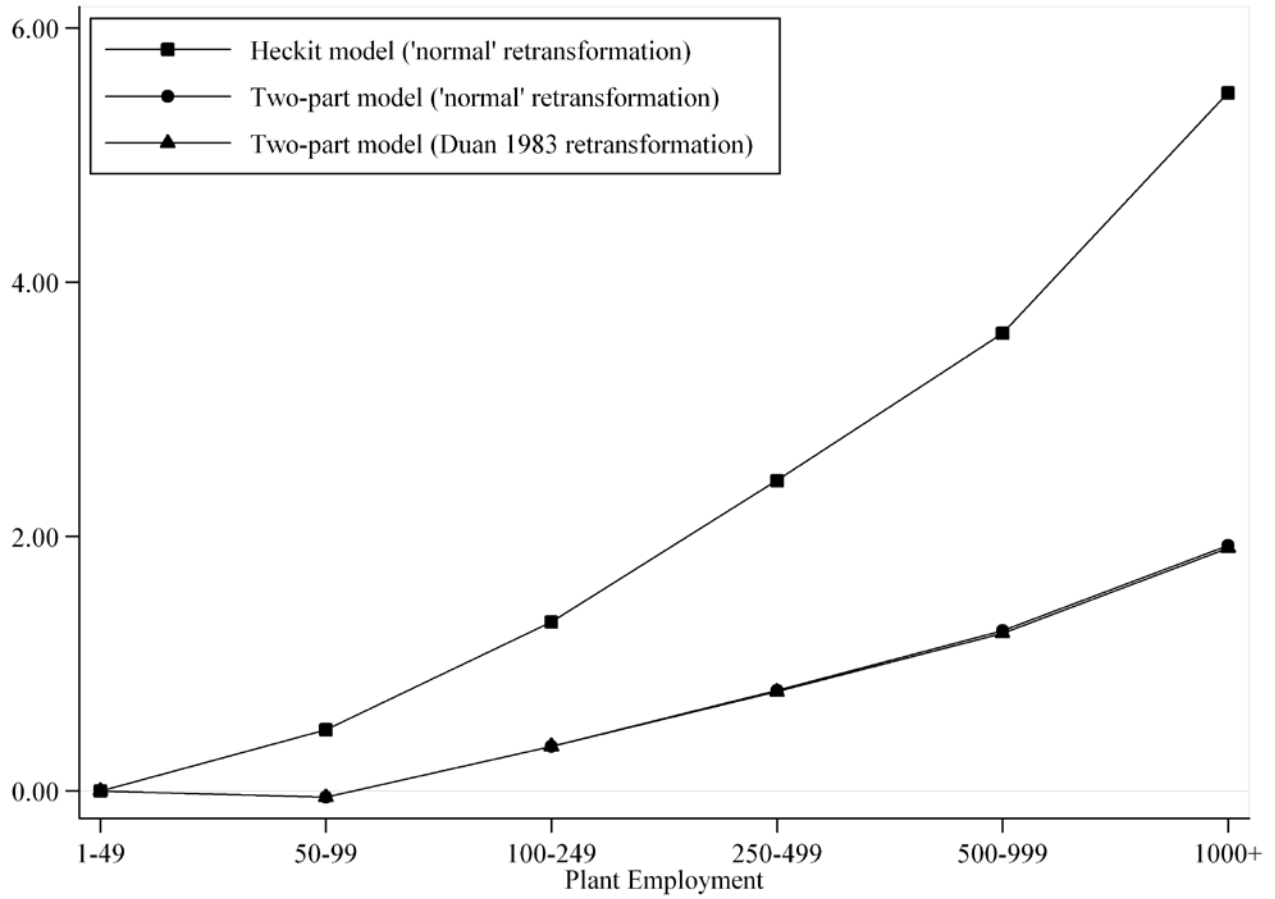(in dollars of PAOC per $1000 of output)
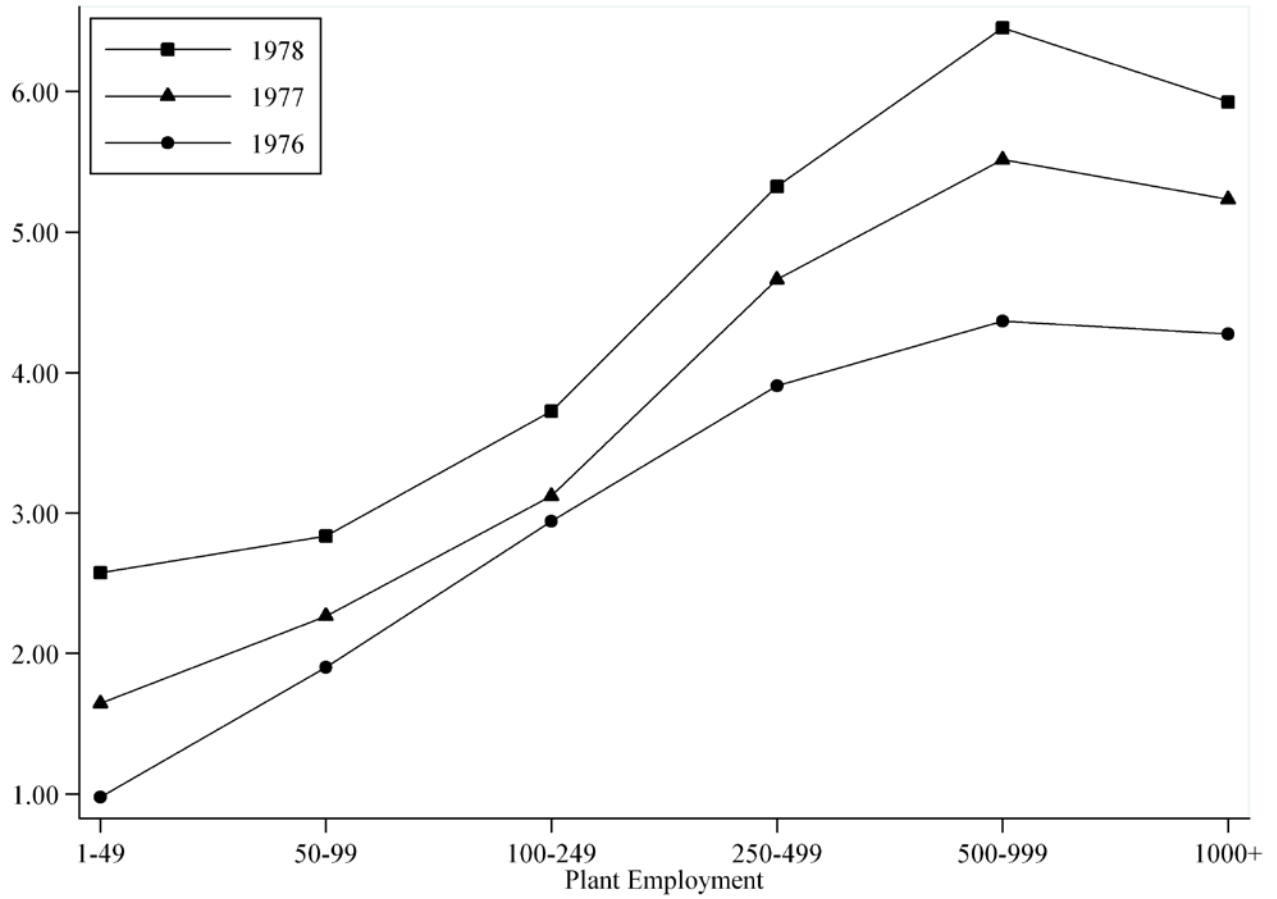
34

