Productivity, Export, and Environmental Performance: Air Pollutants in the United States

Jingbo Cui, Harvey Lapan, and GianCarlo Moschini

This paper studies the firm-level relationship among productivity, decision to export, and environmental performance. The emerging theoretical and empirical literature suggests that trade has an important role in determining firms' heterogeneity: increased openness to trade induces a reallocation effect that increases within-industry efficiency, thereby linking firms' decisions to export and adopt newer (and cleaner) technology. We argue that this framework provides the following empirically-relevant predictions: there is an inverse relationship between firm productivity and pollution emissions per unit output; exporting firms have lower emissions per unit output; and larger firms have a lower emission intensity. To examine these implications empirically, we have assembled a uniquely detailed dataset of the U.S. manufacturing industry for the years 2002, 2005, and 2008 by matching facility-level air emission data from the U.S. Environmental Protection Agency with the facility's economic characteristics contained in the National Establishment Time Series database. The strategy is to first estimate a facility-level total factor productivity parameter as a plant-specific fixed effect. We then investigate how this estimated productivity parameter correlates with emission intensity on a pollutant-by-pollutant basis. Our empirical findings support the hypotheses suggested by the conceptual model. For each criteria air pollutant considered, we find a significant negative correlation between estimated facility productivity and emission intensity. Conditional on a facility's estimated productivity and other controls, exporting facilities have significantly lower emissions per value of sales than non-exporting facilities in the same industry. We also find that plant size is negatively and significantly related to emission intensity for all pollutants.

Key words: Export, facility-level pollution, heterogeneous firms, total factor productivity, trade and the environment.

JEL codes: F18, Q56.

It is recognized that the relationship between international trade and the environment is a complex one, with several distinct possible effects that pull in opposite directions. Not surprisingly, therefore, no consensus has emerged on whether the overall impact of increased economic integration across national borders increases or decreases pollution; neoclassical trade theory suggests that either outcome is possible (Copeland and Taylor 1994; 1995). Empirical investigations have documented the separate roles of the "scale effect" (i.e., trade tends to expand economic activity, which *ceteris paribus* worsens the environment), the "technique effect" (i.e., trade raises national income, which is presumed to lead to more stringent environmental regulations that benefit the environment), and the "composition effect" (i.e., trade leads to greater specialization and redistribution of factors across industries, with ambiguous effects on the environment). Using panel data of aggregate (country-level) sulfur dioxide (SO_2) pollution, Antweler, Copeland, and Taylor (2001) find that the technique effect tends to dominate the scale effect and the composition

Jingbo Cui is an assistant professor in the Institute for Advanced Study at the Economics and Management School, Wuhan University, China. Harvey Lapan is a university professor in the Department of Economics, Iowa State University. GianCarlo Moschini is a professor and Pioneer Chair in Science and Technology Policy, Department of Economics and Center for Agricultural and Rural Development, Iowa State University. The authors thank editor JunJie Wu and the journal's reviewers for their comments. Correspondence may be sent to: moschini@iastate.edu.

Amer. J. Agr. Econ. 98(2): 447–467; doi: 10.1093/ajae/aav066
Published online December 14, 2015

effect to reduce pollution, so that the overall effect of trade is favorable to the environment. Frankel and Rose (2005), using a cross-section data set and a methodology that deals with the endogeneity of trade, also find that trade reduces SO_2 pollution. The results in Cole and Elliott (2003) are similar with regard to SO_2 emission, but they find that trade may increase carbon dioxide (CO_2) and nitrogen oxides (NO_x) pollution. The results in Managi, Hibiki, and Tsurumi (2009), who rely on a larger sample of countries, are also decidedly mixed, with the impact of trade on the environment varying across pollutants and countries.

The foregoing studies have investigated the link between trade and the environment in the context of inter-industry trade models and with aggregate (country-level) data. Much of the recent trade theory, however, has focused on firm-level decisions and emphasized the role of firm heterogeneity with respect to their productivity (Melitz and Redding 2012). This theoretical framework shows that increased openness to trade can have important consequences for firms' productive decisions, can impact the dynamics of firms' entry and exit, and can lead to a reallocation of resources within an industry. With the increasing availability of micro-level data sets, the observed differences between exporters and non-exporters has been investigated with respect to many dimensions, with findings indicating that exporters tend to be larger, more productive, pay higher wages, and generally perform better than non-exporters (Bernard and Jensen 1999; Tybout 2003; De Loecker 2007; Bernard et al. 2007). More recently, some studies have attempted to extend this framework of analysis to import decisions (Gibson and Graciano 2011) and to study the environmental implications of trade. In this paper, we are particularly interested in the latter. Attention to environmental performance in this setting is important because, with heterogeneous firms, increased openness to trade brings about a reallocation effect that increases within-industry efficiency (Melitz and Trefler 2012). As noted by Kreickemeier and Richter (2014), in the context of the relationship between trade and pollution, this "reallocation" effect is distinct from the standard scale, technique, and composition effects noted earlier, and provides yet another way in which trade can affect the environment.

Whether, and by how much, exposure to trade affects environmental performance in the presence of firm heterogeneity is largely an empirical issue. Using micro-level panel data from different countries and various measures of environmental activities, existing empirical studies find that exporters have better environmental performance than non-exporters. For example, Holladay (forthcoming) investigates toxic pollution emissions from U.S. manufacturing establishments over the years 1990–2006, and finds that exporters emit less toxic emissions than non-exporters when controlling for establishment output and industry characteristics. Further, Girma and Hanley (2015) use a measure of a four-point ordinal response to two surveyed questions concerning the environmental impacts of innovation for UK firms. These authors find that exporters are more likely than non-exporters to denote innovation as having "high" or "very high" environmental effects. Using data from a panel of Irish manufacturing firms, Batrakova and Daves (2012) adopt fuel consumption as a proxy for firms' environmental behavior, and show a negative correlation between export status and fuel expenditures for high fuel intensity firms. Similarly, Forslid, Okubo, and Ulltveit-Moe (2014) construct firm-level CO_2 emissions using data on all types of fuel use, together with emission coefficients from Swedish firms. These authors' findings also suggest a negative correlation between an export dummy and CO_2 emission intensity at the firm level, and also a negative relation between emission intensity and firm productivity. Other recent studies include Barrows and Ollivier (2014), who analyze firm-level CO_2 emissions intensity for Indian manufacturing firms, and Cao, Qiu, and Zhou (Forthcoming), who study investment on pollution abatement technology with Chinese firm panel data.

This paper contributes to this emerging body of literature. The new application that we provide, relative to the existing literature discussed in the foregoing, is to study the environmental performance of U.S. manufacturing firms, with a focus on their air pollution emission. Our conceptual framework suggests two testable hypotheses of interest in the presence of firm heterogeneity: (a) facility productivity is inversely related to emission intensity; and (b) export status is negatively correlated with emission intensity. An ancillary hypothesis is that the size of firms is inversely related to pollution intensity. To test these hypotheses we have compiled a uniquely detailed facility-level dataset of the U.S. manufacturing industry for the years 2002, 2005, and 2008. The dataset is assembled from a variety of sources. The National Emission Inventory (NEI) of the U.S. Environmental Protection Agency (EPA) provides facility-level criteria air pollution data for SO_2, carbon monoxide (CO), Ozone (O_3), and Total Suspended Particulates (TSPs). The facility-level economic characteristics data are obtained from the National Establishment Time Series database (NETS). These two databases are matched through the Data Universal Number System (DUNS), which is a unique facility identifier. We have further augmented our dataset with pollutant-specific county nonattainment/attainment designations under the Clean Air Act Amendments (CAAA) legislation. Whereas the impact of CAAA regulations on industrial activities, per se, has been the object of several existing studies, the main reason for us to consider the plant's exposure to such regulations is secondary.[1] Given our focus on the two hypotheses discussed above, it is important to control for possible other determinants of environmental performance, one of which is the firm's exposure to environmental compliance costs (e.g., CAAA regulation), which may differ across individual plants.

The aim of this paper is largely empirical, and the strategy that we employ involves two main steps. First, a facility-level productivity parameter, which we model as TFP, is estimated as a plant-specific fixed effect from a panel data set using nine years of data (2000–2008). Second, given this estimated productivity parameter, we investigate how plant-level productivity correlates with emission intensity on a pollutant-by-pollutant basis. The additional role of exposure to trade for environmental performance is also investigated in this setting. To further validate the conceptual model that we postulate, the impacts of facility attributes on the probability of selection to export are estimated via a logistic regression of export status on measures of trade costs, facility TFP, and other controls (including a facility's exposure to environmental regulations). Our empirical findings are overall supportive of the hypotheses suggested by the conceptual model. For each criteria air pollutant, that is, SO_2, CO, O_3, and TSPs, we find a significant negative correlation between the estimated facility productivity and emission intensity. Conditional on a facility's estimated productivity and other controls, exporting facilities have lower emissions per value of sales than non-exporting facilities in the same industry. These negative correlations between exporting status and emission intensity are statistically significant for all pollutants we track. We also find that plant size is negatively and significantly related to emission intensity for all pollutants. There is also some evidence that polluters located in CO, O_3, or TSPs nonattainment counties have lower emission intensity than those residing in attainment areas.

