Public Pressure and the Heterogeneous Effects of Voluntary Pollution Abatement

Ruohao Zhang[a,*], Neha Khanna[a]

[a] *Department of Economics, Binghamton University, Binghamton, NY*

This version: July 2020

Abstract

With widespread environmental awareness, polluters face abatement pressure from two sources: formal regulation pressure and informal public pressure. While the impact of formal regulation on plant emissions is well understood, the role of public pressure in reducing pollution is less clear. We build a conceptual model highlighting the role of public pressure in environmental regulation in the context of voluntary pollution abatement. The launch of a voluntary pollution abatement program changes both regulatory pressure and public pressure albeit differently for participants and non-participants. Our theory describes these changes as well as the plant's emission choices. We show that the effectiveness of a voluntary pollution abatement program depends on the cost from public scrutiny of participating firms and the associated risk of being labeled greenwashers: greater public scrutiny yields fewer program participants who free-ride thereby increasing the effectiveness of the program. Our model, which provides a framework for reconciling the mixed empirical results on the effectiveness of voluntary pollution abatement programs, is supported by data from the EPA's 33/50 program.

Key words: Environmental Regulation, Public Pressure, Greenwashing, Free-riding, 33/50 program, Threshold effect

JEL codes: D21 D22 H41 Q52 Q53

[*] Corresponding author. E-mail: rzhang47@binghamton.edu.

1 Introduction

Mandatory environmental regulation in the United States relies on costly government interventions such as emissions monitoring and pollution source inspections. However, with widespread environmental awareness, polluters also face public pressure from stakeholders to reduce pollution. With limited and declining resources for traditional monitoring and inspection efforts, innovative policy changes have enhanced the role of public pressure in environmental regulations such as through information disclosure and voluntary pollution abatement programs. While there is a substantial literature on the impact of traditional regulation on firm environmental behavior, the impact of public pressure is less well understood. In this paper, we model the role of public pressure in environmental regulation in the context of voluntary pollution abatement. Our theory implies that the key factor influencing the effectiveness of voluntary pollution abatement is the public scrutiny of participating firms and the associated risk of being labeled a greenwasher, which is an important component of public pressure. By highlighting the power of public evaluation of participating firms' environmental performance, we provide a consistent explanation for the mixed empirical results on the effectiveness of voluntary pollution abatement programs found by the previous literature.

Voluntary pollution abatement programs are widely used. In the United States, voluntary pollution abatement comes up as an solution when mandatory regulation is difficult to apply, or serves as a complement when mandatory regulation is absent or not stringent enough.[1, 2] While voluntary pollution abatement programs are associated with very limited or no direct financial benefit, they also lack sanctions to motivate firm participation that generates pollution abatement in excess of the mandated level. This begs the question of what factors motivate firms to participate in voluntary pollution abatement, and the mechanism through which voluntary pollution abatement programs affect firm environmental performance. Our model answers the question through the fact that the launch of a voluntary pollution abatement program is accompanied by changes in both traditional regulation pressure and stakeholder public pressure. In

[1] The difficulties of applying mandatory regulation may come from many aspects, such as mounting and increasingly complex legislation, technical innovation and scientific discoveries, regulatory budget cuts and increased use of and effectiveness of citizen lawsuits (Brouhle et al., 2005).

[2] There are three types of voluntary pollution abatement programs: negotiated agreements between government and industry; public voluntary programs created by a regulatory authority such as the EPA's 33/50 program; and industry-led initiatives such as chemical industry's Responsible Care Program (Brouhle et al., 2005).

particular, for participating firms, environmental performance is publicly known and scrutinized. For example, the United States Environmental Protection Agency (EPA) issued annual reports highlighting the environmental accomplishments of firms participating in its signature 33/50 voluntary pollution abatement program; the American Chemistry Council reports the performance of member companies through its Responsible Care program (http://reporting.responsiblecare-us.com/Reports/Members/Envrm_Cmpny_Rpt.aspx). Such information disclosure provides stakeholders with a tool for identifying greenwashers: firms that commit to environmental improvement but fall short of that commitment in terms of measured outcomes (or worse still, whose performance deteriorates despite the positive environmental signal).

In the absence of a voluntary pollution abatement program, there is pre-existing mandatory regulation pressure and background public pressure. Supported by empirical evidence, the launch of the voluntary pollution abatement program shifts regulation resources from participants to non-participants (Innes and Sam, 2008; Li and Khanna, 2018), thus decreasing the marginal emission cost for participants and raising it for non-participants. The change in public pressure is, however, twofold. First, because voluntary pollution abatement program participation provides a positive environmental signal, it lowers public pressure for participants but increases it for non-participants, further reducing the marginal emission cost for participants and increasing it for non-participants. Second, we expect that public pressure grows with emissions, but the growth rate is higher for participants than non-participants. The difference in the growth rate of public pressure arises because participating in a voluntary pollution abatement program comes with additional public scrutiny of participants' environmental performance, and results in the risk of being labeled a greenwasher (Lyon and Maxwell, 2011; Kim and Lyon, 2011). We show that for an effective voluntary pollution abatement program, such risk increases rapidly with emissions and quickly raises the participants' marginal emission cost so that the previously described decrease in the participant's marginal emission cost is eventually outweighed. That is, our theory indicates that as compared with the pre-program level, non-participants have higher marginal emission cost at any emission level, whereas participants have lower marginal emission cost only at relatively lower emission levels. Eventually, the marginal emission cost of participants exceeds that of the non-participants at some threshold emission level, which depends on how quickly the public scrutiny of participants' environmental performance and the associated risk of being labeled a greenwasher rises with emissions. We assume a firm's marginal abatement cost is determined exogenously by firm characteristics and abatement technologies. Firms choose their optimal emission levels where the marginal emission cost equals the marginal abatement cost.

We find that while the voluntary abatement program attracts participation by firms on either side of the threshold, firm emissions must be close to the threshold for participation to be profit maximizing. However, in terms of the effectiveness of the program, it matters whether participating firms have emissions above or below the threshold. Thus, we find that for firms with emissions higher than the threshold emission level, less polluting firms are more likely to participate in the voluntary pollution abatement; and by participating, they reduce their emissions relative to not participating. For the group of firms with emissions lower than the threshold emission level, more polluting firms are more likely to participate in the voluntary pollution abatement program; and by participating, they *increase* their emissions relative to non-participants. These participants are free-riders. Greater public scrutiny of participants' environmental performance yields a lower threshold emission level, and fewer program participants who free-ride. However, compared with the pre-program level, both free-riders with relatively higher emissions and all non-participants also lower their emissions. Only less polluting free-riders increase their emissions relative to the pre-program level under a voluntary pollution abatement program.

We use the EPA's 33/50 program and plant level data from Zhou et al. (2020) to test our theory. The 33/50 program is a voluntary pollution abatement program initialized in 1991 that targeted the emissions of 17 toxic chemicals. As the first voluntary pollution abatement program sponsored by the US EPA, it received much attention and was considered a huge success by the EPA, which reported that the quantitative program goals where reached a year ahead of schedule. Nonetheless, the academic literature has debated the success of the program (see the next section for details). We use two different samples in our analysis: the full sample with all plants, and a preferred sample, which is a subsample including first year (1991) participants and all non-participants.[3] We empirically identify the threshold values and estimate the participation effects separately for each sample. While the preferred sample yields a more precise estimate

[3]Plants participated in the program in different years, but a majority of the participants joined at the beginning of the program in 1991.

of the threshold value (since the threshold may vary across different participation years) the results shows that the threshold values are similar between the two samples.

Accounting for the additional public scrutiny of participants and the possibility of free-riding, we find sizable differences in participation effects on either side of the threshold in both samples. For example, for the preferred sample, participants with emissions below the threshold are free-riders and increase their emissions by 113.4% compared to the non-participating counterfactual, while the participants with emissions above the threshold reduce their emissions by 55.3%. While we do not detect a free-riding phenomenon in the full sample, possibly because the threshold value is not precisely identified, we find that compared to the non-participating counterfactual, plants above the threshold reduce emissions by 61.6%, which is nearly 4 times as much as plants below the threshold. Without considering free-riding, we find that for the full sample, participants on average reduce their emissions by 42.7%; for the preferred sample, participants on average reduce their emissions by 24.9%; which is similar to the results of Innes and Sam (2008) and Bi and Khanna (2012).

These results validate our conceptual framework. More importantly, by comparing the results with and without free-riding, we can point to a powerful explanation for the disparity in the empirical findings in the voluntary pollution abatement literature that is couched in terms of public pressure and the risk of participants being labeled greenwashers.

The paper has the following structure: In section 2, we summarize the previous literature related to voluntary pollution abatement and polluters' abatement pressure. Section 3 describes our conceptual framework. We present the formal model structure in section 4. We then propose the econometric model in section 5, and analyze the 33/50 data in section 6. Section 7 includes several robustness checks of our empirical analysis. Section 8 concludes.

2 Previous Literature

There is a large literature debating the effectiveness of voluntary pollution abatement programs. Several studies focus on the EPA' s 33/50 program and conclude that voluntary pollution abatement programs are effective. Khanna and Damon (1999) and Innes and Sam (2008) found that the program is effective at the firm level during the first several years. Bi and Khanna (2012) found that the program is effective at the plant level but not the firm level.

Zhou et al. (2020) incorporate technology spillovers, and they find an even larger effect of the 33/50 program than previous studies. On the other hand, Vidovic and Khanna (2007, 2012) and Gamper-Rabindran (2006) found that participation in the 33/50 program is not associated with a decline in emissions. Brouhle et al. (2009) investigated the Strategic Goals Program, a voluntary pollution abatement program in the metal-finishing industry, and found that participation leads to little additional reduction in emissions. Welch et al. (2000) found the Climate Challenge Program is ineffective. Gamper-Rabindran and Finger (2013) who examine the early years Responsible Care Program found that participants increase their emissions compared to non-participants. Vidovic et al. (2019) found that the Responsible Care program is not effective even with the modification of third-party certification. In fact, third-party certification may even increase participants' emissions. King and Lenox (2000) who also studied the Responsible Care program, found opportunism makes the program difficult to maintain without explicit sanctions. They hypothesize that this is caused by much greater public pressure on non-participants to reduce emissions relative to participants.

There are only a small handful of studies examining the mechanism through which voluntary pollution abatement programs affect firm behavior. Decker (2002)'s model shows that extra investments in environmental performance not only reduce a plant's regulation burden, it also transfers additional burden to other peer plants. Zhou et al. (2020) argue that voluntary pollution abatement programs advance information about environmentally-friendly technologies, and the technology spillover drives the effectiveness of the programs. They, along with Lyon and Maxwell (2007) and Borck and Coglianese (2009) conclude that failure to account for technology spillovers from participants to non-participants explains why some empirical studies find voluntary pollution abatement to be ineffective. However, the literature has not identified an explanation for why participants in a voluntary pollution abatement program may increase emissions relative to non-participants.

To explain why voluntary pollution abatement programs are effective in some cases but counterproductive in the others, we consider three factors in our model that may affect firm behavior with respect to voluntary pollution abatement (Khanna and Damon, 1999; Lyon and Maxwell, 1999; Welch et al., 2000; Anton et al., 2004; Rivera, 2004; Vidovic and Khanna, 2007; Zhang et al., 2008; Innes and Sam, 2008; Carrion-Flores et al., 2013): firm characteristics, mandatory regulation pressure, and public pressure. Firm characteristics such as production and abatement technologies, managerial skills, firm size, etc., determine a firm's marginal abatement cost. Regulation pressure affects firm behavior though expected violations and penalties. Participating in voluntary pollution abatement programs defers tougher regulatory enforcement (King and Lenox, 2000; Decker, 2002; Gamper-Rabindran and Finger, 2013), lessens the probability of receiving inspections, and also lowers the expected violation penalties (Gamper-Rabindran, 2006; Li and Khanna, 2018). Conversely, non-participants may receive additional regulation pressure transferred from participants (Decker, 2002; Li and Khanna, 2018).

Public pressure is another major factor affecting firm decisions, and whether a firm participates in a voluntary pollution abatement program or not leads to changes in public pressure (Lyon and Maxwell, 1999; Khanna and Damon, 1999; King and Lenox, 2000; Anton et al., 2004; Rivera, 2004; Gamper-Rabindran, 2006; Carrion-Flores et al., 2013; Bi and Khanna, 2012; Gamper-Rabindran and Finger, 2013; Zhou et al., 2020). Public pressure comes from many aspects, including "green" consumers and investors (Khanna and Damon, 1999; Lyon and Maxwell, 1999; Welch et al., 2000; Gamper-Rabindran, 2006; Rivera, 2004; Flammer, 2015); stock market response (Wang et al., 2019), peer firms or industry associations (King and Lenox, 2000; Gamper-Rabindran and Finger, 2013; Rivera, 2004); local communities (Welch et al., 2000; Anton et al., 2004; Gamper-Rabindran, 2006; Zhang et al., 2008; Bi and Khanna, 2012; Gamper-Rabindran and Finger, 2013; Zhou et al., 2020); NGOs and environmental groups (Lyon and Maxwell, 1999; Maxwell et al., 2000; Rivera, 2004); media and press (Rivera, 2004; Wang et al., 2019) and other stakeholders.

There is a very small literature studying the effect of public evaluation of firms' environmental performance as an important part of public pressure. Lyon and Maxwell (2011) and Kim and Lyon (2011) argued that firms are hesitant to disclose their environmental efforts because it comes with the risk of being labeled greenwashers. In our framework, such risk only affects the voluntary pollution abatement program participants through additional public scrutiny of their environmental performance, since declaring participation is equivalent to announcing the firm's environmental efforts. Lyon and Maxwell (2011) and Kim and Lyon (2011) claimed that being labeled a greenwasher is costly because it may trigger consumer boycotts, lawsuits, negative media coverage, and attack campaigns.

3 Conceptual Framework: A Narrative

In this section we introduce the conceptual narrative underlying our theoretical framework. In the absence of formal regulation or public pressure, profit maximizing firms do not consider the social costs of their emissions. Environmental taxes, inspections, penalties and other forms of government intervention is one way to internalize the social cost of pollution. Public pressure from stakeholders through "conflicts between corporations and society" (Heal, 2005) is a second way. For example, "green" consumers may boycott polluting companies (Khanna and Damon, 1999; Wang et al., 2019); a local community may demand compensation for local health impacts caused by firms' pollution. These conflicts with stakeholders can generate an emission cost to firms that is higher than the cost of abatement (Wang et al., 2019). We define public pressure as the variety of methods that the public uses to convert the social cost of emissions to firms' private cost through Coasian bargaining (Heal, 2005).

