

Claremont Colleges Scholarship @ Claremont

CMC Faculty Publications and Research

CMC Faculty Scholarship

1-1-2015

Standardization and the Impacts of Voluntary Program Participation: Evidence from Environmental Auditing

Mary F. Evans

Claremont McKenna College

Lirong Liu

Sam Houston State University

Sarah L. Stafford

College of William and Mary

Recommended Citation

Evans, Mary F., Lirong Liu, and Sarah Stafford, Standardization and the Impacts of Voluntary Program Participation: Evidence from Environmental Auditing, *International Review of Law and Economics* 43: 10-21, 2015.

This Article - preprint is brought to you for free and open access by the CMC Faculty Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in CMC Faculty Publications and Research by an authorized administrator of Scholarship @ Claremont. For more information, please contact scholarship@cuc.claremont.edu.

Standardization and the Impacts of Voluntary Program Participation: Evidence from
Environmental Auditing

March 30, 2015

Mary F. Evans^a, Lirong Liu^b, and Sarah L. Stafford^c

Abstract:

We explore how limits to our insight about the underlying decision-making structure of firms may affect the conclusions we draw about the likely impacts of participation in voluntary environmental programs. We develop a theoretical model to examine the conditions under which a multi-facility firm chooses to employ a standardized adoption policy for a voluntary program. We test this model empirically using a firm-level dataset on the adoption of a voluntary environmental auditing program and find that, consistent with the theoretical model, a standardized auditing outcome is less likely among firms with more heterogeneous portfolios of facilities. We also examine the effect of environmental auditing on facility compliance using both firm-level and facility-level controls. We find that the estimated effect depends on whether or not measures that proxy for firm incentives for standardization are included in the analysis. These findings suggest caution in drawing conclusions on the effectiveness of voluntary programs in improving compliance based on analyses that assume standardization on the part of the firm.

Keywords: voluntary programs, environmental auditing, compliance, standardization, multi-facility firm

JEL codes: Q53, Q58, L29

^a Corresponding author, The Robert Day School of Economics and Finance, Claremont McKenna College, 500 E. Ninth Street, Claremont, CA 91711, mevans@cmc.edu, tel. 909.607.3423, fax 909.621.8249

^b Department of Economics and International Business, Sam Houston State University, Huntsville, TX 77341, lx1004@shsu.edu

^c Department of Economics, College of William and Mary, P.O. Box 8795, Williamsburg, VA 23187
slstaf@wm.edu

Standardization and the Impacts of Voluntary Program Participation: Evidence from Environmental Auditing

1. Introduction

In recent years, participation in voluntary environmental programs and other forms of industry self-regulation has received increased emphasis from both regulators and scholars as a potential means of improving compliance with environmental regulations.¹ As noted by Khanna and Broule [3], the number voluntary programs sponsored by the U.S. Environmental Protection Agency (EPA) grew dramatically from 28 programs in 1996 to 54 in 1999 and to 87 in 2005. Although the growth of these programs has slowed in the Obama administration, EPA continues to support a large number of voluntary initiatives, with more recent programs focused on energy and climate change.² A number of studies have examined the incentives for participation in voluntary programs and/or the effectiveness of participation on environmental performance.³ The empirical evidence on program effectiveness is mixed with some studies showing modest success and others showing no discernable impacts, even among analyses that focus on the same program. One possible explanation for these mixed results is that many voluntary programs focus on encouraging firms to participate whereas compliance with formal environmental regulations is usually assessed at the facility level.

¹ This appears to have been motivated, at least in part, by significant reductions in enforcement resources at the Environmental Protection Agency (EPA), which decreased by 5 percent in real terms between 1997 and 2006 [1]. Kotchen [2] cites the political feasibility of agency-sponsored voluntary programs as one factor explaining their proliferation.

² See <http://www.epa.gov/climatechange/EPAactivities/voluntaryprograms.html> for information on these programs.

³ See Khanna and Brouhle [3] and Prakash and Potoski [4] for reviews.

Non-mandatory approaches to environmental protection encompass a diverse set of programs but, relative to formal regulation, most voluntary programs allow participating entities a greater degree of flexibility in implementation. For many voluntary programs, a participating firm can implement the program at some of its facilities but not at others or can implement a different set of program-related actions at each of its facilities. However, this decision making process is generally unobserved. In most instances, only the firm-level decision about whether or not to participate is known. Thus empirical efforts to examine the effects of voluntary program participation on environmental performance often require the researcher to make an assumption about the nature of program participation and, ultimately, the process of decision-making within multi-facility firms. Constraints imposed by available data usually drive this assumption as many empirical studies only have data on participation at the firm level and assume, explicitly or implicitly, a standardized application of the voluntary program across all facilities owned by a firm. For the most part these empirical analyses have not accounted for the possibility of heterogeneous application of the voluntary program across the firm's various facilities.

The goal of this paper is to better understand how firms make decisions about facility-level participation in voluntary programs and how this process impacts our ability to make valid inferences about the effects of participation on environmental performance. We develop a theoretical model that identifies the conditions under which a multi-facility firm chooses to employ a standardized adoption policy for a voluntary program rather than a heterogeneous policy in which some of the firm's facilities adopt the program while others do not. The model suggests that the incentives for a firm to employ a

standardized adoption policy weaken as the degree of heterogeneity among facilities owned by a firm increases. This intuitive result holds whether the firm uses a centralized or decentralized decision-making structure and calls into question the frequently invoked assumption that the decision to adopt a voluntary program is uniformly applied to all of a firm's facilities. We then test this model empirically to see whether our theoretical predictions of the factors that encourage firms to standardize adoption decisions hold. In particular, we use a firm-level dataset on the adoption of a voluntary environmental program, an environmental audit, across a firm's facilities. We also examine the effect of adoption on facility compliance using both firm-level and facility-level controls. Since the incentives for adoption of an environmental auditing program are similar to those for adoption of other voluntary environmental programs, we believe that the results of our study can provide insight for evaluating a wide range of voluntary environmental programs.

