Draft: 02.26.09

Direct and Indirect Effects of Voluntary Certification:

Evidence from the Mexican Clean Industry Program

Andrew Foster

Emilio Gutiérrez

Brown University

This work supported in part by a Hewlett/PRB grant on the effects of Health and Demographic Change on Economic Growth. We wish to thank Naresh Kumar who processed and provided us with the satellite-based AOD measures for Mexico over the relevant period. We also with to thank Raul Tornel, Jaime Garcia Sepulveda, and Jose Domingo Morales of PROFEPA who provided us with the inspection and certification data and described in some detail the nature of their programs and the Mexican Ministry of Economics, which provided the firm-level data. We also thank seminar participants at Brown, Vanderbilt, University of Illinois, Guanajuato, ITAM, Université de Montréal and Banco de México for their comments and suggestions. All errors are ours.

Abstract

In this paper we develop a model of environmental regulation in a developing country that integrates firm 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 firms 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 firms with relatively low cost 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 firms 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 firms. Thus, according to our model, the reductions in pollution emissions levels should not only be observed amongst participating firms, but also amongst non-certified firms in industrial sectors with a high percentage of certified firms. 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 firm emissions. We surmount this problem by integrating newly developed satellite based measure of suspended particulates and a firm-level data set with geographical identifiers. As predicted by the model we find that particulate matter concentrations are significantly lower both in areas with certified firms and in those with a large fraction of non-certified firms in high-certification sectors.

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. It has been noted, for example, that the 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. But similar concerns arise with respect to attempts to cap emissions through direct regulation of firms— monitoring costs relative to output may be high and the government agency in question may have limited capacity to monitor firms, particularly given that a substantial fraction of polluting activity may occur among relatively small firms 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 have a different impact on different types of firms—e.g., exporting firms might arguably be asked to meet a different standard than are those primarily producing for domestic markets. But obviously this sort of differentiated policy has the potential to create important market distortions and rent seeking behavior.

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 emissions reductions on firms with low compliance costs and/or in those firms who 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 firms to reduce their emissions levels in both the developed and developing world (OECD, 1999, 2003).

This paper develops a model of environmental regulation in a developing country that integrates firm 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 firms 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 firms with relatively low cost 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 firms 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 firms. Thus, according to our model, the reductions in pollution emissions levels should not only be observed amongst participating firms, but also amongst non-certified firms in industrial sectors with a high percentage of certified firms. 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 firm emissions. We surmount this problem by integrating newly developed satellite based measure of suspended particulates and a firm-level data set with geographical identifiers. As predicted by the model we find that particulate matter concentrations are significantly lower both in areas with certified firms and in those with a large fraction of non-certified firms in high-certification sectors.

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), generally studying these programs in isolation from other environmental programs. The existing literature, which has primarily focused on industrialized countries, is especially concerned with testing whether participating firms are those already in compliance with the emissions standards, or if firms 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 firms participating in the program were those who had reduced their emissions before the implementation of the Program. However, (Vidovic and Khanna, 2007) find the opposite result. According to the latter, a very small percentage of the total emissions by participating firms can be attributed to the program. Some of these studies also try to test if the firms 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).

There is also one study that looks directly at the Mexican Clean Industry Program (Blackman et al., 2007), which is the focus of the present paper. That study shows that firms that have been inspected or fined for not complying with pollution emissions standards in the past are more likely to participate. It argues that this is evidence that the program is contributing to pollution reduction, given that participating firms are more likely than average to be in non-compliance before entering the program, and presumably in compliance with environmental regulations when they graduate from it. But there are alternative possible conclusions including that the original fine caused the firm to become compliant and thus the costs of certification were minimal. No direct measurement of changes in air quality is provided as part of that study, which would help in the adjudication of this issue.

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 firm 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 firm 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 firm level that makes direct regulation difficult also means that one rarely has available firm 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 air-quality in a region. They thus are ill-suited to a systematic evaluation of the effects of a change in policy.

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). 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. Inspections are performed at random, assigning a higher probability of inspection to sectors with higher perceived risk of polluting. If a plant is found to be in non compliance, it is forced to pay a fine, which increases in case of relapse. Relatively small plants are inspected less frequently. 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. The plant is then 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. 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.

The model in this paper integrates these two components by assuming that inspection probabilities are chosen by the regulator to maximize environment benefits net of audit costs borne by the regulator and that firms compare the costs of certification, compliance without certification, and non-compliance given the fines and probability of being inspected. In particular, by imposing some structure to the cost of participation and compliance we derive testable predictions that help establish whether firms 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 firms 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 reduces the cost of inspection in sectors with a high percentage of certified firms. Thus, according to our model, the reductions in pollution emissions levels should not only observed amongst participating firms, but also among non-certified firms in industrial sectors with a high percentage of certified firms. We then test these implications by integrating a newly developed satellite based measure of suspended particulates and a firm-level data set with geographical identifiers.

III. Modeling firms' participation in the Clean Industry Program

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 firms. Firms must decide whether to comply or not comply with emission standards prior to the introduction of 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. The model is built around some key empirical facts that we now proceed to describe.

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-cost 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[1].

[1] 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.

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[2]. At first glance, it is worth noting that the fraction of inspected firms is not significantly related to the level of compliance either before or after the introduction of certification. However, the level of certification is significantly higher in those sectors in which inspection rates are high. 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. A key insight is that the non-compliance results mirrors an important idea that has arisen in attempts to evaluate whether police discriminate based on the race of motorists. (Antonovics and Knight, 2008; Knowles, Persico and Todd 2001). In the present context, the idea is that PROFEPA is imposing sufficiently higher inspection probabilities in sectors with a high cost of compliance, relative to those with a low cost of compliance, such that firms in both types of sectors are equally likely to choose compliance.

III.1. The Firms' Problem

We now turn to the structure of the model. We assume, as noted, that firms have a choice between three different options: Complying with pollution emissions standards without getting certified as clean firms; compliance with emissions standards and obtaining a "Clean Industry Certificate"; and non compliance. Note that all certified firms are assumed to be in compliance, but not all the non-certified firms are non-compliant. Each of the options has a different cost for each firm, depending both on the types of goods they produce (industrial sector) and unobserved firm specific characteristics. Firms will

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

choose the option that has a lowest cost for them. Authorities do not observe the firm specific cost of compliance, but do observe a compliance cost that is common to firms in the same sector.

In particular, the cost of compliance with pollution emissions standards without certification for firm *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 firm specific cost, not observed by the authorities, with a distribution *F*. The cost of certification is assumed to be a linear function of these cost components:

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

where α and β are constants, which for the moment we will assume as unknown, although common to all firms.[3] As can be seen, α multiplies the industrial sector level cost of compliance and β multiplies the firm specific cost of compliance. Effective costs of compliance for those firms that certify may differ from the costs of compliance in the absence of certification because (a) possible marketing benefits (b) reductions in liability (c) the costs of an audit and (d) the grace period provided. 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 firm in sector *j* and *M* is the fine imposed if the firm is found to be in non compliance. *M* is assumed to be fixed and exogenous[4] and P_j is set at the sector level, given that authorities are unable to observe (or unwilling to use) the firms' specific d_{ij}. As

[3] 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.

[4] Given the structure of the model allowing M to vary would not increase the ability of the firm to ensure compliance.

stated above, we assume that the firm specific component of the cost of compliance is only observed by the firm.