In the absence of voluntary pollution abatement, there is pre-existing mandatory regulation pressure and background public pressure. With the launch of a voluntary pollution abatement program, there is a shift of government regulation resources from participants to non-participants. Especially, participation in a voluntary pollution abatement program results in a smaller probability of being inspected by the budget constrained regulator than non-participants (Gamper-Rabindran, 2006; Innes and Sam, 2008; Li and Khanna, 2018). Therefore, as compared with the pre-program situation, participants in a voluntary pollution abatement program face lower regulation pressure while it increases for non-participants. Figure 1 shows the firm's marginal emission cost curve generated jointly by regulation pressure and public pressure. We assume that firms with higher emissions face greater marginal cost due to both types of pressure and the marginal emission cost, MC_0, is convex in emissions. Figure 1 also illustrates how marginal emission cost shifts due to the change in regulation pressure, *ceteris paribus*, where MC'_P is the marginal emission cost of participants and MC'_N is the marginal emission cost of non-participants. Here, we do not assume a parallel shift in marginal emission cost curve. In fact, we assume that the downward shift of MC'_P is greater for smaller emitters than for larger emitters and gradually converges to zero as emission level goes to infinity;[4] likewise, we assume that the upward shift of MC'_N is greater for more polluting firms. Such changes imply that

[4]Intuitively, the regulator will not offer any regulation relief to an extremely polluting firm for its participation.

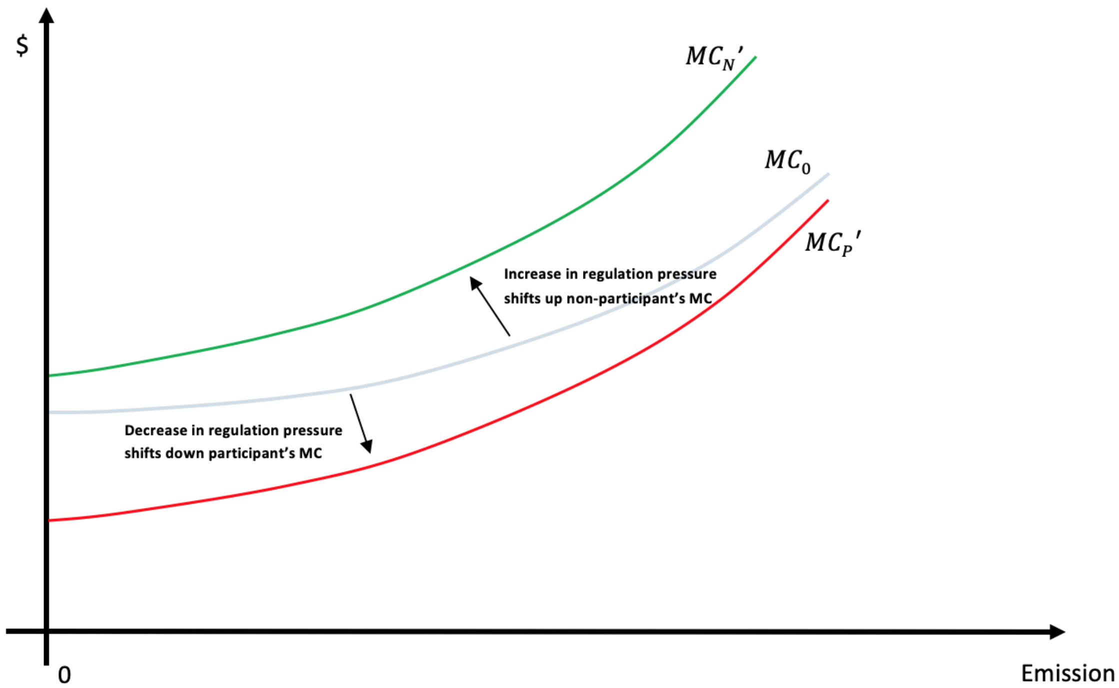


Figure 1: Shift in Marginal Emission Cost Curve with Changes in Regulation Pressure

for any level of emission, participants experience less marginal emission cost from regulation pressure than non-participants.

In Figure 2 we introduce the changes in the marginal emission costs due to the changes in public pressure *ceteris paribus*. Unlike regulation pressure, the change in public pressure is twofold. On the one hand, because public pressure is affected by environmental information (Aerts et al., 2008), participation as positive information reduces public pressure, whereas non-participation as negative information increases public pressure. Accordingly, in Figure 2, MC_P' and MC_N' shifts to MC_P'' and MC_N, respectively. Again, we do not assume a parallel shift in marginal emission cost curve. Instead we let the downward shift of MC_P'' be greater for smaller emitters than for larger emitters and that shift gradually converges to zero as emissions increase to infinity;[5] the upward shift of MC_N' is greater for more polluting firms, further exacerbating the gap in the marginal emission cost between participants and non-participants at any emission level. On the other hand, having publicly signaled their commitment to pollution abatement, program participants' environmental performance is subject to scrutiny. This leads to an addi-

[5]Similar intuition as before: an extremely polluting firm will not receive the benefit from public pressure by participating.

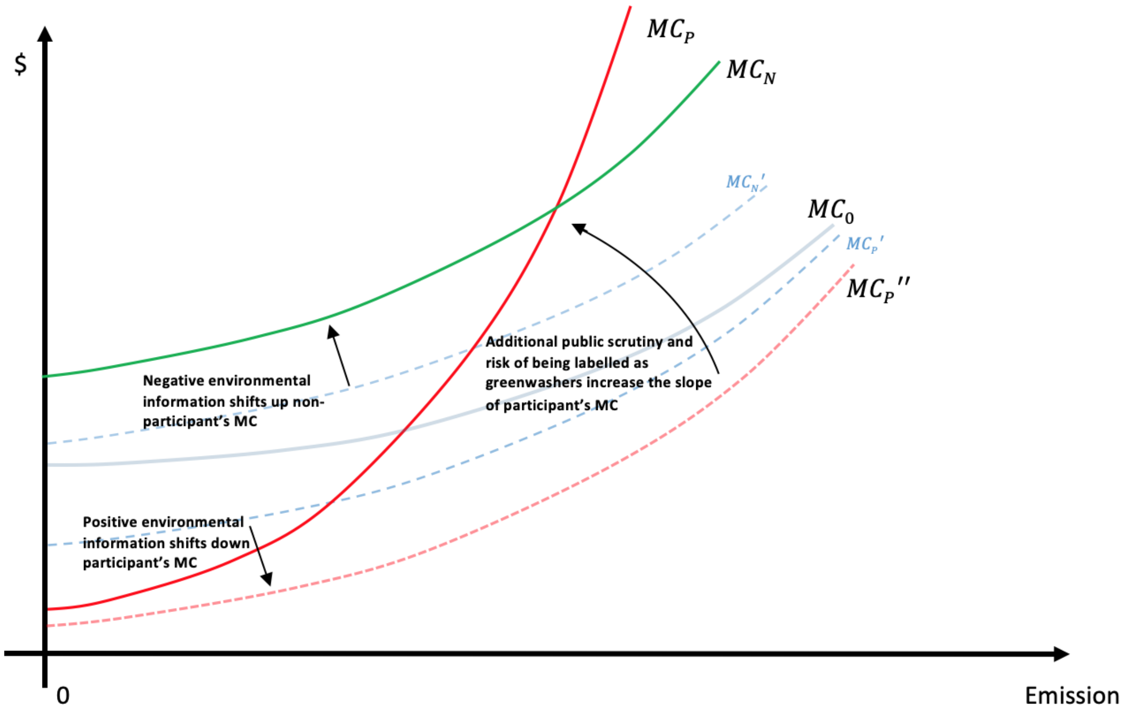
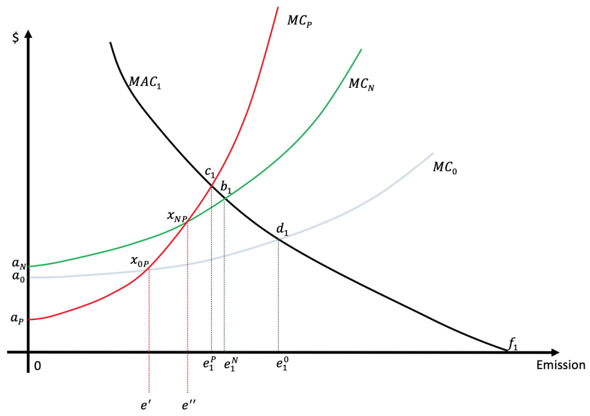


Figure 2: Shift in Marginal Emission Cost Curve with Changes in Public Pressure

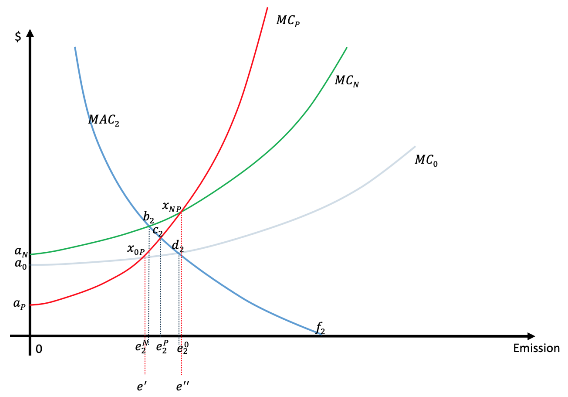
tional emission cost due to the the risk of being labeled a greenwasher, and this cost increases with emissions: a poor environmental performer is more likely to be targeted by stakeholders including environmental activists and NGOs (Delmas and Burbano, 2011). Furthermore, we assume that this cost through public scrutiny increases rapidly so that the downward shift in the participant's marginal emission cost curve is eventually outweighted. That is, as compared with the pre-program level, non-participants receive more public pressure at any emission level, whereas participants receive less public pressure only at relatively low emission levels, plus the public pressure faced by participants grows faster than both the background public pressure and non-participant's public pressure due to the additional scrutiny of their environmental performance under the program. Accordingly, in Figure 2, the slope of participants' marginal emission cost increases rapidly and MC_P'' rises to MC_P, so that participants receive more overall public pressure than both background and non-participants when emissions are relatively high.

Based on the above narrative, we build a model to show how firms change their emissions in response to changes in the marginal emission cost function due to the variations in regulation pressure and public pressure caused by voluntary pollution abatement programs. According to our theory, a firm's marginal emission cost curve consists of two components: marginal emission cost from regulation pressure, and marginal emission cost from public pressure. The regulation pressure component mainly comes from the expected frequency of inspections and expected violation penalties, while the public pressure component comes from the Coasian bargaining (Heal, 2005). In the presence of a voluntary pollution abatement program, the marginal emission cost of non-participants shifts up from the pre-program level due to a larger regulation pressure component (more inspections) and a larger public pressure component (negative environmental information due to non-participation). The marginal emission cost of participants shifts down from the pre-program level at low emissions due to lower regulation pressure (fewer inspections) and public pressure (positive environmental information due to participation). However, because of the additional public scrutiny of participants environmental performance, their public pressure grows faster with emissions relative to both the background (or pre-program) and the nonparticipants' public pressure. Therefore, at higher emission levels, the marginal emission cost of participants becomes even larger than nonparticipants.

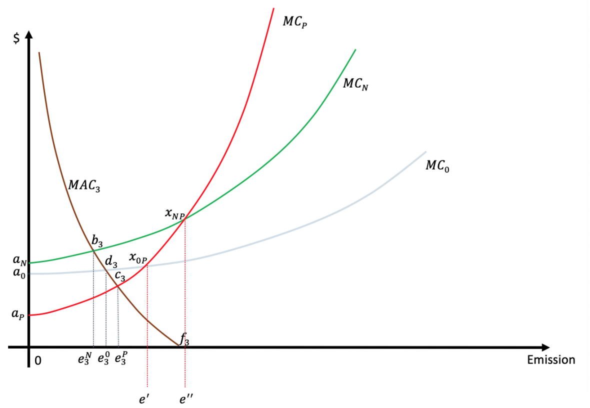
Figure 3 illustrates how the changes in the marginal emission cost affect a firm's emission decision and incentives to participate in the voluntary pollution abatement program. There are 3 types of firms making decisions about participation and emissions. MAC_1, MAC_2 and MAC_3 are the representative marginal abatement cost curves for each of these firm types. MC_0 is the pre-program marginal cost of emission, MC_P and MC_N are marginal cost of emission for participants and nonparticipants respectively. e' is the emission level where MC_P crosses MC_0, which means firms emitting more than e' experience higher marginal emission cost if they participate in the program compared with the pre-program costs. e'' is the emission level where MC_P crosses MC_N, which means firms emitting more than e'' experience higher marginal emission cost when participating than not participating in the program. Type 1 firms have relatively high marginal abatement costs and pre-program emissions. We define type 1 firms as firms whose marginal abatement cost curves cross MC_P and MC_N on the right of e''. Type 2 firms have moderate marginal abatement costs and pre-program emissions. We define type 2 firms as firms with marginal abatement cost curves that cross MC_P and MC_N between e' and e''. Type 3 firms have relatively low marginal abatement costs and pre-program emissions, so that their marginal abatement cost curves cross MC_P and MC_N on the left of e'. Firms choose



(a) Type 1 Firms Emission Decision



(b) Type 2 Firms Emission Decision



(c) Type 3 Firms Emission Decision

Figure 3: Emission Decision of 3 Types of Firms

their optimal emission levels based on the equilibrium condition $MC = MAC$, and choose participation or not by comparing the fixed cost of participation c_F (Zhou et al., 2020) with the difference between the optimized total emission costs under participation and the optimized total emission costs as a non-participant. In order to illustrate the total emission costs, we use $Area(\cdot)$ to identify the area of polygons represented by labeled points shown in Figure 3, which is the sum of total abatement cost and total emission cost.

For a representative type 1 firm, the optimal emissions if not participating is e_1^N where MAC_1 equals MC_N at point b_1; the optimal emissions if participating is e_1^P since MAC_1 equals MC_P at point c_1; and the pre-program optimal emissions is e_1^0 since MAC_1 equals MC_0 at point d_1. The total environment cost for firm 1 is $Area(0a_N b_1 f_1)$ if not participating, and $Area(0a_P c_1 f_1)$ if participating. The difference in total emission cost between non-participation and non-participation is

$$\Delta_1 = Area(0a_N b_1 f_1) - Area(0a_P c_1 f_1)$$
$$= Area(a_P a_N x_{NP}) - Area(X_{NP} c_1 b_1).$$

A type 1 firm will participate if $\Delta_1 \geq c_F$, and will not participate if $\Delta_1 < c_F$. A participating type 1 firm reduces $e_1^0 - e_1^P$ emissions, which is larger than a non-participating type 1 firm's reduction $e_1^0 - e_1^N$.

For a representative type 2 firm, the optimal emissions if not participating is e_2^N (MAC_2 equals MC_N at point b_2); the optimal emissions if participating is e_2^P (MAC_2 equals MC_P at point c_2); and the pre-program optimal emissions is e_2^0 (MAC_2 equals MC_0 at point d_2). The total emission cost for firm 2 is $Area(0a_N b_2 f_2)$ if not participating, and $Area(0a_P c_2 f_2)$ if participating. The difference in total emission cost between non-participation and participation is

$$\Delta_2 = Area(0a_N b_2 f_2) - Area(0a_P c_2 f_2)$$
$$= Area(a_P a_N b_2 c_2).$$

A type 2 firm will participate if $\Delta_2 \geq c_F$, and will not participate if $\Delta_2 < c_F$. A participating type 2 firm reduces $e_2^0 - e_2^P$ emissions, which is less than a non-participating type 2 firm's reduction $e_2^0 - e_2^N$.

