Third Party Certification and the Effectiveness of Voluntary Pollution Abatement Programs: Evidence from Responsible Care

Martina Vidovic
Department of Economics
Rollins College
Email: mvidovic@rollins.edu

Neha Khanna
Department of Economics and Environmental Studies Program
Binghamton University
P.O. Box 6000
Binghamton, NY 13902-6000
Email: nkhanna@binghamton.edu

Michael Delgado
Department of Agricultural Economics
Purdue University
Email: delgado2@purdue.edu

Preliminary Draft
December 27, 2013

Key words: self-regulation, certification, difference-in-difference, semi-parametric model, heterogeneity, American Chemistry Council

JEL codes: Q53, Q58, L60

Acknowledgments: We thank Andrew King for providing information on membership in the Chemical Manufacturers' Association from 1988-2001.

1. Introduction

Industry self-regulation via voluntary pollution abatement has become popular not only with industry groups but also with environmental policy-makers because it gives them a relatively easy to use lever that does not require an act of Congress. There is a substantial academic debate on the effectiveness of such programs, with some authors arguing that these programs are quite effective in reducing pollution (for example, Khanna and Damon 1999 and Bi and Khanna 2012) while others argue, and with equal conviction, that these programs are ineffective at best (Gamper-Rabindran 2006, Vidovic and Khanna 2007, 2012, Carrión-Flores et al. 2013) and counter-productive at worst (King and Lennox 2000, Gamper-Rabindran and Finger 2013). The Achilles heel of this debate as it relates to the United States is that it relies on relatively old data from the 1990s and on programs that are either no longer in existence (for example, the 33/50 Program which ended in 1995) or on early versions of programs that were changed substantively in later years (for example, the American Chemistry Council's Responsible Care, RC, program has been analyzed only through 2001, after which major structural changes were incorporated).

We update the literature on the effectiveness of voluntary pollution abatement in the United States. We use the structural changes in the RC program to ask whether the introduction of independent third party certification from 2005 onwards has yielded lower emissions from RC plants compared to statistically equivalent non-RC plants in the US chemical industry. Our identification strategy relies on the fact that independent third party certification was made mandatory from 2005 onwards. We use a standard difference-in-difference approach to estimate the average treatment effect of third party certification by comparing RC plants before and after third party certification was introduced to other plants in the US chemical industry between 1995 and 2010 who were not remembers of RC and were therefore not subject to the requirement of third party certification.

Similar to Gamper-Rabindran and Finger (2013), we address firms' self-selection into RC and instrument for the RC participation status of a plant's parent firm using attributes of other plants belonging to the same firm which are hypothesized to influence a parent firm's decision to join the RC but do not directly affect pollution at a given plant.

We also explore plant-level heterogeneity in the treatment effect using a semi-parametric model. The advantage of this model is that we can identify heterogeneity in the effect of treatment across plants without imposing *a priori* an ad hoc parametric specification of heterogeneity (see Gamper-Rabindran and Finger 2013 for a parametric example), and so serves as a robustness check of our parametric difference-in-difference model. In addition, we use the semi-parametric model to determine if there are certain types of plants in the chemical industry for which third party certification has been more or less effective, or to the degree to which third party certification was effective across plants.

It is worth mentioning that we have carefully constructed our dataset to mitigate potential problems associated with missing data on facility membership in the RC for three years as well as potential biases in the effect of treatment on emissions that could result from facilities not reporting emissions, being traded between firms, or entering and exiting the program just before or shortly after the treatment takes place.

Preliminary results reveal a statistically insignificant albeit negative average treatment effect. That is, the introduction of third party certification did not lead to a decline in emissions from RC plants compared to other chemical plants that were not a part of RC. Confirming Gamper-Rabindran and Finger's (2013) result for the early years of RC we find the emissions from RC plants are always statistically higher than emissions from non-RC plants. At the same time, while the emissions from both RC and non-RC plants are declining over the time period we study, the introduction of third party certification in 2005 did not result in a statistical change in the decline rate of emissions from RC plants compared to non-RC plants.

2. Self-regulation and third party certification

Over the past three decades, voluntary approaches to environmental management have become equally popular among environmental policymakers, industry groups and non-governmental organizations. The U.S. E.P.A's Partnership Programs website alone lists over 40 programs with more than 13,000 participants (http://www.epa.gov/partners/programs/index.htm). The growing reliance on self-regulatory approaches to environmental protection begs the question whether voluntary programs are able to elicit meaningful changes in environmental performance and whether the signals they send accurately reflect the behavior of their participants. Prior research evaluating the effectiveness of voluntary pollution abatement programs found that participation in such programs was either not associated with promoting superior environmental performance among their participants (Rivera, de Leon and Koerber 2006; Gamper-Rabindran 2006; Vidovic and Khanna 2007, 2012) or has actually led to worse environmental outcomes (King and Lennox 2000, Gamper-Rabindran and Finger 2013). On the other hand, Khanna and Damon (1999), Innes and Sam (2008), Sam et al. (2009), Bui and Kapon (2012) and Bi and Khanna (2012) argue that such programs are quite effective in reducing pollution. Some authors have begun to caution that program design characteristics and lack of performance requirements may be responsible for the failure of voluntary approaches to make a difference (Darnall and Carmin 2005; Potoski and Prakash 2005; Rivera, deLeon and Koerber

2006). Weak performance standards and the absence of effective enforcement can permit firms to free ride and continue to serve their own interests at the expense of other participants and the consumers.

The evidence regarding the ineffectiveness of US voluntary programs in achieving environmental protection primarily rests on the evaluations of the 33/50 Program (Gamper-Rabindran 2006, Vidovic and Khanna 2007, 2012), the Sustainable Slopes program (Rivera, de Leon and Koerber 2006) and the early years of the RC program (King and Lennox 2000, Gamper-Rabindran and Finger 2013); all are programs that relied on self-monitoring and assurance from participants that they adhered to the program requirements. It is not clear whether the participants failed to adopt superior environmental protection practices or the program failed to elicit improvement among the participants. At least in the case of RC, King and Lenox (2000) argue that voluntary programs designed by industry associations lack appropriate implementation, monitoring, and reporting procedures that would initiate superior environmental performance by participants.

Among the voluntary programs that award a label or recognition if certain standards are met, third-party oversight has emerged as a way of providing credibility to the certification system. For example, to ensure integrity and sustainability, the EPA integrated third party verification in its Water Sense and the Energy Star programs. The forest product label from the Forest Stewardship Council and the sustainable seafood label from the Marine Stewardship Council use third party verification to award certification to sustainable management of forests and fisheries. Similarly, third party audits of the ISO 9000 Quality Management System Standard and ISO 14001 Environmental Management System Standard were instituted by the International Organization for Standardization. Recently, the American Chemistry Council

(ACC) incorporated third party certification in its signature RC program.

The studies that analyze whether third party certification improves environmental performance via voluntary approaches are mainly focused on one program, ISO 14001 certification system. Several early studies found that ISO 14001 certified firms reduced waste and use of resources significantly more than non-registrants (Rao and Hammer 1999; Montabon et al. 2000; Melnyk et al. 2002). Unfortunately these studies suffer from some methodological and sample issues and the results should be interpreted with caution.[1] King et al. (2005) found a weak negative effect of ISO 14001 on emissions improvement; certification provides stakeholders mainly with information about the ongoing efforts to improve the performance of an environmental management system but it is not correlated with reductions in emissions. Russo (2009) found that being an early adopter is associated with lower emissions and that emissions fall the longer a facility operates under ISO 14001certification. The two most systematic studies that compare the environmental performance of adopters and non-adopters over time are Potoski and Prakash (2005) and Toffel (2006). Both studies, using different methodologies for comparing adopters to non-adopters, found that ISO 14001 certified facilities reduced their pollution emissions more than non-certified facilities. Based on their findings, the authors suggest that programs whose enforcement mechanisms are based on third-party audits could potentially improve compliance with underlying program commitments even in absence of public disclosure of the audit information.

We add to the existing literature on the effectiveness of voluntary management programs

[1] For example, the data used by Rao and Hamner (1999) is based on information collected from a questionnaire administered to ISO 14001 registrants and there is no information on non-registrants in their dataset. In Montabon *et al.* (2000) and Melnyk *et al.* (2002) both independent and dependent variables are constructed from answers to a survey where the respondents were likely the same people who made decisions regarding their firm's participation in the ISO 14001 and provided opinions on its impact on the firm's performance.

by examining whether the introduction of independent third party certification from 2005 onwards yielded lower emissions from RC plants compared to statistically equivalent non-RC plants in the US chemical industry. The advantage of studying the RC program in this context is that the mandatory certification under RC was modeled on the certification under ISO 14001. Our analysis sheds light on whether third party oversight of voluntary abatement programs makes them a more effective instrument in the US policymaker's environmental toolbox.