In the absence of certificates, it is clear from equation (1) and (3) that only firms 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[5] (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 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. The assumption $\alpha < 1$ implies that most of the benefits from participating in the program are common to all firms within one industrial sector. The assumption $\beta > 1$ implies that the cost of participating in the program is higher for firms with relatively high compliance costs. Firms will get certified if $d < a$. Firms for which $a < d < b$ will be in compliance and not certified, and firms with $d > b$ will be in non compliance. Of course, in the absence of certification, the certification cost is not available to firms and all firms to the left of *b* will be compliant. It is not obvious, however, whether the introduction of certification lowers or raises overall compliance because, as illustrated below, the introduction of certification entails an adjustment in the inspection rate and thus the point at which the compliance and non-compliance cost curves intersect.

---

[5] 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: $MP_j < \frac{(\beta-\alpha)}{(\beta-1)}C_j$.

Figures 2 and 3 plot the same three hypothetical cost schedules, this time for $\alpha > 1$. In Figure 2, β is set to be lower than one but higher than zero, illustrating a situation in which firms with intermediate levels of *d* get certified. Figure 3 shows the extreme case, in which β is negative, implying that the firms that get certified are those with the highest levels of *d*. In both of these cases, at least some of the firms getting certified are firms that would not be in compliance in the absence of the program, at least for a given inspection probability.

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 firms 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 firm and to pay a cost of *B* for every inspection. The regulator maximizes benefits minus costs through the choice of inspection probabilities by sector:

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

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. Thus $L_j = L_{j'}$ for all j, j'. Because the distribution of observed 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. 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-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.

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 firms in sector *j*, given regime (based on Figures 1-3) *k*. 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:

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

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

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 firms 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 firms 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. Let us consider the regimes in reverse order. In regime (3) the compliance rate among non-certified firms is

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

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,

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

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:

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

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 firms in regime 2 is

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

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 firms 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* yields

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

The implication 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 firms is if regime 1 is in place and the unobserved costs is generated according to an exponential distribution with hazard θ. Under these assumptions compliance among non-certified firms will be

$$R_j^1 = 1 - \frac{B}{\theta AM}. \tag{23}$$

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 firms that certify are those that have the lowest cost of compliance within their sector.

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

$$1-\frac{B}{MA}. \tag{24}$$

and the total fraction compliant under certification

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

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

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

and it is not clear how (24) compares to (26). Thus, even though the lowest cost firms within each sector will be compliant regardless of the certification program, it is possible that there are some firms 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 cost 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 firms 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 combine 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 (state-owned) 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.

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. 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 emission 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. These three 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. Firm level data

Given the difficulty of accessing geographically or 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 firms in the industrial sector. It includes each firm's name, exact address (including zip code), NAICS industrial sector, number of employees and dummy variables indicating whether the firm exports or imports. SIEM does not include government owned firms. The geographic coordinates of each of the firm's zip code was obtained from Postal Code World©, which provides geographic coordinates for the 2900 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 firms certified in each sector calculated previously in this paper is also assigned to each firm, given their declared NAICS sector in SIEM.

SIEM does not include all firms in Mexico. For this reason, not all of the certified firms listed by PROFEPA were found in our firm level database. 406 of the 1,266 certificates listed by PROFEPA for firms in industrial sectors were successfully matched to the SIEM data. While other issues about selection of firms into the sample could bias our empirical results, the percentage of certificates matched to the SIEM data in each industrial sector does not seem to be correlated with the percentage of firms certified in that sector.[6]

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.

IV.3. Air quality data

A key issue with 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

[6] The results of this test are available from the authors, upon request.

Aerosol Optical Depth (AOD) at a 5km spatial resolution for cloud-free areas over the whole of Mexican territory for the period between February 1st 2000 to December 31st 2006.[7]

It is perhaps useful to provide some background on the measurement of AOD using remotely sensed images. In particular, aerosols are liquid and solid particles suspended in the air. AOD can be described as the extinction of beam power caused by the presence of these particles in the atmosphere. 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 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), 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. At least one observation for each of the 2668 zip codes is included in the analysis.

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

[7] We gratefully acknowledge the assistance of Naresh Kumar in accessing these data.

and Monterrey), other regions of the country seem to show comparable high 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. High levels of AOD outside large metropolitan areas are then possibly driven by the fact that, while the positive relationship between AOD and suspended particles is generally strong within regions, comparisons across regions are hard to make, given the great variety in geographic and weather characteristics of each of location, especially when dealing with an area that is as variegated as Mexico. 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. An average monthly value of the temperature and dew point were calculated for each weather station. Then, these values were interpolated using an inverse distance weight technique for the whole Mexican territory. In particular , 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.

Table 3 provides descriptive evidence on the basic data. We see first of all that the fraction of plants inspected is sufficiently high that inspection risk is likely to be an important consideration for firms in terms of making decisions about compliance with environmental regulations. In particular, 32 percent of plants were inspected in the 92-95 period, though this declined to 18 percent in the 03-06 period. It is also worth noting that there is substantial variation across sectors in the percent inspected. The standard deviation is .67 in the prior period and .37 in the latter period. It is also evident that the levels of inspection (and the cost of fines for being in violation) are not so high that all plants are in compliance. Averaging across sectors, 82 percent of the plants prior to the introduction of certification were not in compliance. About half of the difference in inspection between the pre- and post-certification periods may be accounted for by the fact that a number of firms were certified (5.12 percent) over this period and thus did not need to be inspected by the agency. The other half presumably reflects some changes in the perceived costs B and benefits A on the part of the regulator of providing inspections that we assume to be fixed across sectors. Table 4 provides means and standard deviations of the zip code level data. As stated, there are 2688 zip codes included in the sample. We find that averaged in this way 2.7 percent of firms were certified, when defining certification from the SIEM data, and 1 percent when defined from the imputed values for each zip code from the census data and certificates records—reflecting the fact that certification is concentrated more in some zip codes than in others.[8] On average the firms had 126 employees with 48 percent of firms having between 10 and 100 employees. We also see that about one quarter of firms in the database exports.

V Results

V.1 Pre-certificate compliance

We now turn to an examination of the relationship at the sector level between non-compliance and the probability of inspection prior to the introduction of certification. This examination establishes whether the results of Table 2 hold up when controlling for other characteristics of the sector. Given the expression for compliance prior to certification (equation 24) we do not anticipate that these other characteristics of the sector should affect compliance. Thus this analysis also provides a kind of specification check for our model—if, for example, the costs of inspection B or perceived benefits A varied by sector and were systematically related to an observable measure that should be evident in such a regression. In particular we control, from the census data, for the log of employees per firm and the log of the percentage of production exported. The regression run is thus:

$$NC_j = \varphi_0 + \varphi_1 P_j + \sum_s \lambda_s X_{sj} + \varepsilon_j \qquad (29)$$

where NC_j is the percentage of inspections resulting in non compliance in sector j (in logs); P_j is the log of the percentage of firms inspected in each sector in the period between 1992 and 1995; and the X_s are a set of s sector level characteristics. As stated, the coefficient of interest is ϕ_1, which we expect to be close to zero.

