INTERNATIONAL POLICY CENTER
Gerald R. Ford School of Public Policy
University of Michigan

IPC Working Paper Series Number 105

Direct and Indirect Effects of Voluntary Certification: Evidence from the Mexican Clean Industry Program

Andrew Foster

September 2010

Direct and Indirect Effects of Voluntary Certification:

Evidence from the Mexican Clean Industry Program

Abstract

We develop a model of environmental regulation that integrates firm and regulator behavior to evaluate a voluntary certification program: the Mexican Clean Industry Program. Imposing some structure on the costs of participation and compliance we establish that plants with lower costs of compliance are the most likely to certify. Moreover, authorities use certification as a screening tool and update their inspection policy, as high certification rates imply lower inspection costs. Empirically, we find that particulate matter (measured from satellite imagery) significantly lowers in areas with non-certified plants in high-certification sectors, but not around certified plants.

JEL codes: Q52, Q56.

There is a clear need for a better understanding of the policy tools available to environmental regulators in low-income countries and, in particular, whether the methods and approaches used in more advanced countries can be readily adapted to lower-income settings. The potential conflict between local demand for poverty reduction and demands for more stringent environmental regulation, that are partly driven by global considerations, have emerged as a major sticking point in negotiations regarding both climate change and the expansion of trade. However, market based interventions that are typically advocated by economists as a mechanism for dealing with unobservable differences in the cost of reducing emissions may not be well suited to a situation in which monitoring costs relative to the value of output are high and in which legal institutions for enforcement of contracts are weak. Similar concerns arise with respect to attempts to cap emissions through direct regulation of firms or plants— monitoring costs relative to output may be high and the government agency in question may have limited capacity to monitor, particularly given that a substantial fraction of polluting activity may occur among relatively small plants or those in the informal sector. Moreover the balance between global and local interests in environmental controls may be different in emerging relative to more advanced economies. The need to mollify trading partners concerned about the potential emergence of pollution havens, for example, may argue for the development of regulations that are targeted towards exporting firms.

An interesting policy tool that may address some of these concerns is that of voluntary certification. In particular, like a market-based system, an appropriately designed voluntary program may have the effect of concentrating emission reductions on plants with low compliance costs and/or in those plants that are most likely to benefit from being able to demonstrate compliance with environmental regulation, such as those involved in international trade. Indeed, voluntary pollution reduction programs are increasingly being used to encourage plants to reduce their emissions levels in both the developed and developing world (OECD, 1999, 2003). Their effectiveness, especially in low and middle-income countries, however, is unclear and may, among other things, depend on the interaction between the information revealed by these programs and other regulatory and non regulatory pressures for improved environmental performance (Blackman, 2008).

This paper develops a model of environmental regulation that integrates plant and regulator behavior and incorporates a combination of voluntary and mandatory controls. The implications of this model are then tested using a data set that has been newly assembled to examine the effects of the Mexican Clean Industry Program, in which plants are provided a Clean Industry Certificate if they are willing to establish via a privately financed audit that they meet the legal emissions standards. In particular, by imposing some structure on the cost of participation and the cost of compliance and drawing out the resulting implications, we are able to establish using data at the industrial sector level that plants with relatively low costs of compliance, conditional on sector, are the ones most likely to participate in the certification program. Moreover, we show that because authorities have the option to update the inspection intensity given the number of plants participating in the certification program, certification serves as a screening tool that reduces the cost of inspection in sectors with a high percentage of certified plants. Thus, according to our model, reductions in pollution emissions levels should not necessarily be observed for participating plants, but rather for non-certified plants in industrial sectors with a high percentage of certified plants.

Testing for these effects is complicated, of course, by the very problem that is faced by environmental regulators—the high cost of direct monitoring of plant emissions. We surmount this problem by integrating newly developed satellite based measures of suspended particulates and a plant-level data set with geographical identifiers. As predicted by the model we find that particulate matter concentrations significantly lower in areas with a large fraction of non-certified plants in high-certification sectors, and do not lower significantly in areas with certified plants.

II. Motivation

The emergence of voluntary certification programs has been followed by a growing body of literature trying to evaluate their effectiveness (Morgenstern and Pizer, 2007; Khanna, 2001). The existing literature for developing countries contexts is limited given the scarcity of performance data and has thus focused on identifying the drivers of participation (Blackman, 2010). A couple of studies attempt do so for the Mexican Clean Industry Program, which is the focus of the present paper. Blackman et al. (2007) show that plants that have been inspected or fined for not complying with pollution emissions standards in the past are more likely to participate. Munoz-Pina et. al. (2006) find that larger plants, exporting plants and those who sell their goods to the government are more likely to participate, while the average income of the community where plants are located, plant's age, fuel and water use intensity, and the amount of past fines have no predictive power on certification for the sample they use. No direct measurement of changes in air quality is provided as part of these studies, so the effects of certification are not measured.

Research for industrialized countries is especially concerned with testing whether participating plants are those already in compliance with the emissions standards, or if plants invest in pollution reduction for reasons not related to the existence of the program. Specifically, the US Environmental Protection Agency 33/50 program has received most of the attention in the literature. Arora and Carson (1996), Gamper-Rabindran (2006) and Sam and Innes (2006), for example, do not find evidence that plants participating in the Program were those who had reduced their emissions before its implementation. However, (Vidovic and Khanna, 2007) find the opposite result. According to the latter, a very small percentage of the total emissions by participating plants can be attributed to the Program. Some of these studies also try to test if the plants with the lowest or highest emissions levels are the ones participating, with no conclusive results. The differences in the findings seem to come from differences in the sample used, the mechanisms used to correct for selection in the participating sample, or the variable used to measure environmental compliance (Alberini and Segerson, 2002).

In any case, in addition to the fact that most studies of voluntary certification focus on the developed world and thus may not be entirely relevant to the developing country context, there are two key limitations to existing studies of voluntary programs. First, these studies have lacked a general equilibrium perspective that would permit an assessment of how these voluntary programs interact with other mechanisms for environmental control, given regulator and plant behavior. By looking at these programs in isolation, existing studies ignore the relevance of the information revealed in the process of certification and how that information can be used by other actors and influence plant behavior. Moreover, in the absence of a reasonable model of who gets certified and why, it is difficult to assess and address possible problems of reverse causality. Second, evaluations of the effects of these programs have been limited by data availability. The high cost of monitoring emissions at the plant level that makes direct regulation difficult also means that one rarely has available plant level information before and after certification. While ground level stations that may permit assessment of changes at an ecological level are available in some areas, these stations tend to be concentrated in relatively few urban areas and placed at strategic points such as busy intersections rather than based on an attempt to elicit a representative picture of variation in air-quality across regions. They thus are ill-suited to a systematic evaluation of the effects of a change in policy on different types of firms.

In this paper we address both of these issues. In particular, we develop and test a model of environmental regulation in a low to middle-income setting that captures key elements of the emissions regulation strategy used by the Mexican Federal Environmental Protection Agency (Procuraduría Federal de Protección al Ambiente, PROFEPA).

II.b. Environmental Regulation in Mexico

There are two key prongs of PROFEPA's regulatory approach. First, the agency is responsible for inspecting plants in order to determine if they comply with the current legal pollution emission standards, which are set in terms of emissions per unit of capacity. Inspections are performed at random, assigning a higher probability of inspection to sectors with higher perceived risk of polluting and to larger firms. If a plant is found not to be compliance, it is forced to pay a fine, which increases in case of relapse. Second, the same agency, in 1997, introduced the Mexican Clean Industry Program (Programa de Industria Limpia), also known as National Environmental Auditing Program (Programa Nacional de Auditoría Ambiental), the main voluntary pollution reduction program in Mexico. Plants participating in this program have to pay for an audit by an independent agency on a list maintained by PROFEPA that determines the actions that need to be taken in order to make the plant compliant with the pollution emissions standards. Formally, the requirements for getting this certification are only that the plant be in compliance with existing standards in terms of emissions per unit of capacity. However, in practice, since auditors examine the technologies in place rather than just current emissions in making their recommendations, certification may involve a greater degree of capital investment than would be the case for a plant that chooses to meet emissions standards simply, for example, through lower utilization of plant capacity.[1] If a plant participates in the program, it is given a grace period during which it is not inspected. After it has been established that the plant meets the pollution standards, it is granted a Clean Industry Certificate, which can be used for marketing purposes and to demonstrate to financial institutions, for example, that it is not subject to a potential adverse shock arising from a failed emissions inspection. If certified, plants are then further exempted from inspections for a given period of time (at least two years). Between 1997 until 2007, 2,568 plants received this certification. However, between 1997 and 2000, only 228 plants received this certification, and many of them were government-owned plants, primarily those belonging to PEMEX, the Mexican Oil Company.

The model in this paper integrates these two components by assuming (a) that inspection probabilities are chosen by the regulator to maximize environment benefits net of audit costs borne by the regulator and (b) that plants compare the costs of certification net of any benefits that accrue such as lower credit costs or more effective marketing of certification to the costs of compliance without certification and to the expected cost of non-compliance given the fines and probability of being inspected. In particular, by imposing some structure to the costs of participation and compliance we derive testable predictions that help establish whether plants with relatively low cost of compliance (conditional on observables to the regulator) are the ones that participate in the certification program. We then use sector-level data to show that patterns of compliance and certification conform to the predictions of the model under the assumption that the plants with the lowest cost, conditional on sector, are the ones that choose to certify. We further establish that under these conditions, certification serves as a screening device that increases the net return, from the perspective of the regulators, to inspection in sectors with a

[1] Appendix 2 discusses this possibility in greater detail.

high fraction of certified plants. Thus, according to our model, the reductions in pollution emissions levels may be observed both among participating plants and among non-certified plants in industrial sectors with high levels of certification. We then test these implications by integrating newly developed satellite-based measures of suspended particulates and a plant-level data set with geographical identifiers.

III. Modeling plants' participation in the Clean Industry Program

As we will explain later, the available data at the industrial sector level for this paper includes the number of inspections performed by the authorities in each industrial sector, and the number of inspections resulting in non-compliance before and after the introduction of the certification program. Taking into account the constraints imposed by these data, our model endogenizes the behavior of the regulatory agency, which selects how it wants to allocate inspection effort across sectors, both before and after certification, and that of plants. Plants must decide whether to comply or not comply with emission standards prior to the introduction of the certification program and may choose to comply and be certified, to comply but not be certified, or to remain non-compliant after the introduction of the certification program.