3. The potential of Responsible Care to improve environmental outcomes

In 1988, the ACC (then known as the Chemical Manufacturers' Association) adopted the RC initiative to promote continuous Environmental, Health, Safety and Security (EHS&S) performance improvement for all of its members. The industry association implemented the program in order to improve public perception about the safety of the chemical industry and in anticipation of more stringent regulatory interventions following the chemical disaster at the Union Carbide plant in Bhopal, India, and the subsequent leak from the Union Carbide's pesticide plant in Institute, West Virginia, in mid 1980s. Participation in Responsible Care was made a condition for membership in the ACC.

Throughout the 1980s and 1990s, the program was structured around a set of codes of EHS&S management practices. In 1996 a voluntary peer-review process called Management System Verification was added to the program. The process served to verify that appropriate systems were implemented to assure ongoing compliance with company's EHS&S performance goals and external regulations. The system was not an audit of a company and did not identify non-compliance with regulations or the level of emissions at a facility.

In 2002 the ACC announced substantial changes to the Responsible Care program recognizing that US regulation of the chemical industry had caught up with RC requirements. In that year 75 percent of the original RC activities were covered by government laws and regulations compared to only 13 percent in 1988 (Phillips 2006). Stakeholders lost support for the program and the companies begun to differentiate themselves from RC because once the program practices were achieved, there was no room to advance performance. As part of its change, the program implemented the Responsible Care Management System (RCMS), a management system approach built on the basic "Plan-Do-Check-Act" philosophy to improve company performance in the key areas: community awareness and emergency response; security; distribution; employee health and safety; pollution prevention; and process and product safety (ACC 2013). To enhance transparency, it adopted a mandatory independent third-party certification of those management systems. Under independent oversight, every Responsible Care company must certify that it has a management system in place and demonstrate progress toward improved performance. To obtain certification, companies must undergo headquarter and facility audits conducted by independent, accredited auditing firms (ACC 2013). The third party certification system was officially launched in 2005 and all members were required to complete third party audits by the end of 2007.

The ACC requires that certification is renewed every three years, and companies can choose to demonstrate conformance either to the RCMS or the RC14001 technical specification which combines Responsible Care and ISO 14001 certification. Recognition and popularity of ISO 14001with stakeholders worldwide prompted companies to seek an approach that would avoid duplicating the RC and ISO 14001 audit processes. The RC14001 technical specification integrates elements of both the RC requirement for third party certification and ISO 14001 allowing a single certification process to fulfill both program requirements (Phillips 2006). In order to obtain the RC14001, organization must conform to the ISO 14001 with respect to environment as well as to health, safety and security requirements within the scope of Responsible Care.

Unlike performance standards which set the level of environmental protection and state requirements for improved environmental performance, certified management standards such as the RC14001 only require firms to establish processes and management systems to ensure that environmental goals are developed, assessed and met. However, certification may still provide information on performance improvement to stakeholders by conveying that an environmental management system exists and whether it leads to improvement. Voluntary programs without third party oversight usually suffer from potential shirking on the part of some participants. Some firms join the program in order to reap the benefits of membership but fail to adhere to the program commitments. If certification is costly and stakeholders are willing to pay more for superior performance, certification may act as a credible signal of superior performance. According to the ACC, the third-party auditing system is part of the association's drive to increase credibility and public confidence in the RC program, since, in the past, RC signatories conducted self-assessment tests to judge their progress to full compliance. Although the ACC always mandated that all firms must adopt RC or they will lose their membership, critics questioned the credibility of the expulsion threat because ACC membership is voluntary and ACC has never expelled a member for non-compliance (Prakash 2000). The certification system is likely to formalize managerial commitment to achieving environmental performance goals (Rondinelli and Vastag 2000), provide accountability and reduce opportunities for participants to behave opportunistically (King and Lenox 2000). In addition, ACC requires public disclosure of environmental information by all ACC members on the ACC website and with governmental agencies. Therefore, we anticipate that following the implementation of third party certification of Responsible Care activities, RC participants improved their environmental performance compared to non-participants in the US chemical industry.

4. Methodology and hypotheses tested

To evaluate whether the adoption of third party certification in 2005 lead to improved environmental performance, we seek the average treatment effect of third party certification for RC plants. The control group consists of other plants in the US chemical industry between 1995 and 2010 that were not members of RC and were therefore not subject to the requirement of third party certification.

There is a well-developed econometric literature advocating the use of treatment effect estimators for program evaluation as such estimators allow for simultaneous control for both observable and unobservable confounding factors. This literature has emphasized particular caution with regard to unobservable variables and the potential for econometric bias in the treatment parameters if such factors are not carefully controlled. Another advantage of the average treatment effect model is that the average treatment effect is the expected causal effect of treatment for all observations, not only for the treated observations.[2] See, for example, Wooldridge (2010) for a thorough review of standard treatment effect models.

Formally, the standard version of the average treatment effect can be defined nonparametrically as follows. For a typical repeated cross-sectional panel dataset, let Y_{it} denote the outcome of observation $i = 1,2,\ldots,N$ in time $t = 1,2,\ldots,T$, W_{it} indicate treatment status for i in t, and let (X_{it}, c_i, τ_t) denote a set of time-varying control variables and fixed effects. Then,

[2] In the simple model, the average treatment effect is identical to the average effect of treatment on the treated population.

the average treatment effect is defined as

$$(1)\ ATE = E[Y_{it}^1 - Y_{it}^0 | X_{it}, c_i, \tau_t] = E[Y_{it}^1 | X_{it}, c_i, \tau_t] - E[Y_{it}^0 | X_{it}, c_i, \tau_t]$$

in which Y_{it}^1 and Y_{it}^0 denote treated and untreated outcomes. The average treatment effect is therefore, conditional on (X_{it}, c_i, τ_t), the difference in expected outcome for the treated and untreated samples. The complexity in identifying the average treatment effect is that Y_{it}^1 and Y_{it}^0 are generally not observed simultaneously – we either observe the outcome for a treated observation or an untreated observation, but never both outcomes for a single observation. Further, this definition of the average treatment effect relies on the selection on observables assumption that treatment is exogenous to the outcome, at least conditional on (X_{it}, c_i, τ_t).

In the case of a standard repeated cross-sectional dataset, a simple linear in parameters regression setup that includes W_{it} as a separate regressor is sufficient for identifying the average treatment effect as the coefficient on W_{it}. Hence, to estimate the effect of RC certification on emissions, we estimate the following equation

$$(2)\ Y_{it} = \alpha + \delta W_{it} + X_{it}\beta + c_i + \tau_t + u_{it}.$$

In the context of RC, Y_{it} is the level of total TRI emissions to the air for facility *i* at time *t*, W_{it} is the post certification dummy equal to 1 for all RC members in years 2005-2010, τ_t represents year fixed effects, c_i captures the facility fixed effects, and u_{it} is the idiosyncratic error term. Given our definition of the average treatment effect, it is simple to verify that the linear in parameters specification above identifies the average treatment effect through δ. Following Bertrand et al. (2004), we use a bootstrap to consistently estimate standard errors for our coefficient estimates.

In our specification, year fixed effects control for changes in regulations and available technologies over time, as well as any general trends in emissions, such as gradual reductions in emissions over time, that should not be erroneously attributed to third party verification (see, for example, Vidovic and Khanna 2007). Facility fixed effects control for differences among facilities that are constant over time. X_{it} is a vector of other covariates hypothesized to affect a facility's emissions: facility to parent firm TRI release ratio, parent firm TRI releases, HAP-TRI release ratio, number of inspections under the Clean Air Act and the number of gases for which the county where a facility is located has been out of attainment with the National Ambient Air Quality Standards (NAAQS). Controlling for observable differences between facilities in the treatment and the control groups improves the efficiency of the treatment effect estimator (Meyer 1995). Moreover, from a practical perspective, inclusion of these controls and fixed effects ensures that we do not omit any confounding factors that may contaminate our estimate of the causal effect of third party verification.

This difference-in-differences approach allows us to study the effect of treatment, in this case third party certification, by comparing the performance of the treatment group pre- and post-treatment relative to performance of the control group pre- and post-treatment. It is through the use of the control group that we are able to control for any unobservable factors that are common to all facilities in our sample. Our main hypothesis is that following the introduction of the third party certification in 2005, RC member facilities lowered their emissions of the TRI air releases by more than the control group of facilities. To the extent that facilities may wish to mitigate the cost and stringency of current and or future mandatory regulation, we anticipate that facilities with greater HAP to TRI release ratio, which captures the exposure of facilities to regulation of HAPs, facilities located in counties classified as being out of attainment with the NAAQS, and facilities with a larger number of government inspections under the Clean Air Act will face an additional incentive to reduce their TRI emissions by participating in RC and have their management system certified by a third party.