Regression results are shown in Table 5. Column 1 shows the coefficient of the regression when including no controls. Column 2 includes the log of the employees per firm and the percentage of the

[8] The certification rate defined from census data is lower than that from the matched certificates to the SIEM data because the census data divides the number of certified firms by all firms in the census. The denominator is higher. Also, the numerator is not weighted by the total number of employees in each certified firm, as we do not have that information for certified firms not found in the SIEM.

production exported in each sector, and Column 3 adds a 2 digit sector level fixed effect. The coefficient for the log of the probability of inspection is close to zero and insignificant for all three specifications. 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.

V.2 Certification

In Table 2 we established a positive relation between certification and percentage inspected pre-certificates. In order to explore this relationship in greater detail, Table 6 shows the results of a set of linear regressions with the log of the percentage of firms certified as the dependent variable, and the probability of inspection pre-certificates as the explanatory variable (excluding 12 sectors where no firms received the certification). Column 1 does not include any controls. Column 2 includes the log of the number of employees per firm and the percentage of the production exported as controls (calculated from the 1999 Mexican census) and, finally, Column 3 also includes 2 digit NAICS sector fixed effects. 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 firms certified before the introduction of certificates (our proxy for the observed cost of compliance at the sector level) is always positive and significant. Note also that those sectors with large average plant sizes are more likely to be certified but those with a high export share are less likely to be certified. If the structure of our model holds, it is the plants with relatively low cost of compliance (within sectors) that are getting certified—

though as noted it is not necessarily the case that all of them would be in compliance in the absence of the certification program.

V.3 Post-certificate compliance

Table 2 shows that the relationship between the probability of inspection pre-certificates and non compliance after the introduction of certificates is close to zero. As this relationship is crucial to determine which firms, within sectors, are participating in the program, Table 7 shows the results for the same specification as Table 5. This time, non compliance between 2003 and 2006 is the dependent variable. As for non-compliance pre-certificates, the coefficient for the log of the probability of inspection is close to zero and insignificant for all three 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 firms that certify are those that have the lowest cost of compliance within their sector.

V.4 Change in inspection probabilities

Our model predicts that, if the authorities actually obtain information about firms' cost of compliance given certification, 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 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).

It is worth noting that the probability of inspection observed by firms after the introduction of the Clean Industry Program is not equal to the total number of inspections performed by the authority divided by the total number of firms, as certified firms are taken out of the inspection pool. However, in our regressions, 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 firms in the census, an overestimate of the total number of firms subject to inspection, and an underestimation of the probability of inspection in sectors with certified firms. 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 and the rate of certification.

Table 8 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 two alternative measures of sector costs, the percent certified and the percent inspected in the US. Table 9 shows that these variables are in fact strongly predictive of percent inspected, with the combined R-squared being 30 percent. When we include these variables in the Table 8 regression directly, instead of initial percent inspected in Mexico, we find the expected positive relationship. The fourth column uses these alternative two measures of sector cost as instruments for initial percent inspected in Mexico and also shows a positive effect. 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. Figures 6 and 7 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 (Figure 6) and the percent inspected in the US (Figure 7). 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 7. 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 firms 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 only be observed amongst participating firms, but also amongst non-certified firms in industrial sectors with a high percentage of certified firms. Among other things, this implies that one cannot use uncertified firms 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 firm level data from SIEM (Sistema de Información Empresarial Mexicano), administered by the Mexican Ministry of Economics, 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 firms certified in an industrial sector will create an incentive for non-certified firms in that sector to reduce their pollution emissions, given the authorities' response in terms of inspection intensity. The model also indicates that it is possible that some certified firms would not have been compliant in the absence of certification, in which case there may be a direct effect on certification as well, especially in sectors with a high percentage of certified firms.

In order to estimate the direct effect of certification, simply correlating certification with changes in pollution would not necessarily capture a causal relationship between these variables. For example, a firm experiencing an exogenous decrease in its cost of compliance during the time period analyzed would make it more likely to participate in the program. 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. In order to isolate the causal direct effect of certification on pollution concentrations, we exploit the fact that firms that participate in the program have to pay for and independent audit, given that the underlying costs of conducting a private audit are likely to predict certification but not compliance net of certification. This theoretical argument 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 firms.

Table 4 shows that the average distance to the first auditor is 54 kilometers, while that to the second is 78 kilometers.

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 discreet change in the inspection probability by firm size. Because PROFEPA prioritizes firms, first by industrial sector, assigning a higher inspection probability to sectors with high polluting potential, and then assigning a much higher likelihood of inspection to larger firms, small firms are at very low risk of inspection.[9] Moreover, the inspection probability seems to discretely increase at some level at which firms start being classified as “medium sized”. Thus, in testing for indirect effects of certification it is necessary to consider the size-distribution of firms in a given area. Although the inspections data provided by PROFEPA does not include information on the size of inspected firms, we can classify firms by size using the SIEM data.

Figure 8 shows a kernel density estimate for the size of firms restricting the sample to firms with less than 100 employees. As can be seen, there is a very high concentration of firms with a small number of employees, and the density seems to flatten considerably after 10 or 15 employees. The high concentration of firms with less than ten employees corresponds to micro enterprises not subject to inspections. As the effective threshold is likely to be somewhere around 5 and 15, we simply classify firms as having less than ten employees, or more.

Table 10 shows the distribution of firms in the SIEM database by size and certification status. In our sample, 68 percent of the firms listed in SIEM declare having less than 10 employees. Out of all of them, 0.1 percent was matched with the certificates list. 24 percent of firms have more than 10 and less than 100 employees, and the matching rate with the certificates list is nearly ten times higher than for

[9] www.profepa.gob.mx

the smallest firms in the sample. 8 percent of firms have more than 100 employees, and 11 percent of them are matched with the certificates list.

In our estimation strategy, we will exploit this change in the probability of inspection in order to isolate the effect of certification on the emissions of non-certified firms as a result of increases in inspections from sector specific trends in pollution emissions, not related to the authorities' inspection policy. If the reductions in emissions are actually related to an increase in the inspection probability given certification, firms with less than ten employees should not be reducing their emissions as a result of certification. The following equation describes the specification that would identify the impact of the program on pollution emissions:

$$\Delta Poll_i = \alpha + \beta_1 C_i + \beta_2 * M_i * CI_i + \beta_3 CI_i + \beta_4 M_i + \sum \gamma_k X_{ki} + \varepsilon_i \tag{30}$$

where $\Delta Poll$ is the change in pollution emissions by plant i, C_i is a dummy variable equal to 1 if the firm has received a Clean Industry certificate (instrumented by the distance to the first and second closest auditor to each zip code), M_i is a dummy variable equal to 1 if the firm has more than 10 employees, which isolates the difference in emissions for firms 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 firm, a dummy variable indicating if the firm exports, a dummy variable indicating if the firm imports, and the interaction of these variables with CI. Finally, CI_i is the percentage of firms certified in firm *i*'s industrial sector which, along with its interaction with the log of the firms size (included as one of the X_{ki}), controls for the changes in emissions correlated with certification, but uncorrelated with the inspection probability. The interaction between the size of the firm and the percentage of firms certified in the sector controls for differences in pollution emissions by different sized firms in sectors with different percentages of certified firms, but that are unrelated to the increases in the inspection probability given certification.

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 for the interaction between the percentage of firms certified in the firm's industrial sector (*CI*) and the dummy variable indicating if the firm has more than 10 employees, and thus is subject to an increase in inspections given certification. Given that our data do not measure pollution emissions by firms, but rather pollution concentrations around firms' zip codes (and that more than one firm 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 firm.