Finally, for a representative type 3 firm, the optimal emissions if not participating is e_3^N; the optimal emissions if participating is e_3^P; and the pre-program optimal emissions is e_3^0. The total emission cost for firm 3 is $Area(0a_N b_3 f_3)$ if not participating, and $Area(0a_P c_3 f_3)$ if participating. The difference in total emission cost between non-participation and participation is

$$\Delta_3 = Area(0a_N b_3 f_3) - Area(0a_P c_3 f_3)$$
$$= Area(a_P a_N b_3 c_3).$$

A type 3 firm will participate if $\Delta_3 \geq c_F$, and will not participate if $\Delta_3 < c_F$. A participating type 3 firm increases emissions from e_3^0 to e_3^P, whereas a non-participating type 3 firm reduces emissions from e_3^0 to e_3^N.

In conclusion, our model suggests that as compared with the pre-program equilibrium, type 1 firms reduce their emissions regardless of whether they participate or not, and the reductions are larger for participants than for non-participants. Type 2 firms also reduce their emissions regardless of their decision to participate or not compared with the pre-program level. But type 2 firms reduce their emissions by a smaller amount when participating than non-participating. Compared with the pre-program equilibrium, non-participating type 3 firms reduce emissions whereas participating type 3 firms increases emissions. Therefore, with given regulation pressure and public pressure, a voluntary pollution abatement program is effective for type 1 firms and type 2 firms regardless of their participation decision, but it is effective for type 3 firms only if they do not participate. While both type 2 and type 3 participating firms are free-riders since they have higher emissions as compared with the corresponding non-participants, only the participating type 3 firms have higher emissions under the program than in the absence of the program.

In addition, our model implies that the key factor determining the effectiveness of voluntary pollution abatement programs is the public scrutiny of participating firms' environmental performance and the associated risk of being labelled greenwashers, which raises the marginal emission cost curve for participants and decreases their optimal emission level. Without this additional public scrutiny of participating firms, all participating firms will free-ride and increase their emissions due to a lower marginal emission cost (shown by the MC_P'' curve in Figure 2).

Our model has important implications for the empirical assessment of voluntary pollution abatement programs. Measuring the effectiveness of the program by comparing the emissions of participants with non-participants underestimates the real program effect. Under such comparison, the program appears to be effective only for type 1 firms, and appears counterproductive for type 2 and 3 firms. In this paper, we label the difference in emissions between participants and non-participants as the "participation effect". By doing so, we are able to identify the heterogeneous participation effects for different types of firms. However, participation effect is different from the overall program effect, which requires comparing emissions for participants and non-participants to their pre-program levels.[6]

[6] To assess the program effect, we need a control group of firms who are not affected by the voluntary pollution abatement program. Our data is from a nation wide voluntary pollution abatement program, so we are unable to identify the overall program effect.

4 Theoretical Model

4.1 Model Overview

We characterize each firm by θ, which represents firm characteristics such as managerial skill, size, energy type, equipment or ownership, and is correlated with the firm's emissions and participation choice. Let a firm's abatement cost be a function of firm characteristics and emissions, i.e, $F(\theta, e)$, where e is the firm's emissions due to its choice of input bundle. Let the firm's marginal abatement cost function be $f(\theta, e) = \frac{\partial F(\theta,e)}{\partial e}$. Abatement costs decrease with emissions, so $f(\theta, e) < 0$. Marginal abatement cost decreases as the amount of abatement decreases, which is equivalent to assuming the marginal abatement cost function is an increasing function of emissions, so $\frac{\partial f(\theta,e)}{\partial e} > 0$. For simplicity of illustration, θ is a scalar and denotes the "abatement efficiency" of the firm.[7] Specifically, a higher value of θ means lower total abatement cost and marginal abatement cost at any emission level, so $\frac{\partial F(\theta,e)}{\partial \theta} \leq 0$ and $\frac{\partial^2 F(\theta,e)}{\partial e \partial \theta} \geq 0$.[8] Because lower e means more abatement, these assumptions imply that both total abatement cost and marginal abatement cost increases with every one additional unit reduction of emission.

The firm's emission cost function is

$$C(x, z, e, p) = R(x, e, p) + M(z, e, p) \tag{1}$$

where p is the firm's participation status with $p = 1$ representing participation in the voluntary program and $p = 0$ representing non-participation. $R(x, e, p)$ is the emission cost due to regulation pressure, and $M(z, e, p)$ is the emission cost from public pressure. e is the firm's emission level, x is a vector of firm and regulator characteristics determining regulation pressure, and z is a vector of industry and stakeholders' characteristics related to public pressure, which includes industry reputation, media and press releases, the share of green consumers and investors, and local communities' characteristics. Similar to θ, for simplicity of model illustration, x and z are scalars representing the effectiveness of regulation pressure and public pressure. A greater value of x and z means the firm faces more stringent regulators and stakeholders with regards to its environmental performance. For the regulation component of emission cost, we assume

[7] See also Zhou et al. (2020).

[8] See Appendix A.1 for details. Because $\frac{\partial F(\theta,e)}{\partial e} < 0$, a lower marginal abatement cost means $\frac{\partial F(\theta,e)}{\partial e}$ becomes less negative, so $\frac{\partial^2 F(\theta,e)}{\partial e \partial \theta} \geq 0$.

that $\frac{\partial R}{\partial x} \geq 0$, $\frac{\partial R}{\partial e} \geq 0$, and $R(x, e, p = 0) \geq R(x, e, p = 1)$. For the public component of emission cost, we assume that $\frac{\partial M}{\partial z} \geq 0$ and $\frac{\partial M}{\partial e} \geq 0$. The relationship between $M(z, e, p = 0)$ and $M(z, e, p = 1)$ is ambiguous depending on the value of e.

We can separate $R(x, e, p)$ and $M(z, e, p)$ into a more specific expression:

$$\begin{aligned} C(x, z, e, p) =& R_0(x, e) + M_0(z, e) \\ & + p(R_P(e) + M_P(e) + G(e)) + (1 - p)(R_N(e) + M_N(e)), \end{aligned} \tag{2}$$

where $R_0(x, e)$ and $M_0(z, e)$ are background (pre-program) emission cost from regulation pressure and public pressure, respectively. $R_P(e)$ and $R_N(e)$ are the changes in emission cost from the changes in regulation pressure, and $M_P(e)$ and $M_N(e)$ are the changes in emission cost from the changes in public pressure due to the firm's participation status, but without considering the cost from public scrutiny of participants. We have $R_P(e) \leq 0$, $R_N(e) \geq 0$, $M_P(e) \leq 0$, and $M_N(e) \geq 0$. $G(e)$ is the additional emission cost from public scrutiny of participants, so $G(e) \geq 0$. This cost increases with the firm's emission level e. We have $G(0) = 0$, because a firm without any emission does not attract public scrutiny of its environmental performance.

We make the following assumptions regarding each component of the emission cost function. First, we assume that the pre-program total and marginal emission cost from regulation pressure increase both in the effectiveness of regulation pressure x and the firm's emission level e, so $\frac{\partial R_0(x,e)}{\partial x} \geq 0$, $\frac{\partial r_0(x,e)}{\partial x} \geq 0$, $\frac{\partial R_0(x,e)}{\partial e} \geq 0$, and $\frac{\partial r_0(x,e)}{\partial e} \geq 0$, where $r_0(x, e) = \frac{\partial R_0(x,e)}{\partial e}$. Second, we assume that the pre-program total and marginal emission cost from public pressure increases in the effectiveness of public pressure z and the firm's emission level e, so $\frac{\partial M_0(z,e)}{\partial z} \geq 0$, $\frac{\partial m_0(z,e)}{\partial z} \geq 0$, $\frac{\partial M_0(z,e)}{\partial e} \geq 0$, and $\frac{\partial m_0(z,e)}{\partial e} \geq 0$, where $m_0(z, e) = \frac{\partial M_0(z,e)}{\partial e}$. Third, recall that participants experience a decrease in regulation pressure (a downward shift to MC'_P in Figure 1), and that the shift is greater for smaller emitters than for larger emitters and gradually converges to zero as emissions go to infinity, so that $\frac{\partial R_P(e)}{\partial e} \leq 0$ and $\frac{\partial r_P(e)}{\partial e} \geq 0$, where $r_P(e) = \frac{\partial R_P(e)}{\partial e}$. Fourth, we assume that without considering the additional public evaluation of participants emissions outcome, there is a decline in public pressure (a downward shift to MC''_P in Figure 2) and the shift is greater for smaller emitters than for larger emitters and gradually converges to zero with emissions, so that $\frac{\partial M_P(e)}{\partial e} \leq 0$ and $\frac{\partial m_P(e)}{\partial e} \geq 0$, where $m_P(e) = \frac{\partial M_P(e)}{\partial e}$. Fifth, assuming non-participants face both higher regulatory pressure and higher public pressure (upward shifts to MC'_N in Figure 1 and MC_N in Figure 2), and both shifts are greater for more polluting firms, we have $\frac{\partial R_N(e)}{\partial e} \geq 0$, $\frac{\partial M_N(e)}{\partial e} \geq 0$, $\frac{\partial r_N(e)}{\partial e} \geq 0$, and $\frac{\partial m_N(e)}{\partial e} \geq 0$, where $r_N(e) = \frac{\partial R_N(e)}{\partial e}$ and $m_N(e) = \frac{\partial M_N(e)}{\partial e}$. Sixth, letting $g(e) = \frac{\partial G(e)}{\partial e}$, we have $\frac{\partial G(e)}{\partial e} \geq 0$ and $\frac{\partial g(e)}{\partial e} \geq 0$, because the cost from additional scrutiny of participants increases as an increasing function of firm emissions. All the assumptions are subject to $\frac{\partial C(x,z,e,p)}{\partial e} \geq 0$, since the marginal emission cost is always non-negative. Therefore, the marginal emission cost can be written as

$$\begin{aligned} c(x,z,e,p) =& r(x,e,p) + m(z,e,p) \\ =& r_0(x,e) + m_0(z,e) + p(r_P(e) + m_P(e) + g(e)) \\ &+ (1-p)(r_N(e) + m_N(e)). \end{aligned} \tag{3}$$

Based on the properties above, we have $\frac{\partial c(x,z,e,p)}{\partial e} \geq 0$, regardless of the participation status p. That is the marginal emission cost is increasing in emissions for both participants and non-participants. For participants, the portion of marginal emission cost generated by the public evaluation of its environmental performance and the associated risk of been labeled a greenwasher increases rapidly with emissions, such that it eventually exceeds the difference between MC_N and MC''_P at some threshold emission level (labeled by e'' in Figure 3). Let $\Delta MC''(e) = r_N(e) + m_N(e) - r_P(e) - m_P(e)$ be the difference between MC_N and MC''_P. Accordingly, we have the following property:

$$\frac{\partial g(e)}{\partial e} > \frac{\partial MC''(e)}{\partial e}, \tag{4}$$

and $c(x,z,e,p=1) > c(x,z,e,p=0)$ if and only if $e > e''$.

The firm makes participation and emission decisions to minimize its total emission cost described above. As mentioned in the previous section, we define a participating firm to be a free-rider of the voluntary pollution abatement program if its emissions are higher than the emissions of non-participants, shown as the type 2 and 3 firms in Figure 3.

4.2 Emission Decision

A firm seeks to minimize its total environmental cost $D(\theta,x,z,e,p)$, which is the sum of total abatement cost, total emission cost, and a fixed cost if the firm decides to participate in the voluntary pollution abatement program:

$$\min_{e,p} D(\theta, x, z, e, p) = F(\theta, e) + C(x, z, e, p) + pc_F, \tag{5}$$

where c_F is the fixed cost of participation. In the first stage the firm identifies its optimal emissions, $e^*(\theta, x, z, p)$, under both participation and non-participation, according to the first order condition

$$f(\theta, e) + c(x, z, e, p) = 0, \ \Rightarrow \ e^* = e^*(\theta, x, z, p). \tag{6}$$

The optimal emission level $e^*(\theta, x, z, p)$ is determined by continuous variables x, z, θ and binary participation status variable p. Based on our assumptions and model properties, we have the following comparative statistics results describing the marginal effect of each variable on the optimal emission level. First, for the continuous variables x, z and θ, it can be shown that $\frac{\partial e^*}{\partial x} \leq 0$, $\frac{\partial e^*}{\partial z} \leq 0$, and $\frac{\partial e^*}{\partial \theta} \leq 0$.[9] In other words, our model indicates that regardless of participation status, firm's emissions decline when there is an increase in regulation pressure, public pressure, or abatement efficiency. The impact of the binary participation variable p on the optimal emission level is the labeled participation effect, and depends on the threshold emission level e''. As shown in Figure 3, there is an emission level e'' at which MC_P crosses MC_N, so that

$$c(x, z, e'', p = 1) = c(x, z, e'', p = 0), \tag{7}$$

which is equivalent to

$$r_N(e'') + m_N(e'') - r_P(e'') - m_P(e'') - g(e'') = 0. \tag{8}$$

Let $e^N = e^*(\theta, x, z, p = 0)$ and $e^P = e^*(\theta, x, z, p = 1)$ be the two potential optimal emission levels for the same firm under different participation status, and $\Delta e^*(\theta, x, z) = e^*(\theta, x, z, p = 1) - e^*(\theta, x, z, p = 0)$ be the difference between these two potential emission levels or the participation effect. It can be shown that $e^N > e''$ if and only if $e^P > e''$, and $e^N \leq e''$ if and only if $e^P \leq e''$.[10] Let $\tilde{e}(\theta, x, z)$ be the observed emission level regardless of the firm's participation decision. In particular, $\tilde{e}(\theta, x, z) = e^P$ if the firm participates, and $\tilde{e}(\theta, x, z) = e^N$ if the firm

[9]See Appendix A.2.
[10]See Appendix A.3.

does not participate. Because $\tilde{e}(\theta, x, z)$ is a special case of $e^*(\theta, x, z, p)$, it has the same properties as $e^*(\theta, x, z, p)$: $\frac{\partial \tilde{e}}{\partial x} \leq 0$, $\frac{\partial \tilde{e}}{\partial z} \leq 0$, and $\frac{\partial \tilde{e}}{\partial \theta} \leq 0$. It can be shown that $\Delta e^*(\theta, x, z) > 0$ if and only if $\tilde{e}(\theta, x, z) > e''$, and $\Delta e^*(\theta, x, z) \leq 0$ if and only if $\tilde{e}(\theta, x, z) \leq e''$.[11] These results lead to our first proposition:

Proposition 1. There exists a threshold emission level e'' such that firms with emissions higher than e'' emit less pollution if participating than not participating in the voluntary pollution abatement program, and vice versa. [12]

Proposition 1 identifies the free-riders of the voluntary pollution abatement program: these are the participating firms with emissions lower than e'' that increase their emissions relative to non-participants.