One potential shortcoming of the linear difference-in-differences specification is that this specification restricts the effect of treatment to be homogeneous across observations or groups (Meyer 1995). Further, the linear specification generally ignores any potential interactive effects between any of the variables (including treatment) in the model. In light of many differences across facilities in our sample (e.g., HAP/TRI, or NAAQS attainment status) these restrictions need not necessarily be justified. Following Wooldridge (2010), we can write $Y_{it} = Y_{it}^0 + (Y_{it}^1 - Y_{it}^0)W_{it}$ and identify the average treatment effect nonparametrically via

$$(3)\ E[Y_{it}|W_{it}, X_{it}, c_i, \tau_t] =$$
$$E[Y_{it}^0|X_{it}, c_i, \tau_t] + \left(E[Y_{it}^1|X_{it}, c_i, \tau_t] - E[Y_{it}^0|X_{it}, c_i, \tau_t]\right)W_{it}$$

in which the last term in parentheses is the average treatment effect by (1). The semiparametric generalization of (2) that yields the conditional mean in (3) is

$$(4)\ Y_{it} = \beta(X_{it}, c_i, \tau_t) + \delta(X_{it}, c_i, \tau_t)W_{it} + u_{it}$$

in which $\beta(X_{it}, c_i, \tau_t)$ and $\delta(X_{it}, c_i, \tau_t)$ are generally unknown, smooth nonparametric functions of the control variables and fixed effects. As in (2), $\delta(X_{it}, c_i, \lambda_t)$ is the effect of third party verification on facility emissions. The advantage of (4) relative to (2) is that the model no longer restricts the treatment effect to be homogeneous across observations. Rather, $\delta(X_{it}, c_i, \tau_t)$ is fully general in that the effect of treatment is allowed to vary across both facility observations and time without general restriction. This specification is convenient because in the absence of treatment (4) collapses into a fully nonparametric function of (X_{it}, c_i, τ_t) that is free from any functional specification biases; hence, $\delta(X_{it}, c_i, \tau_t)$ can be thought of as an adjustment function following treatment. Because we allow for arbitrary forms of heterogeneity in (4), we require a nonparametric estimator to obtain consistent estimates of $\beta(X_{it}, c_i, \tau_t)$ and $\delta(X_{it}, c_i, \tau_t)$.[3]

While we prefer the straightforward treatment effect specification in (2) as our primary specification, we deploy the generalized specification in (4) as a secondary specification through which to assess the robustness of our parametric estimates, and as a lens through which to analyze heterogeneity in treatment by systematically examining variation in our estimate of $\delta(X_{it}, c_i, \tau_t)$. To assess the robustness of our parametric specification to arbitrary forms of parametric model misspecification, we consider the average of our nonparametric estimate of the treatment effect: $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\delta}(X_{it}, c_i, \tau_t)$. The advantage of this average is that we have a single treatment effect estimate analogous to our single estimate of δ from (2), yet free from any parametric model misspecification biases that may be erroneously imposed in the linear setup. For instance, substantial differences in our nonparametric average from $\hat{\delta}$ obtained from (2) provides potential insight that our parametric specification may be overly restrictive.

In order to further check the robustness of our results, we consider that firms may have self-selected into the RC program. Since the ACC never expelled a member if it failed to comply with the RC, it is possible that some firms joined the program in order to gain recognition and improve their public image but did nothing or very little to comply with the program's commitments. Since self-selection is likely to be based on factors that are unobserved to researchers but correlated with the program outcomes, the effect of treatment – introduction of third party certification -- on facility emissions may be biased either towards or against finding that third party certification reduces pollution. Similar to Gamper-Rabindran and Finger (2013)

---

[3] A standard nonparametric kernel estimation approach for the model given in (4) is provided by Li et al. (2002). Specifically, Li et al. (2002) propose the weighted least squares estimator

$$\hat{\gamma} = \left[W'K_h(\tilde{X})W\right]^{-1}W'K_h(\tilde{X})Y$$

in which $\hat{\gamma} = [\hat{\beta}, \hat{\delta}]$, W (and Y) is a matrix (vector) containing W_{it} (and Y_{it}), $\tilde{X}$ is a matrix with the it^{th} row being $\tilde{X}_{it} = (X_{it}, c_i, \tau_t)$, and $K_h(\tilde{X})$ being a product kernel function with bandwidth h that can be chosen via cross-validation. Given that $\tilde{X}$ contains both continuous and discrete regressors, we adopt the generalized product kernel function of Racine and Li (2004). See Li and Racine (2007) for further technical details.

we estimate the treatment effect of third party certification of a facility's emissions conditional on parent firm's self-selection into the RC.

A facility's expected net benefit from participation in the RC, Z_{1it}^{*}, is given by

$$(5)\ Z_{it}^{*} = R_{it}\gamma + \varepsilon_{1it,}$$

where R_{it} is a vector of covariates, γ is a vector of coefficients, and $\varepsilon_{it} \sim N(0,1)$. We do not observe Z_{it}^{*}, only whether the facility participated or not. That is, we observe Z_{it}, a dichotomous variable equal to 1 if the expected net benefit is positive and 0 otherwise. We estimate γ using a random effects probit model. In order to correct for the sample selection bias in (2), we construct the inverse Mills ratio[4] ($\hat{\lambda}_{it}$) using the estimates from (5) and using the full sample, we estimate facility emissions as a function of R_{it} and $\hat{\lambda}_{it}$. Since the inverse Mills ratio is an estimate we bootstrap the standard errors in the facility participation equation.

For the model to be identified there should be at least one right hand side variable that appears in the selection equation (5) but does not appear in the outcome equation (2). In our case the selection equation includes parent firm level variables. The decision to participate in the RC was made by the parent firm and it is the same for all of a firm's facilities, while pollution performance is specific to each facility. Following Gamper-Rabindran and Finger (2013) we include firm level variables that exclude the facility in question, which ensures that the instruments are correlated with the likelihood that a plant belongs to a parent firm that is a member of the RC but that do not directly affect a facility's emissions and are therefore exogenous. As instruments we include the number of other facilities reporting under the parent

[4] $\hat{\lambda}_{it} = \frac{\phi(-R_{it}\hat{\gamma})}{1-\Phi(-R_{it}\hat{\gamma})}$ for facilities that participated in the RC and $\hat{\lambda}_{it} = \frac{-\phi(-R_{it}\hat{\gamma})}{\Phi(-R_{it}\hat{\gamma})}$ for facilities that did not participate in the RC. γ is the estimated parameter vector from the probit estimation of the facility participation equation, R_{it} is the facility *i* set of explanatory variables and $\phi(-R_{it}\hat{\gamma})$ and $\Phi(-R_{it}\hat{\gamma})$ are the normal density function and the cumulative distribution function, respectively.

firm and the other plants' TRI air releases[5]. Additional instruments are dummy variables for the four digit NAICS for the major industries represented in our sample (NAICS 3251, 3252, 3253, 3254, 3255, 3256 and 3259) and the dummy variable that is equal to one if the parent firm is publicly owned.

5. Data description and sources

Our data consist of chemical manufacturing facilities located in the United States that have reported emissions of toxic chemicals to the EPA's Toxic Releases Inventory (TRI). We restrict our sample to facilities that report SIC 28 and/or NAICS 325 as their primary industry, representing the largest single share of the facility's economic activity[6]. Andrew King provided us with a list of Responsible Care participants from 1988 to 2001: this is the same information utilized in King and Lennox (2000). We obtained the list of current American Chemistry Council participants and the information on their certification status between 2005 and 2010 from the American Chemistry Council website (http://reporting.responsiblecare-us.com/Reports/Members/RCMSC_Cmpny_Rpt.aspx, accessed May 14, 2012).

The commitment to RC is reported at the firm level and we assume that all facilities belonging to a participating parent firm participated in the program. We have information on the RC status for each firm in each year between 1988 and 2001. We also have information on whether firms undertook third party certification during the period 2005-2010, and we only

[5] We have also considered the following instruments: the number of government enforcement actions under the Clean Air Act, a dummy variable for whether a facility participated in the 33/50 program, a dummy variable for whether a facility is listed by the Resource Conservation Recovery Act (RCRA) as a large quantity generator, the number of times the facility was inspected for violations of hazardous waste regulations under the RCRA, the number of times the EPA brought an enforcement action against a facility for violations of hazardous waste regulations under the RCRA and county level socio-economic variables. However, none of these instruments were statistically significant and were therefore excluded from the final model.