Table 1 contrasts sectors in which there are both low levels of inspection and of certification and those with both high levels of inspection and certification. While there is not a clean division between these two groups, a rough cut would suggest that the high group includes sectors in which there is a high degree of chemical processing such as cement, pharmaceuticals, synthetic materials, and explosives. The low-certification/inspection sectors are ones in which agricultural products play a key role such as natural fibers, coffee/tea and chocolates, and wood products. Indeed the importance of chemical processing as a criterion for the targeting inspections is clearly articulated on PROFEPA's web page[2].

Table 2 illustrates some interesting facts about the compliance rates, pre and post certification, and the level of certification stratified according to the percent of plants in the sector that are inspected[3]. The surprising result is that, although the level of certification is higher in those sectors in which inspection rates are high, the fraction of inspected plants within the sector is not systematically related to the level of compliance either before or after the introduction of certification. In particular, the fraction of non-compliance prior to certification ranges from 79 to 84 percent but there is no perceptible trend. Post-certification non-compliance may be a bit lower for the lowest inspection sectors but otherwise the range of variation is 65 to 67 percent compliance. By contrast, the certification rate is just 4 percent in the lowest inspection sectors but rises almost monotonically to 10 percent in the highest inspection sectors. The implication of these results, along with the division in Table 1, is that PROFEPA is imposing sufficiently higher inspection probabilities in sectors with a high cost of compliance so that plants in both types of sectors are equally likely to choose compliance.

III.1. The Plants' Problem

We now turn to the structure of the model. We assume, as noted, that plants have a choice between three different options: complying with pollution emissions standards without getting certified; compliance with emissions standards and obtaining a "Clean Industry Certificate"; and non compliance. Note that all certified plants are assumed to be in compliance, but not all the non-certified plants are non-compliant. Each of the options has a different cost for each plant,

---

[2] While the distinction between beer and wine may be less clear it is instructive that beer is only moderately high in terms of inspection, it is among the highest in terms of certification, a possible consequence of the value of certification for this important export for Mexico.

[3] We describe how we constructed these variables in better detail later in the text.

depending both on the types of goods they produce (industrial sector) and plant specific characteristics, unobserved by the authorities. Plants will choose the option that has a lowest cost for them. Authorities do not observe the plant specific cost of compliance, but do observe a compliance cost that is common to plants in the same sector.[4]

In particular, the cost of compliance with pollution emissions standards without certification for plant *i* in sector *j* is:

$$C_{ij}^{c1} = C_j + d_{ij} \tag{1}$$

where C_j is the sector *j* level cost, observed by the authorities and d_{ij} is the plant specific cost, not observed by the authorities, with a distribution *F*, which we assume to be differentiable. The *p.d.f.* of *d* is denoted as *f*.

Net costs of compliance for those plants that certify may differ from costs associated with compliance without certification for a variety of reasons including (a) possible marketing benefits (b) reductions in liability and thus improved credit terms (c) the costs of an audit (d) the grace period provided and (f) a need to upgrade capital in order to meet the terms of an audit. These costs may affect differentially the observable and unobservable components of compliance costs so we assume that the cost of certification (net of the benefits) can be approximated by a linear function of these cost components:

$$C_{ij}^{c2} = \alpha C_j + \beta d_{ij} \tag{2}$$

[4] The set up of the model is equivalent to a situation in which the authorities observe more than the sector specific cost of compliance, but are constrained to set their inspection policy conditioning on industrial sector. The rationale for such an information constrained approach may be that it would be relatively easy to test whether inspectors are following the proscribed policy.

where α and β are constants that are common to all plants.[5] As can be seen, α multiplies the industrial sector level cost of compliance and β multiplies the plant specific cost of compliance. We do not theoretically constrain the values of these parameters given the diversity of factors in play but will try to draw inferences about their values from the data.

Finally, the cost of non compliance is given by:

$$C_{ij}^{nc} = P_j \times M \tag{3}$$

where P_j is the probability that the authorities will inspect a plant in sector *j* and *M* is the fine imposed if the inspected plant is found to be in non compliance. *M* is assumed to be fixed and exogenous[6] and P_j is set at the sector level, given that authorities are unable to observe (or unwilling to use) the plants' specific d_{ij}. As stated above, we assume that the plant specific component of the cost of compliance is only observed by the plant.[7]

Given this setup, in the absence of certificates, it is clear from equation (1) and (3) that only plants with low d_{ij} will comply with pollution emissions standards. However, in the presence of certification the values of α and β will determine who chooses to get certified. While theoretically we do not impose restrictions on the values of these parameters, we restrict attention to cases in which there is an interior solution in each sector[8] (this assumption is supported empirically given that in most sectors all three choices are evident). We see that three such general scenarios are possible. For this purpose, we define *a* as the intersection between

---

[5] In principle we could slightly generalize by adding a constant term to (2). However, from the standpoint of empirical inference this extra term would play little role because it is common across all sectors.
[6] Given the structure of the model allowing M to vary but keeping total expected fines constant would not increase the ability of the regulator to ensure compliance unless the cost
[7] Appendix 1, given that the present study focuses on a voluntary program in a middle-income country, discusses how the setup presented this far can incorporate the possibility for corruption.
[8] This restriction is related to the value of MP_j. For cost schedules 1 and 2, it can be expressed formally as: $MP_j > \frac{(\beta-\alpha)}{(\beta-1)}C_j$. For cost schedule 3, it is: $MP_j < \frac{(\beta-\alpha)}{(\beta-1)}C_j$.

equation (1) and (2), *b* as the intersection between equations (1) and (3), and *c* as the intersection between equations (2) and (3).

Figure 1 plots the cost of compliance, non compliance and compliance with certification for different values of *d*, when $\alpha < 1$ and $\beta > 1$ and assuming an interior solution. $\alpha < 1$ implies that most of the benefits from participating in the program are common to all plants within one industrial sector, and that these benefits outweigh the costs of participation for the plants with lowest *d*. The assumption $\beta > 1$ implies that the cost of participating in the program is higher for plants with relatively high compliance costs[9]. Plants will get certified if $d < a$. Plants for which $a < d < b$ will be in compliance and not certified, and plants with $d > b$ will be in non compliance. Since certification is infra-marginal to compliance, it is evident that in the absence of certification all plants to the left of *b* will be compliant and thus, for given inspection probability, certification will have no effect on compliance. However, because, as illustrated below, the inspection rate adjusts to the presence of certification it is not only possible that overall compliance increases but also that plants that were not in compliance prior to the introduction of certification would choose to certify.

Figures 2 and 3 plot the same three hypothetical cost schedules this time for $\alpha > 1$, which implies that the costs of certification outweigh the benefits for plants with the lowest plant-specific cost of compliance. In Figure 2, β is set to be lower than one but higher than zero, illustrating a situation in which plants with intermediate levels of *d* get certified. Figure 3 shows the extreme case, in which β is negative, implying that the plants that get certified are those

[9] Appendix 2 provides an illustration of why certification costs may rise more rapidly with *d* than compliance costs without certification. The premise of the model is that given that standards are set in terms of emissions per unit of capacity a plant can be in compliance through either capital upgrading or output reduction but certification can only be obtained through capital upgrading. Output reduction turns out to be particularly advantageous for firms with high shares of "dirty" capital..

with the highest levels of *d*. In both of these cases, some of the plants getting certified, for given inspection probability, are plants that would not be compliance in the absence of the program.

III.2. The Regulator's Problem

We now turn to the problem faced by the regulator. In particular, consider first the pre-certification case so that the compliant fraction in sector *j* is just the fraction of plants in that sector for which the expected cost of fines exceeds the cost of being in compliance:

$$L(P_j, C_j) = F(b) = F(MP_j - C_j) \tag{4}$$

The regulator is assumed to receive a benefit *A* for every compliant plant and to pay a cost of *B* for every inspection. The regulator maximizes benefits minus costs through the choice of inspection probabilities by sector[10]:

$$S = A\sum_j N_j L(P_j, C_j) - B\sum_j N_j P_j \tag{5}$$

Differentiating with respecting to P_j and solving yields the result

$$P_j M - C_j = \delta_0 \tag{6}$$

for some constant δ_0 that is invariant across sectors[11]. Thus $L_j = L_{j'}$ for all *j, j'*. Because the distribution of unobserved costs is assumed to be the same by sector, the probability of inspection is set in such a way that the fraction in compliance is the same in all sectors.[12] In particular, as is evident from (6), the probability of inspection is higher in high-cost sectors. Thus the stratification by percent inspection in Table 2 may also be thought of as a ranking by sector-

[10] The implications derived from this setup are the same as if we assumed the authorities maximized compliance given a total budget to perform inspections.
[11] A sufficient condition for the solution to this equation to maximize the authorities budget constraint is that $f'(d)<0$.
[12] This result mirrors an insight from the recent literature on racial profiling (Antonovics and Knight, 2008; Knowles, Persico and Todd, 2001). In particular that literature provides conditions under which conviction rates of motorists who are stopped by police should not differ by race. The basic insight of these models is that inspection probabilities by race adjust optimally to differences by race in the propensity to be in violation of the law.

level cost of compliance. Moreover, the probability of inspection when certificates are not available can be used as a proxy for the sector level fixed cost of compliance, C_j. Indeed, our measure of inspection intensity at the sector level is highly correlated with measures of pollution abatement costs for the US[13].