FIGURE 2
Average Incremental Effect by Establishment Size Category:
Full Sample
(in dollars of additional PAOC per $1000 of establishment output)

Legend:
- Heckit model ('normal' retransformation)
- Two-part model ('normal' retransformation)
- Two-part model (Duan 1983 retransformation)

X-axis: Plant Employment — 1-49, 50-99, 100-249, 250-499, 500-999, 1000+

Source: First three columns of Table III.

FIGURE 3
Average Incremental Effect by Establishment Size Category:
Sample with Outliers Removed
(in dollars of additional PAOC per $1000 of establishment output)

Legend:
- ■ Heckit model ('normal' retransformation)
- ● Two-part model ('normal' retransformation)
- ▲ Two-part model (Duan 1983 retransformation)

X-axis (Plant Employment): 1-49, 50-99, 100-249, 250-499, 500-999, 1000+
Y-axis: 0.00, 2.00, 4.00, 6.00

Source: Final three columns of Table III.

FIGURE 4
Average Incremental Effect by Establishment Size Category
Under Different Treatment of Outliers:
Two-Part Model
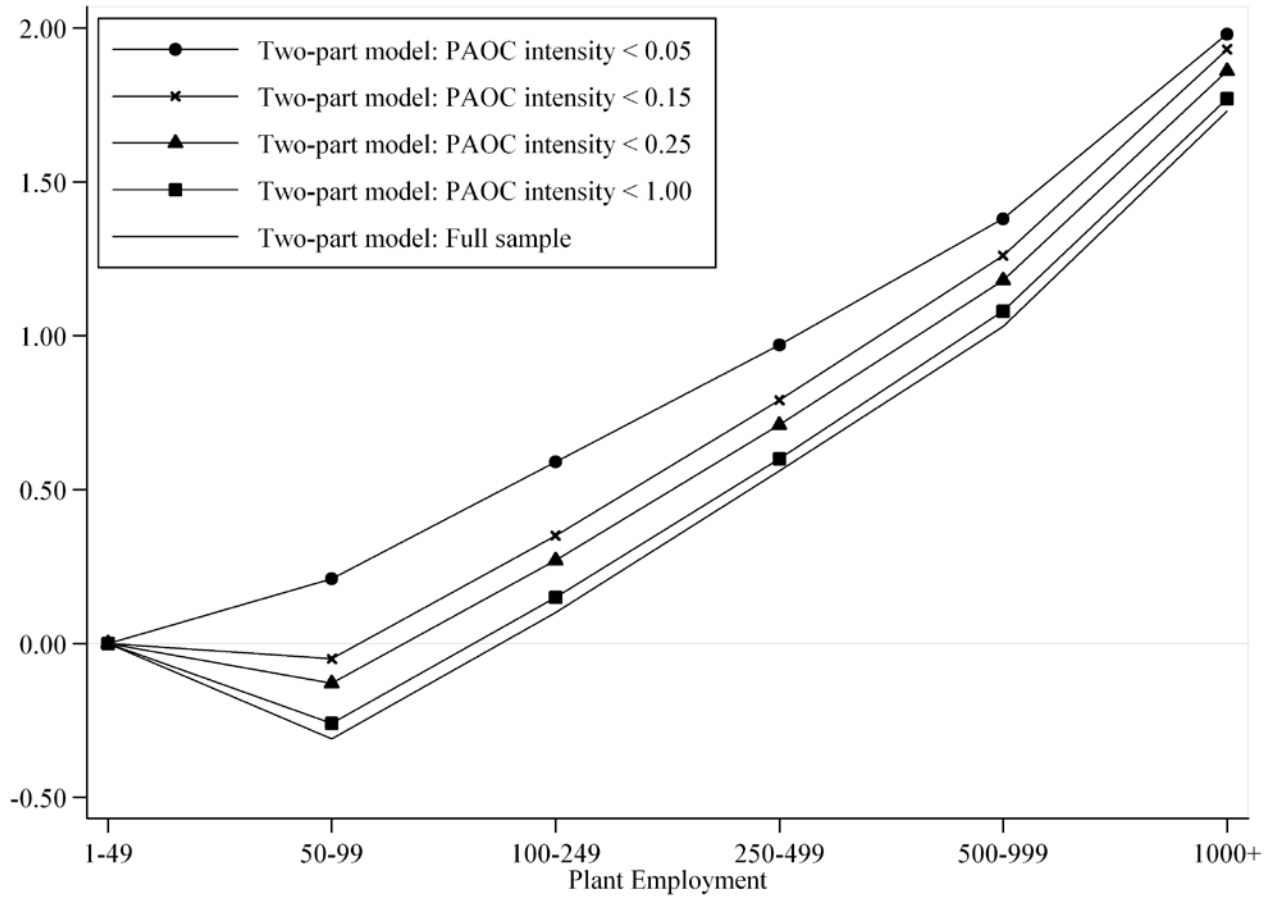(in dollars of additional PAOC per $1000 of establishment output)