Methodological Framework

The key insight in much of the theoretical and empirical literature spawned by Melitz (2003) and Bernard et al. (2003) is that firm heterogeneity is an intrinsic feature of the real world, even within narrowly defined industries, and this feature ought to be an explicit ingredient of both theoretical and empirical analyses. In most papers, this heterogeneity is modeled by assuming that firms are endowed with an exogenously drawn productivity parameter. This firm-specific parameter is meant to be a summary statistic representing various forms of heterogeneity in production that reflect each firm's ability to turn inputs into outputs. Such heterogeneity includes differences in technical efficiency, organizational skill, managerial ability, quality of workers, and corporate culture. These heterogeneities translate into equilibrium outcomes that reflect the economic environment in which firms operate. For example, differential openness to trade across industries and/or countries affects the size of the

[1] Greenstone (2002) finds negative impacts of CAAA regulation on the growth of polluting manufactures in nonattainment counties. Others (Becker 2011; Greenstone, List, and Syverson 2011) find that the CAAA nonattainment designation is associated with drops in total factor productivity (TFP) for surviving polluting plants. There is also a debate on whether the CAAA causes firms to relocate (within the country or abroad). Henderson (1996) and Becker and Henderson (2000) show that the O_3 nonattainment regulation leads to the relocation of polluting plants from more to less polluted areas. Hanna (2010) finds that the CAAA causes regulated U.S.-based multinational firms to increase their foreign assets and outputs. As for the impact of the regulation on pollution cleanup, Greenstone (2004) finds that the SO_2 nonattainment designation plays a minor role in the dramatic decline of county-level ambient concentrations of SO_2 from 1969–1997.

market and hence the equilibrium distribution of firms' efficiency. The larger potential market, as well as the presence of foreign trade costs that exceed those for domestic sales, can lead to an equilibrium in which the most efficient firms export, the least efficient firms are driven from the market, and firms of intermediate efficiency serve only the domestic market.

Whether and how such increased efficiency translates into improved environmental performance—the central empirical question addressed in this paper—depends on the specifics of the model that one postulates. Kreickemeier and Richter (2014) discuss an adaptation of the Melitz model that incorporates pollution as a joint output of production. This illustrates a first channel by which firm heterogeneity can impact environmental performance. Insofar as pollution, as a byproduct of production, is proportional to the quantity of inputs used in the production process, more efficient units (which, by definition, obtain more output from the same bundle of inputs) turn out to have a lower emission intensity (quantity of emitted pollution per unit of output).

Another channel by which firm heterogeneity can affect environmental performance involves technology adoption. This perspective is inspired by Bustos (2011), who studies new technology adoption by heterogeneous firms and finds that exporters are more likely to innovate. Following such an approach, Cui, Lapan, and Moschini (2012) develop a model where firms choose between two alternative technologies, one of which (the upgraded technology) is assumed to be an emission-saving technical change relative to the initial technology; upgrading the technology requires extra fixed costs but yields lower marginal costs.[2] This model predicts that a continuum of heterogeneous firms is partitioned by technology upgrade choice and export status. Productive firms can earn enough revenues to cover the fixed costs of entering the export market, and thus select to be exporters. Moreover, only the most productive exporters upgrade to the emission-saving technology because they are the only ones with profitable incentives. This setting suggests the two hypotheses that are the focus of our empirical analysis: (a) facility productivity is inversely related to emission intensity; and (b) export status is negatively correlated with emission intensity. Forslid, Okubo, and Ulltveit-Moe (2014) also provide an explicit adaptation of the Melitz model that focuses on technology adoption, by allowing heterogeneous firms to emit pollution and to invest in abating pollution. But rather than a discrete technology choice, they assume a continuous investment in pollution abatement.

Naturally, the incentives of firms to invest in pollution abatement equipment or to undertake costly upgrades of their technology specifically to reduce pollution depend upon the economic benefits from doing so. If there is an explicit price (or tax) for emissions, then firms that are larger, either because they are more productive and/or because they face a larger potential market because of exports, will have a greater incentive to undertake these fixed costs. Similarly, even if there are no current emission charges, firms located in non-attainment areas may be required to adopt less emitting technologies, particularly whenever existing capital is replaced or whenever the firm expands. It is also possible that firms that export, particularly to areas like the European Union, may find it beneficial to market their product as being a "clean good," and hence they may have a greater incentive to adopt cleaner technologies as a form of product differentiation. Finally, inasmuch as capital decisions depend upon a long planning horizon, the expectation of some future pollution price (e.g., through cap and trade) may influence current decisions regarding investment in pollution-reducing technologies. Thus, while the "pollution as by-product model" clearly predicts an inverse relationship between firm efficiency and firm emissions per unit output (or sales), the forgoing discussion implies that if firms see any economic benefit from reducing pollution, there should also be an inverse relationship between potential market size and firm emissions per unit output. Consequently, the firm's export status, as well as variables that affect trade such as transport costs, have an explanatory role in addition to that of firm productivity in explaining emissions.

Empirical Model

To test the two main hypotheses suggested by the underlying theoretical framework,

[2] Cui (2014) further explores the effects of openness to trade and the stringency of environmental regulations on firms' technology upgrading choices and exports decisions.

we adopt a two-step approach. First, we obtain an estimate of the TFP parameter at the plant level. Given the estimated facility productivity, we then investigate how productivity is related to the plant's emission intensity, as well as the role of the plant's export status in this setting.

Plant-level productivity measures are notoriously difficult to perform (Van Beveren 2012). This is particularly the case when, as in our case, the database lacks information about capital stock or investment at the plant level: the only productive input for which we have detailed information is labor. To overcome this challenge, and to derive a meaningful estimate of facility-level productivity with the data on hand, we propose a practical approach that identifies a plant's TFP within a fixed effect framework. A key ingredient of our procedure is to assume that all firms in the same industry use the same technology (although they are heterogeneous with respect to productivity), and that this technology can be represented by a homogeneous production function. More specifically, the production function that applies to a facility i in industry j at time t is written as

$$(1) \qquad q_{ijt} = \exp(A_j + \phi_{ij} + e_{ijt}) \cdot h_j(L_{ijt}, \boldsymbol{x}_{ijt})$$

where q_{ijt} is output, L_{ijt} is the amount of labor employed at the plant, $\boldsymbol{x}_{ijt}$ is the vector of all other inputs of production, A_j is an industry-level scaling constant, ϕ_{ij} is the plant-specific productivity parameter, and e_{ijt} is a zero-mean identically and independently distributed (i.i.d.) error term. The plant-level productivity parameter ϕ_{ij} is also presumed to have zero mean, an obvious property given that it is meant to measure a plant's productivity relative to the industry average.

Given that we do not have detailed information on the vector $\boldsymbol{x}_{ijt}$, to proceed we postulate that the production function is homogeneous of degree κ_j, so that we write

$$(2) \qquad q_{ijt} = \exp(A_j + \phi_{ij} + e_{ijt}) \cdot L_{ijt}^{\kappa_j} \cdot h_j(1, \boldsymbol{x}_{ijt}/L_{ijt}).^3$$

This reformulation of the production function is useful because it separates the plant-level labor input L_{ijt}, which is observable in our data, from the input ratios $\boldsymbol{x}_{ijt}/L_{ijt}$ (e.g., the capital/labor ratio), which we do not observe. If we now assume that all firms within the same industry face the same input prices, then the maintained assumption that all firms in the same industry have a common and homogeneous production function (apart from their individual productivity parameter) leads to the conclusion that all firms in the same industry would select the same input ratios $\boldsymbol{x}_{ijt}/L_{ijt}$ as a result of cost minimization.[4] This suggests that the unobservable component $h_j(1, \boldsymbol{x}_{ijt}/L_{ijt})$ in equation (2) is common among all plants in the same industry and can therefore be proxied by a suitable set of industry-by-year dummy variables.

Taking logs, the production function can be written as

$$(3) \qquad \ln q_{ijt} = A_j + \phi_{ij} + \lambda_{jt} + \kappa_j \ln L_{ijt} + e_{ijt}$$

where the degree of homogeneity κ_j of the production function measures the industry-specific degree of returns to scale (which, in this context, can be either increasing or decreasing), and $\lambda_{jt} \equiv \ln h_j(1, \boldsymbol{x}_{ijt}/L_{ijt})$ is an industry-specific time-varying term. The plant-specific parameters ϕ_{ij} can be estimated with the aid of plant-specific dummy variables, and the industry-specific but time-varying term λ_{jt} can be estimated with the aid of industry by year dummy variables. It is now apparent that the TFP parameters ϕ_{ij} can be estimated using a fixed effect model (Pavcnik 2002).

Because we wish to relate productivity to emissions, and we have data on emissions for three distinct years (2002, 2005, and 2008), a question is whether or not a plant's productivity coefficient should be allowed to vary across these three observation points. An essential assumption in estimating equation (3), implicit in the foregoing, is that a plant's own relative productivity is time-invariant over the period used to estimate it. To allow a plant's TFP parameter to differ for the three emission years 2002, 2005, and 2008, therefore, requires one to estimate it based on three years of data only (e.g., 2006–2008 for the 2008 productivity parameter). Hence, for the main results discussed in the paper,

[3] Recall that, for a function $f(z)$ that is homogeneous of degree κ, then $f(tz) = t^{\kappa} f(z)$, $\forall t > 0$.

[4] The homogeneity of the production function, in particular, ensures that cost-minimizing input ratios depend only on price ratios and are not affected by the scale of output.

we elect to estimate the TFP parameters presuming constancy over the period 2000–2008. Alternative TFP parameters that vary across pollution years, however, are also considered and discussed in the robustness section.