The remainder of the paper is organized as follows. In the next section, we discuss previous empirical studies that analyze the impacts of voluntary program participation with a focus on environmental auditing as well as another well-known voluntary program, the EPA's 33/50 program. Section 3 develops a model of program adoption within a multi-facility firm to explore the incentives for employing a standardized versus a heterogeneous adoption policy. In Section 4, we test the model's implications in a firm-level analysis of the decision to adopt a standardized auditing policy and find results broadly consistent with our theoretical predictions. In section 5, we further explore the implications of our model using a facility-level empirical analysis of the effect of audit adoption on compliance. We find that when we include both a facility-level adoption

measure and firm-level controls in our empirical analysis, auditing does not have a significant effect on compliance. However, when the firm-level controls are excluded, we find a positive and significant effect of environmental auditing on facility compliance. Section 6 discusses the implications of our results. In particular, we suggest that researchers use caution in assuming that firm-level participation in a program applies equally to all facilities owned by that firm, as such an assumption may induce systematic measurement error that could affect the results of their analyses and alter policy-relevant conclusions.

2. The environmental performance effects of voluntary program participation:
evidence from environmental auditing and the 33/50 program

Voluntary environmental initiatives can range from a government-sponsored voluntary program (e.g., Energy Star) to programs sponsored by trade associations (e.g., the American Chemistry Council's Responsible Care Program) to unilateral initiatives undertaken by the regulated entity without direct government involvement. Regardless of type, most voluntary programs provide explicit or implicit incentives for participation and involve actions that go beyond what is required by various environmental statutes. As discussed in Lyon and Maxwell [5], incentives for participation mainly fall into three categories: participation may make the regulated entity more attractive to consumers and investors, participation may shape current and future regulatory decisions by reducing enforcement effort or preempting future regulations, and participation may result in increased productivity or efficiency at the regulated entity.

Environmental auditing is one example of a voluntary environmental program. An environmental audit is “...a systematic, documented, periodic and objective review by regulated entities of facility operations and practices related to meeting environmental requirements” and EPA has long suggested environmental auditing programs as a way for regulated firms to improve their environmental compliance as well as environmental performance more generally [6].⁴ EPA’s Audit Policy also provides explicit incentives for firms to voluntarily adopt environmental auditing, as firms that choose to report and correct violations discovered during the course of a voluntary environmental audit are eligible for significant penalty reductions.⁵ Many states also have their own audit policies that actively encourage environmental auditing by granting statutory privilege for environmental audit reports and/or immunity from penalties for violations discovered (and corrected and disclosed) during the course of an audit.⁶

EPA has embraced environmental auditing because, at least in theory, such audits provide a mechanism to improve a regulated entity’s compliance with environmental regulations.⁷ Survey and anecdotal evidence suggests positive impacts of auditing on environmental performance [16, 17]. Additionally, three previous empirical studies

⁴ 50 FR 46504, Section II.A. The academic literature on environmental regulation and compliance (as well as that on compliance more generally) sometimes uses the term audit more generally to mean any periodic inspection or review of an entity’s compliance status. To distinguish reviews initiated by regulators from those initiated by regulated entities and to be consistent with EPA’s terminology, we use the term “audit” to refer to reviews voluntarily undertaken by the regulated entity and “inspection” to refer to reviews initiated by regulators (and thus involuntary from the perspective of the regulated party).

⁵ Pfaff and Sanchirico [7], Short and Toffel [8], and Stafford [9] examine various aspects of EPA’s Audit Policy.

⁶ Forty one states currently have some sort of environmental audit program. See Stafford [10, 11] and Khanna and Widyawati [12] for further discussion of these state policies.

⁷ Most theoretical models of environmental auditing focus on the informational aspect of auditing (see, for example, Mishra et al. [13], Pfaff and Sanchirico [14], Friesen [15]). The theoretical models generally assume that environmental performance or compliance with environmental regulations includes a stochastic element. As a result, regulated entities cannot fully observe their levels of emissions or compliance status *ex ante*. By conducting an environmental audit, a regulated entity learns its true level of environmental performance and, where such performance is below the entity’s optimal level, can undertake corrective actions.

explore the impact of environmental auditing on compliance. Khanna and Widyawati [12] find higher contemporaneous compliance with Clean Air Act (CAA) regulations among facilities whose S&P 500 corporate parent indicates the presence of an environmental auditing program. Evans, Liu and Stafford [18] report no significant influence of auditing on long-term facility compliance with the Resource Conservation and Recovery Act (RCRA) among a sample of hazardous waste generators in Michigan. Earnhart and Harrington [19] find that more frequent auditing of a facility improves compliance with Clean Water Act (CWA) effluent limits for total suspended solids (TSS) but not biological oxygen demand (BOD).⁸

Given their similar focus, the results of Khanna and Widyawati and Evans, Liu, and Stafford at first glance appear contradictory. Khanna and Widyawati find that environmental auditing improves compliance while Evans, Liu and Stafford find the opposite. However, upon closer consideration, the results of these two previous studies are not necessarily at odds because there are a number of differences between the analyses could individually or collectively drive the disparate findings. The studies focus on different environmental media (i.e., air versus hazardous waste) and different time frames (i.e., contemporaneous versus long-term). The samples are distinct with Khanna and Widyawati's sample consisting of a set of facilities whose S&P 500 parent companies responded to the Investor Research Responsibility Center (IRRC) survey on environmental management practices survey and Evans et al.'s sample including large and small hazardous waste generators in Michigan. Lastly, and for our purposes