In the presence of certification we need to distinguish between the percent certified D_j^k and the percent compliant L_j^k, inclusive of both certified and uncertified but compliant plants in sector *j*, given regime *k* (based on Figures 1-3). A key feature of the resulting expressions is that the probability of inspection affects the compliance share in each of the three regimes and the certification probability in the second and third regime, but does not affect the certification probability net of the compliance cost in the first regime. In particular, because the certified group is on the far left in regime 1, the fraction certified depends only on the intersection between the certified and compliant cost curves, with compliance being determined as in the non-certification case:

$$D_j^1(P_j, C_j) = F(a) = F\left(\frac{(1-\alpha)}{(\beta-1)}C_j\right), \tag{7}$$

and

$$L_j^1(P_j, C_j) = F(b) = F(MP_j - C_j) \tag{8}$$

Because in the second regime the certification group is in the middle, the relevant cut points of the certification group are the intersections of the certification line and the other two lines. Compliance is determined by the intersection of the certified and non-compliant curves:

[13] Appendix Figures 1 and 2 show a lowess line relating our measure of inspection intensity in Mexico and the US and the abatement costs as a percentage of the total value of shipments in the US, taken from the PACE survey, at the sector level.

$$D_j^2(P_j, C_j) = F(c) - F(a) = F\left(\frac{MP_j - \alpha C_j}{\beta}\right) - F\left(\frac{(1-\alpha)}{(\beta-1)} C_j\right) \tag{9}$$

$$L_j^2(P_j, C_j) = F(c) = F\left(\frac{MP_j - \alpha C_j}{\beta}\right). \tag{10}$$

Finally, in the third regime, certification is determined by the intersection of the certified and non-compliant lines, while compliance is determined by the intersection of the non-compliant curve with that of the two other groups:

$$D_j^3(P_j, C_j) = 1 - F(c) = 1 - F\left(\frac{MP_j - \alpha C_j}{\beta}\right) \tag{11}$$

$$L_j^3(P_j, C_j) = 1 - F(c) + F(b) = 1 - F\left(\frac{MP_j - \alpha C_j}{\beta}\right) + F(MP_j - C_j) \tag{12}$$

The regulator's objective function given certification reflects the fact that inspections need not be carried out on certified plants because they have already established compliance through a privately financed audit:

$$S^k = \sum_j N_j A L_j^k(P_j, C_j) - \sum_j BP_j (1 - D_j^k(P_j, C_j)) \tag{13}$$

This effect plays a key role in the analysis because it implies that the decision about how to allocate effort by sector is affected by the level of certification within a sector even when the share of plants being certified is not influenced by the inspection probability as in regime 1. Thus, the first order condition for the inspection probability in sector *j* and regime *k* is

$$A\frac{\partial L_j^k}{\partial P_j} - B(1 - D_j^k) - BP_j \frac{\partial D_j^k}{\partial P_j} = 0\,. \tag{14}$$

III.3. Identification of Regime

We now turn to the question of the identification of regime based on the preliminary descriptive evidence from Tables 1 and 2. We consider the regimes in reverse order. In regime (3) the compliance rate among non-certified plants is:

$$R_j^3 = \frac{L_j^3 - D_j^3}{1 - D_j^3} = \frac{F(MP_j - C_j)}{F(\dfrac{MP_j - \alpha C_j}{\beta})}. \tag{15}$$

We note first that if, as suggested in Table 2, certification is increasing in the inspection probability before the introduction of certificates, which is a proxy for the observed (by the authorities) cost of compliance (C_j). Then,

$$\frac{dD_j^3}{dC_j} = -f(\frac{MP_j - \alpha C_j}{\beta})\frac{1}{\beta}(M\frac{dP_j}{dC_j} - \alpha) > 0 \Rightarrow M\frac{dP_j}{dC_j} > 1 \tag{16}$$

where the latter follows from the assumptions on $\alpha > 1$ and $\beta < 0$ necessary for regime 3. But (16) implies that the numerator of (15) is decreasing in C_j:

$$\frac{d(L_k^3 - D_k^3)}{dC_j} = f(MP_j - C_j)(M\frac{dP_j}{dC_j} - 1) > 0 \tag{17}$$

Because the numerator of (15) is increasing in observed costs and the denominator is decreasing in observed costs (certification is increasing in percent inspected), the compliance probability must be increasing in observed costs, rather than constant, as shown in Table 1.

Analogously, the compliance probability among non-certified plants in regime 2 is

$$R_j^2 = \frac{L_j^2 - D_j^2}{1 - D_j^2} = \frac{F(\dfrac{C_j(\alpha - 1)}{1 - \beta})}{1 - F(\dfrac{MP_j - \alpha C_j}{\beta}) + F(\dfrac{C_j(\alpha - 1)}{1 - \beta})} \tag{18}$$

But since the denominator must be decreasing in observed costs and the numerator is clearly increasing in C_j given the assumptions of regime 2 ($\alpha > 1$ and $0 < \beta < 1$), compliance among non-certified plants must be increasing in observed costs.

We now turn to regime (1), in which case the compliance fraction is,

$$R_j^1 = \frac{L_j^1 - D_j^1}{1 - D_j^1} = \frac{F(MP_j - C_j) - F(\frac{C_j(1-\alpha)}{\beta - 1})}{1 - F(\frac{C_j(1-\alpha)}{\beta - 1})} . \tag{19}$$

Solving (19) for $F(MP_j - C_j)$ and substituting into the first-order condition in regime 1,

$$f(MP_j - C_j) = \frac{B}{AM}(1 - F\left(\frac{(1-\alpha)}{(\beta - 1)} C_j\right), \tag{20}$$

yields the differential equation

$$Af(MP_j - C_j)M - B(1 + \frac{F(MP_j - C_j) - R_j^1}{1 - R_j^1}) . \tag{21}$$

The unique solution to this differential equation using the boundary condition *F(0)=0* is

$$F^*(z) = 1 - \exp(-\frac{Bz}{(1 - R_j^1)AM}) . \tag{22}$$

The implication of (21) is that, given the other assumptions of the model and the result that higher cost sectors are assigned higher inspection probabilities, the only way to generate a positive effect of sector level cost of compliance on certification and a zero effect of sector level cost of compliance on compliance among non-certified plants is if regime 1 is in place and the unobserved costs is generated according to an exponential distribution with hazard θ.[14] Under these assumptions compliance among non-certified plants will be

[14] In practice we may think of this is an approximate result given that the non-compliance rate in Table 2 for the lowest inspection rate sectors are slightly lower than in the other sectors.

$$R_j^1 = 1 - \frac{B}{\theta AM}. \qquad (23)$$

Thus, the patterns evident in Table 2 are only consistent with the prediction of regime 1 and it would be appropriate to conclude that, as predicted under that regime, the plants that certify are those that have the lowest cost of compliance within their sector.

Given that certified plants have the lowest cost and lower cost plants are more likely to comply in any case, this finding raises the question about whether certified plants would be in compliance in the absence of certificates. Clearly, overall compliance (certified plus compliant non-certified plants) increases as a result of certification. Under the exponential distribution the fraction compliant in the absence of certification is

$$1 - \frac{B}{MA}. \qquad (24)$$

and the total fraction compliant under certification

$$1 - \frac{B\exp\left(-\frac{\theta C_j(1-\alpha)}{\beta - 1}\right)}{MA}. \qquad (25)$$

The comparison of (23) and (24) indicates that the fraction compliant among non-certified plants can be higher or lower in the presence of certification, depending on whether the hazard of the exponential distribution underlying the costs is less than or greater than one. However, the total fraction compliant must be higher given that the exponentiated term in (25) is negative for regime 1. This result reflects the informational role played by certification. Because certified plants are induced to pay for their own audits, the agency can concentrate its efforts on the non-certified plants and thus induce greater compliance among these plants. However, this result does not necessarily indicate whether certification has a direct effect on compliance at the level of the individual plant. In particular, the fraction certified under this distributional assumption is

$$D_j^1 = 1 - \exp\left(-\frac{\theta C_j (1-\alpha)}{\beta - 1}\right) \tag{26}$$

and it is not clear how (24) compares to (26). Thus, even though the lowest cost plants within each sector will be compliant regardless of the certification program, it is possible that there are some plants who certify that would not be compliant in the absence of certification, and this is more likely in sectors where certification is high.

III.4. Additional implications

The model under regime 1 also yields implications for the relationship between the inspection probability and sector costs before and after the introduction of certification. Examination of this relationship is useful because it can be used to assess whether differential technological change that reduces compliance costs might be responsible for an observed relationship between certification and improvements in air quality. In particular, letting the superscript *c* denote certification and the superscript *nc* denote non-certification,

$$\frac{\partial P_j^c}{\partial C_j} = \frac{\beta - \alpha}{M(\beta - 1)} > 0 \tag{27}$$

and

$$\frac{\partial P_j^c}{\partial C_j} - \frac{\partial P_j^{nc}}{\partial C_j} = \frac{1-\alpha}{M(\beta - 1)} > 0 \tag{28}$$

given the parameters necessary to produce regime 1. Equation (27) confirms that after certification, as was also shown in the non-certification case, the probability of inspection is directly proportional to sector compliance cost. Equation (28) shows that, given the model and assuming that regime 1 is in place, the probability of inspection should increase more in high-cost sectors following the introduction of certification than in low-cost sectors. Conversely, if technological improvements lead to a greater lowering of compliance costs in high-cost sectors over time than in low-cost sectors, and this change were responsible for both higher certification and improvements in air quality in areas with many plants in high-cost sectors, then one would expect a lower increase in inspection probabilities in high relative to low cost sectors.

IV. Data

IV.1. Sector level data

In order to examine in more detail the results presented in Table 2, to test additional predictions derived from the theory, and ultimately to examine the effects of certification on air quality, we combine a variety of different data sets. For the first part of the analysis (inclusive of Tables 1 and 2) we combined three data sets. First, we obtained the total number of plants, employees and the value of production for each four digit NAICS (North American Industrial Classification System) sector from the 1999 Mexican Industrial Census. Second, we obtained from PROFEPA a list of all plants that were granted a Clean Industry Certificate from 1997 until 2006, as well as a yearly list of the total number of inspections performed since 1992 until 2007, by NAICS industrial sector. We also know how many of these inspections found the plants to be in non-compliance each year. Data on the address (including zip code) of the 94 auditors licensed by PROFEPA, used later in the paper, were also obtained. We restrict the sample to 160 manufacturing sectors where at least one inspection took place in the period analyzed, excluding utilities (run by the government) and services.

We define the probability of inspection before the introduction of certificates (our proxy for the observed sector level cost of compliance) as the total number of inspections between 1992 and 1995 in each industrial sector divided by the total number of plants in each sector in the 1999 Industrial Census. The probability of inspection after the introduction of certificates is defined as the total number of inspections between 2003 and 2006 divided also by the total number of plants in each sector in the 1999 Industrial Census. The fraction of plants certified in each sector is simply the total number of certified plants divided by the total number of plants in each sector. Alternative measures incorporated the total number of plants in the 2005 Mexican industrial census as denominator, and the results did not change considerably. In addition, it is worth mentioning that the growth in the number of plants per sector between 2000 and 2005 is uncorrelated with the certification intensity, or inspection probabilities.