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 declared by each firm divided by the total number of employees in each zip code (the sum of the employees of all firms 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 firms not included in the SIEM database. As stated before, SIEM does not list government owned firms. Also, if not all the listed certificates were matched with the SIEM database, there are other private firms 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 firms not included in our sample to be uncorrelated with the industrial composition of each zip code calculated from the SIEM database.

V.5.a. OLS results

The results of the OLS regression (not instrumenting for certification) defining certification in each zip code from the certificates matched to SIEM are presented in Table 11. Table 12 shows the results of the same specification, defining certification in each zip code from census data. In both cases, column 1 is the regression output for the change in AOD at the zip code level against the weighted percentage of firms 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 firms. As can be seen, the coefficient on certification is negative and significant for most specifications, suggesting that certified firms actually do experience a reduction in their pollution emission levels. However, no causal relationship can be yet established.

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 from the fact that it is only firms that are subject to inspection that seem to reduce their pollution emissions as a result of other firms in their sector getting certified. Our coefficient of interest, that of the interaction between the dummy variable indicating if the firm 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.

The positive and slightly significant coefficient on the certification intensity suggests that it is not sectors that are reducing their pollution for reasons different from an increase in inspections that are getting certified. It seems that sectors with high certification intensity, when not subject to inspections, actually increase their emissions. These results provide good evidence supporting our theory. Firms in sectors with a high percentage of certified firms are reducing their pollution emissions given the increase in the observed probability of being inspected resulting from other firms getting certified.

V.5.b. Instrumental variables estimates

Figure 9 shows the geographic location of each of the 94 auditors in Mexico, as well as an interpolation of the fraction of firms certified around the country (defined from census data). While the relationship will be clearer in the first stage regression results, certification does seem higher in geographic regions served by a higher number of auditors.

However, probably given the low number of certificates matched to the SIEM data, the first stage for the fraction of firms certified in each zip code when using the certification variable calculated from SIEM data, shows coefficients consistent in signs but not statistically significant to the ones presented in this section. The instrumental variables analysis is then restricted to the specification where we defined certification at the zip code level from the imputed values from census and certification data.[10]

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 13[11]. Each column includes the same controls as the corresponding column in Table 12. As can be seen, distance variables strongly predict certification with the joint F on these variables being above 7 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. Given a simple model of Cournot

[10] The first stage regression results using the certification variable from SIEM data is available from the authors upon request.

[11] 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.

competition in which distance increases the marginal cost of serving a particular area, the positive zero on distance to the nearest firm is a bit surprising. The pattern can, however, be easily generated by a model in which there are fixed costs of entry and entry is endogenous. In particular, when the close firm has a sufficiently large cost advantage due to its relative proximity, the other firms do not enter the market, thus allowing the close firm to obtain monopoly rents by curtailing supply.

The second stage regression results are presented in Table 14. As can be seen, for all specifications, the coefficient on certification is negative, significantly different from zero, although around eighteen to twenty times higher as the OLS estimate. This larger result may be reflective of measurement error—while certification is well measured we suspect given the nature of the data on firm locations we suspect that the denominator of the certification rate is quite noisy. An alternative explanation is that this result arises from a heterogeneous treatment effect: the model presented earlier in this paper suggests that firms with an intermediate unobserved (to the regulator) cost of compliance will be both more sensitive to variation in auditor cost and more likely to alter emissions after the introduction of certification than are those firms with either low or high costs of compliance. It is worth noting that the sign and significance level of the coefficient measuring the indirect effects of certification on air quality do not differ substantially from those in the OLS regression presented earlier.

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 firm 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 lead to had both direct and indirect effects 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 firms 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 firms with a relatively low cost of compliance thus reducing the overall cost of achieving a given level of compliance. While our results also suggest that many of the firms who choose to certify would already be in compliance in the absence of the certification program, the results on air quality suggest that at least some of the certifying firms were not in compliance prior to the introduction of the program.

Second, in examining the effects of particular programs it is important to keep in mind what other programs are in place and how the implementation of these other programs is likely to be affected by the introduction of this program. In the absence of a systematic scheme to monitor and fine uncompliant 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 programs. The presence of these indirect effects also has implications for the establishment an appropriate control group for evaluating emissions among those firms that choose to certify. In this particular case, for example, the behavior of uncertified firms 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.

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. 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, 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. July 2007. RFP DP 07-36.

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

_______,Gllobal monitoring of air pollution over land from EOS-Terra MODIS. Journal of Geophysical Research 2003;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. Satellite remote sensing of particulate matter and air quality assessment over global cities. Atmospheric Environment 2006;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. An empirical relationship between PM2.5 and aerosol optical depth in Delhi Metropolitan. Atmospheric Environment 2007;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.

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.

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 (quantiles)

Quantile 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

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

Table 4

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 firms certified from the SIEM data, using only those certified firms 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.

Table 5

Determinants of non Compliance 1992-1995 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 6

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 7

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 8

Dependent Variable: Change in the Log of the % of Firms Inspected between 1993 and 1995 and 2003-2005

	OLS	OLS	OLS	IV
Log (% Inspected 1993-1995)	-0.23722			0.33094
	[0.05579]***			[0.12985]**
Log (1+ % Certified)		0.30205		
		[0.09576]***		
Log (% Inspected USA)			0.20781	
			[0.07844]***	
Constant	-1.07129	-0.83926	-0.19585	0.2393
	[0.16201]***	[0.14192]***	[0.16028]	[0.32492]
Observations	160	160	160	160
R-squared	0.1	0.06	0.04	

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

Table 9

First Stage

Dependent Variable: Log of the % of Firms Inspected between 1993 and 1995

Log (1+ % Certified)	0.82516		0.47094
	[0.11609]***		[0.14783]***
Log (% Inspected USA)		0.68976	0.43955
		[0.09336]***	[0.12003]***
Constant	-3.16774	-1.21722	-2.10384
	[0.17205]***	[0.19076]***	[0.33444]***
Observations	160	160	160
R-squared	0.24	0.26	0.3

Standard errors in brackets

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

Table 10

Distribution of firms in SIEM by size and certificartion status

Firm's size	Total	% certified (matched)
Less than 10 employees	21949	0.10
11 to 100 employees	7675	0.94
More than 100 employees	2708	11.04
Total	32332	1.22

Table 11

OLS Regression Results
Fraction Certified defined from SIEM data
Dependent variable: Difference in AOD between 2000 and 2006

Fraction Certified (SIEM)	-0.092	-0.073	-0.073	-0.078
	[0.032]***	[0.033]**	[0.034]**	[0.035]**
Medium*Log Cert. Intensity wt		-0.083	-0.085	-0.087
		[0.034]**	[0.034]**	[0.035]**
Log Size* Log Cert. Intensity wt		0.009	0.01	0.005
		[0.008]	[0.008]	[0.008]
Log Cert. Intensity wt	0.008	0.028	0.028	0.025
	[0.008]	[0.013]**	[0.013]**	[0.013]*
Log Size wt	0.013	0.005	0.012	0.016
	[0.003]***	[0.008]	[0.008]	[0.008]*
% Medium		0.059	0.053	0.051
		[0.034]*	[0.034]	[0.033]
% Exporting			-0.062	-0.058
			[0.024]***	[0.031]*
% Importing			-0.013	-0.048
			[0.023]	[0.029]*
Exporting*Log Cert. Intensity				-0.003
				[0.024]
Importing. * Log Cert. Intensity				0.053
				[0.022]**
AOD 2000	-0.296	-0.297	-0.298	-0.299
	[0.011]***	[0.011]***	[0.011]***	[0.011]***
Constant	-0.654	-0.659	-0.662	-0.663
	[0.027]***	[0.028]***	[0.028]***	[0.028]***
Weather Controls	Yes	Yes	Yes	Yes
Month fixed Effects	Yes	Yes	Yes	Yes
Observations	19849	19849	19849	19849
R-squared	0.18	0.18	0.18	0.18