Next, we wish to investigate the relationship between the plant's estimated TFP measure and the plant's emission intensity, and also to evaluate the role of the plant's export status on emission intensity. If E^{ℓ}_{ijs} denotes a plant's total emission of pollutant ℓ in year s, emission intensity is defined as emission per unit of output, which in log terms is measured by $y^{\ell}_{ijs} \equiv \ln(E^{\ell}_{ijs}/q_{ijs})$. The emission regression of interest, omitting the pollutant index for clarity of exposition, is therefore written as

$$(4) \qquad y_{ijs} = \theta_j + \lambda_s + \gamma_1 TFP_{ij} + \gamma_2 X_{ij} + \sum_k \gamma_{3k} Z^k_{ijs} + \varepsilon_{ijs}$$

where $TFP_{ij} \equiv \hat{\phi}_{ij}|$ is the plant's estimated productivity, $X_{ij} \in \{0, 1\}$ is the "export status" indicator ($X_{ij} = 1$ if the facility is exporting), and Z^k_{ijs} denotes a set of control variables that may be plant-specific (including exposure to regulation). In equation (4), the industry-specific intercept θ_j allows for pollution intensity to differ systematically across industries, λ_s is a year-specific coefficient controlling for possible time trends affecting all industries, and ε_{ijs} is the stochastic error term (assumed i.i.d.).

Data

For the purpose of estimating equations (3) and (4), we have assembled a unique detailed facility-level emission dataset on criteria air pollutants and facility characteristics that pertains to the U.S. manufacturing industry. A "facility" is a place where economic activities that result in air emissions occur. Facility emission data are obtained from the NEI database of the U.S. EPA, and pertain to pollution data for the years 2002, 2005, and 2008.[5] The facility economic characteristics are taken from the NETS database, and pertain to the period 1990–2008. These two databases are matched through the DUNS number assigned by Dun and Bradstreet to identify unique business establishments.

The NETS database, developed through a joint venture with Dun and Bradstreet by Walls and Associates, is a unique database that provides information on a large number of U.S. business establishments for every year since 1990.[6] The data acquired for this study include establishment name, number of employees, value of sales, an export indicator, the DUNS number, geographic location (i.e., latitude and longitude), Zip Code, and five-digit Federal Information Processing Standard (FIPS) county code.[7]

The EPA's NEI database contains information about facilities that emit criteria air pollutants for all areas of the United States. Since 2002, the EPA has released an updated version of the NEI database every three years. The facility-level NEI database acquired for this empirical study includes emission data for four criteria air pollutants, that is, SO_2, CO, O_3, and TSPs during 2002, 2005, and 2008.[8] The Facility Registry System (FRS) of the EPA provides DUNS numbers of these polluting facilities, allowing us to match the NEI with the NETS databases. We first match polluting facilities within the NEI database across years, and then retrieve DUNS numbers for these polluters from the FRS of the EPA. Second, we match these polluters with those appearing in the NETS database through the DUNS number. This matching procedure narrows down our dataset to 16,695 polluting facilities (i.e., with a nonzero emission value for at least one pollutant) in year 2002, 12,022 polluting facilities in year 2005, and 10,144 polluters in 2008, all in the U.S. manufacturing industry as determined by having a four-digit Standard Industrial Classification (SIC) code between 2000 and 4000. This amounts to roughly half

[5] This includes all the years for which such data were available at the time this study was undertaken. Recently, the EPA has released access to NEI data for 2011 as well.

[6] NETS data have been used to study issues related to job creations and destructions, business relocation, and business ownership (Kolko and Neumark 2008, 2010; Neumark, Wall, and Zhang 2011). Neumark, Wall, and Zhang (2011) provide a detailed description of the NETS and an assessment of the quality of the NETS database along many dimensions. Holladay (forthcoming) also relies on the NETS database.

[7] Whereas the plant-level number of employees and value of sales change over time in the NETS database, the export status indicator is time-invariant.

[8] A more detailed discussion of the facility-level NEI database is provided in the appendix, which also discusses some caveats (particularly as they relate to the 2005 data).

of the polluters in the manufacturing industry reported in the NEI database prior to matching.[9] More details on the data matching procedure are provided in the appendix.

The regulatory attainment/nonattainment county status information is obtained from the Green Book Nonattainment Areas for Criteria Pollutants reported by the EPA.[10] A list of variables and data sources used in the paper is summarized in table A1 in the appendix. For each criteria air pollutant that we track, the Green Book indicates whether only part of a county or the whole county is in nonattainment. We assign a county to the nonattainment category for each of four criteria pollutants, that is, SO_2, CO, O_3, and TSPs, if the entire county or part of the county is designated with nonattainment status.[11,12]

Descriptive Statistics

Our merged dataset—used to investigate the relationship among productivity, export status, and emission intensity—consists of an unbalanced panel of polluting facilities in 2002, 2005, and 2008, for a total of 38,192 facility-by-year observations from 18,435 facilities located in 2,017 U.S. counties. There are 7,525 facilities represented throughout the study period. Table 1 provides summary statistics on a number of variables in the merged dataset. The value of sales is deflated by the four-digit SIC industry shipment deflator provided by the NBER-CES manufacturing industry database.[13] It is worth noting that, because of the matching procedure adopted, each facility emits at least one pollutant, but not all facilities have emissions reports for all four criteria air pollutants. In many cases, facilities only have estimates for one pollutant in the NEI database. In addition, the dataset contains a few observations with extremely low emissions, which do not appear credible.[14] These outliers, which only account for a small fraction of total relevant observations, were dropped from the analysis and are not included in table 1.[15]

The last two columns of table 1 summarize the differences between exporters and non-exporters across facility characteristics. Exporters are larger than non-exporters in terms of sales and number of employees. These descriptive results are in line with the growing empirical trade literature on heterogeneous firms. When it comes to environmental performance, exporters emit more SO_2, O_3, and TSPs, but less CO than non-exporters. Pollution intensity measured by emissions per value of sales (tons per thousand dollars), however, is lower for exporters relative to non-exporters for all criteria air pollutants. The differences are persistent for each sample year separately.

According to the EPA's Green Book, in 2002 only a small number of the total of 3,143 U.S. counties were designated as nonattainment: 21 counties in SO_2 nonattainment, 19 counties in CO nonattainment, 251 counties in O_3 nonattainment, and 64 counties in TSPs nonattainment. In 2005, the number of counties with SO_2 or CO nonattainment designations declines to 12 and 11, respectively, while the number of counties with O_3 or TSPs nonattainment status increases drastically to 431 and 259, respectively. In 2008, the number of counties with O_3 nonattainment status substantially dropped from 431 to 293, while the number of counties with other nonattainment status changes slightly. Most nonattainment counties are covered in our merged dataset.

9 Prior to data matching, the NEI database contains 25,574 manufacturing polluters in 2002, 20,948 in 2005, and 21,102 in 2008.

10 For detailed information, see http://www.epa.gov/air/oaqps/greenbk/index.html.

11 The formation of ground-level ozone is a complicated chemical process that involves volatile organic compounds (VOCs) and oxide of nitrogen (NO_x) when these two react in the presence of sunlight. There are separate standards for NO_2, 1-hour O_3, and 8-hour O_3. We classify a county as nonattainment for O_3 if it is in nonattainment for NO_2 or O_3, including both 1-hour and 8-hour standards. Therefore, the pollution of VOCs and NO_x is associated with this combined O_3 nonattainment designation.

12 There exist separate standards for PM10 and PM2.5. We classify a county as nonattainment for TSPs if it is in nonattainment for at least one of these standards. TSPs in this study are primary particulates matters (the sum of primary PM10 and primary PM2.5).

13 This database pertains to the U.S. manufacturing sector for the period 1958–2009, and it is assembled from data obtained from various federal agencies with the goal of providing consistent time series for a large number of industries, including price deflators, capital stocks, and productivity estimates (Becker, Gray, and Marvakov 2013).

14 For example, the smallest facility-level nonzero value of SO_2 in the data was 2.1×10^{-10} tons per year, that is, 0.21 micrograms per year (a microgram is equal to one billionth of one kilogram).

15 Specifically, we adopted the threshold of 0.001 tons per year (i.e., one kilogram) for inclusion in the analysis. The fraction of observations with annual emissions less than 0.001 tons per year are as follows: 7.73% for SO_2, 1.22% for CO, 0.49% for O_3, and 1.81% for TSPs. Empirical estimation with these outliers is considered but not reported in the paper. Accounting for the outliers does not change the empirical results in any significant way.

Table 1. Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min.	Max.	Exporter Mean	Non-exporter Mean
Sales (thousand $)	38,192	31,946.1	79,320.7	0.1	3,404,478	45,021	27,685
Employees	38,192	215.4	479.0	1	15,000	308.4	185.1
SO_2 (tons)	20,221	130.4	880.6	0.001	41,845.2	150.5	123.2
CO (tons)	23,829	141.1	1,593.9	0.001	87,428.9	123.2	147.5
O_3 (tons)	34,980	101.9	464.7	0.001	23,121.4	101.7	102.4
TSPs (tons)	29,272	41.2	198.8	0.001	11,383.1	42.8	40.6
SO_2 per Sales	20,221	0.051	1.413	2.25E−09	139.5	0.011	0.065
CO per Sales	23,829	0.063	2.930	2.46E−09	309.4	0.013	0.080
O_3 per Sales	34,980	0.041	0.946	1.09E−08	89.4	0.027	0.046
TSPs per Sales	29,272	0.016	0.416	2.39E−09	53.3	0.008	0.019
Export Status	38,192	0.246	0.431	0	1	1	0
SO_2 NA	38,192	0.009	0.093	0	1	0.008	0.009
CO NA	38,192	0.079	0.270	0	1	0.064	0.084
O_3 NA	38,192	0.468	0.499	0	1	0.469	0.467
TSPs NA	38,192	0.203	0.402	0	1	0.190	0.207

Note: The acronym NA stands for nonattainment, and is one-year lagged status. Sales are deflated by value of shipment deflator on the basis of four-digit SIC industry level; O_3 is sum of NOx and VOCs, TSPs is the sum of PM_{10}-PRI and $PM_{2.5}$-PRI. The numbers of pollutant-specific polluters with annual emissions less than 1 kg are 1,669 for SO_2, 297 for CO, 175 for O_3, and 533 for TSPs.