4.3 Participation Incentives

A firm makes its participation decision by comparing the optimal total costs under participation versus non-participation. If the firm participates, the optimal emission cost is

$$\begin{aligned} D^P &= F(\theta, e^P) + C(x, z, e^P, p = 1) + c_F \\ &= F(\theta, e^*(\theta, x, z, p = 1)) + C(x, z, e^*(\theta, x, z, p = 1), p = 1) + c_F. \end{aligned} \tag{9}$$

If the firm does not participate, the optimal emission cost is

$$\begin{aligned} D^N &= F(\theta, e^N) + C(x, z, e^N, p = 0) \\ &= F(\theta, e^*(\theta, x, z, p = 0)) + C(x, z, e^*(\theta, x, z, p = 0), p = 0). \end{aligned} \tag{10}$$

Denoting ΔD^* as the difference in the optimal cost between non-participation and participation, we have

$$\Delta D^* = \Big(F(\theta, e^N) + C(x, z, e^N, p = 0)\Big) - \Big(F(\theta, e^P) + C(x, z, e^P, p = 1)\Big) - c_F. \tag{11}$$

[11] See Appendix A.3 for proofs.
[12] See Appendix A.3 for proofs.

A firm will participate if $\Delta D \geq 0$, and will not participate if $\Delta D \leq 0$. A larger ΔD implies a greater participation incentive. Recall that $\tilde{e}(\theta, x, z)$ is the firm's emissions regardless of participation status. Applying the envelope theorem, we obtain the following results. If $\tilde{e} > e''$, we have $\frac{\partial \Delta D}{\partial x} > 0$, $\frac{\partial \Delta D}{\partial \theta} > 0$, and $\frac{\partial \Delta D}{\partial z} > 0$. If $\tilde{e} \leq e''$, we have $\frac{\partial \Delta D}{\partial x} \leq 0$, $\frac{\partial \Delta D}{\partial \theta} \leq 0$, and $\frac{\partial \Delta D}{\partial z} \leq 0$.[13] In conclusion, our model gives the following proposition:

Proposition 2. If the outcome emission level $\tilde{e} \leq e''$, then a marginal decrease in "abatement efficiency", regulation pressure or public pressure, *ceteris paribus*, increases the probability of participation, and vice versa.

Based on the two propositions presented in this section, we conclude that if the threshold emissions e'' is empirically identifiable, then the theory leads to testable hypotheses regarding the disparity in firm behavior across the threshold. In the remainder of the paper, we propose and then estimate an econometric model that tests the empirical validity of our theory. We use plant level data and assume that the participation and emission decisions are made by individual plants (Bi and Khanna, 2012; Zhou et al., 2020).

5 Econometric Model

Our econometric model follows the previous literature by using a two-stage estimation approach to solve the endogeneity from self-selection in participation. In the first stage, we estimate the participation incentives using a probit model, because participation status is a binary outcome. In the second stage, we estimate the participation effect on emissions. We include presumably exogenous variables in the first stage so that the probit model may serve as an instrument for endogenous participation in the second stage. The main empirical challenge is to identify the threshold emission level e''. We identify the threshold in the first stage using a grid search to find the e'' value that achieves the best goodness of fit (in terms of log likelihood) in a probit model. A similar method is used in Hansen (1999, 2000).

[13]See Appendix A.4 for details.

5.1 Threshold Identification

Hansen (1999, 2000) consider the following model

$$y_{it} = \beta_1 X_{it} \mathbb{1}_{q_{it} \leq \gamma} + \beta_2 X_{it} \mathbb{1}_{q_{it} > \gamma} + \mu_i + \epsilon_{it}, \tag{12}$$

where y_{it} is the dependent variable, μ_i is the fixed effect, and X_{it} is a vector of independent variables. q_{it} is an observed threshold variable. $\mathbb{1}_{q_{it} \leq \gamma}$ and $\mathbb{1}_{q_{it} > \gamma}$ are indicator functions. $\mathbb{1}_{q_{it} \leq \gamma} = 1$ if the threshold variable is less or equal to γ, and $\mathbb{1}_{q_{it} \leq \gamma} = 0$ if the threshold variable is greater than γ. $\mathbb{1}_{q_{it} > \gamma}$ is defined similarly. γ is the unknown threshold value. Let $\hat{\gamma}$ be the estimated value of γ, Hansen (1999, 2000) estimate $\hat{\gamma}$ via a grid search on all possible values of γ, so that the estimated $\hat{\gamma}$ minimizes the sum of square errors. In our empirical analysis, we identify the threshold value in the first stage probit regression via a grid search on all possible values of the threshold to find the value that maximizes the log likelihood.

It is difficult to derive the statistical inference of γ because it has a non-standard asymptotic distribution. Hansen (1999, 2000) use the "non-rejection region" method to form the confidence interval around $\hat{\gamma}$, using the likelihood ratio test on the null hypothesis $\gamma = \gamma'$, where γ' can be any possible threshold value. A 95% confidence interval consists of all values of γ' where the null hypothesis is not rejected at the 5% significance level. In our case, we calculate the following likelihood ratio statistics

$$LR(\gamma') = 2(L(\hat{\gamma}) - L(\gamma')), \tag{13}$$

where $L(\gamma)$ is the likelihood of the probit model with γ threshold. Hansen (1999, 2000) proved the statistics $LR(\gamma)$ asymptotically follow an exponential distribution ζ with cumulative density function[14]

$$P(\zeta \leq x) = (1 - exp(-\frac{x}{2}))^2. \tag{14}$$

Therefore, the critical value of the test statistics at the 5% significance level is 7.3523, and any γ' with $LR(\gamma') < 7.3523$ are in the 95% confidence interval of $\hat{\gamma}$.

[14] Please see Hansen (1999) appendix, proof of Theorem 1; and Hansen (2000) appendix, proof of Theorem 2 for the proof.

5.2 First stage, participation incentives

The first step is to identify the participation incentives. Following Bi and Khanna (2012) and Zhou et al. (2020), we assume that plants make the participation decision and choose emissions based on lagged information. According to proposition 2, when lagged emissions are lower than the threshold, participation incentives decrease in lagged "abatement efficiency", lagged regulation pressure and lagged public pressure, and vice versa. However, in the absence of detailed plant level data that can reliably isolate each of these three factors, we use lagged plant emissions to capture the relationship. As shown in Appendix A.2, holding everything else constant, the lagged emission level decreases in lagged "abatement efficiency", lagged regulation pressure and lagged public pressure. Therefore, we have the first testable hypothesis:

Hypothesis 1. If the lagged emission level $e_{t-1} \leq e''$, then a marginal increase in e_{t-1} increases the probability of participation. If the lagged emission level $e_{t-1} > e''$, a marginal increase in e_{t-1} decreases the probability of participation.

We use a probit model to estimate the plant's participation probability, allowing the probit model coefficients to be different on either side of the unknown threshold. Following Hansen (1999, 2000), we consider the threshold probit model:

$$\begin{aligned} p_{ijt} =& \beta_1 X_{ijt} \mathbb{1}_{e_{ij,t-1} \leq e''} + \beta_2 X_{ijt} \mathbb{1}_{e_{ij,t-1} > e''} + \zeta_1 Z_{ijt} \mathbb{1}_{e_{ij,t-1} \leq e''} + \zeta_2 Z_{ijt} \mathbb{1}_{e_{ij,t-1} > e''} \\ &+ Ind_j + u_t + \epsilon_{ijt}, \end{aligned} \tag{15}$$

where p_{ijt} is the net benefit from participation for plant i in industry j that joins the voluntary pollution abatement program in year t, X_{ijt} is a vector of covariates. Z_{ijt} is a vector of variables correlated with the participation incentives but not included in the second stage regression, and are presumably uncorrelated to plant emissions. $e_{ij,t-1}$ is lagged log emissions, e'' is the threshold value with the same unit as $e_{ij,t-1}$, Ind_j is the industry fixed effect, u_t is the year fixed effect, and $\epsilon_{ijt} \sim N(0,1)$. Plant i in industry j participates if $p_{ijt} > 0$, and does not participate if $p_{ijt} \leq 0$. $\mathbb{1}_{e_{ij,t-1} \leq e''}$ and $\mathbb{1}_{e_{ij,t-1} > e''}$ are indicator functions. $\mathbb{1}_{e_{ij,t-1} \leq e''} = 1$ if plant's lagged log emissions are less or equal to e'', and $\mathbb{1}_{e_{ij,t-1} \leq e''} = 0$ if plant's lagged log emissions are greater than e''. $\mathbb{1}_{e_{ij,t-1} > e''}$ is defined similarly. This model allows heterogeneous coefficients for plants with emissions on either side of the threshold e''.

The grid search process used to identify the threshold e'' is visualized in Figure 4, which plots the relationship between the threshold value and the log likelihood from the first stage for two different analysis samples. As the figures show, the estimated threshold maximizes the log likelihood. The horizontal dashed lines show the 95% confidence interval.

5.3 Second stage, participation effect in emissions

Following proposition 1, we obtain the following testable hypothesis:

Hypothesis 2. If the lagged emission level $e_{t-1} \leq e''$, then a participating plant reduces emissions by a smaller magnitude than non-participant (and is a free-rider). If the lagged emission level $e_{t-1} > e''$, then a participating plant reduces pollution by a larger amount than non-participant.

To test hypothesis 2, we use the second stage model from Bi and Khanna (2012) and Zhou et al. (2020), but allow the program effect to be heterogeneous on either side of the threshold e''. Consider the following dynamic panel model

$$\begin{aligned}\Delta e_{ijt} = &\rho \Delta e_{ij,t-1} + \beta \Delta X_{ijt} + \gamma_1 Part_{ijt} \mathbb{1}_{e_{ij,t-1} \leq e''} + \gamma_2 Part_{ijt} \mathbb{1}_{e_{ij,t-1} > e''} \\ &+ \Delta Ind_j + \Delta u_t + \Delta Ind\_Trend_{jt} + \Delta State\_Trend_{jt} + \Delta \epsilon_{ijt}.\end{aligned} \tag{16}$$

$\Delta e_{ijt} = e_{ijt} - e_{ij,t-1}$ is the first difference of log emissions for plant i in industry j in year t, $\Delta e_{ij,t-1}$ is the one year lag of Δe_{ijt}, X_{ijt} is the vector of covariates, $Part_{ijt}$ is the participation dummy, and Ind_Trend_{jt} and $State\_Trend_{jt}$ are the industry and state specific time trends. The threshold value e'' is identified in the first stage. The participation probability estimated in the first stage is used to instrument $Part_{ijt}$, and we follow Arellano and Bond (1991) and use the third-year lag of emissions $e_{ij,t-3}$ as the instrument for $\Delta e_{ij,t-1}$ to avoid serial correlation. This dynamic panel model can be estimated via a generalized method of moments (GMM) approach, where we use Hansen's J-statistics to test for over-identification, and Cragg-Donald Wald F-statistics to test weak identification.[15]

[15] We use the same GMM estimation method with published STATA code from Zhou et al. (2020), so the results are comparable.

6 Empirical Analysis

6.1 Data and Specifications

We obtain data on the EPA's 33/50 voluntary pollution abatement program from Zhou et al. (2020), which contains plant level information from the Toxics Release Inventory (TRI). The 33/50 program was the EPA's first voluntary pollution abatement program, and was initiated in 1991. The program goal was to reduce total emissions of 17 toxic chemicals by 33% by 1992 and by 50% by 1995, compared to the 1988 baseline (Vidovic and Khanna, 2007). Like Zhou et al. (2020), we assume the participation and emission decisions are made at individual plants. The 33/50 program lasted 5 years from 1991 to 1995, and new participants could join any time during the program period.

Our analysis sample and model specification closely follow Zhou et al. (2020). We define our study period to be 6 years starting from the first year of program initiation: 1996 is included to capture emission reduction in 1995. The full sample includes 39,201 annual observations for 8,670 TRI plants across 11 industries from 1991 to 1996. To estimate the participation effect (second stage), we use log annual emissions of the 17 chemicals covered by 33/50 program as the outcome variable. Our control variables include the ratio of hazardous air pollutants to all TRI emissions (HAP/TRI ratio), number of AFS inspections[16] related to air pollution, county non-attainment status (number of pollutants under non-attainment designation), state LCV score[17], and county income per capita. HAP/TRI ratio and number of inspections account for regulation pressure, while state LCV score and county income per capita account for public pressure. County non-attainment status is related to both regulation pressure and public pressure. We include three additional variables as instruments in the first stage participation incentive model: parent firm's first invitation group status, parent firm's pre-program 33/50 emissions changes, and plant's share in parent firm's total TRI emissions. Zhou et al. (2020) and Vidovic and Khanna (2007, 2012) suggest that a plant whose parent firm was in the first invitation group, or whose parent firm has greater pre-program 33/50 emissions reduction, is more likely to

[16] AFS inspections are mandatory air regulations inspections obtained from EPA's Aerometric Information Retrieval System (AIRS) Facility Subsystem database. Inspections are from both U.S. EPA and/or state and local air regulatory agencies.

[17] League of Conservation Voters Score, the average rating assigned to the officials in the Senate and the House of Representatives by the League of Conservative Voters (LCV, National Environmental Scorecard, www.lcv.org)

Table 1: Plants statistics by participation status

Year	New Participants.	New Invited Participants	New Non-invited Participants	All Participants	All Invited Participants	All Non-invited Participants
1991	396	298	98	396	298	98
1992	388	68	320	758	354	404
1993	86	23	63	769	347	422
1994	59	18	41	739	331	408
1995	7	1	6	682	311	371
1996	0	0	0	647	295	352
Total				936	408	528

Year	Non-Participants	Invited Non-Participants	Non-invited Non-Participants	Total
1991	7,217	2,698	4,519	7,613
1992	6,774	2,560	4,214	7,532
1993	5,999	2,290	3,709	6,768
1994	5,439	2,081	3,358	6,718
1995	5,050	1,913	3,137	5,732
1996	4,731	1,801	2,930	5,378
Total	7,734	2,898	4,836	8,670

participate in the program.[18] A plant with larger share of TRI emissions within its parent firm is also more likely to participate.[19]

Table 1 summarizes the plant participation numbers by year. We find the majority of participating plants joined the program in 1991. We do not observe any participants who quit the program, so we assume that once plants participate in the program, they do not quit. Accordingly, in the first stage, we exclude plants after they join the program. In the second stage, the instrumental variable (the estimated participation probabilities) holds constant after

[18]With parent firm in the first invitation group, plants may experience a greater relief in regulation pressure by participation. With parent firm having greater pre-program 33/50 emissions reduction, plants are more likely to achieve the 33/50 emission reduction target and have lower risk of being labeled greenwashers.

[19]A plant with larger share of TRI emissions within its parent firm may also have a greater relief in regulation pressure by participation.

Table 2: Summary Statistics

Variable	Participants	Non-participants	All	Participants vs Non-participants
Plant 33/50 Emissions (million lbs.)	0.114	0.079	0.082	0.035***
	(0.281)	(0.405)	(0.296)	
First Invitation Group	0.485	0.379	0.390	0.106***
	(0.500)	(0.485)	(0.488)	
Plant Share in Firm TRI Emissions	0.494	0.326	0.343	0.168***
	(0.448)	(0.401)	(0.409)	
Pre-program 33/50 Emissions Change (million lbs.)	-0.877	-0.387	−0.437	−0.490***
	(2.323)	(1.222)	(1.383)	
Lag HAP/TRI Ratio (%)	74.732	76.586	76.397	−1.639***
	(30.923)	(30.677)	(30.707)	
Lag AFS Inspection Number	0.504	0.401	0.412	0.108***
	(1.293)	(1.103)	(1.124)	
County Non-attainment Status	1.124	0.958	0.975	0.166***
	(1.130)	(1.140)	(1.140)	
State LCV Score	99.951	95.753	96.181	4.197***
	(38.854)	(40.250)	(40.129)	
County Income per Capita (Thousand $)	22.050	20.847	20.970	1.203***
	(4.823)	(4.609)	(4.645)	
Number of Plants	971	8,229	8,670	
Number of Observations	3,991	35,210	39,201	

Note: *p<0.1; **p<0.05; ***p<0.01
Column "Participants vs Non-participants" reports the difference in mean statistics between first two columns using a t-test.