Robust standard errors clustered at the zip code level in brackets
= * significant at 10%; ** significant at 5%; *** significant at 1%

Table 12

OLS Regression Results Fraction Certified defined from Industrial Census Dependent variable: Difference in AOD between 2000 and 2006				
Fraction Certified (Census)	-0.594	-0.575	-0.504	-0.502
	[0.298]**	[0.302]*	[0.308]	[0.308]
Medium*Log Cert. Intensity wt		-0.083	-0.085	-0.086
		[0.034]**	[0.034]**	[0.034]**
Log Size* Log Cert. Intensity wt		0.008	0.009	0.003
		[0.008]	[0.008]	[0.008]
Log Cert. Intensity wt	0.005	0.03	0.03	0.027
	[0.008]	[0.013]**	[0.013]**	[0.013]**
Log Size wt	0.012	0.004	0.011	0.014
	[0.003]***	[0.008]	[0.008]	[0.008]*
% Medium		0.062	0.057	0.055
		[0.034]*	[0.033]*	[0.033]*
% Exporting			-0.06	-0.059
			[0.024]**	[0.031]*
% Importing			-0.013	-0.046
			[0.023]	[0.029]
Exporting*Log Cert. Intensity				0.002
				[0.025]
Importing. * Log Cert. Intensity				0.048
				[0.023]**
AOD 2000	-0.295	-0.297	-0.298	-0.298
	[0.011]***	[0.011]***	[0.011]***	[0.011]***
Constant	-0.645	-0.653	-0.656	-0.657
	[0.027]***	[0.028]***	[0.028]***	[0.028]***
Weather Controls	Yes	Yes	Yes	Yes
Month fixed Effects	Yes	Yes	Yes	Yes
Observations	19849	19849	19849	19849
R-squared	0.18	0.18	0.18	0.18

Robust standard errors clustered at the zip code level in brackets

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

Table 13

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	-0.002	-0.002
	[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	0.005	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	7.37	7.19	7.26	7.35

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 14

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)	-10.723	-10.249	-9.792	-9.941
	[3.428]***	[3.411]***	[3.319]***	[3.323]***
Medium*Log Cert. Intensity wt		-0.086	-0.087	-0.081
		[0.036]**	[0.036]**	[0.037]**
Log Size* Log Cert. Intensity wt		0.01	0.01	0.003
		[0.008]	[0.008]	[0.008]
Log Cert. Intensity wt	0.022	0.039	0.039	0.039
	[0.010]**	[0.015]**	[0.015]**	[0.016]**
Log Size wt	0.02	0.014	0.018	0.022
	[0.005]***	[0.010]	[0.010]*	[0.010]**
% Medium		0.045	0.043	0.038
		[0.038]	[0.037]	[0.037]
% Exporting			-0.032	-0.052
			[0.028]	[0.034]
% Importing			-0.01	-0.029
			[0.025]	[0.032]
Exporting*Log Cert. Intensity				0.031
				[0.029]
Importing. * Log Cert. Intensity				0.031
				[0.026]
AOD 2000	-0.303	-0.303	-0.304	-0.305
	[0.012]***	[0.012]***	[0.012]***	[0.012]***
Constant	-0.591	-0.6	-0.604	-0.605
	[0.034]***	[0.035]***	[0.035]***	[0.035]***
	19849	19849	19849	19849
Weather Controls	Yes	Yes	Yes	Yes
Month fixed Effects	Yes	Yes	Yes	Yes
Observations	19849	19849	19849	19849

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. $\alpha < 1, \quad \beta > 1.$

