High Priority Violations and Intra-firm Pollution Substitution

Binish Rijal and Neha Khanna

Department of Economics, Binghamton University

Email: brijal@binghamton.edu (Rijal); nkhanna@binghamton.edu (Khanna)

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

High Priority Violations and Intra-firm Pollution Substitution

Binish Rijal and Neha Khanna

Department of Economics, Binghamton University

PO Box 6000, Binghamton, NY 13902

Email: brijal@binghamton.edu (Rijal); nkhanna@binghamton.edu (Khanna)

June 18, 2017

Draft: Do not quote or cite without permission

Abstract

We examine the sign and the magnitude of spillovers associated with the High Priority Violations Policy (HPVP) under the Clean Air Act (CAA), which caps the quantities of emissions that a polluting facility can generate per period of time or per unit of fuel input. Using data containing over 181,000 plant-year level observations for 24,048 polluting facilities across all industries in the U.S., we find evidence of negative spillovers associated with the HPVP. On average, a compliant facility increased its air emissions by about 43 percent if it had at least one other sister-facility, within the same 6-digit NAICS industry code and belonging to the same parent firm, under violation. The magnitude of such intra-firm pollution substitution was stronger towards compliant facilities with no history of CAA violation.

JEL: Q52, Q53

Keywords: pollution substitution, pollution leakage, CAA, spillovers, nearest neighbor matching

1. Introduction

Environmental regulations are largely associated with positive increases in social welfare (Chay and Greenstone, 2003; Environmental Protection Agency, 2011). However, the magnitude of the welfare gains due to environmental regulations can be difficult to pin down. One of the main reasons why estimating the magnitude of net welfare gains is so challenging is due to the existence of spillovers. A spillover is an unintended consequence of an intended regulation, which can affect the overall welfare gains that can be attributed to the regulation. For example, if a pollution reduction policy in a certain location leads to a decrease in pollution in that location but, unintendedly, also leads to increase in pollution elsewhere, then the welfare impact of the pollution regulation is likely going to be positive in the location where it was implemented and negative where there were unintended consequences (spillovers). As a whole, the net welfare gains associated with such a policy would depend on the magnitude of the welfare gains due to the intended decrease in pollution as well as the magnitude of the unintended spillovers. As spillovers can either be negative or positive, they can cause the estimates of net welfare gains due to environmental regulations to be biased upward or downward, respectively. A reliable welfare analysis of any pollution regulation, thus, warrants a closer examination of potential spillovers. In this paper, we examine the potential spillovers associated with one policy in particular – the High Priority Violations Policy (henceforth, HPVP) under the Clean Air Act (CAA).

Implemented in 1999, the HPVP mandates inspection agencies to label certain violations of CAA requirement as "high priority" so as to ensure that these violations are addressed and resolved in a timely and appropriate manner. Figure 1 provides basic guidelines on how high priority violations are identified. A violation is labeled as "high-priority" (i) if it fits within one of the ten "General HPV Criteria," (ii) if it is associated with emission levels that fit within the "HPV Matrix Criteria," or (iii) on a discretionary basis (EPA, 1999). We are interested in those violations that either fit the eighth general HPV criterion or any of the five HPV matrix criteria because the HPVP can have some unintended consequences for these particular types of violations. For example, if a polluting facility is found to be in a high priority violation

that fits the abovementioned criteria, then the parent firm is required to decrease its emissions in the violating facility as per the policy mandate[1]; however, such requirements could also create incentives for that firm to substitute its emissions away from the violating facility towards other facilities that are not in violation with the CAA requirements. In that case, the sign of the spillover associated with the HPVP is negative, and the existence of such intra-firm pollution substitution, if not accounted for, can result in an upward bias in the estimates of the overall welfare gains due to the HPVP. Conversely, a cap on emissions at a violating facility could also trigger positive spillovers: the parent firm might be inclined to simultaneously decrease emissions in other facilities that are not in violation so as to avoid a situation where it has multiple facilities under violation, and hence, multiple fines and penalties. Positive spillovers might alternatively be triggered by the desire to maintain a decent public image of the company. Since all the information on HPVs are publicly available in the EPA's National Enforcement and Compliance History Online (ECHO) website, this is indeed a plausible hypothesis. Regardless, the magnitude and net sign of the spillovers associated with mandatory pollution reduction policies such as the HPVP is an important empirical question with profound policy implications.

To estimate the sign and magnitude of the spillovers associated with the HPVP under the CAA, we use data from EPA's Toxic Release Inventory (TRI) and the Integrated Compliance Information System for Air Pollution (ICIS-AIR). ICIS-AIR contains compliance and permit data for stationary sources of air pollution regulated by the EPA, state and various local air pollution agencies. Using one-to-one nearest neighbor matching for identification, we check whether a compliant facility belonging to a multi-facility firm increases or decreases its total TRI air emissions while it has at least one other sister-facility (i.e. facility belonging to the same parent firm) within the same 6-digit NAICS code under a high priority violation. Our treatment group comprises facilities belonging to multi-facility companies that have at least one sister-facility concurrently under high priority violation, while the control group

[1] Other HPV criteria on their own do not require facilities to decrease their emissions. For example, HPV criterion #1 is "failure to obtain new source review (NSR) permit". Similarly, HPV criterion # 10 is "violation of CAA Section 112(r)", which requires facilities to submit their accidental release prevention/risk management plan. These types of violations can be lifted with appropriate actions, which does not necessarily entail decreasing emissions.

consists of facilities belonging to multi-facility firms that do not have any sister-facilities concurrently under violation. To establish causality between the assignment of high priority violation at a facility and the change in emissions at other compliant sister-facilities, we match our treated observations with the nearest control based on various facility-level and firm-level characteristics such as the growth rate of the facility's total TRI emissions, the growth rate of the parent firm's total TRI emissions, a proxy for the age of the facility, the projected growth in the plant's total TRI emissions, and the projected growth in the parent firm's total TRI emissions. We also make sure that the nearest neighbor in our matched sample has an exact match on variables such as the reporting year, the 6-digit NAICS industry code, the facility's location (county/state), and the facility's participation in pollution prevention (P2).