⁸ An important feature of their study is its focus on a group of facilities that regularly conduct self-audits (i.e., chemical manufacturing facilities regulated under the National Pollution Discharge Elimination System (NPDES)). Thus, the results of their study cannot speak to whether the adoption of an audit program improves compliance compared to non-adoption.

potentially most important, the analyses measure environmental auditing at different levels of the firm's decision-making structure. Khanna and Widyawati observe a firm-level auditing measure, the presence or absence of an auditing program as indicated by the parent company, and assume uniform application of the auditing program to all facilities owned by the firm.⁹ That is, Khanna and Widyawati assume the firm adopts a standardized auditing policy (i.e., the firm either audits no facilities or audits all facilities in its portfolio). In contrast, Evans et al. observe auditing outcomes at the facility-level and are therefore not obliged to assume standardization.

Similar issues about the use of firm-level vs. facility-level participation measures have arisen in the context of another voluntary environmental program, the government-sponsored 33/50 program. The 33/50 program was developed by the EPA in the early nineties and set a goal of reducing aggregate emissions of 17 chemicals by 33% in 1992 and by 50% in 1995, relative to the participant's 1988 baseline. While regulated firms were invited to join the 33/50 program, not all facilities owned by such firms were active participants in the program.¹⁰ The 33/50 program is perhaps the most widely studied of all voluntary environmental programs, although to date there remains little consensus on the effect of participation in the 33/50 program on toxic emissions. Of course, the empirical studies that examine this question differ along multiple dimensions that could explain the discrepancy in results. Like the studies of environmental auditing, two important dimensions are the level at which the participation outcome is measured and the level at which environmental performance is measured.

⁹ If the parent company does not indicate the presence of an auditing program, then Khanna and Widyawati assume no facilities owned by the parent company conduct an environmental audit.

¹⁰ See Bi and Khanna [20] and Vidovic and Khanna [21] for a more thorough discussion of facility vs. firm participation.

Three studies, Khanna and Damon [22], Vidovic and Khanna [23], and Innes and Sam [24], use a firm-level measure of program participation provided by EPA (whether or not at least one facility committed to the program) as well as a firm-level measure of environmental performance (calculated by aggregating some measure of emissions across all facilities owned by a participating firm). Khanna and Damon and Innes and Sam find a reduction in emissions as a result of program participation while Vidovic and Khanna do not. In contrast, Gamper-Rabindran [25] conducts a facility-level analysis of emissions using EPA's measure of firm-level participation and finds that participation decreases emissions only in a few select industries. The two most recent contributions to this literature, Bi and Khanna [20] and Vidovic and Khanna [21], use both a facility-level measure of 33/50 participation and a facility-level measure of performance. Interestingly, the facility-level participation data suggest that just over one-fifth of the facilities owned by 33/50 firms (i.e., parent company participants in 33/50) actually participated in the program and furthermore, most 33/50 firms had only a few participating facilities [20]. However, even with data at similar measurement levels, Bi and Khanna (2012) find a significant negative effect of facility participation on emissions of 33/50 chemicals while Vidovic and Khanna [21], looking at a more restrictive sample, find no significant effect of facility participation on facility emissions.

3 Conceptual Model

In this section we present a model of voluntary program adoption within a multi-facility firm. A multi-facility firm may adopt various decision-making structures. Under a centralized structure, a single location (e.g., headquarters) makes and implements the

program adoption decision for all facilities. Under a decentralized structure, each facility independently chooses whether to adopt the voluntary program. A structure in between these two extremes might have one facility make adoption decisions for itself and for other nearby facilities (e.g., the firm may house an environmental compliance department in one facility for each region in which the firm operates). Regardless of the adopted structure, the decision-making may result in a standardized outcome in which either all facilities adopt or no facilities adopt. In this case, we say that the firm has a standardized adoption policy. Alternatively, adoption outcomes may be characterized by adoption at some facilities and not at others; we refer to this case as a heterogeneous adoption policy.

Our model draws on the literature on the centralization-decentralization decision for multi-unit firms, itself a subset of the literature on organizational structure. Papers in this literature consider different incentives for a firm to centralize or decentralize its decision-making structure. For example, in Sah and Stiglitz [26] the optimal organizational structure depends on the relative cost of Type I and Type II errors about the benefits of various projects. In Van Zandt [27] the optimal structure depends on the delay associated with distributing information across units relative to the benefits from increased information. Chang and Harrington [28] focus on the level of positive spillovers and mutual learning across units as well as the interplay between incentives of local managers to maximize unit profit and headquarters managers to maximize firm profit. The model we develop is most like the model presented in Hunnicutt [29]. Hunnicutt examines the decision-making process of a multi-plant firm that must weigh the relative benefits of making production decisions at a single location and shipping bundles to various outlets relative to the benefit of having each outlet make its own

production decision. Under centralization, the firm achieves economies of scale but can make mistakes about the optimal amount of product for each outlet. Under decentralized production, mistakes are minimized but economies of scale are not achieved. Importantly, Hunnicutt also examines the conditions under which the firm finds it optimal to standardize the amount of product for each outlet.