To examine if inspection probabilities are primarily being driven by technological features of the sector as posited by our model rather than, say, being targeted based on political or other factors, we also incorporated a data set on inspections and compliance in the US, and used data at the sector level from the Pollution Abatement Costs And Expenditure Survey 2005 (PACE), also for the US.

In particular, the US Environmental Protection Agency (EPA) publishes the Enforcement and Compliance History Online (ECHO). ECHO is a Web-based tool that provides public access to compliance and enforcement information for approximately 800,000 EPA-regulated facilities. ECHO gives access to permit, inspection, violation, enforcement action, and penalty information covering the past five years in the United States. The site includes facilities regulated as Clean Air Act stationary sources, Clean Water Act direct dischargers, and Resource Conservation and Recovery Act hazardous waste generators/handlers. From this system, we obtained the number of pollution emissions inspections conducted in each industrial sector in the US, from 2002 through 2007. The probability of inspection in the United States is then defined as the total number of inspections reported in ECHO, divided by the total number of establishments in the US Industrial Census, for each 4 digit NAICS industrial sector.

The PACE survey contains information on pollution abatement capital expenditures and operating costs associated with compliance with environmental regulation in the United States. We constructed a measure of compliance costs by NAICS industrial sector (the 143 sectors for which data is available), by dividing the total cost of compliance in each sector by the total value of shipments reported in the Annual Survey of Manufactures (ASM).

These data sets were linked at the sector level and we thus considered information from 1992 to 2007 for 160 four digit NAICS industrial sectors.

IV.2. Plant level data

Given the difficulty of accessing geographically identified plant-level data from the census, our plant-level and zip-code level analysis uses data from the SIEM (Sistema de Información Empresarial Mexicano), administered by the Mexican Ministry of Economics. These data contain information on 32,332 plants in the industrial sector. It includes each plant's name, exact address (including zip code), NAICS industrial sector, number of employees and dummy variables indicating whether the plant exports or imports. SIEM does not include government-owned plants. The geographic coordinates of each of the plant's zip code was obtained from Postal Code World©, which provides geographic coordinates for the 2668 zip codes in SIEM. The SIEM data were then linked using names and addresses to the PROFEPA data on Clean Industry Certificates. The percentage of plants certified in each sector is also assigned to each plant, given their declared NAICS sector in SIEM.

SIEM does not include all plants in Mexico. For this reason, not all of the certified plants listed by PROFEPA were found in our plant level database. 406 of the 1,266 certificates listed by PROFEPA for plants in industrial sectors were successfully matched to the SIEM data. While other issues about selection of plants into the sample could bias our empirical results, the percentage of certificates matched to the SIEM data in each industrial sector is uncorrelated with the percentage of plants certified in that sector.[15]

However, given the low number of certificates matched with SIEM, we will show two sets of empirical results. First, we will calculate the direct impact of certification on air quality by measuring the change in air quality in zip codes with a plant in SIEM listed as certified by PROFEPA. Second, we will assign to each zip code the percentage of the total number of plants certified in the municipality where they are located.[16]

IV.3. Air quality data

A key issue with the evaluation of the effects of air quality regulations in developing countries, as noted, is the absence of systematically collected data on emissions. A significant contribution of this paper is that it is among the first by economists to use remotely sensed data on air quality. In particular, spectral data on reflectance from the Moderate Resolution Imaging Spectroradiometer (MODIS onboard the Terra Satellite) were acquired from the NASA's Goddard Space Flight Center Earth Sciences Distributed Active Archive Center (DAAC). These data were used to construct daily measures of Aerosol Optical Depth (AOD) at a 5km spatial resolution for cloud-free areas over the whole of Mexican territory for the period between March 1st 2000 to December 31st 2006.

AOD has been shown to be a very good predictor of levels of suspended particles in the atmosphere (Chu et al., 2002; Gupta et al., 2006; Kumar et al., 2007). It is worth noting, however, that while AOD is likely to provide a measure of air quality, it does not allow us to carefully distinguish between different kinds of particles. An estimated measure of average AOD monthly (from March through December 2000 and 2006) levels for each zip code was

[15] The results of this test are available from the authors, upon request.
[16] Means and standard deviations for the two data sets are available as Appendix Tables A1 and A2.

constructed from the 5km pixel-level images. Using GIS technology, the observed measures of AOD from the satellite images were overlapped with the area around each of the zip codes. Daily measures of AOD were first calculated for each of the areas, and then the estimated AOD daily value for each zip code was averaged for each month in the sample, only considering those days for which we had an AOD measure. We then assigned ten measures of AOD for each year to each zip code, one for each month between March and December. As our regression estimates will be looking at the within month changes in AOD levels, given the unavailability of AOD measures on cloudy days, out of 26,680 possible observations (ten months for each of the 2668 zip codes with plants in the SIEM database), our sample, which considers any zip code with an AOD measure for a month in both years as an observation, was reduced to 19,849.

Figure 4 shows a map of the calculated levels of AOD for the whole Mexican territory, in October, 2006. While AOD levels seem higher around metropolitan areas (Mexico City, Guadalajara and Monterrey), other regions of the country seem to show comparable levels of AOD. Location of polluting industries, or geographic conditions that could facilitate the accumulation of particulate matter in specific areas could explain this.

However, in addition to the fact that AOD measures cannot be calculated on cloudy days, weather conditions, particularly dew point and temperature, can influence satellite based measurement of AOD and its relationship with suspended particles. In addition, the empirical relationship between the ground measurements of suspended particles and AOD can vary regionally, given that the composition of aerosols is different in each geographic region. We address these issues by focusing on changes in AOD levels within zip codes, and by adding monthly measures of the temperature and dew point in each zip code in the regressions. A map of the change in the logarithm of AOD between October 2000 and 2006 (the dependent variable in our empirical analysis) is shown in Figure 5.

IV.4. Weather data

The weather data were obtained from the US National Climatic Data Center, which publishes the Global Surface Summary of Day Data providing daily information for the 2000-2006 period for over one hundred weather stations spread around the Mexican territory. Average monthly values of the temperature and dew point were calculated for each weather station. A weighted average of a variable for each pixel in the map was assigned using weights that are an inverse function of the distance between that point and each of the points for which a measure of the variable exists (in this case, each of the weather stations). The mean monthly temperature and dew point for each zip code were then estimated by averaging over the interpolated data within each zip code's boundaries.

V. Results

V.1 Relationship between inspections, certification and compliance.

We now turn to a more formal analysis of patterns observed in Table 2. In particular, Table 3 tests the relationship between the probability of inspection before the introduction of the certification program and non compliance. The coefficient for the log of the probability of inspection is close to zero and insignificant. We also see no evidence that the other sector-level observables predict the non-compliance rate. This result is consistent with our hypothesis that authorities are assigning a higher inspection probability to sectors that face high compliance costs, thus imposing on them a higher incentive to invest in reducing pollution emissions. The inspection intensity prior to the introduction of certificates is then likely to provide us with a good proxy for the observed sector level cost of compliance.

The second set of regressions (Table 4) tests for the relationship between inspection intensity pre-certificates and certification rates at the sector level, While the magnitude of the coefficient goes down with the introduction of control variables, the relationship between the probability of inspection and the percentage of plants certified before the introduction of certificates (our proxy for the observed cost of compliance at the sector level) is always positive and significant. Interestingly we find that net of inspection probabilities, sectors in which there is a higher percentage of exports are less likely to certify.[19]

The third set of regressions (Table 5) tests the relationship between inspection intensity pre-certificates, and non compliance post-certificates. As for non-compliance in the pre-certification period, the coefficient for the log of the probability of inspection is close to zero and insignificant for all specifications. We also see no evidence that the other sector-level observables predict the non-compliance rate after the introduction of certificates.

Thus, the patterns evident in Table 2, robust to a more careful regression analysis presented this far, are only consistent with the prediction of regime 1 and it would be appropriate to conclude that, as predicted under that regime, the plants that certify are those that have the lowest cost of compliance within their sector.

V.2 Change in inspection probabilities

Our model predicts that, if the authorities obtain information about plants' cost of compliance from the certification process, they should update the inspection probabilities. The derivative of the probability of inspection with respect to the sector level fixed cost of compliance should be higher in a context in which certificates are available. In our industrial sector level data, we have information about the number of inspections performed since 1992

[19] These equations should not be interpreted causally as inspection probability and certification are jointly determined in our model.

until 2007. Given that certificates were introduced in 1997, we can compare the calculated inspection probabilities before and after the introduction of the program (1992-1995 and 2003-2006).

We define the probability of inspection after the introduction of the certificates as the total number of inspections performed, divided by the total number of plants in the census, an overestimate of the total number of plants subject to inspection, and an underestimation of the probability of inspection in sectors with certified plants. The coefficients of the change in inspections given certification will be an underestimate of the actual ones, given this fact.

Also, for this section, it is worth recalling that because we cannot directly observe the sector-level cost of compliance, we are in effect using the probability of inspection as a proxy for the fixed cost of compliance in each sector. If the probability of inspection is a noisy measure of the sector level fixed cost of compliance, correlating the probability of inspection in the 1992-1995 period against the change in the probability of inspection before and after the introduction of the certificates will produce a downward biased estimate of the derivative of the inspection probability with respect to the fixed cost of compliance. We thus consider several other possible proxies for the underlying sector-level costs, including the US inspection probabilities, the rate of certification, and the measure of abatement costs at the sector level in the US taken from PACE.

Table 6 shows the results of regressions of the change in probability of inspection at the sector level on these alternative measures of the cost of compliance. As expected, given the possibility of measurement error, the regression of the change in probability of inspection on the initial probability of inspection yields a negative coefficient. We thus examine three alternative measures of sector costs: the percent certified, the percent inspected in the US, and the measure of abatement costs taken from the PACE survey. Table 7 shows that these variables are in fact strongly predictive of percent inspected, with the combined R-squared being 34 percent. When we include these variables in the Table 6 regression directly, instead of initial percent inspected in Mexico, we find the expected positive relationship. The fifth column uses the percent certified, and the percentage inspected in the US as instruments for initial percent inspected in Mexico. The sixth column includes also our measure of abatement costs taken from the PACE survey as an instrument. As can be seen, both regressions show a positive relationship between pre-certificates inspection rates and the change in inspections around the introduction of the program. We thus conclude that, as anticipated, high-cost sectors saw a higher increase in the probability of inspection following the introduction of certification.