We find strong and statistically significant evidence of negative spillovers or intra-firm pollution substitution associated with the HPVP. On average, a compliant facility increased its total TRI air emission by roughly 35 percent if it had at least one other sister-facility within the same 6-digit industry code under violation. Furthermore, we find that the magnitude of the intra-firm pollution substitution is much stronger towards compliant facilities with no history of high priority violation as compared to those with some history of high priority violations.

Our research contributes to the literature on pollution leakage[2] (Levinson and Taylor, 2008; Fowlie, 2009; David and Kahn, 2010; Gibson, 2015.) The empirical question we attempt to answer has important policy implications. If multi-facility firms substitute emissions among facilities within the company, then pollution reduction policies directed at protecting public health must not only consider the intended emission reductions at the targeted facility but also regulation induced unintended change in emissions across all other facilities within the firm. The magnitude of the elasticity of pollution substitution would be a vital input for future policy design.

[2] Pollution leakage or the spatial leakage of pollution, is the transfer of pollution from one geographical location to another, usually (but not always) due to differential stringency of environmental regulations in the two locations.

The rest of the paper is organized as follows: Section 2 provides a brief literature review on pollution leakage and other spillovers associated with environmental regulations. Section 3 contains some more background information on the Clean Air Act and the HPVP. Section 4 provides a description of the data and the summary statistics. Section 5 contains a discussion of the identification strategy. Section 6 contains a discussion of the results, and Section 7 concludes.

2. Pollution Leakage: A Brief Review of Literature

Pollution leakage has mostly been discussed in the literature related to the spillovers associated with environmental regulation (Carrada-Bravo, 1995; Henderson, 1996; List et. al., 2003; Dean et. al., 2009; Hanna, 2009). However, much of the "leakage" discussed in the literature has been the transfer of pollution across national boundaries as opposed to pollution leakage within national boundaries. Levinson and Taylor (2008) examine the effect of regulatory intervention on the flow of trade and pollution emissions across national borders. Using U.S. regulations and net trade flow data for 130 manufacturing industries from 1977 to 1986 in the U.S., Canada and Mexico, they show that industries whose abatement costs increased due to additional environmental regulation experienced the largest increase in net imports. The results were indicative of a geographical substitution of pollution towards "pollution havens" due to the scale effect.[3] Hanna (2010) also finds that multinational firms when subject to environmental regulations in the United States increase their foreign assets by 5.3 percent and increase their foreign output by 9 percent as compared to their non-regulated counterparts. Overall, the results in these papers are consistent with the view that U.S. environmental regulations cause U.S. firms to move capital, jobs, and pollution to locations abroad.

[3] The scale effect considers the role of output levels on total emissions: holding the composition of clean vs dirty industries and the production technique employed constant, total emissions can increase in pollution havens because more output is produced there.

In examining the impact of environmental regulation on the geographical substitution of mobile-source air pollution, David and Kahn (2010) exploit the trade deregulation that took place between United States and Mexico under the North American Free Trade Agreement (NAFTA). Their results show that traded vehicles leaving the U.S. [following NAFTA] were higher-emitting per mile than the stock of vehicles in the United States, but lower-emitting per mile than the stock of vehicles in Mexico. The study also shows that used vehicles exported to Mexico from the United States were more likely to have failed regulatory emissions testing; hence, supporting the pollution leakage hypothesis.

Some studies have also analyzed the negative spillovers of environmental regulation under complete regulation versus incomplete regulation. Fowlie (2009), in particular, proposes that when regulated producers are less polluting than their unregulated counterparts, the aggregate level of emissions that would have occurred in the absence of regulation could be exceeded by the level of emissions under incomplete regulation. Furthermore, when regulated firms are relatively more polluting, aggregate emissions under complete regulation could exceed aggregate emissions under incomplete regulation. These propositions were empirically tested and confirmed using simulations on California's electricity sector under a source-based cap-and-trade program.

Environmental regulations can have other potential spillovers in that they can also significantly impact where the polluting firms decide to construct new plants or even move existing ones. Condliffe and Morgan (2009), for example, examine the impact of the 1977 Clean Air Act (CAA) amendments on the location decision of pollution-intensive manufacturing firms. Using panel data on manufacturing plant births and county-level NAAQS attainment status for the federal standard of ozone concentration, they find that enhanced regulatory scrutiny and more stringent local level environmental regulations can significantly deter new plant birth in that area. These results are consistent with the findings from Becker and Henderson (1996) who also show that the designation of non-attainment status as per the CAA mandate reduces expected birth rate of plants by 40-50% in the polluting industries, ceteris paribus.

Gibson (2015) is perhaps the only paper in the pollution leakage literature that attempts to examine the spatial substitution of pollution within the United States. Using EPA's Toxic Release Inventory (TRI) data, he examines whether CAA regulations create incentives for multi-plant firms to substitute air pollution away from plants located in close proximity to air quality monitors that are in violation with the national ambient particulate matter (PM_{10}) standards to plants located in counties where the air quality monitors are in compliance with PM_{10} NAAQS. He finds that regulation in an average plant increases air emissions at unregulated plants belonging to the same 6-digit NAICS code owned by the same firm by roughly 17 percent, resulting in a net-emission increase. Adding state level linear time trends and accounting for the number of treated plants (plants located within 2 km radius of PM non-attainment air quality monitors), the estimates were slightly smaller at 12 percent. Gibson's results indicate that although stringent measures taken to improve the ambient air quality around non-attainment monitors can be effective, the regulatory pressure to decrease emissions can also trigger spatial substitution of pollution towards other locations where regulatory scrutiny might be relatively lax. Our work adds to Gibson's in three major ways. First, unlike in Gibson's work, where treatment is assigned at the air quality monitor level, in our analysis treatment occurs at the plant level. As much of his identification rests on the assumption about the exogeneity of treatment on plant-level outcomes, this is not likely a concern in our analysis as it is extremely unlikely for emissions from one plant to affect the treatment status in another. Second, since we consider many more toxic gases than just PM_{10}, our results on pollution substitution are more generalizable than Gibson's pollutant specific result. Lastly, Gibson focuses on estimating cross-media pollution substitution.[4] We are not interested in such cross-media substitution of pollution within a plant but rather the substitution of TRI air emissions across plants within a parent firm. We also attempt to examine the heterogeneity in the magnitude of intra-firm air pollution substitution across several plant and firm level characteristics, which is absent in Gibson's work.