Our primary interest lies in understanding when a firm is likely to adopt standardized policy with respect to adoption of a voluntary program rather than a heterogeneous policy. Our focus on the standardization decision is in part due to the fact that the underlying decision-making structure of the firm (e.g., centralized, decentralized) remains unobservable, while the adoption of a standardized or heterogeneous policy is observable (with appropriate data). Thus we develop a model of the decision to standardize for a multi-facility firm that owns N total facilities of $M \geq 1$ different types. Let γ_m denote the number of type- m facilities owned by the firm such that $\sum_{m=1}^M \gamma_m = N$. We refer to the firm as single-type if $M = 1$, or multiple-type if $M > 1$. Following Hunnicutt, we examine the decision to standardize under two organizational structures, centralization and decentralization. We assume that under centralization, the firm incurs lower adoption costs but observes the true benefit of adoption at each facility type with noise. Under decentralization, the firm observes the true benefit of adoption at each of its facilities but faces higher adoption costs.

Denote the benefit of adoption at a type- m facility as b_m . For $M > 1$, we assume $b_1 > b_2 > \dots > b_M \geq 0$ so that the benefits of adoption at a type-1 facility are highest. For a given M , the larger is the difference between adoption at a type-1 facility (i.e., the highest benefit) and a type- M facility (i.e., the lowest benefit), the more diverse are the M

facility types owned by the firm. We use this difference, $b_1 - b_M$, as a measure of the degree of heterogeneity among the facilities in a firm's portfolio.

Under decentralization, each facility observes its own benefit of adoption. The cost of adoption at a facility under decentralization is constant and equal to c . Under centralization, the single-type firm observes the true benefit of adoption at each of its facilities.¹¹ In contrast, the multiple-type centralized firm lacks full information about each facility's operations due to agency problems. As a result, the multi-facility firm fails to observe the true benefit of adoption at each facility type, which is either high (b_m^H) or low (b_m^L) with $b_m^H \geq b_m^L$ for each facility type m . The multiple-type centralized firm has prior probability λ_m that $b_m = b_m^H$ for facility type m and knows the number of facilities of each type. The firm observes a signal of the benefit of adoption at each facility type (i.e., the same signal for all facilities of a given type).¹² Let z_m denote the signal the firm receives for a type- m facility where z_m equals b_m^H or b_m^L according to the following probability distributions:

$$P(z_m = b_m^H | b_m = b_m^H) = p_m$$

$$P(z_m = b_m^L | b_m = b_m^H) = 1 - p_m$$

$$P(z_m = b_m^H | b_m = b_m^L) = q_m$$

$$P(z_m = b_m^L | b_m = b_m^L) = 1 - q_m.$$

¹¹ This assumption is consistent with a single facility making decisions for all facilities. This "headquarters facility" might, for example, house the single environmental compliance department for the firm. In this case, even in the presence of agency problems arising through interactions with the headquarters facility and other facilities, the headquarters facility still has full information about its own benefit of adoption. Since facilities are identical, the benefit of adoption at the headquarters facility is equal to the benefit of adoption at any other facility.

¹² For example, the centralized firm may rely on reports provided by facility managers based on information the managers compile from various divisions within their facilities. The organizational structure of the firm can create incentives for misreporting at various levels. For example, Evans et al. [30] show that compensation determined by a rank-order tournament may encourage such malfeasance on the part of managers.

where $p_m > q_m$.

The firm benefits from centralization through a lower unit cost of adoption. The cost of adoption at a facility under centralization is equal to αc with $0 < \alpha < 1$. The lower adoption cost under centralization may stem from the elimination of duplicated activities or from returns to specialization (e.g., using a single team of trained environmental compliance staff is less costly than providing such specialized training to staff onsite at each facility).

3.1 The single-type firm's decision

We first consider the adoption decision of a single-type firm (i.e., $M = 1$). Let b denote the benefit to the single-type firm of adoption at one of its facilities. The single-type firm, under centralization or decentralization, will either adopt the program at all facilities or adopt for no facilities. That is, it chooses a standardized adoption policy.

Under centralization, if the marginal benefit of adoption at a facility, b , exceeds the marginal cost of adoption, αc , then the single-type firm maximizes its net benefits by adopting the program at all N facilities. Otherwise, the centralized single-type firm does not adopt the program at any facility. Under centralization, the net benefits of program adoption at all N facilities for the single-type firm are given by

$$NB_c^1(M = 1) = bN - \alpha cN.$$

The subscript denotes the firm structure, c for centralized and d for decentralized, while the superscript indicates the single-type firm.

Under decentralization, each facility will adopt the program provided $b > c$. The net benefits to the decentralized single-type firm of adoption at all facilities is then

$$NB_a^1(M = 1) = bN - cN.$$

Notice that the total benefit of adoption at all facilities is the same under both structures. Given this and since $\alpha < 1$, the net benefits of adoption are higher under centralization for the single-type firm (i.e., $NB_c^1(M = 1) > NB_a^1(M = 1)$). Thus, a firm that owns identical facilities chooses a centralized structure and follows the straightforward adoption decision rule discussed above. Intuitively, the single-type firm enjoys lower costs of adoption under centralization without the associated risk of misestimating the benefits of adoption. For the single-type firm, the decision to adopt a program for all facilities versus no facilities is driven entirely by the benefits and costs of adoption under centralization.