Empirical Results

By using the data discussed in the foregoing section, we test the two main hypotheses that we have postulated: first, that productivity is inversely related to emission intensity; second, that there exists a negative correlation between export status and emission intensity.

Productivity Estimates

We begin by estimating the facility-level productivity following the fixed-effect methodology discussed earlier, based on the time series of economic characteristics for each facility included in the merged dataset. We consider two alternative productivity estimates. As noted earlier, the implicit assumption is that a plant's own productivity is time-invariant over the period used to estimate it. In the baseline, the results of which are presented here, equation (3) is estimated with data from 2000–2008. Hence, the maintained assumption is that plant-level productivity parameters are the same for the three years for which we have data on emissions (2002, 2005, and 2008).

The estimated plant-level TFP parameters are too numerous to report individually, but figure 1 illustrates the distribution of TFP estimates. By construction, our TFP measure is relative to the industry average, and hence this distribution has zero mean. The standard deviation of 0.599 suggests considerable heterogeneity across manufacturing plants. This is consistent with existing studies.[16] Perhaps not surprisingly, given that these are log productivity parameters, this distribution appears rather symmetric. It should be recalled, however, that the maintained assumption in equation (1) is that the plants' multiplicative productivity terms are $\exp(\phi_{ij})$.[17]

We note that, as a by-product of our TFP estimation procedure, we obtain the coefficient of returns to scale $\hat{\kappa}_j$ for all 136 three-digit SIC industries in our sample. The distribution of such estimates is illustrated in figure 2. On average, the manufacturing industries in our sample display essentially constant returns to scale (the simple average over all industries is 0.972), with some variation (the standard deviation is 0.097).

For robustness checks, productivity parameters were also estimated with an alternative procedure that allows plant-level TFP to change over time. Specifically, for each of the years that we have plant-level emission data (2002, 2005, and 2008), equation (3) is estimated with a three-year window of data up to and including the year of interest. To estimate TFP parameters for the year 2002, therefore, we use data for the period

[16] For example, for labor productivity, Bernard et al. (2003) find the standard deviation of within-industry log productivity to be 0.6.

[17] If the estimated log productivities in figure 1 were normally distributed, for example, then the implied multiplicative productivity terms would be log-normally distributed.

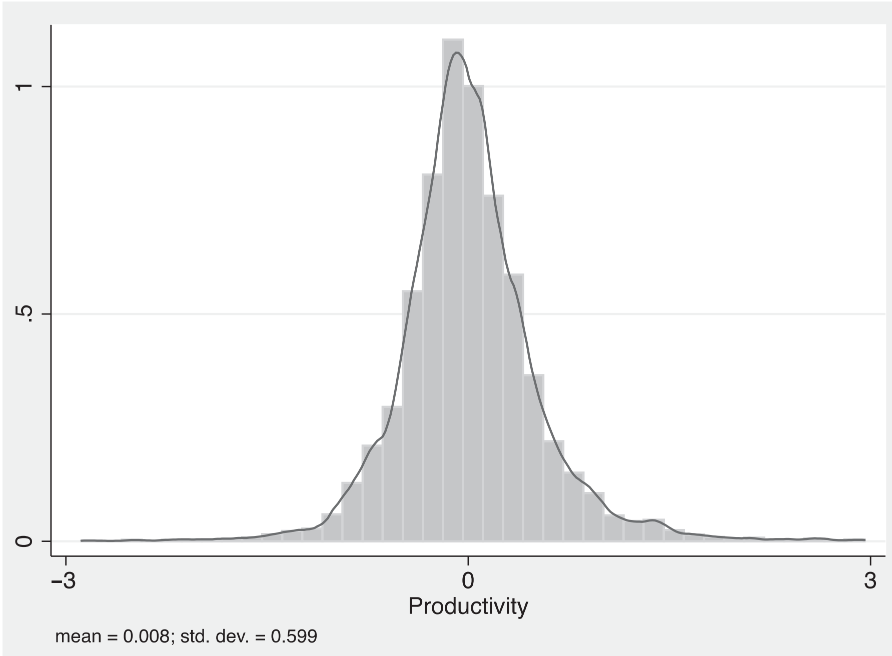

Figure 1. Density distribution for the estimated TFP parameters

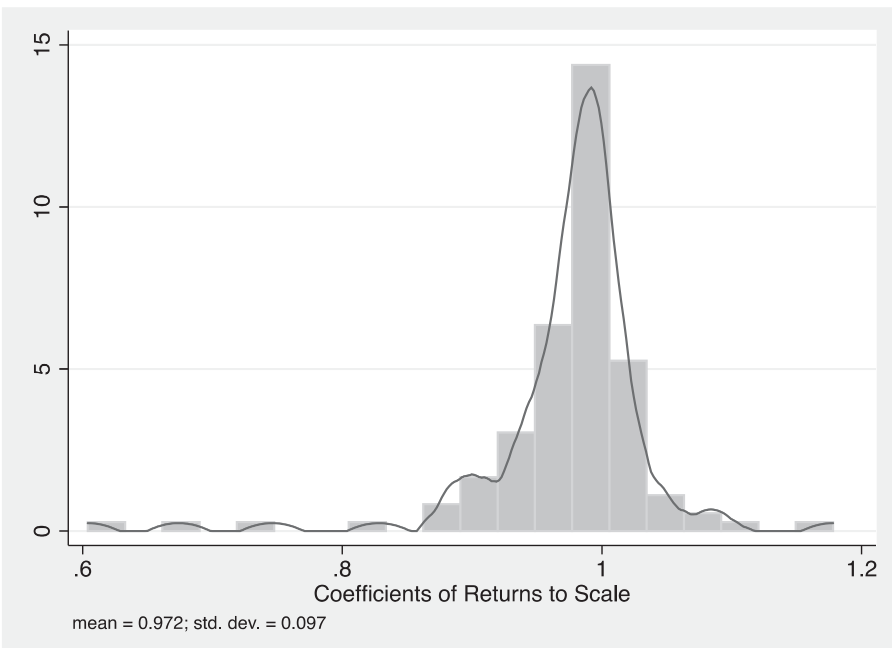

Figure 2. Density distribution for the coefficients of returns to scale

2000–2002. Similarly, TFP parameters for 2005 are estimated with 2003–2005 data, and TFP parameters for 2008 are estimated with 2006–2008 data. In any event, the distribution of TFP estimates thus obtained is similar to those of the baseline, although somewhat more dispersed: the standard deviation is 0.564 for 2002, 0.596 for 2005, and 0.659 for 2008. As illustrated in the supplementary online appendix, there is substantial correlation between the baseline and the alternative TFP estimates. Specifically, the correlation coefficients are 0.882 for 2002, 0.908 for 2005, and 0.856 for 2008. It turns out that the results discussed below are extremely robust to the choice of TFP estimates.[18]

Emission Intensity

To assess the relationship among productivity, export status, and emission intensity—the main subject of interest of this paper—we estimate the model in equation (4). The left-hand side variable here is the log of the emission per unit of output (deflated value of sales). The explanatory variables, in addition to the indicator of export status (a dummy variable) and the plant-level TFP, include other relevant control variables. The rationale for expecting emission intensity and export status to be related to plant-level productivity was articulated earlier. As for other control variables, we include dummy variables that flag whether the plant is located in a county classified by the EPA as being in nonattainment status for the year of interest.[19] This status is pollutant-specific, of course, and equals one if the facility is located in a nonattainment county for that pollutant (zero otherwise). The inclusion of these variables is for control purposes. Presumably the compliance requirements for plants located in nonattainment counties affect both their emission and their productivity, and differentially so relative to counties that do not face such regulatory pressure. Given the focus of this paper, it is important that the connection we establish between productivity, export status, and emission be conditioned on the regulatory pressure that might differ across plants.

The model of equation (4) is estimated pollutant-by-pollutant, with the emission intensity computed separately for each criteria air pollutant. In addition to the variables discussed in the foregoing, we also control for the plants' location by including state fixed effects, and for possible systematic differences across industries by including industry fixed effects. We also include year fixed effects (recall that here we are pooling the observations for the three years 2002, 2005, and 2008). The results are reported in table 2. The sample size of polluting facilities varies with pollutant type. Standard errors reported in parentheses are clustered at the three-digit SIC industry level.[20] For each of the four pollutants considered, we report the results of two regressions, one of which includes the plant's size (measured as the log of the number of employees) on the right-hand side of the table (in addition to the variables discussed in the foregoing).

From table 2 it is apparent that we find a strong negative relationship between estimated productivity and emission intensity. Consider first the regressions without the size variable. In all cases, the estimated coefficient associated with productivity is different from zero at the 1% significance level. The magnitude of this estimated coefficient varies across pollutants, ranging from -1.015 (for SO_2) to -0.757 (for CO), values that are not too far from one in absolute value. This finding is suggestive of the hypothesis, noted earlier, by which productivity may affect emission intensity because pollution is a by-product of input use. Recall that the postulated production function for our model was given in equation (1), where $h_j(L_{ijt}, \boldsymbol{x}_{ijt})$ is an industry-specific production function, and L_{ijt} is the amount of labor employed at plant i in year t, and $\boldsymbol{x}_{ijt}$ is the corresponding vector of all other inputs of production. If the amount of plant total emission E^{ℓ}_{ijt} for pollutant ℓ were proportional to input use, for example, $E^{\ell}_{ijt} = \alpha \cdot h_j(L_{ijt}, \boldsymbol{x}_{ijt})$ for some $\alpha > 0$, then the structure in equation (1) would imply that the coefficient linking pollution intensity $\ln(E^{\ell}_{ijt}/q_{ijt})$ and productivity ϕ_{ij} would be exactly -1.