[4] Cross media pollution substitution is the substitution of pollution within a plant from one medium (such as air) to other media (such as water, landfills, etc.).

Positive spillovers associated with environmental regulations are relatively elusive in the literature. Nevertheless, that is not to say that they are non-existent. Liu (2012), for instance, shows that facility-level inspections under the CAA have a positive and significant effect on facility compliance with Reservation and Conservation Recovery Act (RCRA) which deals with hazardous and non-hazardous solid waste. Furthermore, Dechezlepretre and Sato (2014) argue that there is now ample evidence linking strong environmental regulations with innovations in clean technologies and lower levels of research and development in conventional polluting technologies. They state that innovations in clean technologies induce larger economic benefits due to knowledge spillovers than innovations in dirtier technologies which they replace. Jiang *et. at.* (2015) attempt to explicitly estimate the magnitude of the positive spillovers associated with a policy designed to decrease SO_2 emissions in China. They show that positive spillovers could be caused by the spread of pollutants (or in their case, the lack of) through air or water across different cities within China. In other words, pollution reduction in one location would spread less pollution to other nearby locations; hence, causing the spillover associated with implementing the SO_2 reduction policy to be positive in sign. Taking into account the positive spillovers, they were able to attribute about 21 percent decrease in SO_2 concentration to the pollution reduction policy, whereas without taking into account the positive spillover, the estimated reduction in SO_2 concentration as a result of the policy was only about 13 percent. Lastly, Evans (2016) has also suggested that provision of information about a firm's environmental record, such as the record of plants entering into and exiting from HPV status, may encourage the firm to make improvements in its environmental performance as they adjust to market responses (e.g., from consumers or stockholders) or other pressures such as from local community groups.

3. Background: The High Priority Violations Policy (HPVP)

Implemented in 1999, the HPVP was designed by the EPA as a management tool for oversight over specific categories of violations under the CAA.[5] The EPA considers a CAA violation a high priority violation if (i) it is likely to result in a significant risk to human health and/or the environment from either direct or indirect release of air pollutants or (ii) it may harm the ability of the state, local, territorial and tribal agencies (henceforth, the enforcement agencies) to implement CAA programs.[6] An HPV is always a significant violation of a federally enforceable CAA program by either a major or synthetic minor stationary source of air pollution.[7] Although the EPA considers all CAA violations important, the HPVs warrant additional scrutiny to ensure that enforcement agencies respond to these types of violations in a timely and appropriate manner and have access to federal assistance (EPA, 1998, 2014).

The day a violation is identified as a high priority violation is called "Day Zero". The enforcement agencies are responsible for reporting all data about HPVs to ICIS-AIR, which is the Air Module of EPA's national compliance information database. The enforcement agency must also advise the polluting source that an HPV has been identified within no more than 45 days of Day Zero. Enforcement agencies can address the HPVs through the following actions: (1) by issuing a legally enforceable order that requires immediate action to come into compliance with the requirement violated; (2) by issuing a legally enforceable order that imposes penalties, where the source has demonstrated that it is currently complying with the requirement violated; (3) by issuing a legally enforceable order that imposes a schedule on the source to comply with the requirement violated and penalties for violation; or (4) transferring the matter to another organization with the authority to initiate a civil or criminal judicial action against the violator[8] (EPA, 2014). The enforcement agencies can take

[5] HPVs include not only the regulations implementing the CAA, but also other state, local, tribal, or territorial regulations approved by the EPA so as to implement the requirements of the CAA.

[6] Figure 1 provides a flowchart on HPV applicability determination.

[7] In our sample, plants have been assigned HPV status due to violation of several federally enforceable CAA programs. Some examples are CAAMACT, CAANSPS, CAAPSD, CAASIP, CAANESH, CAANSR, etc.

[8] This is known as "a referral."

any of the above actions to address the HPV but it needs to be done within 180 days of Day Zero. For those HPVs that are not addressed within 180 days, the enforcement agency is required to schedule a case-specific consultation with the regional EPA to discuss the matter within no less than 270 days from Day Zero.

A high priority violation that is not addressed or resolved in a timely manner gets listed under the "watch list". The EPA has maintained the watch list since 2004 to track HPVs of the CAA that are not addressed or resolved within 9 months of a facility entering violation. It is worth noting that the listing of HPVs on the watch list depends both on the behavior of the regulated facilities and the behavior of the regulator. An HPV can be listed on the watch list if either the facility fails to come into compliance or if the regulator fails to address the HPV with an appropriate action.

The watch list, by design, implies increased regulatory scrutiny and increased pressure for a facility to return to compliance (Evans, 2016). Furthermore, the penalties for failure to resolve high priority violations are also sufficiently high such that they effectively deter the polluting sources from deliberately refraining to comply. In cases where no penalty is levied, or where the assessed penalty is different from the collected penalty, the enforcement agency must provide an explanation to the regional authority (EPA, 2014).

4. Description of the dataset

A. Emissions Data

The plant level emissions data come from the EPA's Toxic Release Inventory (TRI) 2000-2014. The TRI records emissions on an annual basis for more than 500 chemicals (measured in pounds or grams) for U.S. facilities across many industries. Typically, only plants that emit more than 10,000 pounds of TRI pollutants (or 5,000 pounds in some cases) are obliged to report their emissions. Reported emissions are estimated using various engineering models,

rather than being obtained through direct measurements. Under TRI there are serious penalties for false reporting. In the past, the EPA has fined up to $27,000 per day for reporting problems (Gamper-Rabindran, 2009; Gibson, 2015). Therefore, we believe that firms have little incentive to under-report their emissions to the TRI and, despite their self-reported nature, the TRI is probably the most widely used source for data on toxic emissions in the US. TRI also contains the Duns & Bradstreet unique 9-digit identification number (DUNS number) indicating the parent company that a facility belongs to. As the EPA started collecting the TRI data since 1987, we also have information on the first year a facility reported its emissions. We use this information as a proxy for the facility's age. In addition to current emissions, the facilities are also required to report their projected TRI emissions for the following year and the second following year, which we make use of in our analysis.