3.2 The multiple-type firm's decision

Now consider a firm that owns $M > 1$ types of facilities (i.e., a multiple-type firm). Under decentralization, the multiple-type firm adopts the program at all facilities only if the benefit of adoption a type- M facility exceeds the cost, $b_M \geq c$. Recall that the type- M facility has the lowest benefit of adoption so the decentralized multiple-type firm finds adoption at all facility types optimal only if the net benefits of adoption at a type- M facility are positive. The decentralized firm does not adopt the program at any facility if the cost of adoption exceeds the benefit of adoption at a type-1 facility (i.e., the highest benefit facility), $c > b_1$. If $b_1 \geq c > b_M$, then the multiple-type firm prefers a heterogeneous adoption policy under decentralization. In other words, for standardization to be optimal for the multiple-type firm under decentralization, either $b_1 > b_2 > \dots > b_M \geq c$ or $c > b_1 > b_2 > \dots > b_M$ must hold. This implies that as the difference in

benefits to adoption at the highest and lowest benefit facility types grows, the range of adoption costs for which a multiple-type firm finds standardization optimal under decentralization gets smaller. Thus, under decentralization, multi-facility firms whose portfolios of facilities exhibit higher degrees of heterogeneity are ceteris paribus less likely to employ standardized adoption policies.

A more concrete two-firm example where the voluntary program in question is environmental auditing illustrates this result. Assume decentralized Firms A and B each own 100 total facilities of ten different types. Each firm owns 10 facilities of each type so that $\gamma_m = 10$ for $m = 1, \dots, 10$. For firm A, a type-1 facility has a benefit of auditing equal to 10 and a type-10 facility has an auditing benefit of 1. The benefits of auditing for facility types 2 through 9 take integer values between 9 and 2 for firm A. For firm B, a type-1 facility has a benefit of auditing equal to 6 and a type-10 facility has an auditing benefit of 5. The benefits of auditing for facility types 2 through 9 take equally distributed values between 6 and 5. Given these assumptions, the mean benefit of auditing a facility is 5.5 for both firms but the difference between auditing the highest benefit facility and the lowest benefit facility (i.e., $b_1 - b_M$) is only 1 for Firm B and is 9 for Firm A. That is, Firm B exhibits a lower degree of heterogeneity than Firm A. Firm B finds standardization optimal when the cost of auditing is less than 5 or greater than 6 whereas Firm A chooses to standardize only when the auditing cost is below 1 or above 10. Thus, the range of auditing costs for which Firm A, the firm with a higher degree of heterogeneity, finds standardization optimal is less than the range of costs for which Firm B chooses a standardized auditing policy.

Now consider the decision facing the multiple-type firm under centralization. For a multiple-type firm in which (i) $b_m^H \geq b_m^L \geq \alpha c$ for all facility types $m = 1, \dots, M$, or (ii) $\alpha c > b_m^H \geq b_m^L$ for all facility types, a standardized adoption policy is optimal. If (i) holds, then the firm adopts the program at all facilities. If (ii) holds, then the firm forgoes adoption at all facilities. The lack of full information has no impact on such a firm's adoption decisions.

The more interesting case is a firm that owns some facility types for which the benefits of adoption satisfy $b_m^H \geq \alpha c > b_m^L$ (i.e., adoption is optimal for the “high” m -type but not for the “low” m -type). In what follows, we restrict attention to this case. The centralized firm bases its adoption decision for facility types with $b_m^H \geq \alpha c > b_m^L$ on the expected benefits of adoption conditional on the signal the firm receives. The ranking of facility types from highest to lowest based on the *true* benefits of adoption need not coincide with the ranking of facility types from highest to lowest based on the *conditional expected* benefits of adoption. The conditional expected benefit of adoption at a type- m facility depends on the posterior probability distribution given by:

$$P(b_m = b_m^H | z_m = b_m^H) = \frac{p_m \lambda_m}{p_m \lambda_m + q_m (1 - \lambda_m)}$$

$$P(b_m = b_m^L | z_m = b_m^H) = \frac{q_m (1 - \lambda_m)}{p_m \lambda_m + q_m (1 - \lambda_m)}$$

$$P(b_m = b_m^H | z_m = b_m^L) = \frac{(1 - p_m) \lambda_m}{(1 - q_m)(1 - \lambda_m) + (1 - p_m) \lambda_m}$$

$$P(b_m = b_m^L | z_m = b_m^L) = \frac{(1 - q_m)(1 - \lambda_m)}{(1 - q_m)(1 - \lambda_m) + (1 - p_m) \lambda_m}.$$

Conditional on observing the signal $z_m = b_m^H$, the expected benefit of adoption at a type- m facility is given in expression (1):

$$E(b_m | z_m = b_m^H) = b_m^H \frac{p_m \lambda_m}{p_m \lambda_m + q_m (1 - \lambda_m)} + b_m^L \frac{q_m (1 - \lambda_m)}{p_m \lambda_m + q_m (1 - \lambda_m)}. \quad (1)$$

If instead the centralized firm observes $z_m = b_m^L$, then the conditional expected benefit of adoption at a type- m facility equals:

$$E(b_m | z_m = b_m^L) = b_m^H \frac{(1 - p_m) \lambda_m}{(1 - q_m)(1 - \lambda_m) + (1 - p_m) \lambda_m} + b_m^L \frac{(1 - q_m)(1 - \lambda_m)}{(1 - q_m)(1 - \lambda_m) + (1 - p_m) \lambda_m}. \quad (2)$$

The centralized multiple-type firm adopts the program at all type- m facilities if

$E(b_m | z_m) \geq \alpha c$ where $z_m = b_m^L, b_m^H$. Otherwise, the firm forgoes adoption at all type- m facilities.