An obvious question that arises in this context is whether this trend in probabilities of inspection was in place prior to the introduction of certification. Appendix Figures 3 and 4 address this issue by plotting at the sector level the change in probabilities prior to the introduction of certification (1992-1994 and 1995-1997) and the change in probabilities before and after the certification program (1992-1995 and 2003-2006) as a function of the percent certified (Appendix Figure 3) and the percent inspected in the US (Appendix Figure 4). The corresponding lowess lines show a clear pattern—the dashed lines (around 1994) are in each case flat, while the solid lines (around 1997) show a pronounced trend as suggested by the regression results in Table 6. Thus it appears that the systematic changes in the probability of inspection with the cost of compliance were initiated after the introduction of the certification program.

This result is consistent with the idea that authorities are able to screen between plants with high and low costs of compliance within sectors as a result of the introduction of the Clean Industry Certificates. Thus, reductions in pollution emissions levels as a result of the program will not necessarily be observed amongst participating plants, but rather among non-certified plants in industrial sectors with a high percentage of certified plants. Among other things, this implies that one cannot use uncertified plants as a comparison group for examining the effects of certification.

V.5 Effects on air quality

We now turn to the analysis of the effects of the program on air quality. For this purpose, we make use of the plant level data from SIEM, the data on air quality and weather, and the information on the geographic location of the licensed auditors. As stated, our empirical strategy will try to estimate the impact of certification on pollution concentrations in the atmosphere. Our model predicts that a high percentage of plants certified in an industrial sector will create an incentive for non-certified plants in that sector to reduce their pollution emissions, given the authorities' response in terms of inspection intensity.

In order to estimate the indirect effects of certification related to the increase in inspection intensity related to the information revealed by certification, we exploit the difference in inspection probability by plant size. Because PROFEPA prioritizes plants, first by industrial sector, assigning a higher inspection probability to sectors with high polluting potential, and then assigning a higher likelihood of inspection to large plants, smaller plants are at very low risk of inspection.[20] Thus, in testing for indirect effects of certification it is necessary to consider the size-distribution of plants in a given area. Although the inspections data provided by PROFEPA does not include information on the size of inspected plants, we can classify plants by size using the SIEM data and match the resulting data to information on certification.

It is clear that from Table 8 that certification rates are higher in larger firms as should be expected given that PROFEPA differentially targets large firms for inspection. In our sample, 68

[20] www.profepa.gob.mx

percent of the plants listed in SIEM report having less than 10 employees. Out of all of them, 0.1 percent was matched with the certificates list. 24 percent of plants have more than 10 and less than 100 employees, and the matching rate with the certificates list is nearly ten times higher than for the smallest plants in the sample. 8 percent of plants have more than 100 employees, and 11 percent of them are matched with the certificates list.

In our estimation strategy, we will look at the direct effects of certification as well as indirect effects of certification that operate at the sector level through changes in the probability of inspection. We isolate the latter effect from sector specific trends in pollution emissions by assuming that these trends are similar across large and small firms and noting that small firms are unlikely to be inspected. If the reductions in emissions are actually related to an increase in the inspection probability given certification, plants with less than ten employees should not be reducing their emissions as a result of certification. The following equation describes the specification that identifies the impact of the program on pollution emissions:

$$\Delta Poll_i = \alpha + \beta_1 C_i + \beta_2 * M_i * CI_i + \beta_3 CI_i + \beta_4 M_i + \sum \gamma_k X_{ki} + \varepsilon_i \tag{29}$$

where $\Delta Poll$ is the change in pollution emissions by plant i, C_i is a dummy variable equal to 1 if the plant has received a Clean Industry Certificate, M_i is a dummy variable equal to 1 if the plant has more than 10 employees, which isolates the difference in emissions for plants with more than ten employees, regardless of the certification intensity, and the X_{ki} are a set of k control variables, including the log of the size of the plant, a dummy variable indicating if the plant exports, a dummy variable indicating if the plant imports, and the interaction of these last two variables variables with CI. Finally, CI_i is the percentage of plants certified in plant i's industrial sector, which controls for the changes in emissions correlated with certification but uncorrelated with the inspection probability. Plant size controls for differences in pollution emissions by different sized plants.

We have two main coefficients of interest: the one measuring the direct effect of certification, the coefficient on the certification dummy; and the one measuring the indirect effect of certification, the coefficient on the sector-specific certification intensity for plants with more than 10 employees. Given that our data do not measure pollution emissions by plants, but rather pollution concentrations around plants' zip codes (and that more than one plant are usually located in each zip code), we assume that the pollution concentrations in each county are a weighted average of the pollution emissions by each plant.

A weighted average of each of the variables in equation (30), including all of the interaction terms, is then calculated for each zip code, using the total number of employees reported by each plant divided by the total number of employees in each zip code (the sum of the employees of all plants in the SIEM database in each zip code) as the weight for each of the observations. The dependent variable is the change in the log of AOD between 2000 (the first point in time for which we have information on the pollution concentration) and 2006. The regressions are then run at the zip code level. Given that we constructed a measure of monthly AOD in each zip code from our data, we pool all calendar months (from March through December), and run the regression including calendar month fixed effects and cluster the standard errors of the coefficients at the zip code level. Controls for the differences in the temperature and dew point in each zip code between 2000 and 2006 are also included.

Relatively strong assumptions have to be made in order for the zip code level regressions to correctly estimate the impact of the Clean Industry Program on particulate matter concentrations. In particular, one of the main concerns is the location of plants not included in the SIEM database. As stated before, SIEM does not list government owned plants. Also, there are other private plants that are not included in the data. For our regressions to correctly estimate the impact of the Clean Industry Program on pollution concentrations, we need the geographic location of plants not included in our sample to be uncorrelated with the industrial composition of each zip code calculated from the SIEM database.

Estimating the direct effect of certification is a difficult task. Simply correlating certification with changes in pollution, as in the previous specification, would not necessarily capture a causal relationship between these variables. For example, a plant experiencing an exogenous decrease in its cost of compliance during the time period analyzed would be more likely to participate in the program, and have lower certification costs. In this case, in an OLS regression, the coefficient of certification on changes in emissions would be negative, but the relationship would not be causal. This selection problem is of course common to all studies evaluating the direct effectiveness of voluntary programs.

To address the selection problem we augment the theory by assuming that the underlying cost of conducting a private audit varies across firms and influences only the certification cost (e.g., it does not affect the cost of compliance without certification) and it does not affect emissions net of certification and/or compliance. This theoretical argument, along with the assumption that auditor´s location is exogenous, justifies the use of regional variation in the market supply of auditors available for certification as an instrument for certification in an assessment of the effects of certification on compliance. In particular, from a data set including all 94 auditors accredited by PROFEPA, with information on their geographic location, we constructed estimates of the distance (in km) between each zip code and each of the two closest environmental auditing plants. The average distance to the first auditor is 54 kilometers, while that to the second is 78 kilometers.

V.5.a. OLS results

The results of the OLS regression defining certification in each zip code from the certificates matched to SIEM are presented in Table 9. Column 1 is the regression output for the change in AOD at the zip code level against the weighted percentage of plants certified in each zip code, and the weighted certification intensity given the sector composition of each zip code, as well as the weighted average size of the plants. As can be seen, the coefficient on certification is not significantly different from zero, suggesting that certified plants do not experience a reduction in their pollution emission levels.

Columns 2-4 include the full specification trying to identify the indirect effects of certification, with an increasing number of control variables. The zero coefficient on the certification intensity in Column 1 seems to be driven by the fact that it is only plants that are subject to inspection that reduce their pollution emissions as a result of other plants in their sector getting certified. Our coefficient of interest, that of the interaction between the dummy variable indicating if the plant is big enough to be subject to inspections by PROFEPA and the certification intensity given the zip codes' industrial composition is negative and highly significant (Columns 2-4). The positive and significant coefficient on the certification intensity suggests that in the absence of a rise in inspection there would have been increased emissions (or less decrease) among firms in sectors with a high cost of compliance relative to those in sectors with a low cost of compliance.

Given the incomplete matching rate with the SIEM data we considered an alternative specification in which zip-code level certification was imputed based on county-level certification apportioned across zip codes according to the composition of industry in the different zip codes. Table 10 shows the results of the same specification as in Table 9 (still not instrumenting for certification). As can be seen, although the coefficient on the percentage of certified firms in each zip code changes in magnitude, it is still not significantly different from zero. All the other coefficients are similar in sign and significance level to those in Table 9.

V.5.b. Instrumental variables estimates

The first stage regression results, with the distance in hundreds of kilometers from each zip code to the first and second closest auditors are shown in Table 11[22]. Each column includes the same controls as the corresponding column in Table 10. As can be seen, distance variables predict certification with the joint F on these variables being above 6.5 for all specifications. Interestingly, we find that increased distance to the nearest auditor is uncorrelated with certification, while distance to the second closest auditor, as expected, results in lower certification. [23]

The second stage regression results are presented in Table 12. As can be seen, for all specifications, the coefficient on certification is negative, but not significantly different from zero. Given the moderate predictive power of the instruments we also constructed conditional likelihood ratio 95% confidence intervals, which are known to be robust to the presence of weak

[22] We also ran the IV specification using the distance to the first five auditors as instrumental variables, and the results were consistent with the ones presented here. The results for that specification are also available from the authors upon request.

[23] This result is consistent with a model in which auditors bid to serve each specific plant. As the cost of an audit is increasing in distance, if all auditors, on average face the same cost function, at the equilibrium, the closest auditor will bid the lowest price, which will be equal to the marginal cost of an audit plus the distance between the second closest auditor and the plant that is being served. While we agree that it is still possible for the auditing plants to be choosing their location endogenously (responding to demand), this result would hold if auditor´s location were exogenous.

instruments (Moreira 2003). The resulting confidence interval for the fourth column is -9.9 to 2.6. This range is at least consistent with the hypothesis that at the margin between certification and compliance without certification that is affected by auditor cost, certification does not result in substantially lower emissions than does compliance. At the extreme negative end of the confidence interval, however, the results indicate a substantial effect, with a 10% point increase in certification resulting in a 1.23 standard deviation decrease in AOD. This corresponds roughly to the difference in air quality between a central location in Mexico City and that in a low-industry rural area.