[6] NAICS were adopted starting with the 2006 reporting year for use within TRI instead of SIC; submissions from previous years of TRI reporting were also assigned NAICS codes based on their 2006 reporting, if any, and on their SIC codes.

count firms and their plants as participants in RC if they have obtained third party certification at the headquarters and at a sample of facilities during the periods 2005-2007 and 2008-2010. However, we do not have data on RC participation for the intervening years, i.e. 2002, 2003, 2004, and we assume that firms that were members in both 2001 and in 2005 remained members through the three years for which we have missing membership information.[7]

Since RC and non-RC facilities could be systematically different, for identification purposes we classify facilities strictly either as RC-members or as non RC-members during the period of analysis, 1995-2010. That is, to avoid contamination of our treatment and control groups we only consider facilities that do not switch between these two groups. For example, if a facility belonged to an RC member firm between 1995 and 1999 and then it was traded to a non-RC firm in 2000, this facility is excluded from our dataset. However, if a participating facility was traded in any year to another parent firm that was also a member of RC, we continue to count this facility as a member of RC. The same is true for a facility that was not a member of RC. We also require that each facility, starts reporting to the TRI by 2003 and that once it reports, it continues reporting until the end. .

We obtain data on emissions of the total TRI air releases, HAP air releases, names of parent firms, and facility names and locations from the TRI (www.rtknet.org/new/tri). Information on the number of inspections under the CAA is from the Integrated Data for Enforcement Analysis database (www.epa-echo.gov/echo/index.html); county nonattainment status with the CAA is from the EPA's Green Book (www.epa.gov/oar/oaqps/greenbk).

We define facility emissions of the HAP and TRI chemicals as annual releases to air. We use air emissions of the 1995 core chemicals which have been reported to the TRI throughout our

[7] Based on the historical data from 1988-2001 we find that firms tend to maintain continuous membership till they choose to opt-out of the program.

period of analysis. Firm emissions are the sum of emissions for all facilities reporting to each parent company in each year.

County non-attainment status is the count of pollutants for which a whole or a part of the county has been designated by the EPA to be out of attainment with the NAAQS. The EPA will designate a county to be in nonattainment whenever air pollution levels persistently exceed the NAAQS for six pollutants: ozone, lead, carbon monoxide, sulfur dioxide, nitrogen dioxide and particulate matter. Non-attainment counties are under pressure to reduce emissions and this may provide an additional incentive for facilities located in these counties to lower their air emissions reported to the TRI (see also Vidovic and Khanna 2012, Bi and Khanna 2012, Gamper-Rabindran and Finger 2013).

To construct our sample we first searched the TRI to identify facilities that operate primarily in the chemical manufacturing sector. This resulted in 6,563 facilities in the continental United States. We successfully matched 4,245 facilities to 1,929 parent companies by parent firm name. We further restricted the sample to facilities that belong to multi-plant firms in order to be able to instrument for a facility's participation in the RC with the characteristics of other plants belonging to the same parent firm. Allowing for a one lagged year of data, our analysis uses an unbalanced panel of 935 facilities that belong to 352 parent firms between 1996 and 2010. Out of the 935 facilities, 409 facilities belonging to 102 parent firms were members of RC and 526 facilities belonging to 250 parent firms were not member of RC leading to 12,999 facility year observations over the period 1996-2010.

Table 1 summarizes our data. Comparing facilities in the chemical industry that adopted RC to facilities that did not adopt RC, we find that on average the adopters have higher total TRI air releases, parent firm TRI air releases and number of inspections. On the other hand, the adopters have lower facility to firm TRI air release ratio, HAPs to TRI emissions ratio and were on average located in counties that are less out of attainment with the NAAQS.

6. Main results and discussion

Figure 1 illustrates the preliminary comparison of mean TRI emissions between the treatment and the control groups using a basic linear regression difference-in-difference estimator with no covariates and ignoring the panel nature of our data. That is, before exploring our difference-in-differences specification, we first consider a simple pooled cross-sectional setup that provides basic insight into the general effect of third party verification on emissions. While this simple setup omits potentially important controls, we point out that these results are changes in emissions for treated facilities relative to any changes in emissions experienced by the control group, and as such are more than simple correlations.

The results indicate that although facilities in the treatment group had higher emissions before and after the treatment than the facilities in the control group, the difference between the treatment and control group facilities' emissions decreases by $5.8*10^4$ pounds (average treatment effect) due to third party certification. This result is statistically significant at the 1 percent level. This is shown in Panel A in Figure 1. Panel B shows that the decline rate of emissions from facilities that participated in Responsible Care before and after third party certification is no different than the decline rate of emissions from the facilities that did not participate in Responsible Care. The average treatment effect is 0.076 but it is not statistically significant at conventional levels of significance.

In Table 2 we further examine the effect of third party certification on TRI emissions using a more robust parametric difference-in-difference model that includes both time-varying controls and fixed effects. In Models 1 and 2, the dependent variable is TRI air emissions measured in pounds. In Models 3 and 4, the dependent variable is the log of TRI air emissions. In models 3 and 4 we use the log of parent firm TRI emissions. We add one to the annual sums of emissions before taking the log to accommodate zero values. To minimize the possibility of endogeneity, we lag all time varying variables by one year relative to the year in which a facility's TRI emissions are measured. We estimate all models using robust standard errors, bootstrapped and clustered by facility. In Models 2 and 4 we interact time dummies with the treatment indicator in order to allow the effect of the policy to change over time. These parametric interactions provide preliminary insight into the potential importance of our secondary, nonparametric treatment effect estimator.

The coefficient on the treatment dummy (δ) is negative and statistically significant at the ten percent level in the first model where the dependent variable is TRI emissions in pounds and we do not interact the treatment variable with the year dummies, indicating that on average facilities that were third party certified under RC reduced their emissions of the TRI chemicals compared to facilities that did not participate in RC and were not independently certified. Once we interact the treatment dummy with the year dummies in Model 2, thus allowing the average treatment effect to vary over time, the coefficient on the treatment dummy, which now represents the average treatment effect for 2005, is no longer statistically significant (albeit still negative); nor are the coefficients on the interaction terms between the treatment dummy and the year dummies, except for the final interaction term (treatment*year2010) which is negative and significant at the 10% level. The coefficient on the treatment dummy is not statistically significant in the last two models where the dependent variable is the log of TRI emissions. This suggests that subsequent to certification, the average decline rate of RC facility emissions was no different than the decline rate of emissions from non-RC facilities.

Comparing Models 1 and 3 to Model 2 and 4 we conclude that while the introduction of third party verification did not have a strongly negative average treatment effect between 2005 and 2010, the treatment effect seems to gather some momentum in the later years (2009-2010) compared to 2005.

Both Toffel (2006) and Russo (2009) found that early adopters of ISO 14001 experienced better environmental performance than later adopters. They argued that environmental leaders tend to move quickly when a new opportunity arises that can differentiate them from competitors in terms of environmental performance. Based on their findings we anticipated that RC certification would lead to greater reductions in emissions in the early years of the program. On the contrary, the coefficients on the interaction terms between treatment and year dummies indicate that the benefit of the change in the program structure may have strengthened in later years. Nonetheless, we do not reaffirm Toffel's overall finding that after being certified to a voluntary program by an independent third party (ISO 14001 in his case, RC in our case), adopters reduce TRI emissions more than non-adopters.

In terms of the control variables, we find that larger facilities as measured by the total TRI releases and more polluting facilities within a firm as measured by the facility to parent firm TRI ratio had higher TRI air releases, as well as the change in TRI releases. Our year indicators are negative and highly statistically significant in each of our models. This is interesting and important for several reasons. First, this finding constitutes robust evidence that air emissions were gradually falling over the entire sample period, regardless of RC (and treatment) status. This trend can also be seen in Figure 2 in which we plot average facility emissions over time for treated and untreated facilities separately. There is clearly a downward trend that is exogenous to treatment. This result is not new – Vidovic and Khanna (2007) found a similar trend in emissions reductions for the 33/50 Program that, if not controlled for, confounds the estimate of the program evaluation parameter. The apparent robustness of these results in our regressions underscores the importance of controlling for these factors so as to not bias our estimated effect of third party verification. Second, the fact that our treatment indicator in Model 2 does not remain significant after controlling for year effects (and other observable, unobservable, and interactive effects) suggests that our finding of a negative and significant effect of third party verification is not robust. However, we do find that the effect of treatment is significant in 2010 in Models 1 and 2 and in Models 3 and 4 in 2009 and 2010. In Figure 2, this is reflected a in a parallel slope for both treated and untreated facilities up until 2008 when there is decrease in the slope of average emissions for treated firms, relative to the upward slope of untreated firms.