The intuition for establishing the conditions under which the centralized firm chooses a standardized adoption policy are similar to that for the decentralized firm with one important difference. The decentralized firm bases its adoption decision on the known benefits of adoption whereas the centralized firm uses the conditional expected benefits of adoption. Since the ranking of facility types based on the *true* benefits of adoption may differ from the ranking based on the *conditional expected* benefits of adoption, the set of facilities adopting the program under centralization may differ from those that adopt under decentralization.

The centralized firm adopts the program at all facilities if it owns no facility types such that $\alpha c > b_m^H \geq b_m^L$ and, among those facilities with $b_m^H \geq \alpha c > b_m^L$, the conditional expected benefit of adoption at the facility type with the lowest conditional expected benefit exceeds the adoption cost. The centralized firm adopts the program at no facilities if it owns no facility types such that $b_m^H \geq b_m^L \geq \alpha c$ and, among those facilities with $b_m^H \geq \alpha c > b_m^L$, if the conditional expected benefit of adoption at the facility type with

the highest conditional expected benefit is less than the cost of adoption. Finally, if neither of these holds, then the centralized firm opts for a heterogeneous adoption policy.¹³ This implies that as the difference in conditional expected benefits of adoption at the facility type with the highest conditional expected benefit and at the facility type with the lowest conditional expected benefit grows, the range of adoption costs for which a multiple-type centralized firm finds standardization optimal gets smaller. Since the conditional expected benefit for facility type m is increasing in b_m^L and b_m^H , the centralized multi-facility firm whose portfolio of facilities exhibits a higher degree of heterogeneity is *ceteris paribus* less likely to adopt a standardized adoption policy.

Overall, the model generates two primary hypotheses with respect to firm characteristics and the likelihood that a firm will employ a standardized adoption policy. First, the model implies that multi-facility firms are less likely to employ a standardized adoption policy if their facilities exhibit substantial heterogeneity. Second, a standardized adoption policy is more likely among multi-facility firms for which the costs of adoption are either prohibitively high or negligible, relative to the benefits of adoption. These intuitive results hold whether the firm has adopted a centralized or decentralized decision making structure.

4. Firm-level analysis of standardization in auditing outcomes

Testing the hypotheses about standardization that arise from our conceptual model requires facility-level data on voluntary program adoption. For our analysis, we consider adoption of an environmental audit program. In general, national data on environmental

¹³ As mentioned above, standardization is also optimal when (i) $b_m^H \geq b_m^L \geq ac$ for all facility types $m = 1, \dots, M$, or (ii) $ac > b_m^H \geq b_m^L$ for all facility types.

auditing are difficult to obtain due in part to the fact that EPA does not require regulated entities to indicate the presence of audit programs. Fortunately, Michigan, like many other states, has its own environmental auditing program. Under the provisions of Michigan's audit policy, which began in 1997, audit documents are privileged and voluntary disclosures are eligible for immunity from penalties.¹⁴ One rather unique feature of Michigan's audit policy is that regulated entities must provide advance notice of an audit to receive penalty immunity.¹⁵ Specifically, the facility must file an "intent-to-audit" notice that identifies the facility at which the audit will be conducted, indicates the time frame for the audit, and states the general scope of the audit.¹⁶

We obtained a list of the intent-to-audit notices filed with the DEQ between 1998 and 2003. The data include the company and facility name, a mailing address, and the date the notice was filed. We used this information to match each facility to EPA's Facility Registry System (FRS) to identify the federal facility identification number and to determine the media programs under which the facility is regulated. Our empirical analysis focuses on facilities regulated under the Clean Air Act (CAA) in Michigan and the firms that own these facilities. We obtained data on facility characteristics from EPA's Aerometric Information Retrieval System (AIRS) Facility Subsystem (or AFS) database.¹⁷ Using DUNS numbers (when available) and owner names from the FRS,

¹⁴ In addition to Michigan, 17 other states provide audit privilege and immunity.

¹⁵ To our knowledge the only other state that requires prior notification is Texas.

¹⁶ While it is possible that a facility might conduct an environmental audit without first notifying the DEQ, Michigan provides strong incentives for facilities to file intent-to-audit notices; a primary benefit of auditing is the potential for penalty mitigation and this benefit is available only to auditing facilities that submit the required intent-to-audit notice. Additionally, since Michigan grants legal privilege to all environmental auditing documents, entities should not be reluctant to disclose environmental audits. See Evans et al. [18] for further discussion of these data.

¹⁷ The AFS database was downloaded from EPA's Envirofacts System in August 2007. AFS also lists 3 "portable sources" located in Michigan, but we excluded them from this analysis since they do not have a fixed facility location.

along with company name from the AFS, we matched CAA-regulated facilities in Michigan that were owned by the same firm.¹⁸ In order to empirically explore the role of heterogeneity, we restrict attention to multi-facility firms, of which we have 171 represented in our data. These 171 firms own a total of 730 facilities.

Our first empirical analysis explores the factors that lead a firm to choose a standardized adoption policy. Thus the dependent variable in the analysis is *Standardization*, which equals one if either all facilities owned by a firm audit or if none of the facilities owned by the firm audit. Otherwise, *Standardization* equals zero. We observe standardization for 155 firms (about 90%), which accounts for 590 facilities (81%). For the remaining 10 percent of firms, we observe heterogeneous auditing among their 140 facilities. Note that since our data contain facilities located in a single state, the degree of standardization we observe is likely to be higher than if our dataset included facilities located in several states. Since policies on audit privilege and audit immunity vary by state, firms that operate in multiple states may be more likely to adopt a standardized audit policy at their facilities *within* a given state (versus adopting a standardized policy at all of their facilities).