The coefficient on certification intensity is significant and positive and suggests that overall there was less decrease in emissions in high cost sectors than in low cost sectors. In particular, a zip code with only small plants (and thus not subject to an increase in inspection) in sectors at the low end of certification in Table 2 (.04) is predicted to have a 3.4 percent lower increase in AOD than an otherwise comparable zip code with small plants in sectors at the high end of certification (.10). The significant negative coefficient on the medium x fraction certified interaction indicates that in firms in high cost sectors subject to inspection there was no concomitant increase in emissions. In particular, for zip codes with only medium sized plants this difference in certification in results in a 4.1 percent decrease in AOD relative to that in zip codes with only small plants. This gap is about one quarter the difference between downtown Mexico City and a low-industry rural area. To put this result in context it is useful to point out that in previous work Foster, Gutierrez and Kumar (2009) looked directly at the question of respiratory effects and found that the elasticity of infant mortality with respect to AOD in Mexico is approximately 4. Thus to the extent that the primary difference between the trends in medium and small firms over the certification period arises from the difference in inspection probabilities by firm size, there is a 16.4 percent reduction in infant mortality associated with the indirect effect of certification.

VI. Conclusions

The specific focus of this paper has been on the effects on air quality of a voluntary pollution reduction program in Mexico. We develop and test a simple model of regulator and plant behavior that incorporates observable (to the regulator) sectoral variation in the cost of compliance with environmental regulations as well as unobserved variation in compliance costs within sectors. Our results suggest that those plants that certify are those with the lowest cost of compliance within sector and that certification provides an informational benefit that increases the efficiency of regulator monitoring of plant behavior. The model is then used to structure an analysis of the effects of the certification on a measure of suspended particulates using the zip code as the level of analysis. Our analysis suggests that the program primarily had an indirect effect on air quality.

In addition to these conclusions this paper has some more general implications for the analysis of regulatory behavior. First, the results suggest that the voluntary certification programs can be an important tool for reduction of emissions in low and middle income countries. By shifting the cost of auditing to the plants while at the same time providing some sort of tangible benefit, the regulating agency can more efficiently target its limited resources and thus induce higher levels of compliance. In addition, the results from this paper suggest that a voluntary certification program can be designed in such a way that it is most attractive to those plants with a relatively low cost of compliance thus reducing the overall cost of achieving a given level of compliance.

Second, in examining the effects of particular programs it is important to keep in mind what other programs are in place and how their implementation of these other programs is likely to be affected by other policies. In the absence of a systematic scheme to monitor and fine non-compliant plants, the effects of the Clean Industry Certification program would likely have been quite different. By the same token the experience of Mexico might not readily generalize to other settings. Even in the presence of experimental variation in access to the certification program it would be difficult to interpret measured effects of the program without a clear understanding of the interaction of different types of policy tools. The presence of these indirect effects also has implications for the establishment an appropriate control group for evaluating emissions among those plants that choose to certify. In this particular case, for example, the behavior of uncertified plants in terms of level of compliance is importantly affected by the presence of the certification program as a result of the endogenous response from the regulator.

Third, our results suggest that remotely sensed measures of air quality can provide a useful tool for the evaluation of emissions regulations in developing countries. As noted, few low and middle income countries have systematically collected ground level data on emissions. These countries also in general lack the capacity to monitor emissions of more than a small fraction of plants, particularly given the presence of a large informal and small-scale manufacturing sector. This lack of data, which is an important constraint for those wishing to control plant emissions and/or to implement a system for trading permits, also presents a problem for the evaluation of alternative programs. The technology for the processing of remote images to construct measures of AOD is still in infancy and much needs to be done over time to both evaluate and improve the accuracy of these measures. But the present work adds to a growing body of evidence that suggests that the technology has a great deal of potential.

VII. References

Alberini, A., and K. Segerson. 2002. “Assessing Voluntary Programs to Improve Environmental Quality.” *Environmental and Resource Economics* 22(1–2): 157–84.

Antonovics, Kate and Brian G. Knight. 2008. “A New Look at Racial Profiling: Evidence from the Boston Police Department.” *Review of Economics and Statistics*, 91(1): 163–177.

Arora, S., and T. Cason. 1995. “An Experiment in Voluntary Regulation: Participation in EPA’s 33/50 Program.” *Journal of Environmental Economics and Management* 28: 271–86.

Blackman, A. 2008. Can Voluntary Environmental Regulation Work in Developing Countries? Lessons from Case Studies. *Policy Studies Journal* 36(1): 119–41.

__________ .2010. “Alternative Pollution Control Policies in Developing Countries”. *Review of Environmental Economics and Policy*, forthcoming.

Blackman A, Lahiri B, Pizer W, Rivera Planter M and Munoz Pina, Carlos. 2007. “Voluntary Environmental Regulation in Developing countries: The Mexican Clean Industry Program.” *Discussion Paper. Resources for the Future.* RFP DP 07-36.

Chu DA. 2006. “Analysis of the relationship between MODIS aerosol optical depth and PM2.5 over the summertime US.” *Proceedings of SPIE*.

_______(2003) “Global monitoring of air pollution over land from EOS-Terra MODIS.” *Journal of Geophysical Research* 108(D21):1-18.

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.

Gupta P, et al. 2006. “Satellite remote sensing of particulate matter and air quality assessment over global cities.” *Atmospheric Environment* 40:5880-5892

Jiménez, O. 2007. “Voluntary Agreements in Environmental Policy: An Empirical Evaluation for the Chilean Case.” *Journal of Cleaner Production* 15: 620–37.

Khanna, M. 2001. “Economic Analysis of Non-Mandatory Approaches to Environmental Protection.” *Journal of Economic Surveys* 15(3): 291–324.

Khanna, M., and L. 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.

Knowles, John, Nicola Persico and Petra Todd. 2001. “Racial Bias in Motor Vehicle Searches: Theory and Evidence.” *Journal of Political Economy*. 109(1): 203-229.

Kumar N, Chu AD, Foster A. 2007. “An empirical relationship between PM2.5 and aerosol optical depth in Delhi Metropolitan.” *Atmospheric Environment*. 41:4492–4503.

Morgenstern, Richard and William Pizer (eds.) 2007. *Reality Check: The Nature and Performance of Voluntary Environmental Programs in the United States, Europe, and Japan*. Resources for the Future.

Organization for Economic Cooperation and Development (OECD). 1999. *Voluntary Approaches for Environmental Policy: An Assessment.* Paris: OECD Environment Directorate.

———. 2003. *Voluntary Approaches for Environmental Policy: Effectiveness, Efficiency and Usage in Policy Mixes.* Paris: OECD Environment Directorate.

Munoz-Pina, Carlos, Marisol Rivera-Planter and Jose Vicente Ruiz-Arredondo. 2006. "Incentivos Economicos de las Empresas a Participar en Acuerdos Ambientales Voluntarios", Instituto Nacional de Ecologia working paper.

Sam, A., and R. Innes. 2006. "Voluntary Pollution Reductions and the Enforcement of Environmental Law: An Empirical Study of the 33/50 Program." *Working Paper. Columbus: Ohio State University Department of Agricultural, Environmental, and Development Economics.*

Videras, J., and A. Alberini. 2000. "The Appeal of Voluntary Environmental Programs: Which Firms Participate and Why?" *Contemporary Economic Policy* 18(4): 449–61.

Vidovic, M., and N. Khanna. 2007. "Can Voluntary Pollution Control Programs Fulfill Their Promises? Further Evidence from EPA's 33/50 Program." *Journal of Environmental Economics and Management* 53: 180–95.

VIII. Tables.

Table 1

Select Sectors with Low and High Inspection and Certification Intensity	
Low Inspection/Low Certification	High Inspection/High Certification
Natural fibers	Synthetic fibers
Wine	Beer
Shoes	Explosives
Printing	Ink for Printing
Wooden Furniture	Paint
Office Supplies	Cleaning Products
Paper	Glue
Coffee/Tea Industry	Pharmaceuticals
Chocolates	Edible Oil
Wooden Construction Supplies	Cement

Table 2

Descriptive Statistics

Sector Level variables by Pre Certificates Inspection Probability (quartiles)

Quartile Inspection Intensity pre certificates	Range Fraction Inspected 92-95	Fraction Non Compliance 92-95	Fraction Non Compliance 03-06	Fraction Certified
1	Less than 2%	0.84	0.50	0.04
2	2-10%	0.79	0.65	0.03
3	10-30%	0.81	0.65	0.06
4	More than 30%	0.83	0.67	0.10

Table 3

Determinants of non Compliance at the sector level Dependent variable: Log of the percentage of inspections resulting in non-compliance pre certificates			
Log % Inspected (92-95)	-0.00338	-0.00435	-0.00081
	[0.00663]	[0.00828]	[0.00842]
Log Employees per firm		-0.00041	-0.00411
		[0.01215]	[0.01217]
% production exported		0.00014	0.00009
		[0.00011]	[0.00012]
Constant	-0.21166	-0.22002	-0.17073
	[0.01937]***	[0.05648]***	[0.06107]***
Observations	160	160	160
R-squared	0	0.01	0.05

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4

Determinants of Certification Dependent variable: Log of the percentage of firms certified			
Log % Inspected (92-95)	0.75164	0.30488	0.25081
	[0.07014]***	[0.05796]***	[0.05697]***
Log Employees per firm		1.14113	1.1957
		[0.08737]***	[0.08362]***
% production exported		-0.00292	-0.0022
		[0.00084]***	[0.00085]**
Constant	1.97504	-3.031	-3.67436
	[0.19940]***	[0.40507]***	[0.42746]***
Observations	138	138	138
R-squared	0.46	0.76	0.79

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5

Determinants of non Compliance (2003-2006) at the sector level Dependent variable: Log of the percentage of inspections resulting in non-compliance post certificates			
Log % Inspected (92-95)	0.01069	0.0176	0.01518
	[0.01156]	[0.01429]	[0.01452]
Log Employees per firm		-0.01296	-0.02091
		[0.02092]	[0.02103]
% production exported		-0.0002	-0.00034
		[0.00020]	[0.00021]*
Constant	-0.4136	-0.34133	-0.36191
	[0.03359]***	[0.09729]***	[0.10536]***
Observations	160	160	160
R-squared	0.01	0.02	0.05