Table 3: Participants 33/50 emissions by participation years

Participation Year	1991	1992	1993	1994	1995
Participants Mean 33/50 Emissions (million lbs.)	0.166	0.069	0.084	0.089	0.045
	(0.538)	(0.219)	(0.278)	(0.150)	(0.060)

plants joined the program. This approach is used by previous studies on the 33/50 program as well (Khanna and Damon, 1999; Gamper-Rabindran, 2006; Innes and Sam, 2008; Zhou et al., 2020).

Table 2 reports the summary statistics for other key variables. We find that, on average, participants have relatively higher 33/50 emissions than non-participants, a larger share of their parent firms' total TRI emissions, lagged HAP/TRI ratio, pre-program 33/50 emissions reduction and lagged number of inspections. Compared to non-participants, on average, participants are located in a county with more pollutants under non-attainment designation and higher income per capita, in a state with higher state LCV score, and more likely to have their parent firms in the first invitation group.

6.2 Results

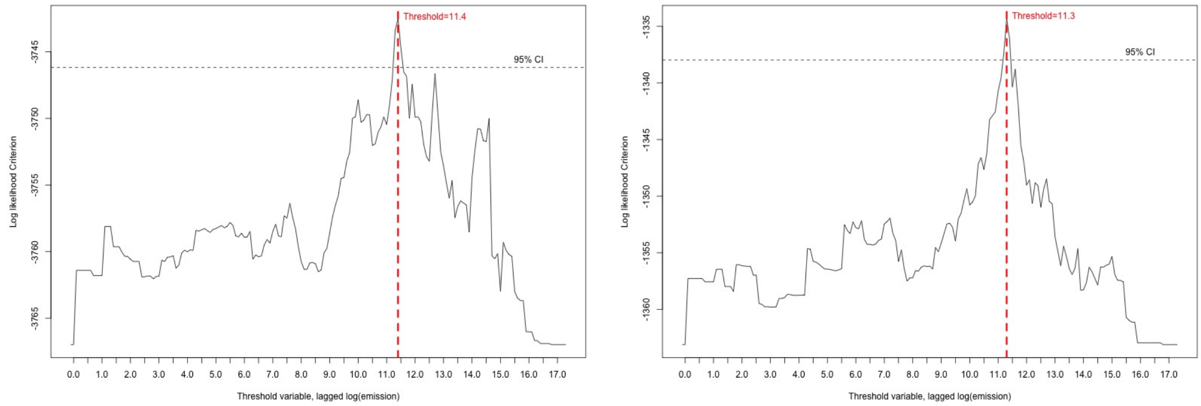
We consider two different samples in the analysis. The full sample includes all participants regardless of the year in which they joined the program. However, it is possible that regulation and public pressure varied over time, so that the threshold value may also vary across program years. To address this possibility we repeat the analysis with our preferred sample, including only the first year (1991) participants. That is, we estimate the threshold e'' and the heterogeneous program effects for the whole sample and our preferred sample separately. We define our preferred sample in such a way for two reasons. First, almost 50% of all participants joined the program in 1991. Second, compared to the later participants, the 1991 participants have much higher 33/50 emissions, as shown in Table 3. Thus, we expect that these plants faced the most significant and sustained public pressure during the program period.

First stage, participation incentives

We first consider the full sample with all participants from 1991 to 1995. Recall that participants never quit the program, so that, in the first stage, we drop participating plants' observations after they join the program, and keep all non-participating plants' observations. Because plants joined the program in different years, the year fixed effect is included in the probit regression. Figure 4a is the plot of the log likelihood as a function of the threshold parameter, generated by the threshold value grid search. It shows that there is a well defined global maximum at $e'' = 11.4$ (0.089 million lbs.), with 95% confidence interval $[11.3, 11.5]$. There are 26,470 observations from 7,815 plants with lagged log 33/50 emissions lower than the threshold and 6,376 observations from 2,220 plants with lagged log 33/50 emissions higher the threshold.

In Table 4, the first three columns show the regression results. Column 1 is the simple probit regression without considering the heterogeneity, column 2 and 3 are the threshold probit results, allowing heterogeneity on all covariates. By allowing heterogeneity in coefficient estimates, the log likelihood value improves from $-3,766.985$ to $-3,742.499$, though with a 10 degree of freedom loss. The threshold probit model also yields a better AIC than the simple probit model.

Comparing the estimated coefficients on either side of threshold, we find some sizable magnitude changes in "*First Invitation Group*" and "*County Income per Capita*", indicating that relatively polluting plants with lagged emissions higher than the threshold have greater participation motivation from their parent firms being in the first invitation group and from being located in a higher income community. Furthermore, non-attainment status of the county only affects the participation probability of plants with lagged emissions above the threshold, whereas the participation incentives for only plants with lagged emissions below the threshold is sensitive to the state LCV score. The estimated coefficients on "*Lag log 33/50 emissions*" for plants with emissions below the threshold is significantly positive, whereas it is insignificant for plants above the threshold. This implies that lagged emissions have a positive marginal effect on the net benefit of participation for plants with $e_{ij,-1} \leq e''$, and no effect for plants with $e_{ij,-1} > e''$. These results are broadly consistent with (ie, do not violate) Hypothesis 1.



(a) MLE criterion of threshold parameter, 33/50 program participation incentive in 1991-1995

(b) MLE criterion of threshold parameter, 33/50 program participation incentive in 1991

Figure 4: Threshold estimation and 95% confidence interval

In our preferred sample, we take a subsample of 1991 observations, and estimate the participation incentives based on plants' participation decisions only in 1991. Plants that do not participate in 1991 (even they participated in the later years) are included in the sample. We drop the year fixed effect from the probit models described in equation 15, because there is only one participation year (1991) in the sample. Figure 4b is the plot of log likelihood as a function of the threshold parameter. We obtain a threshold estimate that is quite similar to the one estimated with the whole sample: $e''_{1991} = 11.3$ (0.081 million lbs.), with 95% confidence interval $[11.2, 11.4]$. In Table 4, column 4 to column 6 show the regression results for the 1991 participation decision. As before, the threshold probit model yields a better fit. We find that there are some sizable magnitude changes in "*First Invitation Group*" and "*Plant Share in Firm TRI Emissions* ". In addition, only plants with lagged emissions lower than the threshold have a negatively significant coefficient on "*Lag HAP/TRI Ratio*", and only plants with lagged emissions higher than the threshold have a positively significant coefficient on "*County Income per Capita*". Notably, the first year participation results show that the marginal relationship between lagged 33/50 emission and net benefit of participation is significantly positive for plants with $e_{ij,-1} \leq e''$, and significantly negative for plants with $e_{ij,-1} > e''$, which strongly supports Hypothesis 1.

Appendix B shows the results for the first stage analysis for each of the later years. While the

Table 4: First stage: probit model of 33/50 program participation incentive

	Dependent variable: Participation Status					
	Full Sample	Full Sample		Preferred Sample	Preferred Sample	
		Before threshold	After threshold		Before threshold	After threshold
First Invitation Group	0.305^{***}	0.158^{***}	0.611^{***}	0.749^{***}	0.494^{***}	1.346^{***}
	(0.040)	(0.048)	(0.077)	(0.069)	(0.082)	(0.144)
Plant Share in Firm TRI Emissions	0.708^{***}	0.653^{***}	0.799^{***}	0.570^{***}	0.328^{***}	1.050^{***}
	(0.043)	(0.048)	(0.094)	(0.081)	(0.095)	(0.154)
Pre Program Reduction in 33/50 Emissions	-0.089^{***}	-0.100^{***}	-0.076^{***}	-0.160^{***}	-0.162^{***}	-0.176^{***}
	(0.011)	(0.014)	(0.017)	(0.015)	(0.020)	(0.023)
Lag HAP/TRI Ratio	0.000	0.000	-0.000	-0.002^{**}	-0.002^{**}	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Lag AFS Inspection Number	0.010	-0.000	0.006	0.026	0.018	0.016
	(0.017)	(0.025)	(0.023)	(0.029)	(0.043)	(0.041)
Lag Log 33/50 Emissions	0.032^{***}	0.021^{***}	0.003	0.046^{***}	0.032^{**}	-0.088^{**}
	(0.006)	(0.007)	(0.022)	(0.011)	(0.014)	(0.037)
County Non-attainment Status	0.017	0.001	0.060^{**}	0.007	-0.013	0.050
	(0.017)	(0.017)	(0.030)	(0.025)	(0.031)	(0.044)
State LCV Score	0.009^{***}	0.011^{***}	0.004	0.002	0.005	-0.004
	(0.002)	(0.002)	(0.004)	(0.004)	(0.004)	(0.006)
State LCV Score Square	-0.000^{***}	-0.000^{***}	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
County Income per Capita	0.013^{***}	0.009^{**}	0.027^{***}	0.015^{*}	0.000	0.048^{***}
	(0.004)	(0.005)	(0.008)	(0.008)	(0.010)	(0.014)
Year FE	Y	Y		N	N	
Industry FE	Y	Y		Y	Y	
Threshold log emission	-	11.4		-	11.3	
Threshold 95% C.I.	-	(11.3, 11.5)		-	(11.2, 11.4)	
Plants	8,880	8,880		8,401	8,401	
Observations	32,846	32,846		8,401	8,401	
Plants before threshold	-	7,815		-	6,438	
Observations before threshold	-	26,470		-	6,438	
Plants after threshold	-	2,220		-	1,963	
Observations after threshold	-	6,376		-	1,963	
Log Likelihood	$-3,766.985$ (df=32,821)	$-3,742.499$ (df=32,811)		$-1,363.122$ (df=8,380)	$-1,334.306$ (df=8,370)	
Akaike Inf. Crit.	7,583.970	7,554.999		2,768.244	2,730.613	

Note: *p<0.1; **p<0.05; ***p<0.01

threshold values for later years are less precisely estimated, the results are generally consistent with Hypothesis 1.

Second stage, participation effect on emissions

In Table 5, the first three columns report the regression results for full sample. Column 1 shows the results without considering the heterogeneity, column 2 and 3 show the results of heterogeneous participation effect above and below the estimated threshold of 11.4. Comparing the two models, we find that except for participation status, all other covariates have very similar coefficients in terms of their signs and magnitudes. Specifically, we find lagged emissions is positively correlated with current emissions, and the HAP/TRI ratio is negatively correlated with current emissions. Emissions are not significantly correlated with LCV scores, county non-attainment status and county income per capita. The coefficients on these covariates are qualitatively identical to Zhou et al. (2020)'s results. Note that we do not include the number of AFS inspection as an independent variable because it captures the variation in regulation pressure, but including it does not significantly affect the magnitudes and statistical significance of the coefficients on other variables.

In terms of the participation effect, we find that without allowing for heterogeneity, the average participation effect is -0.556. Allowing for heterogeneity, although the participation effect is still significantly negative for all plants, the estimated coefficient for plants below the threshold is much smaller than that for plants above the threshold. The significant differences in magnitude across the threshold implies that compared to the non-participants, participating plants with emissions above the threshold reduce their emissions by a much larger amount than participating plants with emissions below the threshold, thus suggesting the possibility that there are free-riders among the latter plants. However, these results do not provide direct evidence, which may be because the threshold value varies across participation years, so that the first stage analysis on the full sample provides a relatively inaccurate threshold value.

In the second stage, our preferred sample includes all non-participants and only first year joiners (1991 participants), but excludes plants who participate in 1992 – 1995. Table 5, column 4 to 6 show the results. Once again we obtain very similar coefficients on the covariates other than participation status between the two models, and they are also similar to the full sample

Table 5: Main analysis: 1991-1996 participation effect (benchmark results)

	Dependent variable: log(33/50 emissions lbs.), first difference					
	Full Sample	Full Sample		Preferred Sample	Preferred Sample	
		Before threshold	After threshold		Before threshold	After threshold
Participation Status	-0.556^{***}	-0.183^{***}	-0.956^{***}	-0.287^{***}	0.758^{***}	-0.806^{***}
	(0.053)	(0.062)	(0.104)	(0.072)	(0.190)	(0.109)
Lag Log 33/50 Emissions	0.526^{***}	0.464^{***}		0.610^{***}	0.562^{***}	
	(0.052)	(0.049)		(0.062)	(0.058)	
HAP/TRI Ratio	-0.011^{***}	-0.010^{***}		-0.013^{***}	-0.012^{***}	
	(0.001)	(0.001)		(0.001)	(0.001)	
State LCV Score	0.003^{*}	0.003^{*}		0.004^{*}	0.003^{*}	
	(0.002)	(0.002)		(0.002)	(0.002)	
State LCV Score Square	-0.000	-0.000		-0.000	-0.000	
	(0.000)	(0.000)		(0.000)	(0.000)	
County Non-attainment Status	0.021	0.022		0.037	0.035	
	(0.030)	(0.030)		(0.033)	(0.032)	
County Income per Capita	0.018	0.014		0.031	0.030	
	(0.020)	(0.019)		(0.022)	(0.021)	
Industry Specific Time Trends	Y	Y		Y	Y	
State Specific Time Trends	Y	Y		Y	Y	
Threshold log emission	-	11.4		-	11.3	
Plant number	8,670	8,670		7,650	7,650	
Plant number before threshold	-	7,804		-	6,774	
Plant number after threshold	-	2,289		-	2,264	
Observations	39,201	39,201		35,385	35,385	
Observations before threshold	-	31,305		-	27,494	
Observations after threshold]	-	7,896		-	7,888	
Over-identification test (Hansen J Statistics P Value)	0.2794	0.3242		0.2105	0.2729	
Weak identification test (Cragg-Donald Wald F Statistics)	69.611	70.083		49.133	47.385	

Note: *p<0.1; **p<0.05; ***p<0.01

All regressors except participation status are in first differences.

IV's are first stage estimated participation probability and the third-year lag of log 33/50 emissions.

results.

Without heterogeneity, the preferred sample gives an average participation effect of -0.287. It is smaller than the full sample, but similar to Innes and Sam (2008) and Bi and Khanna (2012)'s results.[20] Accounting for heterogeneity, we find that the participation effect is significantly negative after the threshold but significantly positive before the threshold. That is, compared to the non-participants, participating plants with lagged emissions above the threshold reduce their emissions, whereas participating plants below the threshold increase their emissions. This results provide strong evidence in support of Hypothesis 2, suggesting that participants with low emissions are free-riders.