We estimate a simple probit regression for *Standardization*. The first testable implication of our conceptual model is that higher levels of heterogeneity among a firm's facilities will decrease the likelihood that the firm uses a standardized adoption policy. Thus, we need to operationalize the concept of heterogeneity from our theoretical model. In the model heterogeneity is represented as the difference in the benefits of auditing the

¹⁸ As a first pass, we electronically matched facilities owned by the same firm using DUNS numbers and company names but since the owner and company name fields do not have standardized formats (e.g., no standard abbreviations or punctuation rules), we made many of the matches manually for facilities with missing DUNS numbers.

facilities that have the highest benefit from auditing and the facilities that have the lowest. Because we don't observe the benefits of auditing different facility types, our empirical measures of heterogeneity rely on observable facility characteristics that may proxy for underlying differences in the benefits of auditing different facilities.¹⁹ We create three firm-level measures of heterogeneity. Each measure captures a different dimension of heterogeneity and takes a value between zero and one where zero denotes a homogeneous portfolio of firms with respect to that characteristic.²⁰ The first measure, *Industry heterogeneity*, focuses on the nature of production at each facility as measured by the up to three 4-digit Standard Industrial Classification (SIC) codes associated with the facility. If the sequence of 4-digit SIC codes is identical for all facilities in the firm's portfolio, then *Industry heterogeneity* equals zero. If each facility in the firm's portfolio has a unique sequence of 4-digit SIC codes, then *Industry heterogeneity* is equal to one.²¹ If the firm's portfolio contains a mix of unique and repeated sequences of SIC codes among its facilities, then *Industry heterogeneity* takes some value between zero and one. In general, a higher value (closer to one) of *Industry heterogeneity* indicates more diversity among the firm's facilities in terms of their production activities.

Our second measure, *Regulatory heterogeneity*, examines the nature of the regulatory environment at the facilities owned by a firm. While our analysis only includes facilities regulated under the CAA, our dataset contains information on other environmental programs to which the facility is subject. We create five indicator

¹⁹ In our theoretical model, we chose to model heterogeneity in benefits while assuming constant costs across facilities. However, similar insights arise in a model with heterogeneity in costs and constant benefits.

²⁰ See the appendix for a more detailed description of these measures.

²¹ Our results are qualitatively unchanged if we use an alternative industry heterogeneity measure created with the sequence of 2-digit SIC codes.

variables from this information. The first two variables indicate that the facility is subject to the provisions of the Toxics Release Inventory (TRI) or the Resource Conservation and Recovery Act (RCRA), respectively. The remaining variables indicate that the facility shows up in an EPA database, the Permit Compliance System (PCS), the Integrated Compliance Information System (ICIS), or the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS), respectively. PCS and ICIS designate facilities that are registered with EPA's federal enforcement and compliance and that hold National Pollutant Discharge Elimination System permits. A facility in CERCLIS is on (or being considered for) Superfund's National Priorities List. *Regulatory heterogeneity* measures the diversity among facilities in terms of the sequence of these five indicator variables. If the sequence of these five indicator variables is the same for all facilities in the firm's portfolio (i.e., the facilities face the same environmental regulatory environment), then *Regulatory heterogeneity* equals zero. A value closer to one indicates more diversity among the firm's facilities in terms of their regulatory exposure.

Our final measure, *Size heterogeneity*, assesses the degree of heterogeneity in terms of the size of the facilities owned by the firm, as measured by the number of employees at the facilities.²² To construct this measure, we first create a categorical variable to characterize the number of employees at the facility as less than or equal to 10, greater than 10 but less than or equal to 100, greater than 100 but less than or equal to 500, or greater than 500. *Size heterogeneity* equals zero if all facilities in the firm's

²² The AFS provides a variable described as the number of employees at the facility. However, this variable takes a value of zero for a large fraction of facilities. If we exclude our size heterogeneity measure from the firm-level analysis below, our results with respect to the other included variables are unchanged. Alternatively, if we assume these zeros are in fact missing values, while our sample size decreases, our results are qualitatively similar.

portfolio have the same value for this categorical variable and one if each facility in the firm's portfolio has a unique value for this categorical variable. Thus, values of *Size heterogeneity* closer to one indicate more diversity among the firm's facilities in terms of facility size. If the degree of heterogeneity discourages adoption of standardized auditing policies, then we expect negative and significant coefficients on these three heterogeneity measures in our probit models.

Our final measure of heterogeneity is intended to capture other sources of unobserved heterogeneity. We argue that a firm whose portfolio includes a larger number of facilities is more likely to exhibit heterogeneity among its facilities than a firm owning a small number of facilities. We acknowledge that *# facilities* is a rough proxy since the number of facilities owned by a firm may also affect the firm's incentives for standardization through other channels (e.g., by affecting the benefits or costs of auditing). This means that our empirical results with respect to this variable will not allow us to confirm or refute our theoretical prediction regarding heterogeneity.

Recall that our model also predicts that the costs and benefits of auditing affect incentives to standardize. In particular, sufficiently high or low benefits relative to the costs of auditing may make auditing all firms or no firms optimal (i.e., may encourage standardization). Unfortunately, the benefits and costs of auditing are unobservable to outsiders. Therefore, to proxy for the relative benefits and costs of auditing, we use a dummy variable, *Public*, that equals one if the firm is publicly traded and zero otherwise. We argue that publicly traded firms face pressure from consumers and investors to comply with environmental regulations that privately held firms do not feel. This is consistent with Henriques and Sadorsky's [31] finding that firms are more likely to

formulate an environmental plan if they face pressure on environmental issues from shareholders. Similarly, Earnhart and Harrington [19] argue that publicly traded firms may have stronger incentives to comply with environmental regulations because they enjoy greater access to external financing, which may reduce costs, and/or face greater pressure from investors for good environmental performance, which may increase benefits. However, since standardization equals one when the firm audits all firms or audits no firms, our model does not offer a clear prediction on the expected sign of the estimated coefficient on *Public* in our probit model.