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6

Second Stage						
Dependent Variable: Change in the Log of the % of Firms Inspected between 1993 and 1995 and 2003-2005						
	OLS	OLS	OLS	OLS	IV-1	IV-2
Log (% Inspected 1993-1995)	-0.23722				0.33094	0.19762
	[0.05579]***				[0.12985]**	[0.11517]*
Log (1+ % Certified)		0.30205				
		[0.09576]***				
Log (% Inspected USA)			0.20781			
			[0.07844]***			
Log (Abatement Costs USA)				0.20326		
				[0.09920]**		
Constant	-1.07129	-0.83926	-0.19585	2.15733	0.2393	-0.02097
	[0.16201]***	[0.14192]***	[0.16028]	[1.28414]*	[0.32492]	[0.28588]
Observations	160	160	160	143	160	143
R-squared	0.1	0.06	0.04	0.03		
Standard errors in brackets						
* significant at 10%; ** significant at 5%; *** significant at 1%						

Table 7

First Stage					
Dependent Variable: Log of the % of Firms Inspected between 1993 and 1995					
Log (1+ % Certified)	0.82516		0.47094		0.43617
	[0.11609]***		[0.14783]***		[0.15624]***
Log (% Inspected USA)		0.68976	0.43955		0.65163
		[0.09336]***	[0.12003]***		[0.14783]***
Log (Abatement Costs USA)				0.31443	-0.20078
				[0.14080]**	[0.13505]
Constant	-3.16774	-1.21722	-2.10384	1.81024	-4.41243
	[0.17205]***	[0.19076]***	[0.33444]***	[1.82259]	[1.71414]**
Observations	160	160	160	143	143
R-squared	0.24	0.26	0.3	0.03	0.34
Standard errors in brackets					
* significant at 10%; ** significant at 5%; *** significant					

Table 8

Distribution of plants in SIEM by Size and Certification Status		
Plant Size	Total	% Certified (matched)
Less than 10 employees	21949	0.10
Between 11 and 100 employees	7675	0.94
More than 100 employees	2708	11.04
Total	32332	1.22

Table 9

OLS Regression Results

Fraction Certified defined from SIEM data

Dependent variable: Difference in AOD between 2000 and 2006

Fraction Certified (SIEM)	-0.003	0.015	0.017	0.013
	[0.027]	[0.028]	[0.028]	[0.028]
Medium*Log Cert. Intensity wt		-0.045	-0.043	-0.047
		[0.014]***	[0.014]***	[0.015]***
Log Cert. Intensity wt	0.009	0.034	0.036	0.034
	[0.007]	[0.010]***	[0.010]***	[0.010]***
Log Size wt	-0.002	0	0.008	0.008
	[0.003]	[0.006]	[0.006]	[0.006]
% Medium		0.008	0.003	0.004
		[0.024]	[0.024]	[0.024]
% Exporting			-0.004	0.007
			[0.019]	[0.024]
% Importing			-0.062	-0.078
			[0.019]***	[0.023]***
Exporting*Log Cert. Intensity				-0.014
				[0.018]
Importing. * Log Cert. Intensity				0.022
				[0.017]
Constant	-0.084	-0.094	-0.095	-0.094
	[0.021]***	[0.021]***	[0.021]***	[0.021]***
Weather Controls	Yes	Yes	Yes	Yes
Month fixed Effects	Yes	Yes	Yes	Yes
Observations	19849	19849	19849	19849
R-squared	0.07	0.07	0.07	0.07

Robust standard errors clustered at the zip code level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10

OLS Regression Results

Fraction Certified defined from Industrial Census

Dependent variable: Difference in AOD between 2000 and 2006

Fraction Certified (Census)	-0.254	-0.246	-0.198	-0.19
	[0.281]	[0.285]	[0.290]	[0.290]
Medium*Log Cert. Intensity wt		-0.043	-0.042	-0.046
		[0.014]***	[0.014]***	[0.015]***
Log Cert. Intensity wt	0.009	0.034	0.036	0.034
	[0.007]	[0.010]***	[0.010]***	[0.010]***
Log Size wt	-0.002	0.001	0.009	0.009
	[0.003]	[0.006]	[0.006]	[0.006]
% Medium		0.006	0.001	0.002
		[0.023]	[0.023]	[0.023]
% Exporting			-0.003	0.007
			[0.019]	[0.024]
% Importing			-0.062	-0.078
			[0.019]***	[0.023]***
Exporting*Log Cert. Intensity				-0.014
				[0.018]
Importing. * Log Cert. Intensity				0.022
				[0.017]
Constant	-0.082	-0.093	-0.094	-0.093
	[0.021]***	[0.021]***	[0.021]***	[0.021]***
Weather Controls	Yes	Yes	Yes	Yes
Month fixed Effects	Yes	Yes	Yes	Yes
Observations	19849	19849	19849	19849
R-squared	0.07	0.07	0.07	0.07

Robust standard errors clustered at the zip code level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 11

First Stage Regression Results

Fraction Certified defined from Industrial Census

Dependent variable: Fraction Certified

Distance to first auditor	0.002	0.002	0.001	0.002
	[0.002]	[0.002]	[0.002]	[0.002]
Distance to second auditor	-0.002	-0.002	-0.001	-0.001
	[0.001]**	[0.001]**	[0.001]**	[0.001]**
Distance to first auditor squared	0.001	0.001	0.001	0.001
	[0.000]	[0.000]	[0.000]	[0.000]
Distance to second auditor squared	0	0	0	0
	[0.000]	[0.000]	[0.000]	[0.000]
Constant	0.006	0.006	0.006	0.006
	[0.001]***	[0.001]***	[0.001]***	[0.001]***
Observations	19849	19849	19849	19849
R-squared	0.03	0.03	0.04	0.04
F-statistic	6.81	6.68	6.82	6.87

Regressions include all controls in second stage regression.

Robust standard errors clustered at the zip code level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 12

IV Regression Results

Fraction Certified defined from Industrial Census

Instruments: Distance to first 2 auditors (and their squares)

Dependent variable: Difference in AOD between 2000 and 2006

Fraction Certified (Census)	-2.334	-2.062	-1.831	-2.11
	[2.614]	[2.639]	[2.616]	[2.620]
Medium*Log Cert. Intensity wt		-0.042	-0.041	-0.045
		[0.014]***	[0.014]***	[0.015]***
Log Cert. Intensity wt	0.012	0.037	0.038	0.037
	[0.008]	[0.010]***	[0.011]***	[0.011]***
Log Size wt	-0.001	0.004	0.01	0.01
	[0.003]	[0.007]	[0.006]	[0.006]*
% Medium		0.001	-0.003	-0.001
		[0.025]	[0.024]	[0.024]
% Exporting			0.002	0.009
			[0.021]	[0.024]
% Importing			-0.062	-0.074
			[0.019]***	[0.024]***
Exporting*Log Cert. Intensity				-0.008
				[0.019]
Importing. * Log Cert. Intensity				0.018
				[0.017]
Constant	-0.144	-0.157	-0.16	-0.158
	[0.024]***	[0.025]***	[0.025]***	[0.025]***
Weather Controls	Yes	Yes	Yes	Yes
Month fixed Effects	Yes	Yes	Yes	Yes
Observations	19849	19849	19849	19849
R-squared	0.07	0.07	0.07	0.07

Robust standard errors clustered at the zip code level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Figure 1
Cost Schedule 1. $\alpha < 1, \quad \beta > 1.$

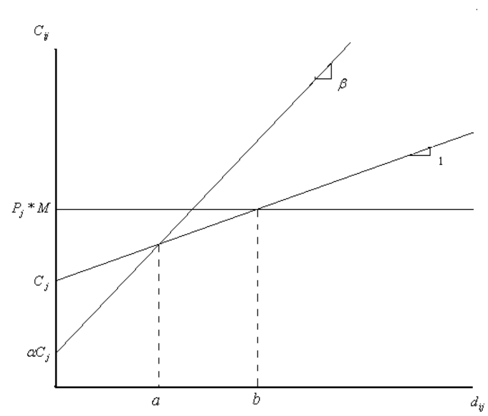


Figure 2
Cost schedule 2. $\alpha > 1, \quad \beta < 1.$

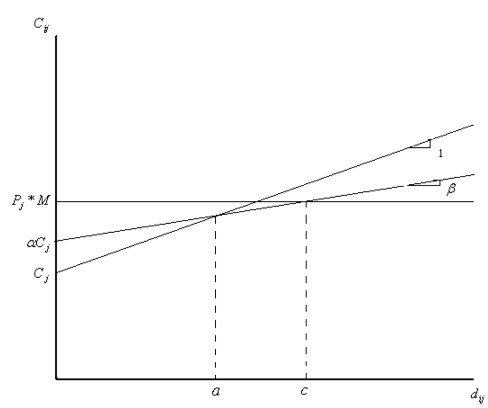


Figure 3
Cost schedule 3. $\alpha > 1$, $\beta < 0$.

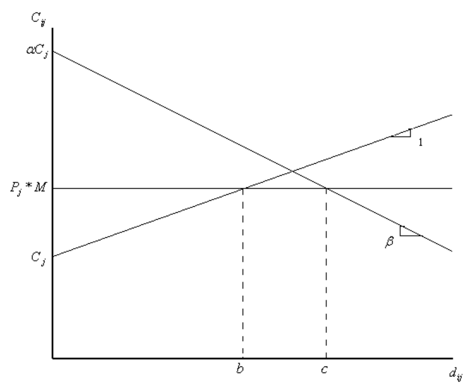


Figure 4

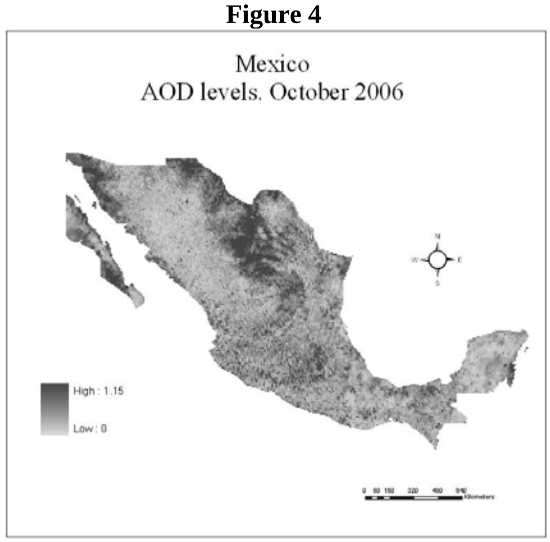


Figure 5


Mexico
Changes in Ln AOD between 2000 and 2006
(October)
High : 4.8
Low : -7.1

Appendixes (Not for Publication)

Appendix I. Incorporating corruption.