6.3 Interpreting the Average Treatment Effect

Since our outcome variable is measured in logs and as a first difference, we next translate the estimated coefficients into the participation effect in terms of emission levels. Let e^P_{ijt} and e^N_{ijt} be the potential emissions of plant i under participation and non-participation status, respectively. According to equation 16, we have

$$\begin{aligned} e^P_{ijt}/e_{ij,t-1} &= e^N_{ijt}/e_{ij,t-1} \times exp\Big(\gamma_1 \mathbb{1}_{e_{ij,t-1} \leq e''} + \gamma_2 \mathbb{1}_{e_{ij,t-1} > e''}\Big), \\ e^P_{ijt} &= e^N_{ijt} \times exp\Big(\gamma_1 \mathbb{1}_{e_{ij,t-1} \leq e''} + \gamma_2 \mathbb{1}_{e_{ij,t-1} > e''}\Big), \end{aligned} \tag{17}$$

where $e_{ij,t-1}$ is the previous year's 33/50 emissions, regardless of the previous year's participation status. Negative γ's mean participants have lower emissions than non-participants, which indicates that the program is relatively effective. Positive γ's mean participants have higher emissions than non-participants, which indicates free-riding. The results for the preferred sample show that the emissions of participants below the threshold are double those of non-participating plants (increase by 113.4%), and the participants above the threshold have emissions that are 45% of non-participating emissions (decrease by 55.3%). The results for the full sample show that, the participants below the threshold have emissions that are 83.3% of non-participating plants' emissions, which implies that the participants below the threshold reduce their emissions by 16.7% compared to the non-participants. The emissions of participants above the threshold are only 38.4% of non-participating plants' emissions, which is equivalent

[20] Bi and Khanna (2012) find the participation effect coefficient between -0.188 and -0.237. Innes and Sam (2008) find the 1992 participation effect coefficient is around -0.35.

Table 6: Robustness Check: 1991-1996 participation impact, no threshold in the first stage

	Dependent variable: log(33/50 emissions lbs.), first difference					
	Full Sample	Full Sample		Preferred Sample	Preferred Sample	
		Before threshold	After threshold		Before threshold	After threshold
Participation Status	−0.673***	−0.409***	−1.027***	−0.289***	0.669***	−0.852***
	(0.056)	(0.067)	(0.101)	(0.075)	(0.191)	(0.114)
Lag Log 33/50 Emissions	0.507***	0.451***		0.609***	0.554***	
	(0.052)	(0.048)		(0.062)	(0.057)	
HAP/TRI ratio	−0.011***	−0.010***		−0.013***	−0.012***	
	(0.001)	(0.001)		(0.001)	(0.001)	
State LCV Score	0.003*	0.003*		0.004*	0.003*	
	(0.002))	(0.002)		(0.002)	(0.002)	
State LCV Score Square	−0.000]	−0.000		−0.000	−0.000	
	(0.000)	(0.000)		(0.000)	(0.000)	
County Non-attainment Status	0.021	0.022		0.037	0.035	
	(0.030)	(0.030)		(0.033)	(0.032)	
County Income per Capita	0.018	0.015		0.031	0.030	
	(0.020)	(0.019)		(0.022)	(0.021)	
Industry Specific Time Trends	Y	Y		Y	Y	
State Specific Time Trends	Y	Y		Y	Y	
Threshold log emission	-	11.4		-	11.3	
Plant number	8,670	8,670		7,650	7,650	
Plant number before threshold	-	7,804		-	6,774	
Plant number after threshold	-	2,289		-	2,264	
Observations	39,201	39,201		35,385	35,385	
Observations before threshold	-	31,305		-	27,494	
Observations after threshold	-	7,896		-	7,888	
Over-identification test (Hansen J Statistics P Value)	0.2756	0.3084		0.2109	0.2707	
Weak identification test (Cragg-Donald Wald F Statistics)	71.858	73.695		49.429	49.443	

Note: *p<0.1; **p<0.05; ***p<0.01
All regressors except participation status are in first differences.
IV's are first stage estimated participation probability and the third-year lag of log 33/50 emissions.

to reducing emissions by 61.6%.

We repeat the second step analysis for other year's participants as well. We find that the difference between the full sample results and the preferred sample results are mainly driven by the group of 1992 joiners, which has a very unstable threshold value. We include these results and their discussion in Appendix B.

7 Robustness Check

7.1 Sensitivity to alternative first stage specifications

We conduct several robustness checks to assess the reliability of the heterogeneous pattern in program effects that we describe above. First, we examine if the participation effect is sensitive to different first stage specifications. In Table 6, we use the estimated probabilities from the simple probit model without considering the threshold heterogeneity in the first stage as the instrument in the second stage. The preferred sample results show that for both models, the participation effects are identical to those reported in Table 5. The participation effects for the full sample, while larger in magnitude, are quantitatively identical to the results in Table 5.

In addition, following Zhou et al. (2020), we exclude lagged emissions from the covariates in the participation stage. The results are in Table 7. We find that for both the full and the preferred samples, the simple model without the threshold effects yields estimated coefficients that are similar to the benchmark analysis in Table 5, though the magnitude of the participation effect is smaller. For the model with threshold effects, the participation effect is significantly negative for plants above the threshold but significantly positive below the threshold. These results support our hypothesis even more clearly than our benchmark results. We conclude that our results are robust and heterogeneous participation effects can be detected with different first stage specifications.

Table 7: Robustness Check: 1991-1996 participation impact, no threshold and lagged emissions in the first stage

	Dependent variable: log(33/50 emissions), first difference					
	Full Sample	Full Sample		Preferred Sample	Preferred Sample	
		Before threshold	After threshold		Before threshold	After threshold
Participation Status	-0.120^{**}	0.249^{***}	-0.840^{***}	-0.117^{***}	0.892^{***}	-0.816^{***}
	(0.049)	(0.070)	(0.094)	(0.076)	(0.204)	(0.116)
Lag Log 33/50 Emissions	0.598^{***}	0.484^{***}		0.636^{***}	0.563^{***}	
	(0.056)	(0.049)		(0.062)	(0.057)	
HAP/TRI ratio	-0.012^{***}	-0.011^{***}		-0.013^{***}	-0.012^{***}	
	(0.001)	(0.001)		(0.001)	(0.001)	
State LCV Score	0.003^{*}	0.003^{*}		0.004^{*}	0.003^{*}	
	(0.002))	(0.002)		(0.002)	(0.002)	
State LCV Score Square	−0.000]	−0.000		−0.000	−0.000	
	(0.000)	(0.000)		(0.000)	(0.000)	
County Non-attainment Status	0.022	0.023		0.037	0.035	
	(0.031)	(0.030)		(0.033)	(0.032)	
County Income per Capita	0.018	0.012		0.032	0.031	
	(0.020)	(0.020)		(0.022)	(0.021)	
Industry Specific Time Trends	Y	Y		Y	Y	
State Specific Time Trends	Y	Y		Y	Y	
Threshold log emission	-	11.4		-	11.3	
Plant number	8,670	8,670		7,650	7,650	
Plant number before threshold	-	7,804		-	6,774	
Plant number after threshold	-	2,289		-	2,264	
Observations	39,201	39,201		35,385	35,385	
Observations before threshold	-	31,305		-	27,494	
Observations after threshold	-	7,896		-	7,888	
Over-identification test (Hansen J Statistics P Value)	0.2862	0.3582		0.2108	0.2800	
Weak identification test (Cragg-Donald Wald F Statistics)	60.456	68.286		47.882	49.530	

Note: *p<0.1; **p<0.05; ***p<0.01
All regressors except participation status are in first differences.
IV's are first stage estimated participation probability and the third-year lag of log 33/50 emissions.

7.2 Sensitivity to technology spillovers

Zhou et al. (2020) highlight the importance of technology spillovers in the 33/50 program. Therefore, as a second robustness check, we incorporate technology spillovers in our analysis. Specifically, we use Zhou et al. (2020)'s first stage spillovers model to instrument endogenous participation, and include technology spillovers in the second stage regression as well. That is, we estimate the following first stage model:

$$p_{ijt} = \alpha X_{ijt} + \sigma Z_{ijt} + \eta S_{j,t-1} + Ind_j + u_t + \epsilon_{it}, \tag{18}$$

where, as before, p_{ijt} is the net benefit of participation, X_{ijt} is a vector of covariates, and Z_{ijt} is a vector of variables that are not included in the second stage regression. $S_{j,t-1}$ is the lagged fraction of plants in industry j that participate in the program, capturing the benefit of technology spillovers from other participants in the same industry. For details see Zhou et al. (2020).

We again allow heterogeneity in the second stage model, and use the threshold values identified in Table 4.[21] The estimation equation for the second stage is as follows:

$$\begin{aligned}\Delta y_{ijt} =& \rho \Delta y_{ij,t-1} + \beta \Delta X_{ijt} + \gamma_1 Part_{ijt} \mathbb{1}_{e_{ij,t-1} \leq e''} + \gamma_2 Part_{ijt} \mathbb{1}_{e_{ij,t-1} > e''} \\ &+ \psi_1 (1 - Part_{ijt}) D_{ij,t-1} + \Delta \sum_{v=1991}^{t} \psi_v (1 - Part_{ijt})(1 - D_{ij,t-1}) S_{j,t-1} \\ &+ \Delta Ind_j + \Delta u_t + \Delta Ind\_Trend_{jt} + \Delta State\_Trend_{jt} + \Delta \epsilon_{ijt}.,\end{aligned} \tag{19}$$

where $D_{ij,t-1} = 1$ if a plant i's parent firm has at least one participating plant in the same industry j in year $t-1$, and 0 otherwise. $(1 - Part_{ijt})D_{ij,t-1}$ and $(1 - Part_{ijt})(1 - D_{ij,t-1})S_{j,t-1}$ captures intra-firm and inter-firm technology spillovers respectively, ψ_1 and ψ_v are the corresponding spillover effects in emissions.

Table 8 reports the results with technology spillovers. Consistent with the benchmark analysis, we find a heterogeneous pattern in the participation effect. In addition, the spillover effects in Table 8 are identical to Zhou et al. (2020)'s results. We conclude that our results are not sensitive to technology spillovers.

[21] The threshold value is exogenously determined by the nature of the voluntary abatement program, so its value should not be sensitive to the first stage model.

Table 8: Robustness Check: 1991-1996 participation impact, considering spillovers

	Dependent variable: log(33/50 emissions, first difference)					
	Full Sample	Full Sample		Preferred Sample	Preferred Sample	
		Before threshold	After threshold		Before threshold	After threshold
Participation Status	-0.348^{***}	0.019	-0.995^{***}	-0.380^{***}	0.418^{***}	-1.166^{***}
	(0.084)	(0.095)	(0.113)	(0.102)	(0.157)	(0.136)
Lag Log 33/50 emissions	0.603^{***}	0.489^{***}		0.636^{***}	0.531^{***}	
	(0.055)	(0.049)		(0.061)	(0.055)	
Intra-firm spillover years since sibling's participation	-0.316^{***}	-0.349^{***}		-0.305^{***}	-0.389^{***}	
	(0.090)	(0.087)		(0.090)	(0.092)	
Inter-firm spillover 1992	-4.010^{***}	-4.473^{***}		-3.862^{**}	-3.801^{***}	
	(1.500)	(1.427)		(1.506)	(1.450)	
Inter-firm spillover 1993	-2.607^{***}	-2.276^{**}		-2.465^{**}	-2.339^{**}	
	(0.969)	(0.916)		(0.989)	(0.957)	
Inter-firm spillover 1994	-3.723^{***}	-3.829^{***}		-4.014^{***}	-3.530^{***}	
	(0.834)	(0.793)		(0.871)	(0.847)	
Inter-firm spillover 1995	-1.415^{*}	-1.027		-1.258	-0.986	
	(0.846)	(0.812)		(0.903)	(0.883)	
Inter-firm spillover 1996	-2.308^{***}	-3.107^{***}		-2.744^{***}	-3.189^{***}	
	(0.689)	(0.834)		(0.853)	(0.922)	
HAP/TRI ratio	-0.012^{***}	-0.011^{***}		-0.011^{***}	-0.012^{***}	
	(0.001)	(0.001)		(0.001)	(0.001)	
State LCV Score	0.003^{*}	0.003^{*}		0.004^{*}	0.003^{*}	
	(0.002)	(0.002)		(0.002)	(0.002)	
State LCV Score Square	-0.000	-0.000		-0.000	-0.000	
	(0.000)	(0.000)		(0.000)	(0.000)	
County Non-attainment Status	0.021	0.021		0.036	0.033	
	(0.031)	(0.030)		(0.033)	(0.032)	
County Income per Capita	0.018	0.012		0.033	0.028	
	(0.020)	(0.020)		(0.022)	(0.021)	
Industry Specific Time Trends	Y	Y		Y	Y	
State Specific Time Trends	Y	Y		Y	Y	
Threshold log emission	-	11.4		-	11.3	
Plant number	8,657	8,657		8,091	7,650	
Plant number before threshold	-	7,729		-	6,774	
Plant number after threshold	-	2,288		-	2,264	
Observations	39,187	39,187		35,385	35,385	
Observations before threshold	-	31,292		-	27,494	
Observations after threshold	-	7,895		-	7,888	
Over-identification test (Hansen J Statistics P Value)	0.1792	0.2306		0.1204	0.1641	
Weak identification test (Cragg-Donald Wald F Statistics)	33.824	40.079		27.006	31.133	

Note: *p<0.1; **p<0.05; ***p<0.01
All regressors except participation status are in first differences.
IV's are first stage estimated participation probability and the third-year lag of log 33/50 emissions.

8 Conclusion

Voluntary abatement programs have been widely adopted by regulators in situations where traditional mandatory regulations are insufficient or absent. Unlike mandatory regulation which requires a huge commitment of resources to implement monitoring and possible inspections of polluters with correspondingly high pollution abatement costs of the regulated entities, voluntary abatement programs are associated with lower costs for both polluters and regulators (Brouhle et al., 2005). We show that public pressure, especially the additional public evaluation of firms' environmental outcomes and the associated risk of being labeled greenwashers by stakeholders, plays an important role in voluntary pollution abatement programs. We develop a theoretical model illustrating a firm's incentives to participate in a voluntary pollution abatement program and outline the firm's emission decisions. Our model shows that the scrutiny of participating firms' environmental outcomes by the public is the key factor in determining whether voluntary pollution abatement program participants free-ride or reduce emissions in concert with the goal of the program. Specifically, because public scrutiny of participants' environmental performance and the associated risk of being labeled greenwashers rises rapidly with emissions, the largest polluters, i.e., participants with emissions above an empirically identified threshold, are incentivized to reduce their emissions. It is only the relatively less polluting participants that have an incentive to free-ride and increase their emissions under the program. In contrast, in the absence of the cost of public scrutiny and the associated risk of being labeled as greenwashers, all participants will free-ride and increase their emissions under the program. We also show that because of an increase in both formal regulation pressure and stakeholder public pressure, all non-participants lower their emissions relative to the pre-program level. Thus, in aggregate terms, whether the voluntary abatement program is effective in reducing the targeted emissions depends on the distribution of participating plants.