On the other hand, the coefficients on the number of inspection, HAP-TRI ratio and county non-attainment status are not statistically significant providing no evidence that the anticipation of more stringent mandatory regulation may have had a negative effect on emissions of the TRI chemicals. These results generally suggest that these control variables do not significantly determine facility emissions, given our other findings of significance. Emissions are highly dependent on productivity, which in turn is dependent on both aggregate demand and facility-specific factors that may not be observable. Such fluctuations are captured in our year indicators and facility fixed effects; in other words, given our set of fixed effects and their apparent significant, it is not surprising that we find less significance of our other control variables.

7. Robustness checks and heterogeneity in treatment

As mentioned earlier, we consider the fact that firms may have self-selected in RC and that the treatment effect of third party certification may be correlated with the decision to participate in RC. Table 3 presents the results from the selection equation on facility participation in RC. Among the right hand side variables, we include all of the factors that affect a facility's TRI emissions. Additional variables such as the number of other facilities reporting, other facilities' TRI releases, a dummy variable for whether the parent firm is publicly traded and dummy variables for the four digit NAICS for the most representative industries serve as instruments for selection. The results indicate that the two instruments pertaining to other facilities belonging to the same parent firm as the facility in question are positive and statistically significant at the 1 percent level. Facilities owned by publicly traded companies are more likely to join the RC. All of the other factors (with the exception of the NAICS dummies) do not seem to be important in determining a whether a facility was an RC participant.

Table 4 examines the effect of third party certification on TRI emissions using a difference-in-difference model that includes the inverse Mills ratio to control for self-selection into RC. The inverse Mills ratio was calculated using the estimates from Model 2 in Table 3. We find that sample selection is not a confounding factor – the Inverse Mills Ratio is not statistically significant in any of the four models in Table 4. As a results, all coefficients remain similar in magnitude and statistical significance compared to those in Table 2, including on the treatment variables.

An alternative way to account for the effect of facility self-selection into the RC on the relationship between third party certification and TRI emissions is by applying the instrumental variable approach. The results are shown in Table 5. In Models 1 and 3 we instrument directly for treatment using the lagged releases by other facilities and the number of other facilities reporting, while in Models 2 and 4 we instrument for treatment using the predicted probability of participation from Model 2 in Table 3. The dependent variable in Models 1 and 2 is the aggregate TRI releases while in Models 3 and 4, the dependent variable is the log of TRI releases. We find a statistically significant negative effect of treatment on the aggregate TRI emissions only in Model 1. In all other models the effect is statistically insignificant.

To be filled in once the rest of the nonparametric results come in…

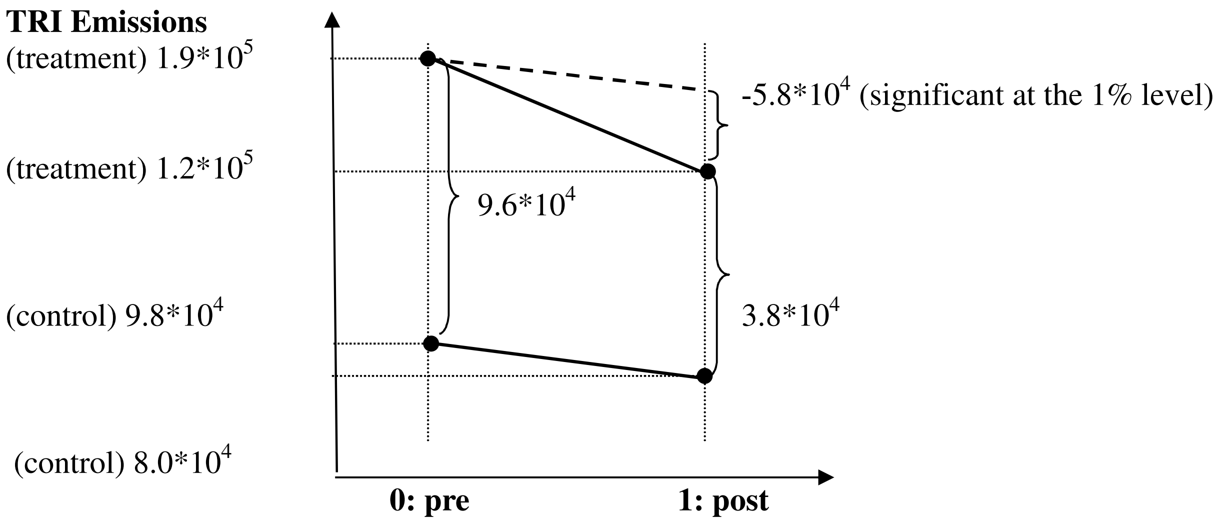
8. Conclusion

In this paper, we estimate the causal effect of third party verification on facility emissions using a novel dataset and a wide array of econometric treatment effect models. We do not find robust evidence that third party verification has a causally negative effect on facility emissions. **More here once the rest of our analysis comes in…**

We point out that our research is of the first of its kind, using rigorous econometric techniques to estimate the causal effect of third party verification on facility emissions. Much econometric research has focused on voluntary pollution abatement programs in general, often finding mixed conclusions regarding the effectiveness of such policy measures. While third party verification is an important policy modification designed to overcome potential criticisms of other voluntary abatement efforts, little econometric attention has been paid to the effect of third party verification.

Figure 1: Simple Comparison of Mean TRI Emissions Before and After Third Party Certification between Treatment and Control Groups