B. Violations Data

Data on the plant specific HPVs were collected from the National Enforcement and Compliance History Online (ECHO) website. In particular, we obtained the ICIS-AIR data which contains emissions, compliance and enforcement data on stationary source air pollution. The sources in the dataset cover a wide spectrum of polluters from large industrial facilities to relatively small operations and industrial plants. The plant level ICIS-AIR data treats the entire facility as one unit rather than looking at individual emission points, processes, or stacks.

From the plant level HPV data, we omit those violations that were resolved before the year 2000 or the ones that were issued after 2014. We are only interested in the HPVs that were active between 2000 and 2014 due to the change in the HPV policy in August of 2014. To analyze the impact of HPVs related to [over]emission of air pollution (i.e. violations that fit the general HPV criterion # 8 or any of the 5 HPV matrix criteria) on the plant's total TRI air emissions, we discard the HPVs that are related to carbon monoxide, sulfur dioxide, nitrogen

dioxide, or other pollutants that cannot be traced to the TRI 2000 core chemical list.[9] Out of the 2,390 high priority violations in our dataset, 219 violations fit this category and thus had to be excluded from our analysis.

C. Analysis Sample and Summary Statistics

The TRI emissions data are merged with the HPV data using the plants' Facility Reporting System (FRS) identification number. Our analysis sample is constrained to only those facilities that have a TRI Facility Identification (TRIF ID) number. Several HPVs that were active between 2000 and 2014 had to be omitted from our analysis because they were not listed as a TRI facility and, hence, lacked information on facility level annual emissions. We place two additional restrictions on our data: we exclude facilities that emit less than 10 pounds of total air emissions per year as well as those facilities that produce more than 100,000 pounds of total air emissions per year, thus excluding very small and exceptionally large facilities, some of which produce hundreds of thousands of pounds of TRI air emissions per year. The histograms showing the distribution of total air emissions in our analysis sample before and after trimming are presented in figure 2. All information on the county non-attainment status was retrieved from the EPA's Green Book.

The summary statistics are presented in table 1. We provide the mean and the standard deviation for treated and the control groups separately. We also include the normalized difference between the control and the treated group. The normalized difference is a natural measure of the difference between the distributional locations in our treated and control arms (Imbens and Rubio 2016). It is a scale-free equal to the difference in means scaled by the square root of the average of the two within-group variances. A normalized difference closer to zero suggests a greater similarity in the distribution of the covariate between the two opposite treatment arms.

[9] Lead is one of the TRI chemicals, and several other TRI chemicals can be mapped into either the PM category or the VOC category. See Greenstone (2003) for more details.

On average, we see that our treated group is a larger producer of total air emissions (measured in pounds). The participation rate in pollution prevention (P2) does not seem to be much different across the two groups; neither does the non-attainment status of the country where the facilities are located. While 39 percent our control facilities are located in counties that are in non-attainment for O3, VOCs or Lead, the percentage of the treated facilities located in non-attainment counties is 42. Overall, we see not too big a difference between the distributional locations in our treated and control arms for matching covariates except: the number of additional compliant sisters, the lag of total air emissions, and the expected growth in the parent firm TRI emissions. While the control facilities tend to have on average about 3 other compliant sister-facilities, the treated facilities tend to have about 7, which is more than double compared to the control. We can also see that while the one-year lag of total air emissions for our control facilities is on average 10,325 lbs. of air emissions, the one-year lag of total air emissions for our treated facilities is 14,226 lbs.

5. Identification Strategy

We use one-to-one nearest neighbor covariate matching to estimate the magnitude of intra-firm pollution substitution triggered by sister-plants entering high priority violations. Our treatment group comprises compliant facilities belonging to multi-facility firms that have at least one sister-facility within the same 6-digit NAICS industry code concurrently under high priority violation; and the control group consists of compliant facilities belonging to multi-facility firms that do not have any sister-facilities under the same 6-digit NAICS industry code concurrently under violation. As mentioned in table 1, we only have 3,126 treated observations in our analysis sample, whereas the number of control observations is 34,038. Due to this stark difference in the number of observations in opposite treatment arms, it is not feasible for us to estimate the average treatment effect via one-to-one nearest neighbor matching as we lack sufficient treated observations which can serve as plausible counterfactuals for the controls. We could allow for matching with replacement so as to get the sample average treatment effect (SATE), but that is not likely going to give us good

quality matches given the degree of the difference in the relative size of the opposite treatment arms in our case.

We thus estimate an alternative average causal effect – the sample average treatment effect on the treated (SATT). This causal estimate gives us the average effect of the treatment only on the group that receives the treatment:

$$\tau_T^S = \frac{1}{N_T} \sum_{i:\, W_i=1} [\, Y_i(1) - \, Y_i(0)].$$

τ_T^S here, is the SATT, N_T is the number of treated units, W_i is the treatment status (1 if treated, 0 otherwise), $Y_i(1)$ is the realized outcome and $Y_i(0)$ are the potential outcome under control treatment. Estimating the SATT is useful for our purpose because we can infer from it how different the observed outcome for our treated facilities would have likely been in the absence of the treatment. In other words, our SATT estimates can tell us how much less (or more) total TRI air pollution would have been emitted by the treatment facilities had they not had at least one sister-facility under the same 6-digit NAICS industry code contemporaneously under high priority violation.

Employing nearest neighbor matching to estimate the SATT also serves several useful purposes in our case. We match our treated observations with the controls based on exact matches for variables including the reporting year, the location of the facility (county or state), the non-attainment status of the county (if we are matching facilities within the same state but not necessarily in the same county), the NAICS 6-digit industry code to which the facility belongs, and an indicator for the facility's participation in the EPA's voluntary pollution prevention (P2) program. By specifying exact matches for these variables, we can prevent our SATT estimates from being driven by industry specific differences between the treated and the control facilities, other industry-year level shocks that are likely going to differentially affect facilities in the two treatment arms, or other unobservables at the aggregate year level that might be correlated with a facility's production of air emissions.