To test our theory, we estimate an econometric model based on Zhou et al. (2020), the most recent study in voluntary pollution abatement literature published in this journal. We empirically identify the threshold emission level as well as the heterogeneous participation effects on either side of the threshold. We use plant level TRI data on participation in the EPA's 33/50 program from Zhou et al. (2020). Our empirical results are generally consistent with our theory and indicate that participants with emissions lower than the threshold double their emissions, compare to their counterfactual non-participating emissions; whereas the participants with emissions higher than the threshold significantly reduce their emissions.

Our study provides a plausible explanation for the long standing conundrum in the literature about the effectiveness of voluntary pollution abatement programs. The heterogeneity in the effectiveness of voluntary pollution abatement programs for large and small polluters suggests that the inconsistent results reported in the literature might be explained by sample differences, which lead to different threshold values and emission distributions. In addition, because voluntary pollution abatement programs affect the emissions of both participants and non-participants, we point to a possible under-estimation of program effects by the previous literature (see also Zhou et al. (2020)).

Our study has important policy implications and highlights the power of public pressure. It suggests that public pressure, such as through the public scrutiny and the associated risk of being labeled greenwashers, can be effectively leveraged to complement traditional regulation and may be sufficient to keep firms from free-riding in voluntary pollution abatement programs. This points to the role of media coverage and environmental information disclosure to the public, NGOs, and other relevant stakeholders.

A Mathematic Appendix

A.1 Properties of abatement cost function

Let a be the abatement quantity, and $\bar{e}$ be the maximum emissions, we have $e = \bar{e} - a$. The abatement cost function can be written as

$$F(\theta, e) = F(\theta, \bar{e} - a).$$

The abatement cost increases by abatement quantity but decreases by "abatement efficiency", so $\frac{\partial F(\theta,\bar{e}-a)}{\partial a} \geq 0$ and $\frac{\partial F(\theta,\bar{e}-a)}{\partial \theta} \leq 0$. We assume the marginal abatement cost increases in abatement quantity, but decreases in "abatement efficiency", so $\frac{\partial^2 F(\theta,\bar{e}-a)}{\partial a^2} \geq 0$, $\frac{\partial^2 F(\theta,\bar{e}-a)}{\partial a \partial \theta} \leq 0$. Because $e = \bar{e} - a$, $\frac{\partial a}{\partial e} < 0$, we have the following equations:

$$\begin{aligned}
f(\theta, e) = & \frac{\partial F(\theta, e)}{\partial e} = \frac{\partial F(\theta, \bar{e} - a)}{\partial a} \frac{\partial a}{\partial e} \leq 0, \\
& \frac{\partial f(\theta, e)}{\partial e} = \frac{\partial^2 F(\theta, \bar{e} - a)}{\partial a^2} \Big(\frac{\partial a}{\partial e} \Big)^2 \geq 0, \\
& \frac{\partial F(\theta, e)}{\partial \theta} = \frac{\partial F(\theta, \bar{e} - a)}{\partial \theta} \leq 0, \\
& \frac{\partial^2 F(\theta, e)}{\partial e \partial \theta} = \frac{\partial^2 F(\theta, \bar{e} - a)}{\partial a \partial \theta} \frac{\partial a}{\partial e} \geq 0.
\end{aligned}$$

A.2 Comparative statistics for continuous variables

Taking the first derivative on equation 6 with respect to continuous variables (x, z, θ,), we get:

$$\begin{aligned}
\theta: & \quad \frac{\partial c(x, z, e^*, p)}{\partial e^*} \frac{\partial e^*}{\partial \theta} + \frac{\partial f(\theta, e^*)}{\partial e^*} \frac{\partial e^*}{\partial \theta} + \frac{\partial f(\theta, e^*)}{\partial \theta} = 0 \\
x: & \quad \frac{\partial c(x, z, e^*, p)}{\partial e^*} \frac{\partial e^*}{\partial x} + \frac{\partial f(\theta, e^*)}{\partial e^*} \frac{\partial e^*}{\partial x} + \frac{\partial r_0(x, e^*)}{\partial x} = 0 \\
z: & \quad \frac{\partial c(x, z, e^*, p)}{\partial e^*} \frac{\partial e^*}{\partial z} + \frac{\partial f(\theta, e^*)}{\partial e^*} \frac{\partial e^*}{\partial z} + \frac{\partial m_0(z, e^*)}{\partial z} = 0.
\end{aligned} \tag{20}$$

We have the following result:

- $\frac{\partial c(x,z,e,p)}{\partial e} \geq 0$, $\frac{\partial f(\theta,e)}{\partial e} > 0$, and $\frac{\partial f(\theta,e)}{\partial \theta} \geq 0$. $\rightarrow$ $\frac{\partial e^*}{\partial \theta} < 0$.
- $\frac{\partial c(x,z,e,p)}{\partial e} \geq 0$, $\frac{\partial f(\theta,e)}{\partial e} > 0$, and $\frac{\partial r_0(x,e)}{\partial x} \geq 0$. $\rightarrow$ $\frac{\partial e^*}{\partial x} < 0$.
- $\frac{\partial c(x,z,e,p)}{\partial e} \geq 0$, $\frac{\partial f(\theta,e)}{\partial e} > 0$, and $\frac{\partial m_0(z,e)}{\partial z} \geq 0$. $\rightarrow$ $\frac{\partial e^*}{\partial z} < 0$.

A.3 Comparative statistics for participation status and proof of Proposition 1

Because p is a binary variable, we use the following integral to show how each term in equation 6 changes due to the change of p between 0 and 1:

$$\begin{aligned}
&1.\ p = 0 \rightarrow p = 1: \\
&\int_{e^N}^{e^P} \frac{\partial c(x,z,e,p=0)}{\partial e} de + \int_{e^N}^{e^P} \frac{\partial f(\theta,e)}{\partial e} de + \left[c(x,z,e^P,p=1) - c(x,z,e^P,p=0)\right] = 0, \\
&2.\ p = 1 \rightarrow p = 0: \\
&\int_{e^P}^{e^N} \frac{\partial c(x,z,e,p=1)}{\partial e} de + \int_{e^P}^{e^N} \frac{\partial f(\theta,e)}{\partial e} de + \left[c(x,z,e^N,p=0) - c(x,z,e^N,p=1)\right] = 0,
\end{aligned} \tag{21}$$

Because $\frac{\partial c(x,z,e,p)}{\partial e} \geq 0$ and $\frac{\partial f(\theta,e)}{\partial e} > 0$, $\Delta e^*(\theta,x,z) = e^P - e^N < 0$ if and only if

$$\begin{aligned}
&\int_{e^N}^{e^P} \frac{\partial c(x,z,e,p=0)}{\partial e} de + \int_{e^N}^{e^P} \frac{\partial f(\theta,e)}{\partial e} de < 0, \\
&\textbf{or} \\
&\int_{e^P}^{e^N} \frac{\partial c(x,z,e,p=0)}{\partial e} de + \int_{e^P}^{e^N} \frac{\partial f(\theta,e)}{\partial e} de > 0,
\end{aligned} \tag{22}$$

which is equivalent to

$$\begin{aligned}
&c(x,z,e^P,p=1) - c(x,z,e^P,p=0) \\
&= r_P(e^P) + m_P(e^P) + g(e^P) - r_N(e^P) - m_N(e^P) > 0, \\
&\textbf{or} \\
&c(x,z,e^N,p=0) - c(x,z,e^N,p=1) \\
&= r_N(e^N) + m_N(e^N) - r_P(e^N) - m_P(e^N) - g(e^N) < 0.
\end{aligned} \tag{23}$$

Recall that we use $\Delta MC''(e) = r_N(e) + m_N(e) - r_P(e) - m_P(e)$ to represent the gap between marginal emission cost between participants and non-participants without considering the cost of being labeled a greenwasher. Equation 23 is equivalent to

$$g(e^P) > \Delta MC''(e^P), \quad \textbf{or} \quad g(e^N) > \Delta MC''(e^N). \tag{24}$$

When emissions are at the threshold e'', we have $g(e'') = \Delta MC''(e'')$. Besides, $\frac{\partial g(e)}{\partial e} - \frac{\partial MC''(e)}{\partial e} > 0$. Therefore, for any emission level $\hat{e} > e''$ it is always true that $g(\hat{e}) > \Delta MC''(\hat{e})$, and vice versa. It can be proved that $e^N > e''$ if and only if $e^P > e''$, and $e^N \leq e''$ if and only if $e^P \leq e''$, because two conditions in equation 22, 23 and 24 are equivalent to each other.

Since $e^N > e''$ and $e^P > e''$ are equivalent, and so does $e^N \leq e''$ and $e^P \leq e''$, no matter which one is the observed emission level $\tilde{e}$, we always have $\Delta e^*(\theta, x, z) > 0$ if and only if $\tilde{e} > e''$, and $\Delta e^*(\theta, x, z) \leq 0$ if and only if $\tilde{e} \leq e''$. In other words, $e^N > e^P$ if and only if $\tilde{e} > e''$, and $e^N \leq e^P$ if and only if $\tilde{e} \leq e''$.

A.4 Proof of Proposition 2

Using envelope theorem, we have the following properties:

- $\frac{\partial D^*}{\partial \theta} = \frac{\partial F(\theta, e^*)}{\partial \theta} + [f(\theta, e^*) + c(x, z, e^*, p)]\frac{\partial e^*}{\partial \theta} = \frac{\partial F(\theta, e^*)}{\partial \theta} < 0,$
- $\frac{\partial D^*}{\partial x} = \frac{\partial C(x,z,e^*,p)}{\partial x} + [f(\theta, e^*) + c(x, z, e^*, p)]\frac{\partial e^*}{\partial x} = \frac{\partial R_0(x,e^*)}{\partial x} > 0,$
- $\frac{\partial D^*}{\partial z} = \frac{\partial C(x,z,e^*,p)}{\partial z} + [f(\theta, e^*) + c(x, z, e^*, p)]\frac{\partial e^*}{\partial z} = \frac{\partial m_0(z,e^*)}{\partial z} > 0,$

Let $\Delta D^* = D^*(\theta, x, z, P = 0) - D^*(\theta, x, z, P = 1)$, we have

$$\frac{\partial \Delta D^*}{\partial x} = \frac{\partial R_0(x, e^N)}{\partial x} - \frac{\partial R_0(x, e^P)}{\partial x}, \tag{25}$$

$$\frac{\partial \Delta D^*}{\partial \theta} = \frac{\partial F(\theta, e^N)}{\partial x} - \frac{\partial F(\theta, e^P)}{\partial x}, \tag{26}$$

$$\frac{\partial \Delta D^*}{\partial z} = \frac{\partial M_0(z, e^N)}{\partial z} - \frac{\partial M_0(z, e^P)}{\partial z}, \tag{27}$$

We show in Appendix A.3 that $\tilde{e} > e''$ is equivalent to $e^N > e^P$, and $\tilde{e} \leq e''$ is equivalent to $e^N \leq e^P$. We have the following results:

- Because $\frac{\partial R_0(x,e)^2}{\partial x \partial e} > 0$, we have $\frac{\partial \Delta D^*}{\partial x} > 0$ if $\tilde{e} > e''$, and $\frac{\partial \Delta D^*}{\partial x} \leq 0$ if $\tilde{e} \leq e''$.
- Because $\frac{\partial F(\theta,e)^2}{\partial \theta \partial e} > 0$, we have $\frac{\partial \Delta D^*}{\partial \theta} > 0$ if $\tilde{e} > e''$, and $\frac{\partial \Delta D^*}{\partial \theta} \leq 0$ if $\tilde{e} \leq e''$.
- Because $\frac{\partial M_0(z,e)^2}{\partial z \partial e} > 0$, we have $\frac{\partial \Delta D^*}{\partial z} > 0$ if $\tilde{e} > e''$, and $\frac{\partial \Delta D^*}{\partial z} \leq 0$ if $\tilde{e} \leq e''$.

B Additional Results

B.1 1992 participants

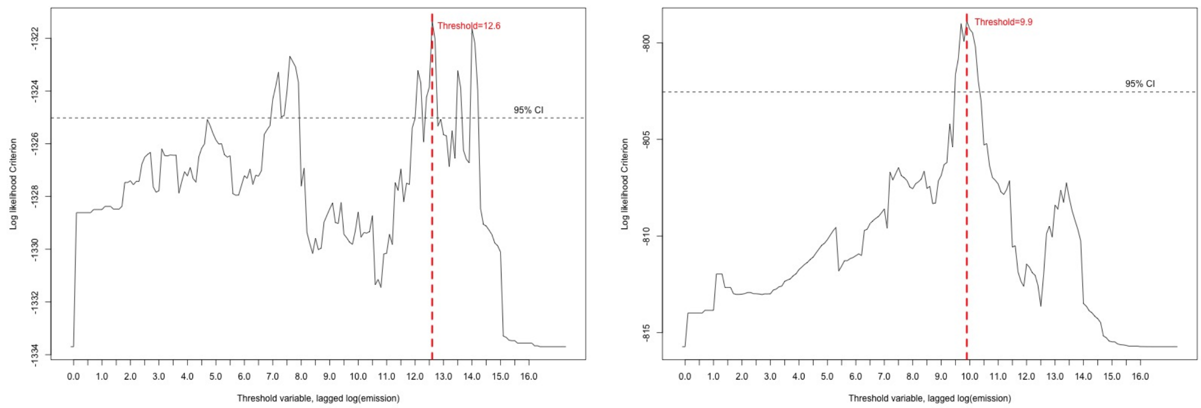
First stage, participation incentives

For the first stage analysis, we use a subsample of 1992 observations and drop all 1991 participants. Then we estimate the participation incentives based on remaining plants' participation decisions only in 1992. All plants that do not participate in 1992 (including those that participate in later years) are included in the sample. We drop the year fixed effect from the probit models described in equation 15, since there is only one year in the first stage sample. Figure 5a is the plot of log likelihood as a function of the threshold parameter. We find that although there is a global maximum at threshold value $e''_{1992} = 12.6$, it comes with a very large 95% confidence interval $[7.0, 14.2]$. This suggests that $e''_{1992} = 12.6$ may not be the unique global maximum.

Table 9 shows the results for the 1992 participation. The threshold probit models always yields a better fit. The sign of estimated coefficients on "*Lag Log 33/50 Emissions*" change across the different sides of the threshold. We find a positive marginal relationship between lagged emissions and net benefit of participation below the threshold (consistent with Hypothesis 1), while the coefficient is not significant above the threshold.

Second stage, participation effect in emissions

Table 10 shows the second stage results using the 1992 sample. In this sample, we include all non-participants and only 1992 joiners, but exclude plants that participate in 1991 and 1993–95. Column (1) and (3) use the first stage estimates that do not consider the threshold, and Column (2) and (4) use the first stage estimates that consider the threshold. The results echo the benchmark analysis using the full sample: although we do not find clear evidence of free-riding participants below the threshold, their participation effect is much lower than the participation effect for participants above the threshold. We believe that the failure to detect the free-riders is related to the unstable threshold value as shown in the first stage, and also drives the results for the full sample benchmark analysis.