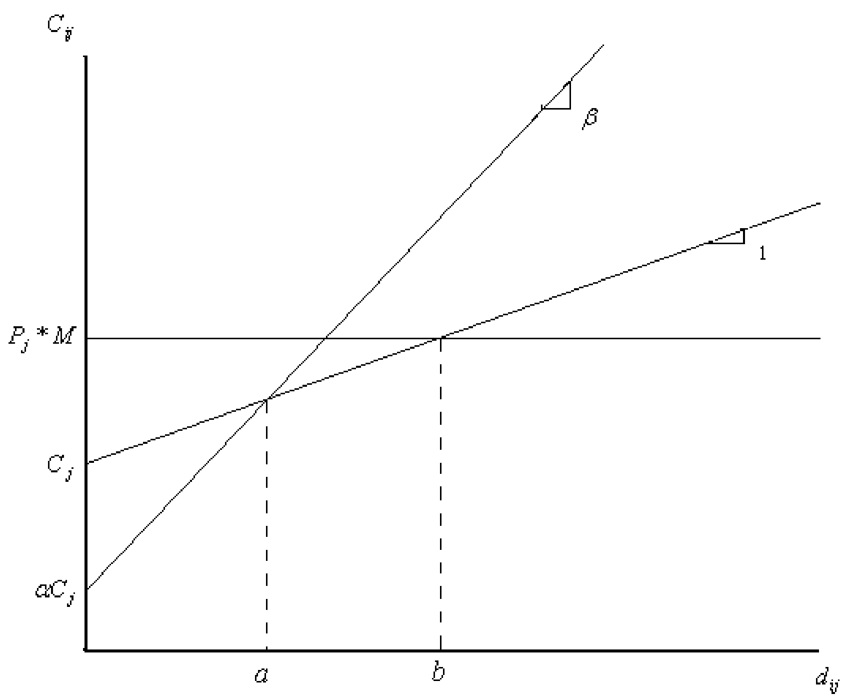


Figure 2

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

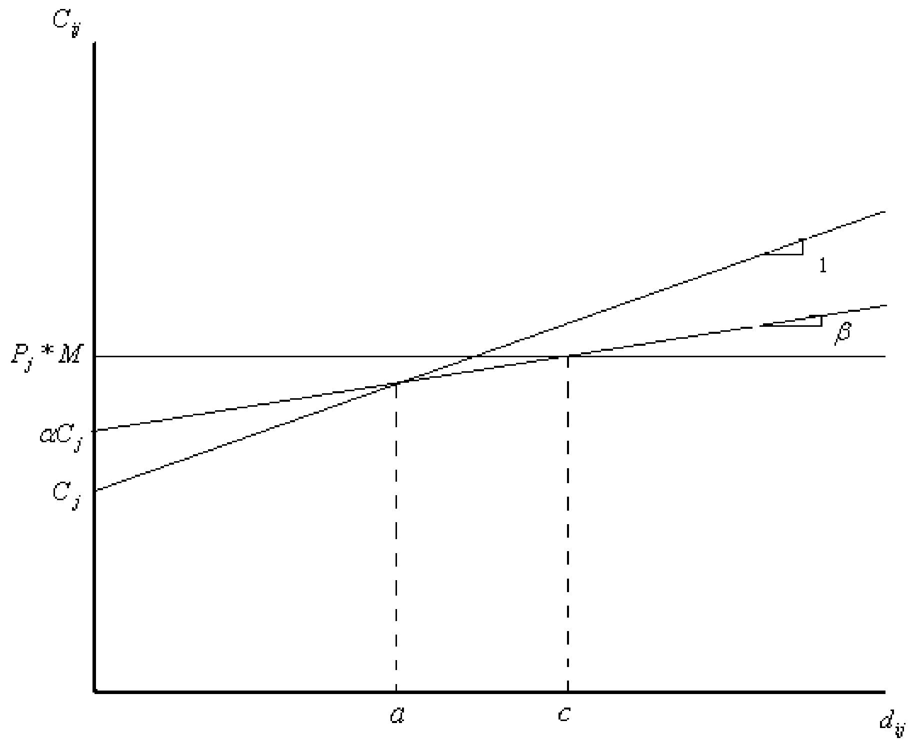


Figure 3

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

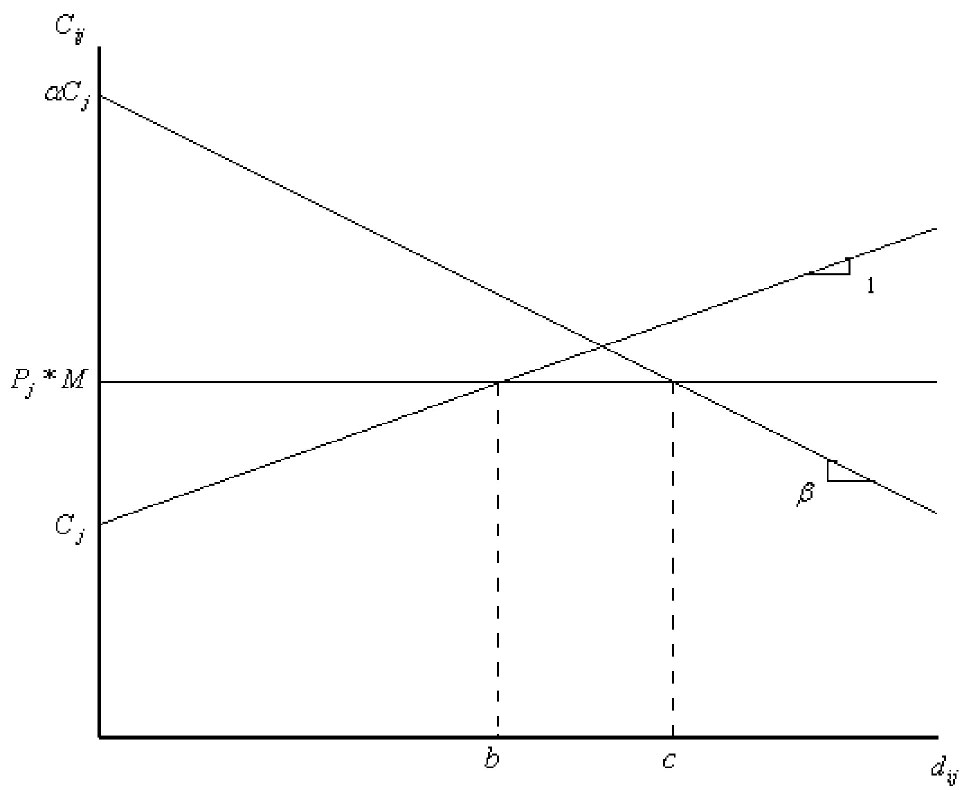


Figure 4

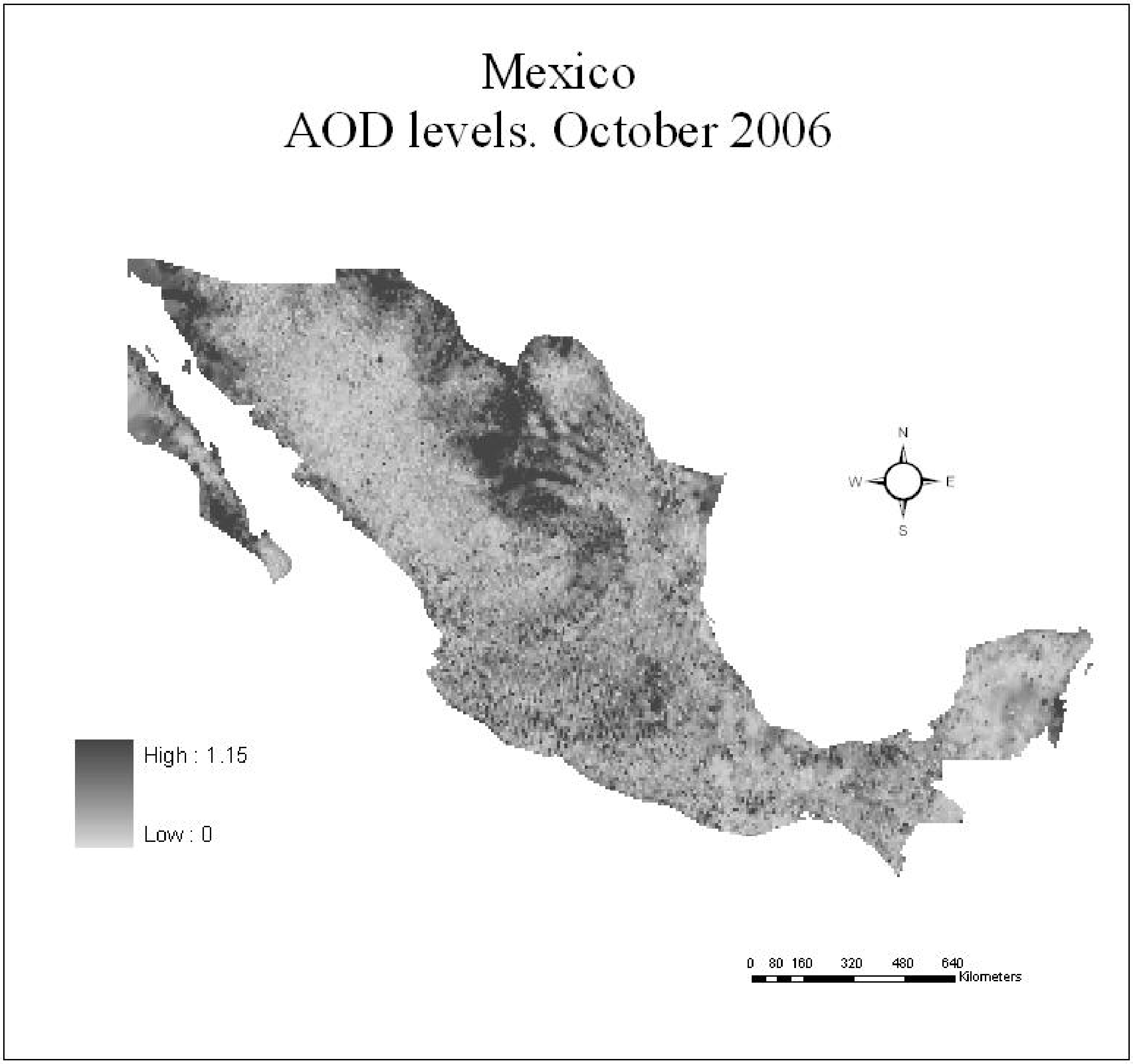

Mexico
AOD levels. October 2006
N
W
E
S
High : 1.15
Low : 0
0 80 160 320 480 640
Kilometers