Furthermore, we make sure that all treated facilities that are P2 participants are matched with control facilities that are also P2 participants, and the non-participant treated facilities are matched with their non-participating control counterparts.

Other continuous variables for which we find the nearest neighbor are: one-year lag of the plant's total air emissions,[10] a proxy for the plant age which is derived from the first year a facility reported its emissions to the TRI, total emissions growth rate of the plant, total emissions growth rate of the parent firm, the expected total emissions growth rate at the plant and the expected total emissions growth rate at the parent firm. Matching based on the plant's total emissions and the firm's total emissions becomes important because comparing treated and control facilities that display differential plant level or firm level emission trends will not allow us to obtain a clean estimate of the treatment effect. For example, if the treated facilities have a greater total emissions growth trend as compared to the controls, perhaps due to a differential increase in demand at the firm level, then our causal estimate of the treatment effect (HPV induced increase in emissions at compliant facilities) is likely going to be biased upward. In that case, much of what we estimate to be the increase in emissions at compliant facilities due to HPV induced intra-firm pollution substitution, might simply be increase in emissions at compliant facilities due to firm level demand shocks. Accounting for the age of the plant also becomes important as older plants tend to be more polluting as compared to their younger counterparts. Furthermore, additional matching variables such as the expected total emissions growth at the plant and the expected total emissions growth at the parent firm not only help us identify facilities that display similar emission trends for comparison, but they also give us a sense of the plant level or the firm level managerial intent to curb total pollution. Hence, the inclusion of these variables would also allow us to match "greener" [or "dirtier"] treatment plants belonging to "greener" [or "dirtier"] firms with other control facilities displaying similar characteristics.

[10] We have information on lagged total TRI air emissions only for those facilities that report emissions for successive years. For the remaining facilities, we use the emissions in the second-latest reporting year in place of one-year lagged emissions. The choice of our lagged variable does not change the results significantly. For facilities that started reporting their emissions to the TRI between 2000 and 2014, we are not able to create any lagged value of total air emissions for their first reporting year and we exclude these observations when obtaining our causal estimates.

When we have multiple continuous variables that cannot be matched exactly, the choice of a distance metric, which puts a weight on the different matching variables, becomes important in determining the nearest neighbor. A very common choice for the distance is the Mahalanobis metric, where the weight matrix is based on the average of the within-treatment-group sample covariance matrices. This metric takes into account correlations across covariates and leads to matches that are invariant to affine transformations of the covariates (Imbens and Rubio, 2015). This property of the Mahalanobis distance metric is useful for us because our pre-treatment variables have no natural scale.

Another potential pitfall of our estimation strategy is that the nearest neighbor matching estimator is likely going to be biased if matching on all the variables is not exact. For example, two facilities with the same one year lagged value of the total air emissions, belonging to the same 6-digit industry, located in the same county and sharing the same P2 participation status, might not be good matches for each other due to other variables that are not perfectly matched. A plant belonging to a firm which has a sharply declining total TRI emissions (perhaps due to a firm level shock) can be considered to be very different from another plant for which the parent firm's TRI emissions are growing, despite the two plants having the exact matches for other variables. To tackle with this problem and avoid getting biased estimates of the treatment effects, we employ simple regression adjustments as prescribed by Abadie and Imbens (2011) on our results using the original continuous matching variables for which we are not likely to get exact matches.[11] All estimates presented in our result tables are therefore regression-adjusted bias-corrected estimates of the sample average treatment on the treated (SATT).

We can see in table 2 that our matched sample tends to show a smaller difference in the distributional locations of the pre-treatment covariates between the two treatment arms. All the variables shown on table 2 were used for nearest neighbor matching in addition to exact matches for reporting year, plant location (state), county's NAAQS non-attainment status, 6-

[11] See Abadie and Imbens (2011) or Imbens and Rubio (2016) for more details.

digit NAICS, and the P2 participation status of the plant. The only normalized difference in mean that exceeds 0.1 is for the 1-yr lag of total air emissions. We acknowledge that we have not used the number of additional compliant sister-facilities as a nearest neighbor matching variable.[12] Nevertheless, a significant difference in the number of other compliant facilities between the treated and the control group can lead to our estimate of the SATT to be biased. If a compliant plant that has a sister-facility concurrently under violation, but also many other sister-facilities that are concurrently not under violation, then it is not likely going to increase its emissions significantly as a result of the treatment. In such cases, the firm might able to cope with the regulatory pressure to decrease its emissions at the violating facility by dissipating the increase in emissions across many compliant facilities. To address this potential source of bias, in our ongoing work, we are separately examining the heterogeneity in the magnitude of intra-firm pollution substitution by the additional number of compliant sister facilities belonging to the treated group.

6. Results

The SATT estimates are presented in table 3A and 3B. In table 3A, we use the natural log of total TRI air emissions as our dependent variable. Doing this we can interpret the SATT estimates as the percentage increase in the facility's total air emissions due to at least one sister-facility under the same 6-digit NAICS industry code being concurrently under violation. In column 1 we match the treated facilities with the controls based on exact matches for the facility county, reporting year, 6-digit NAICS code and facility's P2 participation status. The estimated treatment effect suggests that the treated facilities are likely to emit about 104% more total TRI air emissions as compared to the control counterparts located in the same county. This is the most basic model where we match the treated group with the control based on exact matches for 4 variables, namely, the reporting

[12] Attempting to match the treated facilities with the controls based on the number of additional compliant sister-facilities involves certain trade-offs. For example, this would significantly decrease the number of potential matches for our treatment facilities affecting the statistical power of our analysis. Another trade-off includes inability to get better matches for other pre-treatment variables, which might lead to biased causal estimates.

year, 6-digit NAICS, facility county, and the status of the facility's P2 participation, and nearest neighbor matching based on 3 variables: one year lag of facility total air emission, the facility's total TRI emission growth, and the parent firm's total TRI emission growth. However, by specifying exact matches based on the facility county, we were only able to match 733 treated observations (which is only about 35 percent of the total treated observations in our analysis sample). Therefore, to increase the size of our matched sample and thus get a more accurate SATT, we match the treated facilities with other controls that are not necessarily located in the same county but rather the same state (see column 2). A potential source of bias is that if our matched treatment facilities, as compared to the controls, are disproportionately located in counties that are in non-attainment with the NAAQS for ambient lead, VOCs, or PM, then our estimate of the intra-firm pollution substitution is likely going to be biased upwards.[13] In the opposite case, when matched control facilities are disproportionately located in non-attainment counties, our SATT estimate can be biased downward. Therefore, to control for such regulatory pressures at the county level that might be confounded with our potential outcomes, we specify an additional exact matching variable denoting the county's non-attainment status. As we can see in column 2, we were able to significantly increase the number of matches by doing so. In an attempt to get a more accurate estimate of the SATT, we also include a few additional continuous variables: the plant's age, expected total TRI emissions growth for the plant and expected total TRI emissions growth for the parent firm in order to select the nearest neighbor. Our estimate of the increase in total air emissions at the treated facility after the inclusion of these additional matching variables is about 35% (see column 3).