Concerning the coefficient on the export status variable, the estimates in table 2 consistently show negative correlations between export status and emission intensity for all four criteria air pollutants tracked in the paper, with magnitudes that are fairly close across pollutants (ranging from -0.295 for SO_2 to -0.262 for CO). Also, these negative coefficients differ from zero at the

[18] These robustness results are reported in the supplementary appendix available online.

[19] Hence, all facilities in the same county are assumed to face the same regulatory pressure. County nonattainment/attainment status is officially reclassified every July. In our regression, the attainment/nonattainment status variable of a county for calendar year t is based on the July determination of year $t-1$.

[20] Alternative specifications of standard errors (i.e., cluster at facility level, cluster at industry level, and robust standard errors) were considered but, because of space reasons, are not reported here and are available upon request. These specifications do not alter inference in any significant way.

Table 2. Main Results: Emission Intensity Equation

	SO_2		CO		O_3		TSPs	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Productivity	−1.015***	−0.848***	−0.757***	−0.621***	−0.886***	−0.721***	−0.776***	−0.603***
	(0.118)	(0.117)	(0.093)	(0.085)	(0.124)	(0.102)	(0.072)	(0.064)
Export status	−0.295***	−0.026	−0.262***	−0.027	−0.285***	−0.041	−0.287***	−0.032
	(0.085)	(0.076)	(0.065)	(0.062)	(0.038)	(0.035)	(0.051)	(0.048)
SO_2 NA	0.295	0.222						
	(0.288)	(0.258)						
CO NA			0.154	0.164				
			(0.104)	(0.109)				
O_3 NA					−0.275***	−0.288***		
					(0.054)	(0.060)		
TSPs NA							−0.159**	−0.090
							(0.072)	(0.068)
Size		−0.694***		−0.553***		−0.563***		−0.601***
		(0.034)		(0.037)		(0.025)		(0.025)
Observations	19,826	19,826	23,347	23,347	34,342	34,342	28,672	28,672
Adjusted R^2	0.445	0.499	0.407	0.471	0.319	0.422	0.401	0.469
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y

Note: Dependent variable is log of emissions per value of sales. The industry fixed effect is a set of three-digit SIC industry dummies. Standard errors reported in parentheses are clustered at three-digit SIC industry. Coefficients for the regression constant and fixed effects are suppressed. Asterisks ***indicate significance at the 1% level, **indicate significance at the 5% level.

1% significance level for all pollutants. The empirical findings are in line with the theoretical prediction that export status is negatively correlated with emission intensity. Exporters emit less pollution per sale than non-exporters, with this reduction effect ranging from 26–29%, depending on the pollutant.

The estimated effect of nonattainment designations on pollution emission intensity is positive (but not significant) for both SO_2 and CO, and negative and significant for O_3 (at the 1% level) and for TSPs (at the 5% level). The variety of impacts that we uncover is, perhaps, not surprising. The need to meet stricter environmental regulations in non-attainment counties should lead to lower total emission, *ceteris paribus*. But the costs of meeting such regulations are likely to affect production efficiency (productivity) as well, which might result in increased emission intensity (because of lower output per amount of resources used). It is unclear which of these effects would dominate. At any rate, the inclusion of these regulation variables in the model is for control purposes, as noted earlier, and our focus of interest is elsewhere.

Table 2 also includes, for each pollutant, the estimation results of a version of equation (4) that includes "size" (measured as the log of the number of employees). In this model, size could affect emissions in much the same way as exports: larger size means the fixed costs of adopting cleaner emissions technology is less burdensome and larger size allows for an economy of scale effect with respect to emissions. To see how the latter might work, consider the pollution-as-byproduct illustration articulated in the foregoing. But now allow for returns to scale for emission to be different than for output, for example, $E_{ijt}^{\ell} = \alpha \cdot [h_j(L_{ijt}, \boldsymbol{x}_{ijt})]^{\sigma}$, where $0 < \sigma < 1$ indicates increasing returns to scale with respect to clean air (decreasing returns to scale with respect to emission). Given the assumption of homogeneous industry-level production functions, invoked earlier, one would then obtain $E_{ijt}^{\ell} = \alpha \cdot L_{ijt}^{\kappa_j \sigma}[h_j(1, \boldsymbol{x}_{ijt})]^{\sigma}$. In view of the production function in equation (1), this would suggest that $\ln(E_{ijt}^{\ell}/q_{ijt})$ is related to $\ln(L_{ijt})$ by the coefficient $\kappa_j(\sigma - 1)$. Hence, the hypothesis of an independent economy of scale effect with respect to clean air (i.e., $\sigma < 1$) would imply a negative relationship between the log of emission intensity and the log of employment. This is indeed the result, as reported in table 2. We find that the size variable has a negative effect on emission intensity, with a magnitude ranging from −0.694 (for SO_2) to −0.553 (for CO). In all cases the coefficient of the size variable is significant at the 1% level. Including the size variable does not substantially change the coefficient of the TFP variable, which remains negative, large, and significant at the 1% level. The inclusion of size, however, does change the magnitude of the coefficients associated with export status, which, while retaining a negative sign, are no longer significantly different from zero.

We should note at this juncture that the interpretation of the effects of the size and export status variables on emission intensity is not straightforward. Our empirical model, while inspired by theoretical models of export decisions and technology adoption by heterogeneous firms, is not a structural representation of such models. The results that we find, therefore, need to be interpreted with some care. Unlike plant-level productivity, which following the theoretical literature discussed earlier we take as pre-determined, size and export status are chosen by firms in equilibrium. And, given productivity, size is influenced by export status because the latter permits access to a larger market—indeed, this is the hallmark of Melitz-type models. What the results in table 2 establish is that productivity clearly matters for emission intensity. Largeness matters as well, because firms with the same productivity within the same industry—as defined by production technology—may face differing markets for their (heterogeneous) products. Thus, given technology and export status, firm size might differ because of the market conditions faced by the firm. Exporting firms might emit less, per unit output, because they tend to be larger and there might be economies of scale if emissions are proportional to input use. Alternatively, of course, larger firms might have more of an incentive to invest in cleaner technology, which is another channel by which the results in table 2 could be rationalized. Finally, it is important to note that the conceptual models of trade with heterogeneous firms that we discussed do not imply a direct channel through which exports affect emission intensity. Rather, it is because the ability to export leads firms to modify other decisions—such as size—that access to export markets might affect emission intensity.

Corroborating Evidence

To further validate the theoretical underpinnings of our model, it is of some interest to examine how the facilities' export status is related to the estimated TFP productivity measure. According to the Melitz-type models discussed earlier, export status is endogenous, depending upon productivity, trade variable costs, and other cost parameters. Furthermore, *ceteris paribus,* in our model export status should be positively correlated with productivity. To investigate this property, we seek proxies of trade cost variables. Two proxies employed in this study are facility-specific and industry-specific trade variable costs. The former is measured by the geographical distance of each polluting facility to its nearest U.S. port, and the latter is measured by the ad valorem freight rate at the four-digit SIC industry level.[21] The geographic distance reflects the costs associated with the transportation of goods from manufacturing sites to the port of shipment. The freight rate, constructed by Bernard, Jensen, and Schott (2006), is the markup of the Cost-Insurance-Freight (CIF) value over the Free-on-Board (FOB) value relative to the FOB. This industry-specific freight rate serves as a proxy of the iceberg trade costs associated with ocean or inland waterway transport of the goods to the port of destination. These two measures together are considered as proxies of trade variable costs.

We employ a logistic model to estimate the probability of selecting to export conditional on the estimated facility productivity, the two measures of trade variable costs, and exposure to environmental regulations, controlling for industry characteristics. Regardless of pollutant type, the logistic regression is specified as

$$(5)\qquad \Pr(X_{ijt}=1) = F\Big(\gamma_0 + \gamma_1\,\text{Distance}_i + \gamma_2\,\text{Freight}_{jt} + \gamma_3\,\text{TFP}_{ijt} + \sum_k \gamma_{4k} Z^k_{ijt} + \theta_j + \lambda_t + \varepsilon_{ijt}\Big)$$

[21] According to IHS Global Services, U.S. seaborne trade with the rest of the world accounts for 78.05% by volume (millions of metric tons), and 48.47% by value of total U.S. trade (billions of dollars) in 2008.

Table 3. Export Decision

	(1)	(2)
Distance	−1.133***	−1.130***
	(0.405)	(0.406)
Freight	0.767	0.759
	(1.794)	(1.794)
Productivity	0.171***	0.172***
	(0.049)	(0.049)
Any NA	0.098	0.097
	(0.062)	(0.063)
Single-plant Dummy		0.040
		(0.065)
Observations	30,312	30,312
Pseudo R^2	0.0955	0.0955
Industry FE	Y	Y
Year FE	Y	Y
State FE	Y	Y

Note: Dependent variable is binary export decision (=1 if plant is exporting). "Any NA" is a dummy variable that equals one if any of the pollutant-specific NA dummy variables does so. Coefficients for the regression constant and fixed effects are suppressed. Asterisks ***indicate significance at the 1% level.

where $F(\cdot)$ denotes the logistic function, and Distance_i denotes the distance (in thousands of miles) of a polluting facility to its nearest U.S. port. The World Port Source online database provides geographic locations (i.e., latitude and longitude) for 548 U.S. ports including harbor, river port, seaport, offshore terminal, and pier, jetty, or wharf.[22] For each polluting facility, we compute its distance to all 548 U.S. ports based on the "Haversine" formula, given the latitude and longitude of two points, then pick the shortest distance as the distance to the nearest port.[23] Freight_{jt} indexes the freight rate at the four-digit SIC industry level. The industry-level data on CIF and FOB are acquired from the online data source of U.S. Manufacturing Exports and Imports compiled by Peter Schott (2010).[24] All other variables were defined earlier.