Figure 5

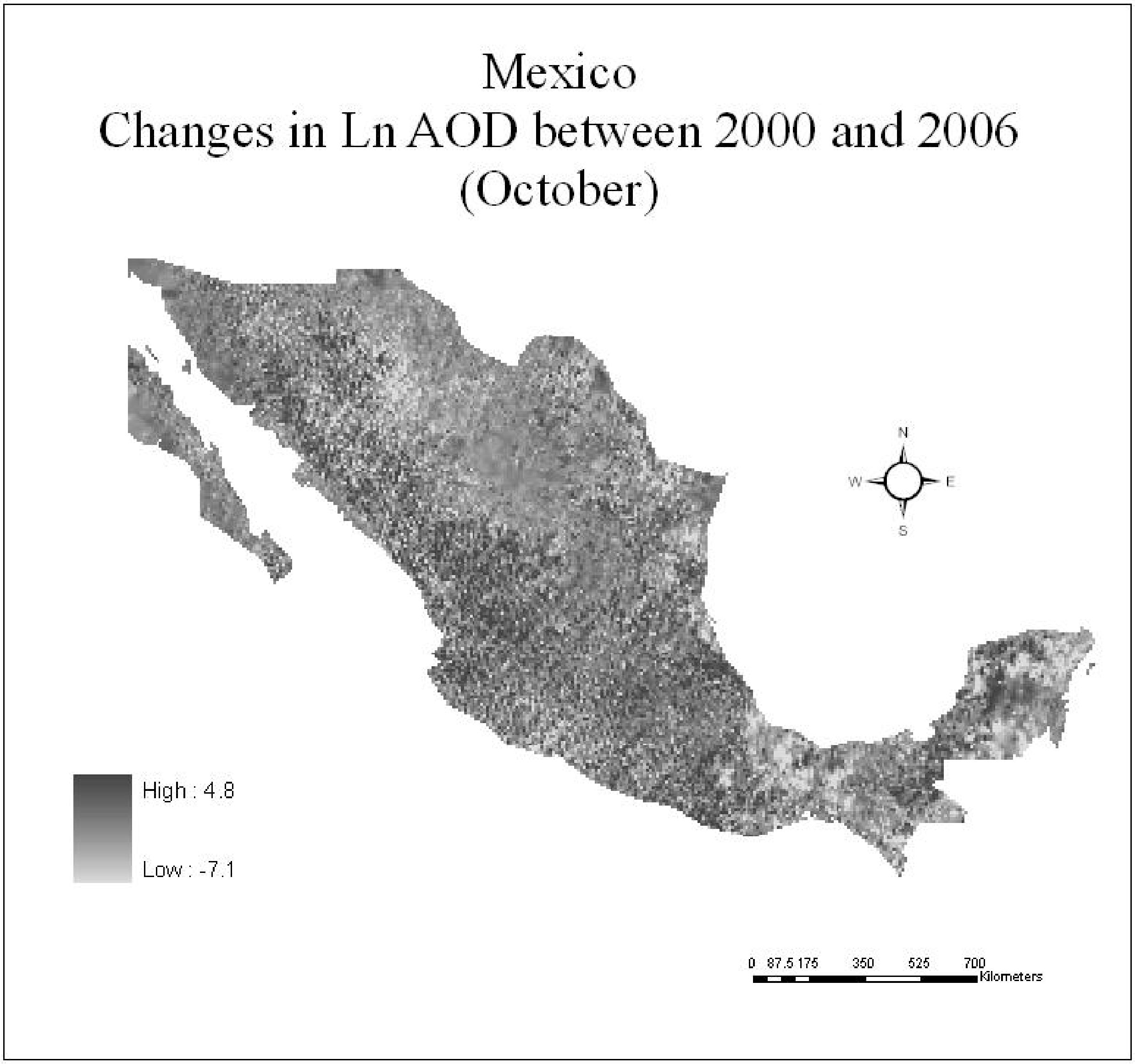
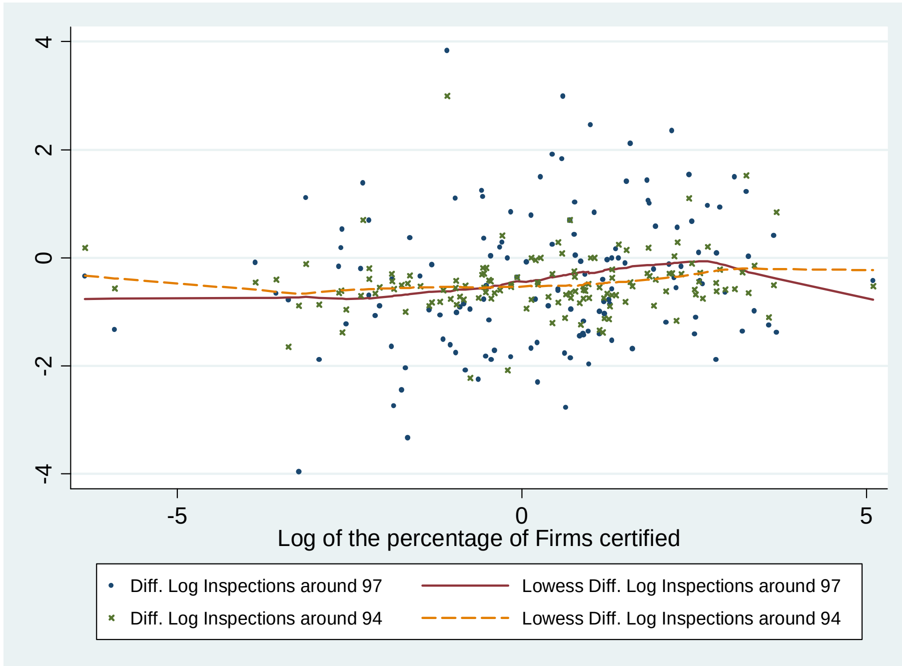

Mexico
Changes in Ln AOD between 2000 and 2006
(October)
N
W
E
S
High : 4.8
Low : -7.1
0 87.5 175 350 525 700
Kilometers

Figure 6



4
2
0
-2
-4
-5
0
5
Log of the percentage of Firms certified
Diff. Log Inspections around 97
Lowess Diff. Log Inspections around 97
Diff. Log Inspections around 94
Lowess Diff. Log Inspections around 94