In table 3B, we carry out the exact same NN matching estimation except with the total TRI air emissions (measured in pounds) as the dependent variable. As we can see in column 3, compliant treated facilities, which have at least one sister-facility within the same industry concurrently under violation, would have likely produced 1,705 pounds of TRI air emission

[13] Other criteria air pollutants like CO, SO2 and NO2 are not included in the TRI chemical list. The non-attainment status triggered by these pollutants is thus not likely going to exert significant pressure on the polluting facility to decrease its TRI air emissions. Lead, on the other hand, is listed as a TRI chemical, and several other TRI chemicals can be categorized as VOCs or PM.

had they not entered treatment. In other words, the impact of the treatment on the treated facilities in our sample is the release of 1,705 additional pounds of air pollution.

In table 4, we can see the heterogeneity in the SATT estimates across top 6 NAICS 3-digit industries for which we were able to get the most number of exact matches. Our results suggest that the high magnitude of intra-firm pollution substitution as shown in tables 3A and 3B are mostly driven by increase in emissions at compliant facilities belonging to 3-digit NAICS industries such as chemical manufacturing (325), primary metals manufacturing (331), and above all, fabricated metal products manufacturing (332).Next, we check the magnitude of intra-firm pollution substitution between treated facilities with no history of HPVs (treated facility is a compliant facility with no history of HPVs that has a sister facility in the same 6-digit NAICS that is concurrently under violation) -- and treated facilities with some history of HPVs (treated facility is a compliant facility with some past history of HPVs with a sister facility in the same 6-digit NAICS that is concurrently under violation). To do show, we run two separate nearest-neighbor matching estimations: one for the treated facilities with some history and another for treated facilities with no history of high priority violation. We also account for the violation history of the nearest neighbor control facilities as inexact matches based on the treated and the control facilities' history of violation could lead to biased estimates of our causal effect. The results are presented in table 5. Column 2 shows that treated facilities with no history of HPV increase their TRI air emissions by over 44 percent when they have a sister-plant within the same industry under violation. This result is statistically significant at the 1 percent level. Column 1 reports the magnitude of HPV induced intra-firm pollution substitution for treated facilities with some history of violation. Although the sign of the SATT estimate is positive, we see that it is not statistically significant. Our results suggest that the effect of the treatment on facilities with no history of HPV is far greater than those with some history of HPV.

7. Conclusion and Discussion

The existence of spillovers makes it challenging for us to estimate the overall welfare gain due to environmental regulations. The magnitude and the sign of the spillovers can cause the estimates of the net welfare gain to be biased in either direction. Thus, from a policy perspective, it becomes important to empirically examine the sign and the magnitude of such spillovers. This, in turn, helps us better access the overall desirability of environmental regulations from a welfare standpoint.

Using the annual TRI emissions data and the High Priority Violations data for all polluting industries across the U.S., we empirically examine the existence of intra-firm pollution substitution, which is a potential source of negative spillover in the HPVP under the Clean Air Act. Our results show that compliant facilities tend to increase total emissions when at least one other facility owned by the same parent firm enters HPV. Our results are statistically significant at the 1 percent confidence level and provide strong evidence supporting the intra-firm pollution substitution hypothesis. These results call for a consideration of pollution leakages in policy design, and the negative welfare impact associated with such pollution substitution presents an interesting topic for future research in pollution control.

In our ongoing work, we are exploring the heterogeneity in the magnitude of intra-firm pollution substitution based on various factors such as the number of violating and compliant facilities owned by the parent firm, the duration of the violations, and whether the parent firm is a publicly traded company. In addition, we also hope to determine the probability that a compliant facility enters violation as a result of intra-firm pollution substitution between violating and compliant facilities. A significant result, in that case, would allow us to directly link regulation induced intra-firm pollution substitution with increase in public health risks.

References

Abadie, A., Imbens, G., (2011), "Bias-Corrected Matching Estimators for Average Treatment Effects", Journal of Business & Economic Statistics, v.29, no.1

Carrada-Bravo (1995) "Trade, Foreign Direct Investment, and Environmental Policy: The Case of Mexico." North American Journal of Economics and Finance 6 (2) pg. 203-210.

Condliffe, S. and Morgan, O.A. (2009) "The Effects of Air Quality Regulation Enforcement and the Firm Location Decision among U.S. Counties." Journal of Regulatory Economics 36 (1) pg. 83-93

Davis, L. W., Kahn, M. E. (2010) "International Trade in Used Vehicles: The Environmental Consequences of NAFTA." American Economic Journal: Economic Policy 2 (4), pg. 54-82

Dean M. J., Lovely E. M., Wang, H. (2009) "Are Foreign Investors Attracted to Weak Environmental Regulations? Evaluating the evidence from China." Journal of Development Economics 90, pg. 1-13.

Dechezlepretre, A., Sato, M. (2014) "The Impact of Environmental Regulations on Competiveness." Granthan Research Institute and Global Green Growth Institute Policy Brief

Dyck, A., Zingales, L., (2002) "The Corporate Governance Role of the Media" Harvard Business School

Environmental Protection Agency (1998), "The Timely and Appropriate (T&A) Enforcement Response to High Priority Violations (HPVs)", EPA Workbook

Environmental Protection Agency (2011), "The Benefits and Costs of the Clean Air Act from 1990 to 2020 Final Report, The Environmental Protection Agency.