Table 3 presents the estimation results for equation (5). A version of the regression equation with a single-plant firm status is also reported in this table. This was suggested by a reviewer to control whether single-plant firms might be less likely to export, *ceteris paribus*. Because this regression is not pollutant-specific, environmental regulation pressure is captured by a single nonattainment variable (which is equal to one if the corresponding county is in nonattainment for any pollutant, and zero otherwise). The results in this table are supportive of our approach and consistent with the prediction of the underlying theoretical model.[25] First, a positive and statistically significant coefficient of the TFP regressor indicates that the higher is a facility's productivity, the more likely it is to export. Second, the estimated coefficients of distance to port are negative and significant at the 1% level. As facilities residing closer to ports are likely to have lower costs associated with transporting the goods from manufacturing sites to the ports of shipment, they are more likely to engage in the export market. Note that this effect is over and above the distance effect that is possibly already captured by the state fixed effect included in all regressions. Having controlled for the distance to port, the measure of freight rates that we have included does not seem to have additional explanatory power (the estimated coefficients are not statistically different from zero).

As for the other variables included in table 3, single-plant firms appear equally likely to export as multi-plant firms, other things being equal. For the variable that codes counties' non-attainment status vis-à-vis the four pollutant categories considered, our model does not have a clear prediction as to their expected sign. In any event, the estimated coefficients for this variable are not statistically different from zero in each of the two versions of the estimated equation reported in table 3.

Robustness Checks

The results of table 2, consistent with the theoretical rationalizations proffered earlier, indicate an average negative correlation between emission intensity and productivity. It is of additional interest to verify whether this relation is monotonic. To test whether indeed the most productive facilities have better environmental performance, we use dummy variables of productivity quintiles to replace the continuous productivity measure

[22] For detailed information, please see http://www.worldportsource.com/states.php.

[23] The "Haversine" formula calculates the great-circle distance between two points, that is, the shortest distance over the earth's surface.

[24] The 2008 industry-level CIF and FOB data, which are not provided in Schott (2010), are simply taken from 2005.

[25] The number of observations drops as compared with the number in table 3 because the Schott (2010) data source that we are using does not contain data for all four-digit SIC industries.

Table 4. Emission and the Quintiles of Productivity Distribution

VARIABLES	SO_2 (1)	CO (2)	O_3 (3)	TSPs (4)
2nd Productivity Quintile	−0.261*	−0.260**	−0.281***	−0.209**
	(0.138)	(0.102)	(0.055)	(0.083)
3rd Productivity Quintile	−0.252	−0.168	−0.330***	−0.148
	(0.161)	(0.108)	(0.073)	(0.098)
4th Productivity Quintile	−0.437***	−0.436***	−0.598***	−0.365***
	(0.163)	(0.118)	(0.065)	(0.087)
5th Productivity Quintile	−1.234***	−0.864***	−1.061***	−0.824***
	(0.173)	(0.110)	(0.077)	(0.099)
Export Status	−0.031	−0.028	−0.045	−0.038
	(0.077)	(0.062)	(0.036)	(0.048)
Size	−0.698***	−0.556***	−0.563***	−0.607***
	(0.036)	(0.037)	(0.025)	(0.024)
F-Statistics for equality of prod.	46.68	33.64	63.87	50.47
Observations	19,826	23,347	34,342	28,672
Adjusted R^2	0.497	0.468	0.415	0.466

Note: Dependent variable is log of emissions per value of sales. See text for definition of quintiles variables. All regressions include three-digit SIC industry fixed effects, year fixed effects, and state fixed effects, as well as pollutant-specific NA variables (just as in table 2). Standard errors reported in parentheses are clustered at three-digit SIC industry. The F-statistics refer to the equality of all productivity quintiles dummies. Asterisks ***indicate significance at the 1% level, **indicate significance at the 5% level, and *indicates significance at the 10% level.

in equation (4). The results are reported in table 4 (due to space constraints, only the versions of equation (4) that include the "size" variable are reported). The results are quite consistent with those of table 2. Relative to the firms with the lowest productivity (the 1*st* quintile), firms in quintiles 2 to 5 have a lower emission intensity, and this gap is different from zero at the 1% significance level for the fourth and fifth quintiles (for all pollutants). Moreover, these productivity coefficients are monotonically decreasing across quintile starting with the third quintile (the productivity coefficients associated with the second and third quintiles are very similar): generally speaking, the higher productivity quintile a facility belongs to, the lower emission intensity it has. Within the same three-digit SIC industry, facilities in the fifth productivity quintile have the best environmental record in terms of the lowest emission intensity. *Ceteris paribus*, the results of table 4 imply that plants in the fifth productivity quintile on average emit only a fraction of the SO_2 emitted by first productivity quintile plants, specifically 29% (this fraction corresponds to 42% for CO, 35% for O_3, and 44% for TSPs). The null hypothesis that the productivity effects are equal across all quintiles is rejected by the appropriate F test at the 1% significance level. From table 4 it is also apparent that representing the productivity impacts in terms of quintiles does not affect the inference about the other variables. In particular, the estimated coefficients, and the standard errors, associated with export status and the size variables are essentially the same as in table 2.

As noted by a reviewer, ideally the empirical characterization of the impact of productivity on emission would need to account for births of new facilities and exits of some plants. Unfortunately, the data that we have is not informative about this dynamic process. As a robustness check, therefore, to guard against the possibility that our results might be affected by the selection bias of ignoring entry and exits of plants, we first re-estimate productivity equation (3) using a balanced panel of data pertaining to all plants that are present for the entire period 2000–2008, and then re-estimate emission intensity equation (4) using the balanced panel sample (by pollutant) in 2002, 2005, and 2008. The results are reported in table 5 (again, to save space, only the versions of equation (4) that include the "size" variable are reported). Table 5 offers considerable support for the corresponding full-sample results of table 2. The estimated coefficients for the productivity variable, in particular, are almost identical.

As suggested by one reviewer, we estimated equation (4) with a subset of observations pertaining to plants in "dirty industry"

Table 5. Emission Intensity Regression (Balanced Sample Results)

VARIABLES	SO_2 (1)	CO (2)	O_3 (3)	TSPs (4)
Productivity	−0.848***	−0.674***	−0.733***	−0.615***
	(0.124)	(0.080)	(0.096)	(0.063)
Export Status	−0.033	−0.024	−0.011	−0.050
	(0.115)	(0.076)	(0.045)	(0.064)
Size	−0.713***	−0.514***	−0.527***	−0.582***
	(0.047)	(0.048)	(0.028)	(0.030)
Observations	10,362	12,843	19,683	15,738
Adjusted R^2	0.524	0.475	0.439	0.482

Note: Dependent variable is log of emissions per value of sales. All variables defined as in table 2. All regressions include three-digit SIC industry fixed effects, year fixed effects, and state fixed effects, as well as pollutant-specific NA variables (just as in table 2). Standard errors reported in parentheses are clustered at three-digit SIC industry. Asterisks ***indicate significance at the 1% level.

Table 6. Pollutant-specific Dirty Industry List

Industry (SIC codes)	SO_2	CO	O_3	TSPs
Pulp and paper (2611–31)	Y	Y	Y	Y
Organic chemicals (2861–69)			Y	
Petroleum refining (2911)	Y	Y	Y	
Rubber and miscellaneous plastic products (30)			Y	
Stone, clay, glass, and concrete (32)	Y		Y	Y
Iron and steel (3312–25, 3321-2)		Y		
Nonferrous metals (333–34)	Y	Y		

Note: Industry classification used for table 7, constructed based on Greenstone (2002).

Table 7. Emission Intensity Regression (“Dirty Industries” only)

VARIABLES	SO_2 (1)	CO (2)	O_3 (3)	TSPs (4)
Productivity	−1.421***	−0.632**	−0.858***	−0.529**
	(0.289)	(0.194)	(0.075)	(0.186)
Export status	−0.152	−0.245	−0.085	0.100
	(0.276)	(0.451)	(0.050)	(0.194)
Size	−0.496***	−0.514***	−0.515***	−0.480***
	(0.071)	(0.086)	(0.080)	(0.065)
Observations	2,365	1,690	6,270	3,563
Adjusted R^2	0.386	0.349	0.385	0.246

Note: Dependent variable, Export Status, Productivity, and Size are as in table 2. All regressions include three-digit SIC industry fixed effects, year fixed effects, and state fixed effects, as well as pollutant-specific NA variables (just as in table 2). Standard errors reported in parenthesis are clustered at three-digit SIC industry. Asterisks ***indicate significance at the 1% level, **indicate significance at the 5% level.

only, that is, industries that are known to be heavy emitters of criteria air emissions. The classification of “dirty industry” used is pollutant-specific, based on Greenstone (2002), and is described in table 6. When restricting the sample to only dirty industries, the fraction of exporters by pollution type drops from 24.6% reported in table 1 of the summary statistics to 14.3% for SO_2 dirty industry (28.3% for CO dirty industry only, 21.5% for O_3 dirty industry only, and 22.7% for TSPs dirty industry only). The results of estimating the emission intensity equation (4) using data from the dirty industries is presented in table 7. The results are quite consistent with what is reported in table 2; in fact, if anything, the productivity coefficients are larger in absolute value than those of table 2 (and remain equally significant).