Panel A



Panel B

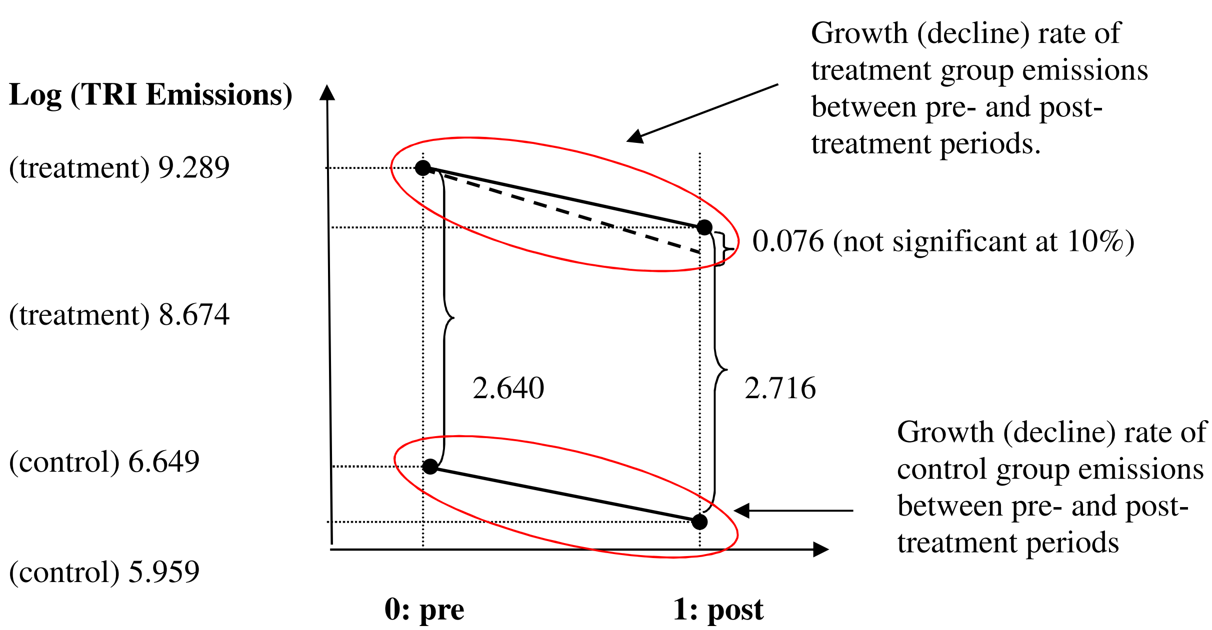


Figure 2: Average Trends in Emissions for Treated and Untreated Facilities

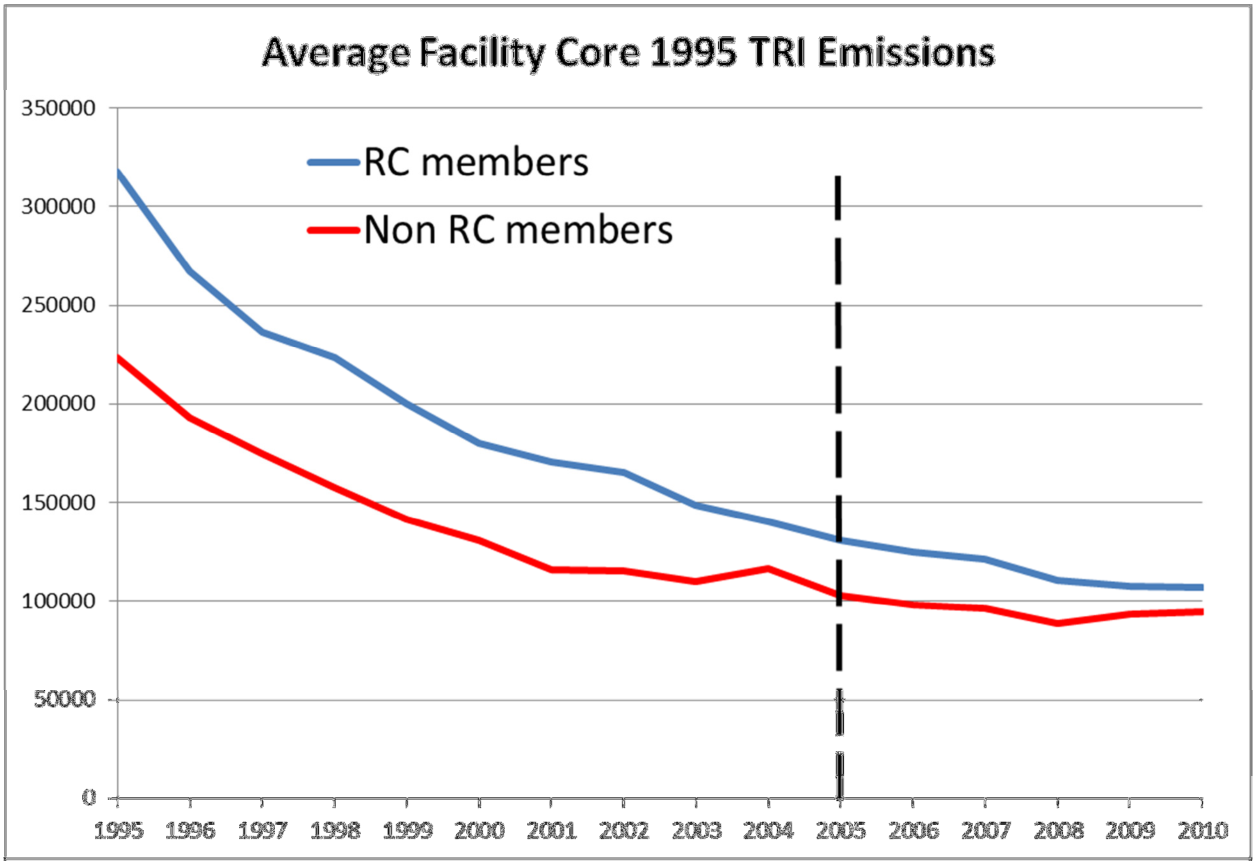


Table 1: Descriptive Statistics

Treatment Group		Control Group		Difference Between Groups	
Variable		**Variable**		**Variable**	
TRI releases		**TRI releases**		**TRI releases**	
Mean	162302.3	Mean	90019	Mean	72283.3
Standard deviation	467647.6	Standard deviation	482150.1		
Median	15571	Median	1349	Median	14222
Facility to firm TRI releases		**Facility to firm TRI releases**		**Facility to firm TRI releases**	
Mean	.1157564	Mean	.2218011	Mean	-.10604
Standard deviation	.2090636	Standard deviation	.3064696		
Median	.0211934	Median	.0560565	Median	-.03486
Parent firm TRI releases		**Parent firm TRI releases**		**Parent firm TRI releases**	
Mean	3024820	Mean	455229.1	Mean	2569591
Standard deviation	4802283	Standard deviation	1496251		
Median	923748	Median	37289.33	Median	886458.7
HAP-TRI ratio		**HAP-TRI ratio**		**HAP-TRI ratio**	
Mean	.736698	Mean	2.092387	Mean	-1.35569
Standard deviation	3.572139	Standard deviation	5.92594		
Median	.71266	Median	.6710102	Median	0.04165
Inspections		**Inspections**		**Inspections**	
Mean	.8107505	Mean	.4457915	Mean	0.364959
Standard deviation	3.665493	Standard deviation	1.421925		
Median	0	Median	0	Median	0
County non-attainment		**County non-attainment**		**County non-attainment**	
Mean	.8796153	Mean	.9014771	Mean	-0.02186
Standard deviation	.9769422	Standard deviation	1.058652		
Median	1	Median	1	Median	0
Facility-year observations	**5823**	**Facility-year observations**	**7176**		

Table 2: Difference-in-Differences Estimate of the Impact of RC Third Party Certification on Facility's TRI Air Releases: Exogenous Treatment

Variable	Model 1 TRI releases	Model 2 TRI releases	Model 3 Log of TRI releases	Model 4 Log of TRI releases
Treatment	-27585.1*	-18020.6	0.047942	0.099728
	(15253.78)	(13746.15)	(0.091559)	(0.087982)
Year 1997	-9882.45*	-9914.75*	-0.15923***	-0.15928***
	(5484.461)	(5483.93)	(0.04352)	(0.043527)
Year 1998	-13798.4	-13846.3	-0.21866***	-0.21857***
	(9015.066)	(9007.186)	(0.051616)	(0.05163)
Year 1999	-25359**	-25416***	-0.25036***	-0.25017***
	9800.58	9791.774	0.058599	0.058616
Year 2000	-28594.6***	-28677***	-0.30296***	-0.30287***
	10277.14	10264.14	0.06791	0.06792
Year 2001	-42806.3	-42896.6***	-0.38366***	-0.38358***
	10072.37	10061.68	0.069047	0.069089
Year 2002	-41045.6***	-41137.7	-0.39043***	-0.39027***
	10629.22	10613.06	0.071257	0.071296
Year 2003	-43825***	-43929.6***	-0.45175***	-0.45176***
	12244.76	12236.03	0.074222	0.07427
Year 2004	-36544.5**	-36660.2***	-0.4685***	-0.46855***
	14335.48	14331.17	0.079559	0.079594
Year 2005	-39744.7***	-44020.6***	-0.51456***	-0.53707***
	14053.35	13355.97	0.088397	0.093644
Year 2006	-41512.5***	-45320.7***	-0.64784***	-0.71567***
	14572.96	14379.98	0.085831	0.090849
Year 2007	-43358.4***	-45315.9***	-0.75237***	-0.82564***
	14194.19	14103.79	0.088856	0.094522
Year 2008	-49652.1***	-48542.9***	-0.67551***	-0.66145***
	14742.65	14468.14	0.092236	0.0965
Year 2009	-40787.7**	-37333.2**	-0.85664***	-0.78701***
	16962.4	18947.05	0.09508	0.099808
Year 2010	-40104.6**	-35109.3*	-0.82663***	-0.74821***
	16654.27	18543.67	0.100352	0.107908
Treatment*year2006	-	-1148.51	-	0.101773
	-	5860.596	-	0.08201
Treatment*year2007	-	-5317.06	-	0.114333
	-	11156.74	-	0.088987
Treatment*year2008	-	-12290.5	-	-0.08305
	-	8583.011	-	0.086474

Treatment*year2009	-	-17715.6	-	-0.21053**
	-	14385.31	-	0.101561
Treatment*year2010	-	-21215.5*	-	-0.23074**
	-	12857.97	-	0.111038
Facility to firm TRI ratio$_{(-1)}$	135847.9***	135935.8***	2.490956***	2.492695***
	22243.88	22241.9	0.174574	0.174528
Parent firm TRI releases$_{(-1)}$	0.020231***	0.020172***	0.199902***	0.20039***
	0.004281	0.004278	0.033504	0.033633
HAP-TRI ratio $_{(-1)}$	6.549403	5.585002	0.000209	0.000201
	1326.299	1313.912	0.042107	0.042348
Number of inspections$_{(-1)}$	-29.5422	-15.7177	-0.00341	-0.00315
	3072.618	3078.01	0.006624	0.006482
County non-attainment$_{(-1)}$	3939.859	3819.164	-0.04585	-0.0465
	5726.393	5741.067	0.053029	0.053255
Constant	99837.66***	100097.6***	5.335431***	5.32981***
	17188.37	17178.11	0.423616	0.425165
Number of observations	12999	12999	12999	12999
Number of groups	935	935	935	935

Note: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Bootstrapped robust standard errors clustered on facilities are in parentheses. In Models 1 and 2 the dependent variable is the pounds of TRI air releases while in Models 3 and 4 the dependent variable is the natural log of TRI air releases. Treatment dummy equals to 1 for all RC participants starting in year 2005. In all models the number of inspections, HAP-TRI, facility to firm TRI, parent firm TRI and the number of gases for which a facility's county is out of attainment with NAAQS are lagged by one year relative to the year in which the dependent variable is measured. The number of observations reflects that our dataset starts in 1995 to allow for lags. Parent firm TRI emissions are measured in natural logs in Models 3 and 4. All other variables are in levels.