(a) MLE criterion of threshold parameter, 33/50 program participation incentive in 1992

(b) MLE criterion of threshold parameter, 33/50 program participation incentive in 1993-1995

Figure 5: Threshold estimation and 95% confidence interval

B.2 1993–95 participants

First stage, participation incentives

There are small number of new participants during 1993–95, so we pool these three years' participation together in the empirical analysis. We take a subsample of 1993–95 observations and drop all 1991 and 1992 joiners. We estimate the participation incentives based on remaining plants' participation decisions from 1993–95. All non-participating plants are included in the sample. Figure 5b is the plot of log likelihood as a function of threshold parameter. We find that for 1993–95 participation, the threshold is $e''_{1993-95} = 9.9$ and 95% confidence interval is $[9.5, 10.3]$.

Table 11 shows the regression results of 1993–95 participation. We find similar results as the previous analysis with a positive and significant coefficient on lagged log 33/50 emissions before the threshold, but an insignificant coefficient above the threshold. The results do not violate hypothesis 1.

Second stage, participation effect in emissions

We report the second stage regression results in Table 12 using the sample of 1993–95 participation. In this sample, we include all non-participants and only 1993–95 joiners, but exclude plants who participate in 1991 and 1992. Column (1) and (3) use the first stage estimates that do not consider the threshold, and Column (2) and (4) use the first stage estimates that consider the threshold. We find some evidence of free-riding, since the participants below the threshold have positive coefficient (10% significant in column (4)), and the participants above the threshold have a statistically significant negative coefficient. However, all regression models fail the over-identification test (Hansen J test), suggesting the endogeneity of deep lag outcome variables as instruments. Therefore, the coefficients of lagged emissions may be biased for this sample.

Table 9: Probit model of 33/50 program participation incentive in 1992

	Simple model	Regression unknown kink model	
		Before Threshold	After Threshold
First Invitation Group	0.000	−0.014	0.832***
	(0.074)	(0.081)	(0.291)
Plant Share in Firm TRI Emissions	1.025***	0.978***	2.265***
	(0.069)	(0.070)	(0.432)
Pre-Program Reduction in 33/50 Emissions	0.040	0.108**	−0.099
	(0.028)	(0.051)	(0.087)
Lag HAP/TRI Ratio	0.001	0.001	0.015**
	(0.001)	(0.001)	(0.007)
Lag AFS Inspection Number	−0.038	−0.036	−0.024
	(0.034)	(0.039)	(0.074)
Lag Log 33/50 Emissions	0.022**	0.028**	−0.116
	(0.010)	(0.011)	(0.087)
County Non-attainment Status	−0.004	−0.001	0.010
	(0.026)	(0.027)	(0.152)
State LCV Score	0.012***	0.013***	−0.005
	(0.004)	(0.004)	(0.015)
State LCV Score Square	−0.000***	−0.000***	0.001
	(0.000)	(0.000)	(0.000)
County Income per Capita	0.018***	0.017**	0.022
	(0.007)	(0.007)	(0.041)
Industry FE	Y	Y	
Threshold log emission	-	12.6	
Threshold 95% C.I.	-	(7.0, 14.2)	
Observations	7,350	7,350	
Observations before threshold	-	6,912	
Observations after threshold	-	438	
Log Likelihood	−1,333.695 (df=7,329)	−1,321.345 (df=7,319)	
Akaike Inf. Crit.	2,709.390	2,704,690	

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: 1991-1996 Program impact, 1992 participants only, participation status instrumented by 1992 participation incentive

	Dependent variable: log(33/50 emissions), first difference					
	(1)	(2)	(3)		(4)	
			Before threshold	After threshold	Before threshold	After threshold
Participation Status	-0.544^{***}	-0.535^{***}	-0.401^{***}	-2.917^{***}	-0.413^{***}	-1.395^{***}
	(0.127)	(0.116)	(0.130)	(0.727)	(0.125)	(0.397)
Lag Log 33/50 Emissions	0.587^{***}	0.588^{***}	0.538^{***}		0.570^{***}	
	(0.058)	(0.058)	(0.055)		(0.057)	
HAP/TRI ratio	-0.013^{***}	-0.013^{***}	-0.012^{***}		-0.012^{***}	
	(0.001)	(0.001)	(0.001)		(0.001)	
State LCV Score	0.004^{*}	0.004^{*}	0.004^{*}		0.004^{*}	
	(0.002)	(0.002)	(0.002)		(0.002)	
State LCV Score Square	-0.000	-0.000	-0.000		-0.000	
	(0.000)	(0.000)	(0.000)		(0.000)	
County Non-attainment Status	0.034	0.034	0.036		0.035	
	(0.034)	(0.034)	(0.033)		(0.034)	
County Income per Capita	0.033	0.033	0.029		0.031	
	(0.021)	(0.021)	(0.021)		(0.021)	
Industry Specific Time Trends	Y	Y	Y		Y	
State Specific Time Trends	Y	Y	Y		Y	
Threshold log emission	-	-	12.6		12.6	
Plant number	7,130	7,130	7,130		7,130	
Plant number before threshold	-	-	6,938		6,938	
Plant number after threshold	-	-	652		652	
Observations	35,060	35,060	35,060		35,060	
Observations before threshold	-	-	32,844		32,844	
Observations after threshold	-	-	2,216		2,216	
Over-identification test (Hansen J Statistics P Value)	0.3200	0.3201	0.3667		0.3401	
Weak identification test (Cragg-Donald Wald F Statistics)	53.910	53.952	53.183		49.954	

Note: *p<0.1; **p<0.05; ***p<0.01
Column (1) and (2) are second stage GMM regression without considering the heterogeneity threshold;
Column (3) and (4) are second stage GMM regression considering the heterogeneity threshold;
The IV used by column (1) and (3) is estimated by the first stage without considering the heterogeneity threshold;
The IV used by column (2) and (4) is estimated by the first stage considering the heterogeneity threshold.
All regressors except participation status are in first differences.
IV's are first stage estimated participation probability and the third-year lag of log 33/50 emissions.

Table 11: Probit model of 33/50 program participation incentive in 1993-1995

	Simple model	Regression unknown kink model	
		Before Threshold	After Threshold
First Invitation Group	0.009	0.191	−0.147
	(0.080)	(0.129)	(0.108)
Plant Share in Firm TRI Emissions	0.388***	0.686***	0.084
	(0.081)	(0.118)	(0.120)
Pre-Program Reduction in 33/50 Emissions	0.082***	0.086	0.083**
	(0.031)	(0.086)	(0.033)
Lag HAP/TRI Ratio	0.002	0.004**	−0.001
	(0.001)	(0.002)	(0.002)
Lag AFS Inspection Number	0.045	−0.060	0.067**
	(0.029)	(0.079)	(0.031)
Lag Log 33/50 Emissions	0.027**	0.033*	0.053
	(0.012)	(0.019)	(0.035)
County Non-attainment Status	0.040	−0.015	0.086**
	(0.030)	(0.046)	(0.042)
State LCV Score	0.017***	0.023***	0.013**
	(0.005)	(0.007)	(0.006)
State LCV Score Square	−0.000***	−0.000***	−0.000*
	(0.000)	(0.000)	(0.000)
County Income per Capita	0.007	−0.012	0.024**
	(0.008)	(0.012)	(0.011)
Year FE	Y	Y	
Industry FE	Y	Y	
Threshold log emission	-	9.9	
Threshold 95% C.I.	-	(9.5, 10.3)	
Plants	6,540	6,540	
Observations	17,095	17,095	
Plants before threshold	-	4,081	
Observations before threshold	-	9,251	
Plants after threshold	-	3,374	
Observations after threshold	-	7,844	
Log Likelihood	−815,728 (df=17,072)	−798.870 (df=17,062)	
Akaike Inf. Crit.	1,677.457	1,663.740	

Note: *p<0.1; **p<0.05; ***p<0.01

Table 12: 1991-1996 Program impact, 1993-1995 participants, participation status instrumented by 1993-1995 participation incentive

	Dependent variable: log(33/50 emission), first difference					
	(1)	(2)	(3)		(4)	
			Before Threshold	After Threshold	Before Threshold	After Threshold
Participation Status	-1.573^{***}	-1.549^{***}	1.084^{**}	-1.391^{***}	0.834^{*}	-1.442^{***}
	(0.466)	(0.414)	(0.532)	(0.483)	(0.451)	(0.465)
Lag Log emission	0.454^{***}	0.458^{***}	0.340^{***}		0.342^{***}	
	(0.065)	(0.062)	(0.065)		(0.065)	
HAP/TRI ratio	-0.011^{***}	-0.011^{***}	-0.008^{***}		-0.008^{***}	
	(0.001)	(0.001)	(0.001)		(0.001)	
State LCV Score	0.002	0.002	0.002		0.002	
	(0.002)	(0.002)	(0.002)		(0.002)	
State LCV Score Square	−0.000	−0.000	−0.000		−0.000	
	(0.000)	(0.000)	(0.000)		(0.000)	
County Non-attainment Status	0.014	0.014	0.021		0.021	
	(0.030)	(0.030)	(0.028)		(0.028)	
County Income per Capita	0.011	0.011	0.005		0.005	
	(0.020)	(0.020)	(0.019)		(0.019)	
Industry Specific Time Trends	Y	Y	Y		Y	
State Specific Time Trends	Y	Y	Y		Y	
Threshold log emission	-	-	9.9		9.9	
Plant number	8,669	8,669	8,669		8,669	
Plant number before threshold	-	-	5,725		5,725	
Plant number after threshold	-	-	5,116		5,116	
Observations	37,358	37,358	37,358		37,358	
Observations before threshold	-	-	19,309		19,309	
Observations after threshold	-	-	18,049		18,049	
Over-identification test (Hansen J Statistics P Value)	0.0225	0.0360	0.0004		0.0006	
Weak identification test (Cragg-Donald Wald F Statistics)	47.658	49.110	41.357		45.780	

Note: *p<0.1; **p<0.05; ***p<0.01

Column (1) and (2) are second stage GMM regression without considering the heterogeneity threshold;
Column (3) and (4) are second stage GMM regression considering the heterogeneity threshold;
The IV used by column (1) and (3) is estimated by the first stage without considering the heterogeneity threshold;
The IV used by column (2) and (4) is estimated by the first stage considering the heterogeneity threshold.
All regressors except participation status are in first differences.
IV's are first stage estimated participation probability and the third-year lag of log 33/50 emissions.

References

Aerts, W., D. Cormier, and M. Magnan (2008). Corporate environmental disclosure, financial markets and the media: An international perspective. *Ecological Economics 64*(3), 643 – 659.

Anton, W. R. Q., G. Deltas, and M. Khanna (2004). Incentives for environmental self-regulation and implications for environmental performance. *Journal of the Association of Environmental and Resource Economists 48*, 632–654.

Arellano, M. and S. Bond (1991, 04). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies 58*(2), 277–297.

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

Borck, J. C. and C. Coglianese (2009). Voluntary environmental programs: Assessing their effectiveness. *Annual Review of Environment and Resources 34*, 305–324.

Brouhle, K., C. Griffiths, and A. Wolverton (2005). The use of voluntary approaches for environmental policymaking in the u.s. *In: Croci E (ed) The Handbook of Environmental Voluntary Agreements. Springer, Dordrecht the Netherlands*.

Brouhle, K., C. Griffiths, and A. Wolverton (2009). Evaluating the role of epa policy levers: An examination of a voluntary program and regulatory threat in the metal-finishing industry. *Journal of Environmental Economics and Management 57*(2), 166–181.

Carrion-Flores, C. E., R. Innes, and A. G.Sam (2013). Do voluntary pollution reduction programs (vprs) spur or deter environmental innovation? evidence from 33/50. *Journal of Environmental Economics and Management 57*(3), 444–459.

Decker, C. (2002, 01). Shifting the burden: Regulatory responses to pollution control. *Journal of Economic Research 7*, 41–75.

Delmas, M. A. and V. C. Burbano (2011). The drivers of greenwashing. *California Management Review 54*(1), 64–87.

Flammer, C. (2015). Does corporate social responsibility lead to superior financial performance? a regression discontinuity approach. *Management Science 61*(11), 2549–2568.

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*(1), 391–410.

Gamper-Rabindran, S. and S. R. Finger (2013). Does self-regulation reduce pollution? responsible care in the chemicals industry. *Journal of Regulatory Economics 43*(1), 1–30.

Hansen, B. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics 93*(1), 345–368.

Hansen, B. (2000). Sample splitting and threshold estimation. *Econometrica 68*(3), 575–603.

Heal, G. (2005). Corporate social responsibility: An economic and financial framework. *The Geneva Papers 30*(3), 387–409.

Innes, R. and M. G. Sam (2008). Voluntary pollution reductions and the enforcement of environmental law. an empirical study of the 33/50 program. *Journal of Law and Economics 51*(2), 271–296.

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

Kim, E.-H. and T. P. Lyon (2011). Strategic environmental disclosure: Evidence from the doe's voluntary green house gas registry. *Journal of Environmental Economics and Management 61*(1), 311–326.

King, A. A. and M. J. Lenox (2000). Industry self-regulation without sanctions: The chemical industry's responsible care program. *Academy of Management Journal 43*, 698–716.

Li, H. and N. Khanna (2018). Does voluntary self-regulation provide regulatory relief? a lesson from the responsible care program in the united states. *Journal of Law and Economics 61*(1), 63–96.

Lyon, T. P. and J. W. Maxwell (1999). Corporate environmental strategies as tools to influence regulation. *Business Strategy and the Environment 33*, 425–439.

Lyon, T. P. and J. W. Maxwell (2007). Environmental public voluntary programs reconsidered. *The Policy Studies Journal 35*(4), 723–750.

Lyon, T. P. and J. W. Maxwell (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics and Management Strategy 20*(1), 3–41.

Maxwell, J. W., T. P. Lyon, and S. C. Hackett (2000). Self-regulation and social welfare: The political economy of corporate environmentalism. *The Journal of Law & Economics 43*(2), 583–618.

Rivera, J. (2004). Institutional pressures andvoluntary environmental behavior in developing countries: Evidence from the costa rican hotel industry. *Society and Natural Resources 17*, 779–797.

Vidovic, M., M. S. Delgado, and N. Khanna (2019). Third-party certification and the effectiveness of voluntary pollution abatement programs: Evidence from responsible care. *Economic Inquiry 57*(4), 1751–1770.

Vidovic, M. and N. Khanna (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*(2), 180–195.

Vidovic, M. and N. Khanna (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 and Resource Economics 52*(3), 369–393.

Wang, Y., M. S. Delgado, N. Khanna, and V. L. Bogan (2019). Good news for environmental self-regulation? finding the right link. *Journal of Environmental Economics and Management 94*, 217–235.

Welch, E. W., A. Mazur, and S. Bretschneider (2000). Voluntary behavior by electric utilities: Levels of adoption and contribution of the climate challenge program to the reduction of carbon dioxide. *Journal of Policy Analysis and Management 9*(3), 407–425.

Zhang, B., J. Bi, Z. Yuan, J. Ge, B. Liu, and M. Bu (2008). Why do firms engage in environmental management? an empirical study in china. *Journal of Cleaner Production 16*, 1036–1045.

Zhou, R., X. Bi, and K. Segerson (2020). Evaluating voluntary environmental programs with spillover effects. *Journal of the Association of Environmental and Resource Economists 7*(1), 145–180.