Several other robustness checks, suggested by the reviewers, were carried out. These checks are omitted here due to space constraints, but are reported in the supplementary appendix online. We explored the performance of our alternative TFP estimates, which allowed variations across years 2002, 2005, and 2008 (as discussed earlier), in explaining emission intensity. Because these TFP estimates are somewhat noisier than those obtained from the 2000–2008 data used in table 2, the corresponding coefficients of the productivity variable are a bit smaller in absolute value, but the sign and significance of these coefficients are the same as those of table 2. Similarly, these alternative estimates of TFP coefficients do not change at all the results and inferences reported in tables 3, 4, and 5. Introducing other control variables (from a limited set of available plant-specific characteristics, such as public status of firms, foreign ownership, gender of CEO, government contract status, and single-plant firm status) also did not change the results of table 2 in any meaningful way. Similarly, omitting the NA variables, or omitting some of the fixed effects (industry, year, or state), did not seem to have impacts on the size, sign, and significance of the coefficients of the productivity variable in table 2.

Conclusion

The relationship among international trade, firm productivity, and environmental outcomes is both intellectually interesting and policy relevant. Contemporary research suggests that firm heterogeneity with respect to productivity is a prominent part of the economic environment and is an important predictor of export status. Also, to the extent that the use of inputs—rather than output per se—is the likely cause of emissions, then an inverse relationship between productivity and emissions per unit output is to be expected. Furthermore, to the extent that large firms have more incentive to adopt newer and cleaner technology, then exporting firms—which face larger potential markets than non-exporters—are more likely to adopt cleaner technologies and thus have lower emissions per unit sales.

In order to test our two basic hypotheses (more efficient firms and exporting firms are likely to have lower emissions per unit sales), we assembled a large and unique data set for the U.S. manufacturing industry. Specifically, we have matched facility-level air pollution data from the U.S. EPA with facility-level economic characteristics data obtained from NETS. The empirical analysis based on these data that we have presented provides strong support for these two hypotheses. We find robust evidence of a negative correlation between the estimated facility productivity and emissions per value of sales. The negative impact of productivity is statistically significant for each criteria air pollutant we track. Furthermore, we present additional evidence that this productivity effect is monotonic: after dividing the firms into quintiles based upon their relative productivity within an industry, we find that the higher the productivity quintile, the lower the firm's relative emissions.

In addition, we find that exporting facilities tend to have less emission per value of sales than competing non-exporters within the same industry, conditional on estimated productivity and on exposure to the CAAA. We also find evidence that, given productivity, larger firms have lower relative emissions, consistent with the notion that larger firms have more incentives to incur costs associated with adopting cleaner technologies. Finally, consistent with the predictions of modern trade theory, we find that facilities with higher estimated productivity are more likely to export.

This empirical evidence, along with empirical work that identifies impacts of trade liberalization on technology adoption (Bustos 2011), have some interesting policy implications. Clearly, the optimal response to pollutants that have global consequences is an internationally coordinated effort to reduce those pollutants through pollution taxes or "cap and trade" programs. Yet, for a wide variety of reasons, efforts to achieve such coordination of environmental policies have not been very successful. On the other hand, while there is broad (if not complete) support among economists for policies that liberalize trade in goods, there is skepticism about the value of these policies among some politicians and the general public. If the predictions that expanded trade leads to an increase in the equilibrium productivity of firms, and if our results showing that increased firm productivity is correlated with lower emissions intensity are correct, then the potential benefits that would come from successful completion of the Trans Pacific Partnership or successful negations for a US-EU free trade area, exceed the purely material benefits. And it should be understood that these benefits are mutual—that is, the expanded productivity in both the United States and in its trading partners should have beneficial environmental consequences. While there are those who fear that globalization will lead to further environmental degradation, the results of this paper in fact provide support for the belief that globalization, largely through its impact on firm-level productivity, may contribute to reducing global pollution. Thus, whereas international trade cannot be construed as a substitute for environmental policies, it is also apparent that it should not be seen as adverse to environmental outcomes.

Supplementary Material

Supplementary material is available at http://oxfordjournals.our_journals/ajae/online.

References

Antweler, W., B. Copeland, and S. Taylor. 2001. Is Free Trade Good for the Environment? *American Economic Review* 91 (4): 877–908.

Barrows, G., and H. Ollivier. 2014. Does Trade Make Firms Cleaner? Theory and Evidence from Indian Manufacturing. Working paper, University of California, Berkley.

Batrakova, S., and R. Daves. 2012. Is There an Environmental Benefit to Being an Exporter? Evidence from Firm Level Data. *Review of World Economics* 148 (3): 449–74.

Becker, R. 2011. Local Environmental Regulation and Plant-level Productivity. *Ecological Economics* 70 (12): 2516–22.

Becker, R., W. Gray, and J. Marvakov. 2013. NBER-CES Manufacturing Industry Database: Technical Notes. *NBER Documentation*, February.

Becker, R., and V. Henderson. 2000. Effects of Air Quality Regulations on Polluting Industries. *Journal of Political Economy* 108 (2): 379–421.

Bernard, A., and B. Jensen. 1999. Exceptional Exporter Performance: Cause, Effect, or Both? *Journal of International Economics* 47 (1): 1–25.

Bernard, A., J. Eaton, B. Jensen, and S. Kortum. 2003. Plants and Productivity in International Trade. *American Economic Review* 93 (4): 1268–90.

Bernard, A., B. Jensen, and P. Schott. 2006. Trade Costs, Firms and Productivity. *Journal of Monetary Economics* 53 (5): 917–37.

Bernard, A., B. Jensen, S. Redding, and P. Schott. 2007. Firms in International Trade. *Journal of Economic Perspectives* 21 (3): 105–30.

Bustos, P. 2011. Trade Liberalization, Exports and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms. *American Economic Review* 101 (1): 304–40.

Cao, J., L.D. Qiu, and M. Zhou. Forthcoming. Who Invest More in Advanced Abatement Technology: Theory and Evidence. *Canadian Journal of Economics*.

Cole, M.A., and R.J.R. Elliott. 2003. Determining the Trade-environment Composition Effect: The Role of Capital, Labor and Environmental Regulations. *Journal of Environmental Economics and Management* 46 (3): 363–83.

Copeland, B., and S. Taylor. 1994. North-South Trade and the Environment. *Quarterly Journal of Economics* 109 (3): 755–87.

———. 1995. Trade and Transboundary Pollution. *American Economic Review* 85 (4): 716–37.

———. 2003. *Trade and the Environment: Theory and Evidence*. Princeton University Press.

Cui, J. 2014. Induced Clean Technology Adoption and International Trade with Heterogeneous Firms. Working paper, School of Economics and Management, Wuhan University.

Cui, J., H. Lapan, and G. Moschini. 2012. Are Exporters More Environmentally Friendly than Non-Exporters? Theory and Evidence. Working Paper No. 12022, Department of Economics, Iowa State University.

De Loecker, J. 2007. Do Exports Generate Higher Productivity? Evidence from Slovenia. *Journal of International Economics* 73 (1): 69–98.

Forslid, R., T. Okubo, and K-H. Ulltveit-Moe. 2014. Why Are Firms that Export Cleaner? International Trade and CO2 Emissions. Discussion Paper No. 8583, Centre for Economic Policy Research.

Frankel, J., and A. Rose. 2005. Is Trade Good or Bad for the Environment? Sorting Out the Causality. *Review of Economics and Statistics* 87 (1): 85–91.

Gibson, M.J., and T.A. Graciano. 2011. The Decision to Import. *American Journal of Agricultural Economics* 93 (2): 444–9.

Girma, S., and A. Hanley. 2015. How Green are Exporters? *Scottish Journal of Political Economy* 62 (3): 291–309.

Greenstone, M. 2002. The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970s and 1977 Clean Air Act Amendments and the Census of Manufactures. *Journal of Political Economy* 110 (2): 353–77.

———. 2004. Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations? *Journal of Environmental Economics and Management* 47 (3): 585–611.

Greenstone, M., J. List, and C. Syverson. 2011. The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing. Washington DC: U.S. Census Bureau, Center for Economic Studies Working Paper No. CES-WP-11-03.

Hanna, R. 2010. U.S. Environmental Regulation and FDI: Evidence from a Panel of U.S. Based Multinational Firms. *American Economic Journal: Applied Economics* 2 (3): 158–89.

Henderson, V. 1996. Effects of Air Quality Regulation. *American Economic Review* 86 (4): 789–813.

Holladay, S. Forthcoming. Exporters and the Environment. *Canadian Journal of Economics*.

Kolko, J., and D. Neumark. 2008. Changes in the Location of Employment and Ownership: Evidence from California. *Journal of Regional Science* 48 (4): 717–44.

———. 2010. Does Local Business Ownership Insulate Cities from Economic Shocks. *Journal of Urban Economics* 67 (1): 103–15.

Kreickemeier, U., and P.M. Richter. 2014. Trade and the Environment: The Role of Firm Heterogeneity. *Review of International Economics* 22 (2): 209–25.

Levinson, A. 2009. Technology, International Trade, and Pollution from US Manufacturing. *American Economic Review* 99 (5): 2177–92.

Managi, S., A. Hibiki, and T. Tsurumi. 2009. Does Trade Openness Improve Environmental Quality? *Journal of Environmental Economics and Management* 58 (3): 346–63.