Table 3: Random Effects Probit Model of Facility Participation in RC

Variable	Model 1	Model 2
Other facilities' TRI air releases$_{(-1)}$	5.65E-07***	5.98E-07***
	6.76E-08	7.26E-08
Number of other facilities$_{(-1)}$	0.3461***	0.238505***
	2.34E-02	2.28E-02
NAICS 3251: Basic chemical manufacturing	5.001502***	3.366943**
	1.52E+00	1.30E+00
NAICS 3252: Resin, Synthetic Rubber, and Artificial	2.710013*	1.858066
Synthetic Fibers and Filaments Manufacturing	1.57E+00	1.33E+00
NAICS 3253: Pesticide, Fertilizer, and Other	-11.964***	-7.8143***
Agricultural Chemical Manufacturing	1.64E+00	1.44E+00
NAICS 3254: Pharmaceutical and Medicine	-5.33212***	-4.17993
Manufacturing	1.68E+00	1.55E+00
NAICS 3255: Paint, Coating, and Adhesive	-12.6396***	-9.73234***
Manufacturing	1.57E+00	1.43E+00
NAICS 3256: Soap, Cleaning Compound, and Toilet	-2.6287	-1.64219
Preparation Manufacturing	2.22E+00	1.56E+00
NAICS 3259: Other Chemical Product and	-6.32944***	-4.36311***
Preparation Manufacturing	1.77E+00	1.34E+00
Year 1997	-	0.019317
	-	5.94E-01
Year 1998	-	0.181101
	-	6.15E-01
Year 1999	-	0.021277
	-	6.55E-01
Year 2000	-	-0.1878
	-	6.86E-01
Year 2001	-	-0.15083
	-	6.63E-01
Year 2002	-	-0.13611
	-	6.95E-01
Year 2003	-	-0.24857
	-	7.08E-01
Year 2004	-	-0.18559
	-	7.07E-01
Year 2005	-	-0.23255
	-	6.76E-01
Year 2006	-	-0.22492

	-	6.94E-01
Year 2007	-	-0.35784
	-	6.86E-01
Year 2008	-	-0.32905
	-	6.95E-01
Year 2009	-	-0.13748
	-	6.61E-01
Year 2010	-	-0.15431
	-	6.23E-01
TRI air releases$_{(-1)}$	3.18E-09	-2.70E-07
	4.38E-07	2.98E-07
HAP-TRI ratio $_{(-1)}$	0.000625	3.81E-05
	6.19E-03	8.87E-03
Number of inspections $_{(-1)}$	0.019327	0.012559
	1.00E-01	8.60E-02
County non-attainment$_{(-1)}$	0.202919	0.207494
	2.10E-01	1.63E-01
Public company	6.015344***	4.664983***
	4.24E-01	3.74E-01
Constant	-7.80446***	-5.4054***
	1.58E+00	1.40E+00
Log likelihood	-508.957	-523.215
AIC	1049.914	1106.431
BIC	1169.476	1330.61

Note: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors are in parentheses. All variables are in levels. All time varying variables are lagged by one year relative to the year in which the dependent variable is measured. Variables that serve as instruments (excluded in the main difference-in-difference specification) are other facilities' TRI releases, the number of facilities reporting and the dummy variables for four digit NAICS.

Table 4: Difference-in-Differences Estimate of the Impact of RC Third Party Certification on Facility's TRI Air Releases: Endogenous Treatment

	Model 1 TRI releases	Model 2 TRI releases	Model 3 Log of TRI releases	Model 4 Log of TRI releases
Treatment	-27383*	-17644	0.03321	0.07461
	16411.9	14325	0.08391	0.08096
Year 1997	-9751.1*	-9789.4*	-0.1505***	-0.1506***
	5300.36	5303.63	0.04492	0.04494
Year 1998	-13538	-13594	-0.2097***	-0.2096***
	8757.11	8750.94	0.05218	0.05219
Year 1999	-24996***	-25063***	-0.2403***	-0.2402***
	9336.13	9331.05	0.05769	0.0577
Year2000	-28389***	-28480***	-0.2973***	-0.2973***
	9780.35	9770.66	0.06532	0.06533
Year 2001	-42682***	-42782***	-0.3787***	-0.3787***
	9892.97	9884.93	0.07006	0.07006
Year 2002	-40698***	-40799***	-0.3882***	-0.388***
	10811.2	10798.9	0.06732	0.06736
Year 2003	-43486***	-43597***	-0.4499***	-0.4499***
	12079.1	12076	0.07221	0.07223
Year 2004	-36119**	-36244**	-0.4664***	-0.4664***
	14171.8	14171.2	0.07408	0.07409
Year 2005	-39628***	-43942***	-0.5215***	-0.5392***
	13142.9	12661.1	0.0834	0.08569
Year 2006	-41386***	-45262***	-0.6499***	-0.7183***
	13980	13660.2	0.08898	0.09569
Year 2007	-43256***	-45242***	-0.7569***	-0.8293***
	14172.6	13683.7	0.09009	0.09547
Year 2008	-49651***	-48392***	-0.677***	-0.6621***
	13941	14142.6	0.08873	0.09462
Year 2009	-40650***	-37203**	-0.8559***	-0.7862***
	15574.7	17622.8	0.09494	0.10038
Year 2010	-40055***	-35044**	-0.8237***	-0.751***
	15385.6	17074	0.09627	0.10534
Treatment *year2006	-	-1094.9	-	0.11484
	-	5843.15	-	0.08065
Treatment *year2007	-	-5389.9	-	0.1243
	-	10695.8	-	0.08693
Treatment *year2008	-	-12840	-	-0.0746
	-	8253.34	-	0.08578

Treatment *year2009	-	-17952	-	-0.2016*
	-	14565.5	-	0.10538
Treatment *year2010	-	-21514*	-	-0.2085*
	-	12564.4	-	0.11644
Facility to firm TRI ratio$_{(-1)}$	139616***	139758***	2.40451***	2.40649***
	24625.9	24634.7	0.16434	0.16404
Parent firm TRI releases$_{(-1)}$	0.02051***	0.02045***	0.20581***	0.20624***
	0.00474	0.00474	0.03072	0.03077
HAP-TRI ratio$_{(-1)}$	6.66508	5.68474	0.0002	0.0002
	851.52	851.521	0.02315	0.02316
Number of inspections$_{(-1)}$	-12.482	0.76752	-0.0029	-0.0026
	2701.14	2707.14	0.00675	0.0066
County non-attainment$_{(-1)}$	4061.55	3918.18	-0.0316	-0.0323
	6461.21	6465.63	0.04958	0.04964
Inverse Mills ratio	2197.5	2108.82	0.07795	0.07752
	5418.74	5418.37	0.07689	0.07704
Constant	99189.7***	99470.7***	5.2776***	5.27265***
	18064.9	18059.1	0.38307	0.38371
Number of observations	12905	12905	12905	12905
Number of groups	930	930	930	930

Note: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Bootstrapped robust standard errors clustered on facilities are in parentheses. In Models 1 and 2 the dependent variable is the pounds of TRI air releases while in Models 3 and 4 the dependent variable is the natural log of TRI air releases. Treatment dummy equals to 1 for all RC participants starting in year 2005. In all models the number of inspections, HAP-TRI, facility to firm TRI, parent firm TRI and the number of gases for which a facility's county is out of attainment with NAAQS are lagged by one year relative to the year in which the dependent variable is measured. The number of observations reflects that our dataset starts in 1995 to allow for lags. Parent firm TRI emissions are measured in natural logs in Models 3 and 4. All other variables are in levels.