Environmental Protection Agency (2014), "The Timely and Appropriate (T&A) Enforcement Response to High Priority Violations (HPVs) – Revision 2014", Memorandum and Workbook

Evans, M., (2016) "The Clean Air Act Watch List: An enforcement and Compliance Natural Experiment.", JAERE, v.3, no. 3, pp-627-65.

Fowlie, M. (2009) "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage." American Economic Journal: Economic Policy 1 (2), pg. 72-112

Gibson, M. (2015) "Regulation Induced Pollution Substitution", Working Paper, University of California, San Diego

Greenstone, M., (2003), "Estimating Regulation-Induced Substitution: The Effect of Clean Air Act on Water and Ground Pollution." American Economic Review: AEA Papers & Proceedings 93 (2) pg. 442-49

Hanna, R. (2010) "US Environmental Regulation and FDI: Evidence from a Panel of US-based Multinational Firms." American Economic Journal: Apllied Economics 2 (3), pg. 158-189

Henderson, J. V. (1996) "Effects of Air Quality Regulation." American Economic Review 86 (4), pg. 789-813

Jiang, L., Ma, S., Zhu, Y. (2015) "The Effect of Environmental Regulation on Industrial Pollution and Industrial Activities: Evidence from China." Working Paper

Kitzmueller, M., Shimshack, J. (2010) "Economic Perspectives on Corporate Social Responsibility" University of Michigan Ann Arbor

Levinson, A., Taylor, M. S. (2008) Unmasking the Pollution Haven Effect." International Economic Review 49 (1), pg. 223-254

Liu, L. (2012) "Spillover Effects Across Environmental Programs: The Case of Hazardous Waste Regulation in Michigan" Environmental Economics 2 (3)

Figures

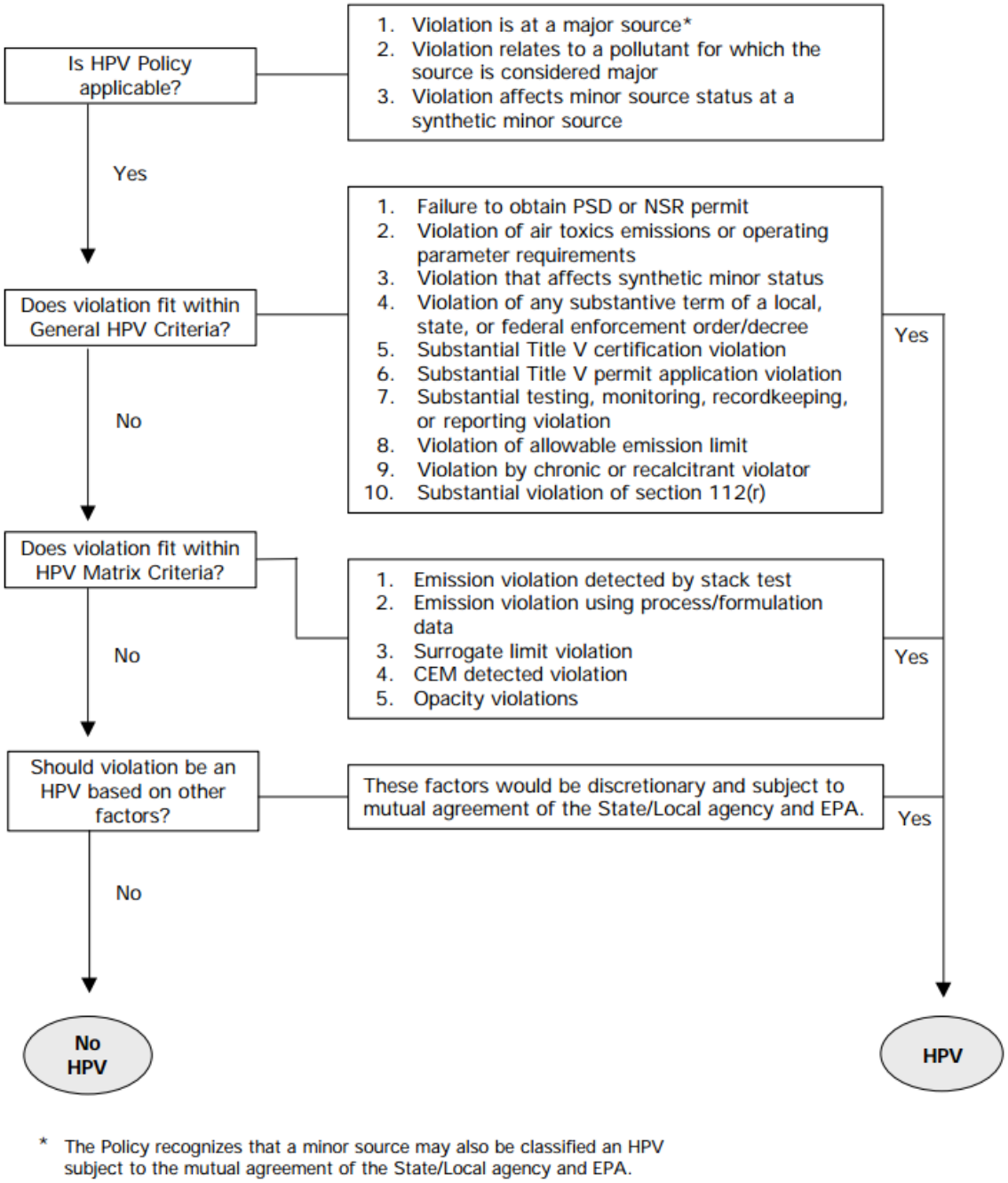

Figure 1: Major and synthetic minor sources are stationary sources of air pollution. A major source is one that either has the potential to emit more than 10 tons of a certain HAP per year or more than 25 tons of total HAPs per year. A synthetic minor source has the potential to emit levels of HAPs equal or greater to the major source thereshold levels, but has accepted federlly enforceable limitations to keep emissions below the threshold level.

Figure 2: Above – the distribution of total air emissions (lbs.) before trimming observations with total air emissions > 100,000 lbs. and <10 lbs. Below - the distribution of total air emissions (lbs.) after trimming.