Corruption can take different forms. First, plants can avoid to pay the fine set by the legal system in case they are found to be in non-compliance by paying the inspector some given sum of money (strictly lower than the fine). In the model, however, as long as the probability of corruption is the same for all firms in the economy, this would simply mean that the value of *M* is strictly lower than the one set by the legal system, and the theoretical predictions of the model would not change substantially. Stronger assumptions have to be made, however, for our empirical analysis to be consistent with a setting in which this type of corruption is possible. As we observe the total fraction of inspections resulting in non-compliance reported by the authorities, inspected plants that avoid being declared as non-compliant by bribing the inspectors, would not be counted in our measure of non-compliance. Our empirical analysis needs then to assume that there is a constant probability that a given plant (regardless on which industrial sector it belongs to), if inspected and found in non-compliance, is unable to corrupt the authorities.

Another possible form of corruption is extortion: inspectors might be able to threaten a compliant plant to declare it as non-compliant every time it is inspected, unless the plant pays some bribe *E*. In this case, the cost equation for a plant to be in compliance with no certificate would be:

$$C_{ij}^{1b} = C_j \quad + d_{ij} + P_j \quad * E$$

Where E is the amount that the plant has to pay to be declared in compliance when inspected and it is complying, and P_j is the probability of being inspected. It is easy to see that, as E increases, compliance goes down. However, E imposes an extra incentive for plants to get certified, as they are exempt from inspections (in the context analyzed), and thus from paying E to inspectors. The conditions for the identification of regime derived in section III.3 are consistent with the conditions derived from a model incorporating this form of corruption.

Finally, there can be corruption in the certification process. Plants might be able to pay some fixed amount G in order to be given a certificate when in non-compliance. However, auditors need to be licensed by authorities. Moreover, although plants are exempt from random inspections if certified, they are still subject to inspections in case other agents in the economy file a complaint to the authorities. As a smaller number of auditors is likely to imply higher monopoly power to certified auditors, each auditor has incentives to file complaints about firms that have been certified by their competition, expecting the authorities to take the license away from corrupt auditors. Thus, we believe that a model that ignores this type of corruption is well suited for the specific context analyzed.

Appendix II. A more explicit model for certification costs.

The purpose of this appendix is to suggest why within-sector differences in costs of compliance may lead to larger differences in certification costs ($\beta > 1$), even though certification costs focus on the same emission targets. The premise of this model is that although the certification process generally results in mandated upgrades to capital, compliance in the absence of certification can be achieved through adjustment along other margins and, in particular, through reductions in output. This insight rests on the fact that legal targets are set in terms of emissions per unit of plant capacity.

We model variation in compliance cost within sector (the *d* component) as driven by differences in the share of "green" capital. Suppose, in particular, that a plant that only uses capital in production and has a 1-*f* share of green capital that produce emissions of 1 per unit capital and an *f* share of dirty capital that produces emissions of *e* per unit of capital with *e*>1. Suppose further that the regulations indicate that emissions per total capital must be less than *n* where *n*>1, and that capital and product markets are sufficiently competitive that the marginal revenue product *p* of a unit of capital is equal to the interest rate *r*. Finally, assume that given *e* and *n*, *f* is sufficiently high that the firm is not in compliance with the regulatory standard if it produces at full capacity.

It is now clear that the firm can be in compliance if it uses only a fraction *d1* of its dirty capital where (1-*f*)+*d1***f***e*=*n*. Solving for *d1* this costs it *p**(*f**(1-1/*e*)-(*n*-1)/*e*) in lost value of output per unit of total capital per period. The discounted cost of an infinite stream of lost earnings is this figure divided by *r*. A unit increase in f thus results in an increase of (*e*-1)/*e*<1 in the cost of compliance.

Conversely, the firm can be in compliance if it replace a fraction *d2* of its dirty capital with clean capital where (1-*f*)+*d2***f*+(1-*d2*)**f***e*=*n* and produces at full capacity. This change costs *f*-(*n*-1)/(*e*-1) in replacement capital per unit of total capital. A unit of increase in *f* in the capital conversion case thus results in a unit increase in the cost of compliance but a less than unit increase in the cost of compliance for output reduction. Intuition maybe gained from the case of *e*=2 and *n*=1. Note that for capital conversion case all dirty capital must be converted so the cost is *f*. But because of the way limits are set, with output reduction, one can continue to use half of the dirty capital and still meet the standard so the cost is *f*/2

In general a firm choosing to comply will choose the cheaper of the two avenues of compliance. But the two cost lines cross at zero and thus for the range of *f* for which the firm is not initially in compliance *f*>(*n*-1)/(*e*-1), reducing output will always be cheaper than converting capital. However if certification leads to a sector specific net reduction in cost ($\alpha < 1$) and must result in the use of the capital upgrading mode of compliance then certification will be the lower cost choice for the lowest cost firms and compliance without certification the lowest cost choices for intermediate cost firms.

Appendix Table 1

Sector Level Descriptive Statistics			
	Observations	Mean	Std. Dev
Number of establishments	160	2128	6590
Number of employees	160	24575	44548
Percentage Certified	160	5.12	14.55
Fraction Inspected (92=95)	160	0.32	0.67
Fraction Inspected (03-06)	160	0.18	0.37
Fraction in non Compliance (92-95)	160	0.82	0.12
Fraction in non Compliance (03-06)	160	0.62	0.55

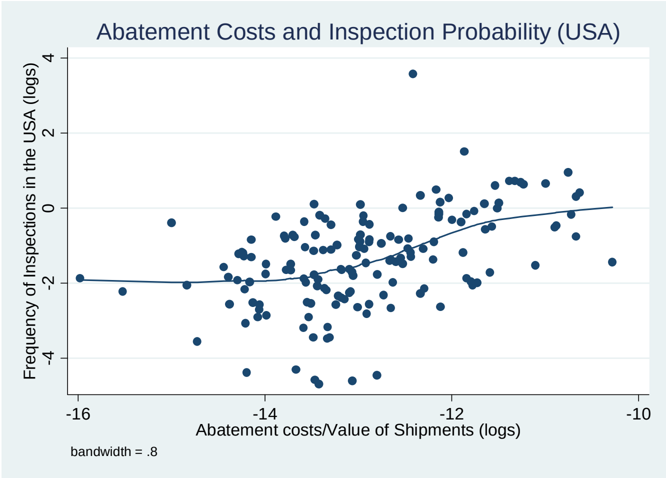
Appendix Table 2

Zip Code Level Statistics			
Variable	Obs	Mean	Std. Dev.
Fraction Certified (SIEM)*	2668	0.027	0.140
Fraction Certified (Census)*	2668	0.010	0.017
Fraction Medium	2668	0.479	0.453
Log Employees Weighted	2668	2.633	1.915
Certification Intensity Weighted	2668	2.585	9.481
Log (1 + Certification Intensity Weighted)	2668	0.515	0.707
Log AOD	2668	-1.328	0.806
Fraction Importing	2668	0.265	0.382
Fraction Exporting	2668	0.202	0.348
Distance to first auditor (100km)	2668	0.541	0.787
Distance to second auditor (100 km)	2668	0.782	0.949
Square Distance to first auditor (100km)	2668	0.912	2.057
Square Distance to second auditor (100 km)	2668	1.513	3.031

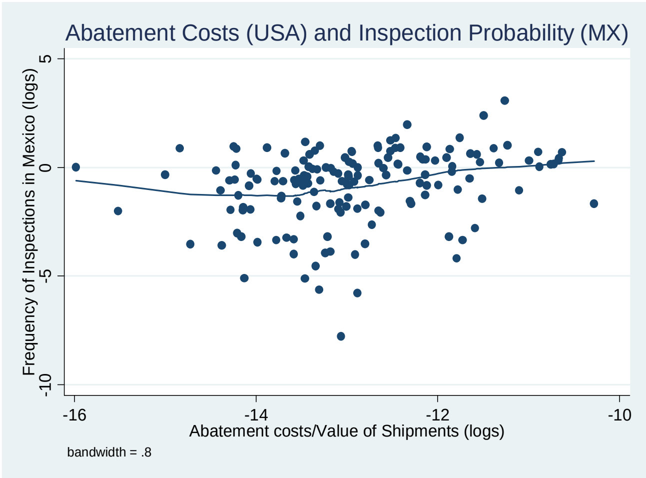
* Fraction certified (SIEM) is defined as the weighted fraction of plants certified from the SIEM data, using only those certified plants found in SIEM.

* Fraction Certified (Census) is the ratio between all certificates granted in the county and the total number of establishments listed in the 2000 Industrial Census in each county.

Appendix Figure 1

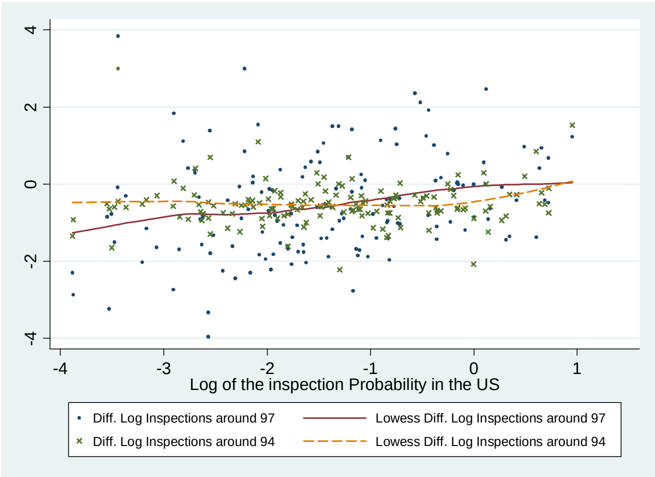


Appendix Figure 2



Appendix. Figure 3

Appendix. Figure 4



Appendix. Figure 5