Figure 7

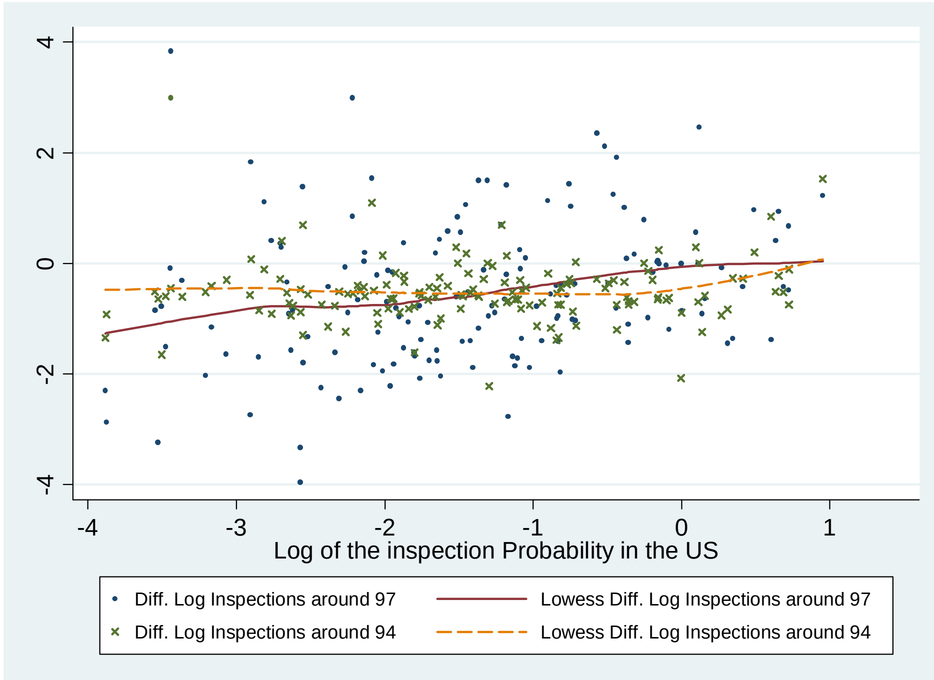


Figure 8

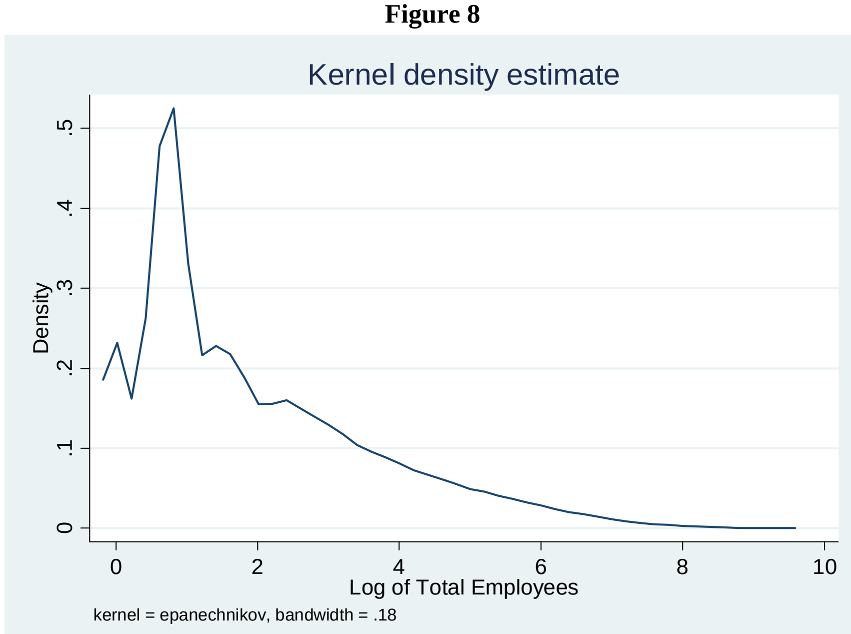
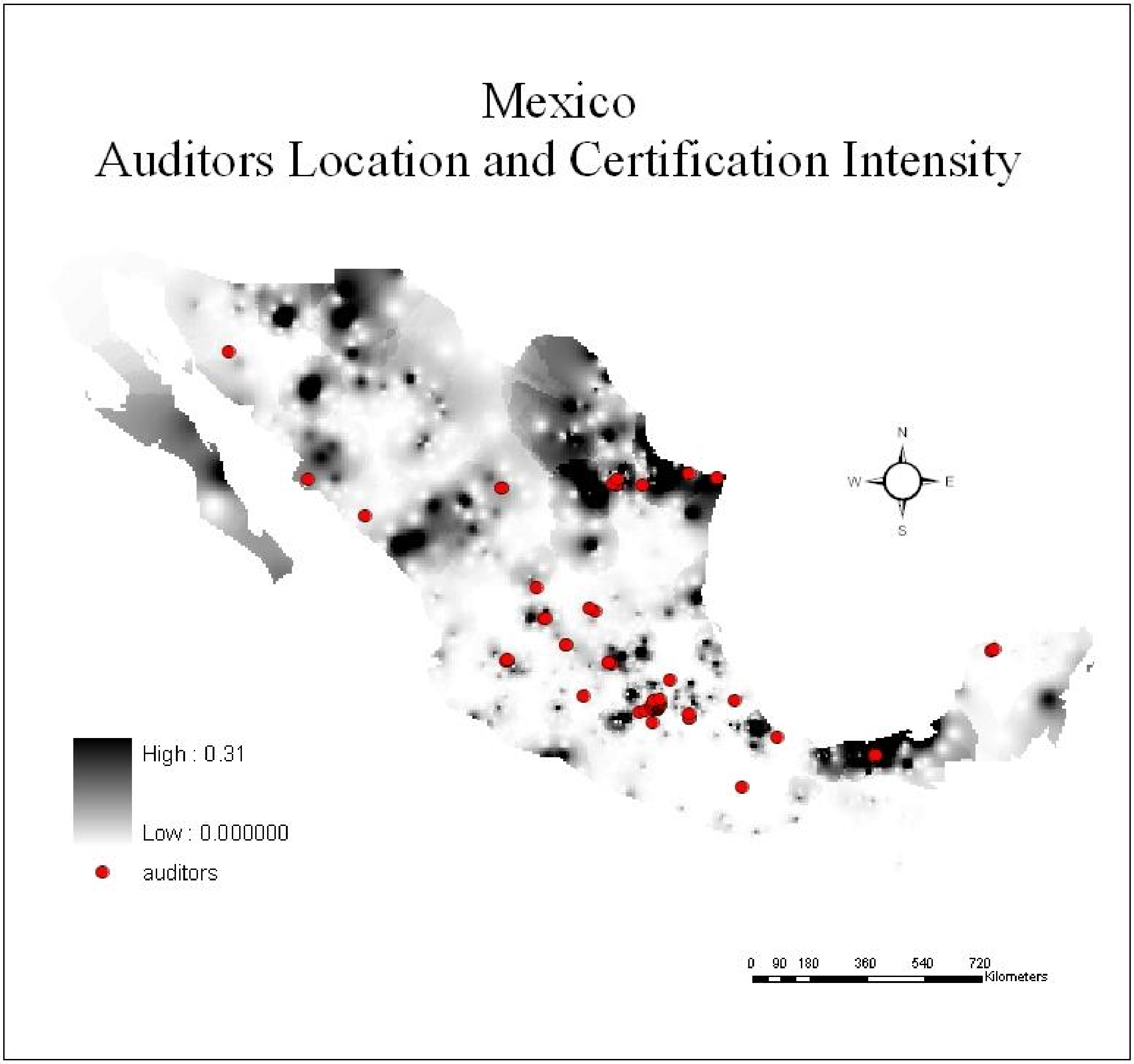


Figure 9


Mexico
Auditors Location and Certification Intensity
N
W
E
S
High : 0.31
Low : 0.000000
auditors
0 90 180 360 540 720
Kilometers