Melitz, M.J. 2003. The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71 (6): 1695–725.

Melitz, M.J., and S.J. Redding. 2012. Heterogeneous Firms and Trade. *NBER Working Paper* 18652.

Melitz, M.J., and D. Trefler. 2012. Gains From Trade When Firms Matter. *Journal of Economic Perspectives* 26 (2): 91–118.

Neumark, D., B. Wall, and J. Zhang. 2011. Do Small Business Create More Jobs? New Evidence for the United States from the National Establishment Time Series. *Review of Economics and Statistics* 93 (1): 16–29.

Pavcnik, N. 2002. Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants. *Review of Economic Studies* 69 (1): 245–76.

Schott, P.K. 2010. U.S. Manufacturing Exports and Imports by SIC or NAICS Category and Partner Country, 1972 to 2005, Yale School of Management & NBER. Available at: http://faculty.som.yale.edu/peterschott/sub_international.htm.

Tybout, J. 2003. Plant- and Firm-level Evidence on New Trade Theories. In *Handbook of International Trade*, ed. E. Kwan Choi and J. Harrigan, 388–415. Blackwell: Oxford.

U.S. Environmental Protection Agency. 2006. Documentation for the Final 2002 Point Source National Emissions Inventory. Technical Report. Washington DC. Available at: http://www.epa.gov/ttn/chief/net/2002inventory.html.

———. 2008. Documentation for the 2005 Point Source National Emissions Inventory. Technical Report. Available at: http://www.epa.gov/ttn/chief/net/2005_nei_point.pdf.

———. 2012. 2008 National Emissions Inventory, version 2 Technical Support Document. Technical Report. Washington DC. Available at: http://www.epa.gov/ttn/chief/net/2008neiv2/2008_neiv2_tsd_draft.pdf.

Van Beveren, I. 2012. Total Factor Productivity Estimation: A Practical Review. *Journal of Economic Surveys* 26 (1): 98–128.

Appendix

A.1. Description of the NEI Database

This section provides a brief introduction of the NEI facility-level emission database and summarized caveats of this database.

The NEI database includes estimates of annual criteria and hazardous air pollutant emissions from sources in the 50 U.S. States, the District of Columbia, Puerto Rico, and the Virgin Islands. Sources are divided into two large categories: stationary and mobile. The former includes point and nonpoint sources, and the latter consists of on-road and non-road sources. The collection and updating of 2002 and 2005 NEI databases follow with the Consolidated Emissions Reporting Rule (CERR). The 2008 NEI is compiled using the Air Emissions Reporting Rule (AERR), rather than its predecessor the CERR.[26] For the case of point sources (polluting facilities) data, both reporting rules require a report on actual emissions for all facilities that emit above certain thresholds, determined by pollutant. State or local pollution control agencies have to comply with the requirement. These agencies report emissions from larger point sources annually, and have a choice to report smaller point sources every three years or one-third of the sources each year. Smaller point source facilities with annual emissions below certain thresholds can be defined as nonpoint area sources. While states are more likely to report major sources as point sources and smaller sources as nonpoint sources, the EPA encourages states to submit small sources to the point inventory.

Some major caveats concerning the NEI database pertaining to point sources can be summarized as follows. First, the EPA developed the 2005 NEI data based on a reduced level of effort. Part of this reduced effort involved using some 2002 NEI data in the 2005 NEI as surrogates for emissions data representing 2005. The 2005 NEI database provides flag variables, “Start Date/End Date” fields, to indicate which data are 2005 emissions and which data are actually taken from 2002 emissions. Around one-third of the observations in the 2005 NEI have a flag variable of “Start Date” referring to year 2002. When it comes to the manufacturing industry, roughly one-quarter of observations in 2005 are duplicates of 2002 emissions. We dropped these observations from our study, as their duplicate nature entails that they do not carry independent information. Second, the 2008 NEI database was built from emissions data in the EIS. Note that this 2008 database uses a new facility identifier, called EIS site ID, rather than the previous NEI site ID. A comprehensive and updated coverage of facility identifiers may be obtained from the Emission Inventory System Gateway. This gateway, however, is only available to EPA staff, EIS data partners responsible for submitting data to EPA, and contractors working for the EPA on emissions-related work. For this study, we rely on the FRS ID reported in the FRS of the EPA to match polluting facilities across sample years. All observations in 2002 and 2005 NEI databases have both records and FRS ID reported in the FRS, and hence can be matched between these two years. However, one-eighth of the 2008 NEI database is missing from the FRS, and roughly 7% of facilities in the manufacturing industry in this database do not have any records in the FRS. These missing manufacturers are discarded in our study. Last but not least, as noted in the EPA technical document (EPA 2012), emission data for filterable and condensable components of particulate matter (i.e., PM10-FIL, PM2.5-FIL, and PM-CON) is not complete and should not be used at any aggregate level. Users interested in PM emissions are suggested to only consider primary particulate matter, which are PM10-PRI and PM2.5-PRI. Following this suggestion, TSPs in our study is the sum of these two pollutants.

A.2. Data Matching

The data matching work consists of two main procedures. First, we match polluting facilities within the NEI database across years, and then retrieve DUNS numbers for these polluters from the FRS of the EPA. Second, we match them with those appearing in the NETS database through the DUNS number.

The 2002 and 2005 NEI databases assign each polluting facility a unique NEI site ID, whereas the 2008 NEI data uses a different facility identifier called Emission Inventory System (EIS) ID. To match these NEI databases across sample years, we retrieve facility FRS ID from the FRS of the EPA. The FRS is a centrally-managed database that identifies facilities, sites, or places subject to environmental regulations or of environmental interests. The EZ Query in the FRS provides data download options for a customized list of facilities, which are

[26] For the CERR, please see http://www.epa.gov/ttn/chief/cerr/cerr.pdf. For the AERR, please refer to http://www.epa.gov/ttn/chief/aerr/final_published_aerr.pdf.

Table A1. Variable List

Variable	Definition	Source
Facility Level		
Sales	Value of sales ($)	NETS
Employees	Number of employees	NETS
Export Dummy	Export indicator, equals 1 if exports, 0 otherwise	NETS
Distance	Distance of a facility to its nearest port (miles)	Calculated
SO_2	Sulfur Oxide (tons)	NEI
CO	Carbon Monoxide (tons)	NEI
VOCs	Volatile Organic Compounds (tons)	NEI
NO_x	Oxide of Nitrogen (tons)	NEI
PM10-PRI	Primary particulate matter less than 10 microns (tons)	NEI
PM2.5-PRI	Primary particulate matter less than 2.5 microns (tons)	NEI
TSPs	Total Suspended Particulates, sum of PM10-PRI and PM2.5-PRI (tons)	Calculated
O_3	Ozone, sum of VOCs and NO_x (tons)	Calculated
SO_2Intensity	SO_2per sales	Calculated
CO Intensity	CO per sales	Calculated
O_3Intensity	O_3 per sales	Calculated
TSPs Intensity	TSPs per sales	Calculated
County Level		
SO_2NA	SO_2 Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
CO NA	CO Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
O_3NA	O_3 Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
TSPs NA	TSPs Nonattainment, equals 1 if nonattainment, 0 otherwise	EPA
Industry Level at Four-digit SIC		
CIF	Cost-Insurance-Freight value of U.S. imports	Peter Schott
FOB	Free-on-Board value of U.S. imports	Peter Schott
Freight Rate	(CIF - FOB)/FOB	Calculated
Deflator	Value of shipment deflator	NBER-CES

associated with NEI or EIS programs.[27] The data obtained from the EZ Query include three different facility identifiers: FRS ID uniquely assigned by the FRS, NEI site ID assigned by the NEI, and EIS facility ID assigned by the EIS. With the NEI site ID contained in the FRS, we are able to match all polluting facilities in the NEI database with those in the FRS through the NEI site ID between 2002 and 2005. However, around 7% of the 2008 NEI database in the manufacturing industry does not have records in the FRS. These observations are dropped in the study. With the FRS ID, facility DUNS numbers are retrieved separately through the Facility Registry System Query.[28] In the end, the facility-level emission dataset we compiled contains criteria air emissions, facility name, FIPS county code, Zip Code, SIC code, facility FRS ID, and DUNS number.

In the next step, we match polluting facilities in the NEI database with those that appear in the NETS Database through the DUNS number. The EPA does not provide further information about how DUNS numbers are reported for polluting facilities and why some of them have missing DUNS numbers in the dataset. Due to an incomplete report on DUNS numbers in the FRS, approximately 80% of polluting facilities in the manufacturing industry collected in the NEI database have associated DUNS numbers. Thus, a pair of facilities from each source is considered as a match if the following series of criteria are satisfied. The facilities share the same DUNS number and are located in the same area in terms of five-digit Zip Code and five-digit FIPS county code. More importantly, for each pair we compare their facility names from each source to ensure the match.

In the matched dataset, it turns out that the number of polluting facilities with zero emissions drops dramatically across years, while the number of polluting facilities with missing values for emission increases accordingly, suggesting a conflation of the two (conceptually distinct) statuses. This pattern

[27] For EZ Query, see http://www.epa.gov/enviro/html/fii/ez.html.
[28] For Facility Registry System Query, please refer to http://www.epa.gov/enviro/html/fii/fii_query_java.html.

actually exists in the original facility-level NEI database prior to matching. We drop from further consideration facilities that show missing values for the emission of all pollutants considered here; and because the distinction between non-emitting facilities from those for which the data are missing does not appear very credible in this dataset, we also drop those facilities with zero emission values for all pollutants.