FIGURE 5
Average Incremental Effect by Establishment Size Category
Under Different Treatment of Outliers:
Heckit Model
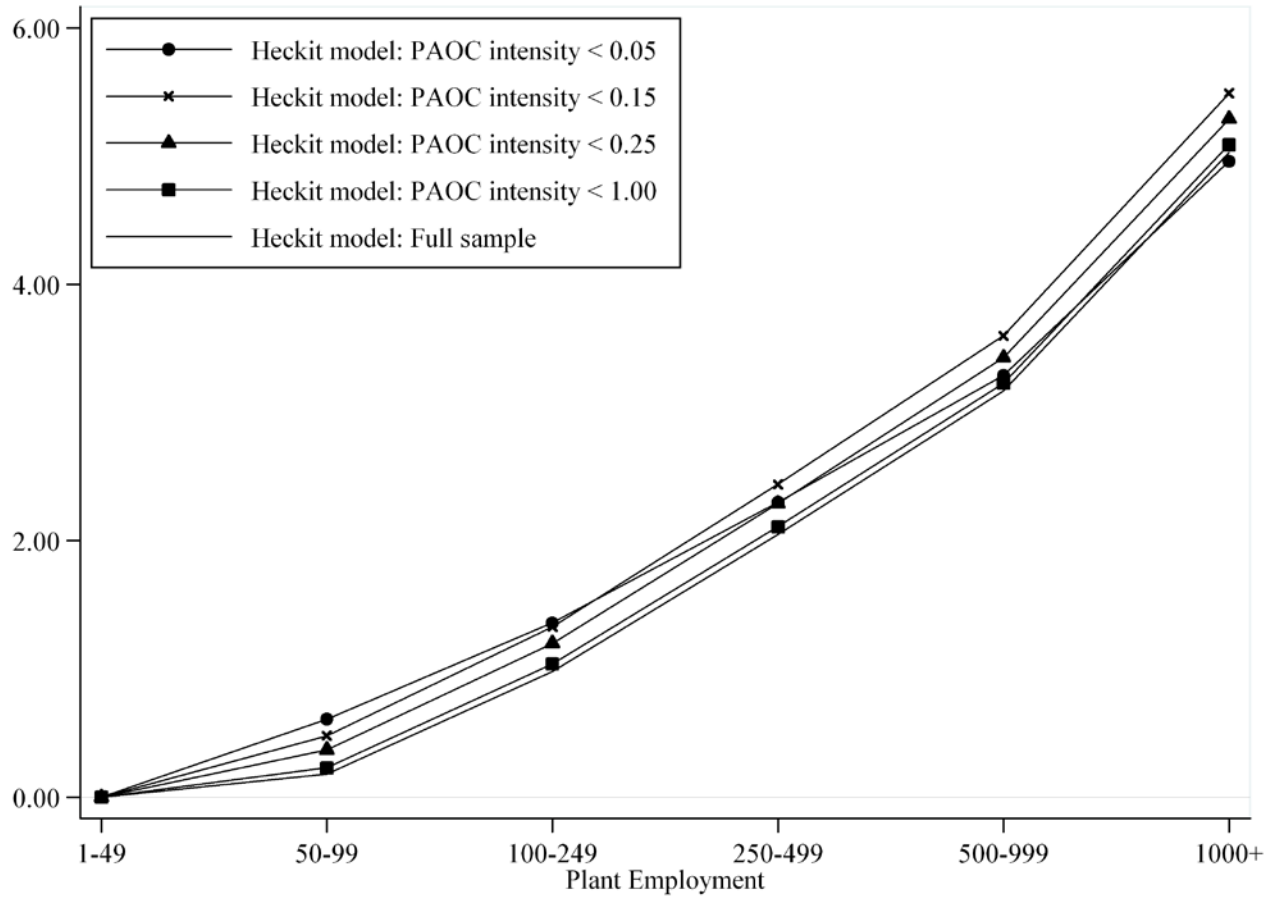(in dollars of additional PAOC per $1000 of establishment output)