Tables

Table 1 – Summary statistics

	Control			Treated			
	Obs.	Mean	SD	Obs.	Mean	SD	Nor. Dif.
Total air emission (lbs.)	34038	10289	18278	3126	14340	20879	0.206
P2 participation [0 or 1]	34038	0.16	0.37	3126	0.12	0.32	-0.139
Non-attainment Status [0 or 1]	34038	0.39	0.49	3126	0.42	0.49	0.051
Plant age (yrs.)	34038	23.53	6.45	3126	23.39	5.97	-0.023
No. of additional com. plants	34038	3.01	7.41	3126	6.92	10.87	0.420
Total firm TRI (1000 lbs.)	34038	358.91	6821.34	3126	344.94	1574.18	-0.003
Plant prior yr. TRI (1000 lbs.)	33589	106.99	4158.61	3096	92.00	660.74	-0.005
Plant current TRI (1000 lbs.)	33589	107.99	4092.35	3096	80.71	526.87	-0.009
Plant next yr. TRI (1000 lbs.)	33589	108.95	4374.49	3096	74.22	446.06	-0.011
Exp. firm TRI growth (%)	26132	0.55	3.26	2295	-0.02	0.44	-0.245
Exp. plant growth (%)	25718	0.184	7.006	2253	0.019	0.717	-0.033
1-yr lag total air emission (lbs.)	27078	10325	17911	2590	14226	20158	0.205

Table 2 – Matched observations

	Control			Treated			
	Obs.	Mean	SD	Obs.	Mean	SD	Nor. Dif.
Plant age (yrs.)	2037	24.05	5.889	2037	23.65	6.05	-0.067
Firm TRI growth (%)	2037	0.455	6.368	2037	9.6	396.992	0.033
Plant TRI growth (%)	2037	0.947	8.693	2037	6.451	207.248	0.038
Exp. firm TRI growth (%)	2037	-0.024	0.229	2037	-0.022	0.4	0.006
Exp. plant growth (%)	2037	-0.013	0.46	2037	0.017	0.683	0.052
1-yr lag total air emission (lbs.)	2037	10316	16448	2037	13944	20237	0.197

Table 3A – Sample average treatment effect on the treated estimates

	Dep. Var. ln(total air emissions)		
	(1)	(2)	(3)
SATT	1.041*** (0.062)	0.377*** (0.057)	0.352*** (0.721)
<u>Exact matching vars.</u>			
County	Y	N	N
State	-	Y	Y
County non-attain	-	Y	Y
Reporting Year	Y	Y	Y
6-digit NAICS	Y	Y	Y
P2 participation	Y	Y	Y
<u>NN matching vars.</u>			
Lag total air	Y	Y	Y
Plant TRI growth rate	Y	Y	Y
Firm TRI growth rate	Y	Y	Y
Plant age	N	N	Y
Exp. firm TRI growth	N	N	Y
Exp. plant TRI growth	N	N	Y
Number of exact matches	733	2037	2037

Standard errors re in parentheses. All SE are robust i.e. we allow for heteroskedastic SE. The unit of observation is plant-year. All SATT estimates are bias-adjusted for imperfect matches on NNM vars.

Table 3B - Sample average treatment effect on the treated estimates

	Dep. Var. Total air emissions (lbs.)		
	(1)	(2)	(3)
SATT	3739.34*** (380.97)	1713.21*** (375.35)	1705.13*** (378.04)
<u>Exact matching vars.</u>			
County	Y	N	N
State	-	Y	Y
County non-attain	-	Y	Y
Reporting Year	Y	Y	Y
6-digit NAICS	Y	Y	Y
P2 participation	Y	Y	Y
<u>NN matching vars.</u>			
Lag total air	Y	Y	Y
Plant TRI growth rate	Y	Y	Y
Firm TRI growth rate	Y	Y	Y
Plant age	N	N	Y
Exp. firm TRI growth	N	N	Y
Exp. plant TRI growth	N	N	Y
Number of exact matches	733	2037	2037

Standard errors are in parentheses. All SE are robust i.e. we allow for heteroskedastic SE. The unit of observation is plant-year. All SATT estimates are bias-adjusted for imperfect matches on NNM vars.

Table 4 – Heterogeniety in the SATT estimates across top 6 NAICS 3-digit industries by number of matches

	Dep. Var. ln(total air emissions)	
3-digit NAICS	SATT	No. of matches
Wholesalers Nondurable goods (424)	0.070 (0.177)	357
Chemical Manufacturing (325)	0.596*** (0.133)	266
Fabricated Metal Products (332)	0.978*** (0.247)	152
Primary Metals Manufacturing (331)	0.615** (0.250)	190
Utilities (221)	0.857 (0.947)	84
Food Manufacturing (311)	0.479 (0.458)	68

Standard errors are in parentheses. All SE are robust i.e. we allow for heteroskedastic SE. The unit of observation is plant-year. All SATT estimates are bias-adjusted for imperfect matches on NNM vars. The exact matching vars. are year, state, 6-digit NAICS, p2 and non-attain. NN matching vars. are lag total air, plant age, plant TRI growth, firm TRI growth, exp. firm TRI growth and exp. plant TRI growth.

Table 5 – Heterogeneity in SATT estimates across plants with HPV history vs. no HPV history

	Dep. Var. ln(total air emissions)	
	History=1	History=0
SATT	0.256 (0.347)	0.446*** (0.073)
<u>Exact matching vars.</u>		
State	Y	Y
County non-attain	Y	Y
Reporting Year	Y	Y
6-digit NAICS	Y	Y
P2 participation	Y	Y
<u>NN matching vars.</u>		
Lag total air	Y	Y
Plant TRI growth rate	Y	Y
Firm TRI growth rate	Y	Y
Plant age	Y	Y
Exp. firm TRI growth	Y	Y
Exp. plant TRI growth	Y	Y
Number of exact matches	271	1573

Standard errors are in parentheses. All SE are robust i.e. we allow for heteroskedastic SE. The unit of observation is plant-year. All SATT estimates are bias-adjusted for imperfect matches on NNM vars.