Table 5: Sensitivity Analysis - Difference-in-Differences Estimate of the Impact of RC Third Party Certification on Facility's TRI Air Releases: Instrumental Variable approach

Variable	Model 1 TRI releases	Model 2 TRI releases	Model 3 Log of TRI releases	Model 4 Log of TRI releases
Treatment	-7098049*	-217360	-0.5801	-18.708
	4087343	10900000	1.77315	177.063
Year 1997	-22270.7**	-10262.6	-0.1587***	-0.1442
	10663.44	27474.59	0.0437	0.1882
Year 1998	-38963.2**	-14553.5	-0.2179***	-0.1996
	17332.87	45481.36	0.05189	0.22099
Year 1999	-59163.7**	-26366.1	-0.2494***	-0.2301
	23103.92	58969.2	0.05891	0.27035
Year 2000	-70301.6**	-29849	-0.3016***	-0.2679
	27608.43	75724.59	0.06821	0.45454
Year 2001	-83046.4***	-44027	-0.3819***	-0.3369
	26877.8	78730.67	0.06941	0.54487
Year 2002	-71659.5***	-41997.1	-0.3876***	-0.3167
	24199.52	76815.14	0.0715	0.7896
Year 2003	-66066.3***	-44552.9	-0.4483***	-0.3559
	22138.08	65356.39	0.07503	1.02582
Year 2004	-51359.4**	-37074.1	-0.4638***	-0.3376
	21842.58	58212.59	0.08054	1.32936
Year 2005	3013914*	42076.27	-0.2365	7.77857
	1781687	4667828	0.78967	76.8334
Year 2006	278265.9	-36210.1	-0.6798***	0.27677
	195496.1	472626.6	0.13409	9.32753
Year 2007	276247.8	-36262.2	-0.7898***	0.16492
	193627.6	468630.1	0.13333	9.27153
Year 2008	275588.8	-39416.9	-0.6252***	0.34121
	194666.8	474825.3	0.13373	9.32281
Year 2009	285748.9	-28236.8	-0.7507***	0.2186
	193017.3	471971.2	0.13696	9.35429
Year 2010	277171.6	-26317	-0.7133***	0.21882
	187542.7	449181.2	0.14063	9.06974
Treatment*year 2006	6205535*	173602.1	0.70364	16.7533
	3575270	9522200	1.5636	156.902
Treatment*year 2007	6208438*	169632.6	0.71661	16.7769
	3580483	9535027	1.56842	157.041
Treatment*year 2008	6207336*	162824.5	0.51975	16.5942

	3584165	9544617	1.57334	157.166
Treatment*year 2009	6173477*	156598.9	0.39207	16.461
	3568705	9482375	1.5698	157.123
Treatment*year 2010	6175489*	153254.1	0.37208	16.447
	3571314	9502970	1.57088	157.159
Facility to firm TRI ratio$_{(-1)}$	48934.94	133486.3	2.48676***	2.32855
	78458.76	329262.7	0.17374	1.44921
Parent firm TRI releases$_{(-1)}$	-0.03843	0.018522	0.19997***	0.1887
	0.034479	0.107134	0.0336	0.12841
HAP-TRI ratio$_{(-1)}$	-29.0771	4.609084	0.0002	9.6E-05
	7058.475	1322.607	0.04075	0.03787
Number of inspections$_{(-1)}$	1684.398	32.14943	-0.003	0.00063
	4858.608	2999.974	0.00666	0.08848
County non-attainment$_{(-1)}$	67235.64	5604.67	-0.04	0.13259
	46724.7	143098.6	0.05481	1.48861
Constant	199272.6***	102889.9	5.32814***	5.28359***
	70232.56	223304.8	0.42469	1.33895
Number of observations	12999	12999	12999	12999
Number of groups	935	935	935	935

Note: *** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. Bootstrapped robust standard errors are in parentheses. In Models 1 and 2 the dependent variable is the pounds of TRI air releases while in Models 3 and 4 the dependent variable is the natural log of TRI air releases. Treatment dummy equals to 1 for all RC participants starting in year 2005. In all models the number of inspections, HAP-TRI, facility to firm TRI, parent firm TRI and the number of gases for which a facility's county is out of attainment with NAAQS are lagged by one year relative to the year in which the dependent variable is measured. The number of observations reflects that our dataset starts in 1995 to allow for lags. Parent firm TRI emissions are measured in natural logs in Models 3 and 4. All other variables are in levels. In Models 1 and 3 we instrument for treatment with the other facilities' TRI releases, the number of other facilities reporting, dummies for NAICS and a dummy variable equal to 1 if the parent firm is publicly owned. While in Models 2 and 4 we instrument for treatment with the predicted probability of participation from Model 2 in Table 3.

References

American Chemistry Council (ACC) 2013. Responsible Care. http://responsiblecare.americanchemistry.com/

Bertrand, M., Duflo, E. and Mullainathan, S., 2004. How much should we trust differences-in-Differences estimates? *Quarterly Journal of Economics*, 119 (1): 249-275.

Bi, X. and Khanna, M., 2012. Re-assessment of the impact of EPA's voluntary 33/50 program on toxic releases, *Land Economics*, 88 (2): 341–361.

Bui, L. and Kapon, S., 2012. The impact of voluntary programs on polluting behavior: Evidence from pollution prevention programs and toxic releases. *Journal of Environmental Economics and Management,* 64: 31-44.

Darnall, N. and Carmin, J., 2005. Cleaner and greener? The signaling accuracy of U.S. voluntary environmental programs. *Policy Sciences,* 38: 71–90.

Carrión-Flores, C. E., Innes, R. and Sam, A. 2013. Do voluntary pollution reduction programs (VPRs) spur or deter environmental innovation? Evidence from 33/50. *Journal of Environmental Economics and Management,* 66: 444–459.

Gamper-Rabindran S., 2006. Did the EPA's voluntary industrial toxics program reduce emissions? A GIS analysis of distributional impacts and by-media analysis of substitution. *Journal of Environmental Economics and Management,* 52: 391-410.

Gamper-Rabindran S. and Finger, S., 2013. Does self-regulation reduce pollution? Responsible Care in the chemicals industry, *Journal of Regulatory Economics,* 43: 1-30.

Innes, R. and Sam, A., 2008. Voluntary pollution reductions and the enforcement o environmental law: an empirical study of the 33/50 program. *Journal of Law Economics,* 51: 271 – 296.

Khanna, M. and Damon, L. A., 1999. EPA's voluntary 33/50 program: Impact on toxic releases and economic performance of firms. *Journal of Environmental Economics and Management,* 37(1): 1–25.

King, A. and Lenox, M., 2000. Industry self-regulation without sanctions: the chemical industry's Responsible Care Program. *Academy of Management Journal,* 43: 698–716.

King, A., Lenox, M. and Terlaak, A., 2005. The strategic use of decentralized institutions: exploring certification with the ISO 14001 management standard. *Academy of Management,* 48: 1091–1106.

Li, Q., Huang, C. J., Li, D. and Fu, T., 2002. Semiparametric smooth coefficient models. *Journal Of Business and Economic Statistics*, 20 (3): 412-422.

Li, Q. and Racine, J. S., 2007. *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.

Melnyk, S. A., Sroufe, R. P., Calantone, R. L. and Montabon, F.L., 2002. Assessing the effectiveness of US voluntary environmental programmes: an empirical study. *International Journal of Production Research,* 40: 1853–1878.

Meyer, B. D., 1995. Natural and quasi-experiments in economics. *Journal of Business and Economic Statistics,* 13(2): 151-161.

Montabon, F., Melnyk, S. A., Sroufe, R. and Calantone, R.J., 2000. ISO 14000: assessing its perceived impact on corporate performance. *The Journal of Supply Chain Management,* Spring: 4–15.

Phillips, D., 2006. RC 14001certification: An integrated management systems approach. Q*uality Digest,* May.

Potoski, M. and Prakash, A., 2005. Covenants with weak swords: ISO 14001 and facilities' environmental performance. *Journal of Policy Analysis and Management,* 24: 745–769.

Prakash, A., 2000. Responsible Care: An assessment. *Business and Society,* 39(2): 183-209.

Racine, J. and Li, Q., 2004. Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics*, 119 (1): 99-130.

Rao, P. and Hamner, B., 1999. Impact of ISO 14000 on business performance. *Working paper*.

Rivera, J., deLeon, P. and Koerber, C., 2006. Is greener whiter yet? The sustainable slopes program after five years. *Policy Studies Journal,* 34 (2): 195–224.

Rondinelli, D. A. and Vastag, G., 2000. Panacea, common sense, or just a label? The value of ISO 14001 environmental management systems. *European Management Journal,* 18: 499–510.

Sam, A., Khanna, M. and Innes, R., 2009. Voluntary pollution reduction programs, environmental management, and environmental performance: An empirical study. *Land Economics,* 85(4): 692-711.

Toffel, M. W., 2006. Resolving information asymmetries in markets: The role of certified management programs. *Harvard Business School Working Paper*.

Vidovic, M. and Khanna, N., 2007. Can voluntary pollution prevention programs fulfill their promises? Further evidence from the EPA's 33/50 program. *Journal of Environmental Economics and Management,* 53: 180-195.

Vidovic, M. and Khanna, N., 2012. Is voluntary pollution abatement in the absence of a carrot or stick effective? Evidence from facility participation in the EPA's 33/50 program. *Environmental Resource Economics,* 52: 369-393.

Wooldridge, J., 2011. *Econometric Analysis of Cross-Section and Panel Data*. 2nd Edition, MIT Press: Cambridge.