FIGURE 6
Average Incremental Effect by Establishment Size Category:
Sample with Outliers Removed
(in dollars of additional PAOC per employee)

FIGURE 7
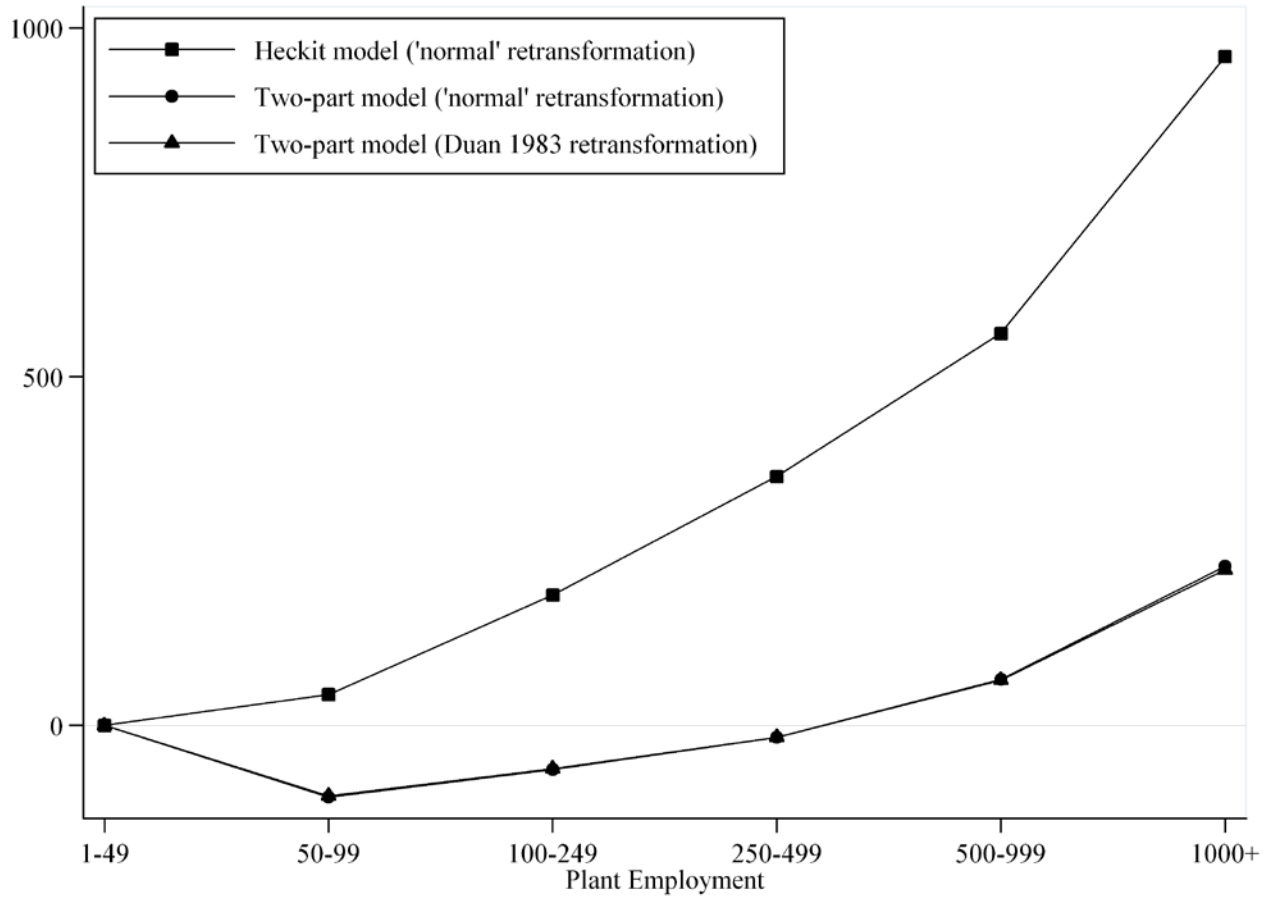**Average Incremental Effect by Establishment Size Category:**
Multi-Unit vs. Single-Unit Firms
(in dollars of additional PAOC per $1000 of establishment output)