Table 1 presents descriptive statistics as well as the results of our probit analysis. The first column reports means and standard deviations. While the sample average for *Industry heterogeneity* is 0.67, 35 firms (about 20%) are homogenous along this dimension and 98 firms (about 57%) have a value of *Industry heterogeneity* equal to one. 52 firms are homogeneous with respect to regulatory heterogeneity and 66 firms have the maximum value of regulatory heterogeneity. With respect to size, 71 firms are homogeneous and 50 firms are maximally heterogeneous. Thirty percent of the firms in our sample are publicly traded and the average firm has just over four facilities in its Michigan CAA portfolio (although the median number of facilities is two).

The third column in Table 1 reports estimated coefficients and robust standard errors from our probit model while the fourth column reports conditional marginal effects and standard errors for the probit obtained using the Delta method. The negative and significant coefficients on *Industry heterogeneity* and *Regulatory heterogeneity* are consistent with the model's prediction that heterogeneity should reduce the likelihood of standardization at a multi-facility firm. The coefficient on *Size heterogeneity* is also

negative, as expected, but not significantly different from zero. The conditional marginal effect for *Industry heterogeneity* suggests that the most heterogeneous firm (i.e., *Industry heterogeneity=1*) is 2.6 percentage points less likely to standardize than the homogeneous firm (p-value=0.058). In terms of regulatory heterogeneity, the most heterogeneous firm is 3.6 percentage points less likely to standardize than the firm with *Regulatory heterogeneity* equal to zero (p-value=0.014). Standardization is also less likely among publicly traded firms; publicly traded firms are 5 percentage points less likely to standardize than privately held firms. If publicly held firms are likely to have significantly high benefits (perhaps because they face investor pressure), then we may expect them to be more likely to audit all firms relative to adopting a heterogeneous auditing policy. However, we would simultaneously expect them to be less likely to audit no facilities relative to auditing some but not all facilities. The negative and significant coefficient on *Public* is consistent with the latter effect dominating. Lastly, the probability of standardization decreases with the number of facilities owned by the firm although the estimated conditional marginal effect is small.

To provide some additional evidence that the variables we include in our probit model are reasonable proxies for the underlying measures from our model, we estimate a fractional logit model [32], in which the dependent variable is the fraction of its facilities the firm audits. We include the same independent variables as in our standardization probit. Results of the fractional logit model are reported in the final column of Table 1. Recall that our model suggests that the firm's decision of *which* facility types to audit *given* a heterogeneous auditing policy is driven exclusively by the benefits (or expected benefits) and the costs of auditing while heterogeneity affects whether a firm will use a

standardized policy. Thus, our theory does not predict a relationship between heterogeneity and the fraction of facilities that audit. As shown in the final column of Table 1, we confirm no statistically significant effect of heterogeneity on the fraction of facilities that audit. However, the positive and significant coefficient on *Public* indicates that, relative to privately held firms, publicly traded firms audit a higher percentage of their facilities, perhaps because the benefits of doing so are higher. The positive coefficient on *# facilities* (p-value = 0.104) suggests that larger firms audit a higher fraction of their facilities, perhaps due to lower auditing costs.

What are the implications of our firm-level analysis for empirical models of environmental auditing? First, our results suggest that the measurement error introduced by assigning the same auditing outcome to all facilities owned by a firm, as is necessary with only a firm-level auditing outcome, is likely to be greater for firms whose facilities are more diverse. To provide further evidence of this, we first explore the degree of misalignment between our facility-level audit measure and two candidate firm-level audit measures. *Firm audit1* equals one if the firm owns at least one facility that submitted an intent-to-audit notice and zero otherwise and *Firm audit2* equals one if all of the facilities owned by the firm submitted intent-to-audit notices and zero otherwise.²³ Specifically, we ask: For how many facilities would we incorrectly assign an auditing outcome were we to substitute facility-level audit data with a firm-level audit measure? For both measures, the firm-level measure is different than the facility-level measure for about

²³ It remains possible that (1) a firm with *Firm audit1* equal to zero audits at one of its facilities located outside the state of Michigan or (2) a firm with *Firm audit2* equal to one chooses not to audit at a facility located in another state. Our two firm-level audit measures are defined based on the sample of facilities we observe, which are all located in Michigan.

10% of the 730 facilities owned by the firms in our analysis.²⁴ Table 2 report means and standard deviations for our three heterogeneity measures by whether or not the facility- and firm-level audit measures are equal. If more heterogeneity is associated with more measurement error as suggested by the theory, then we expect the mean heterogeneity to be higher among observations in which the facility- and firm-level measures differ. The table confirms this.

Ultimately we do not know whether either of these firm-level audit measures would align with a firm-level auditing outcome obtained from a survey of firms, as used in Khanna and Widyawati [12]. Regardless, this exercise reinforces the need for caution in assuming complete standardization in auditing outcomes among facilities owned by the same firm. Additionally, our findings hint at the potential importance of controlling for firm characteristics that might influence the standardization decision in analyses of any voluntary program, including environmental auditing. In the next section, we explore the implications of this second insight in a facility-level analysis of the impact of environmental auditing on compliance with the Clean Air Act (CAA).

5. A facility-level analysis of environmental auditing

Our facility-level analysis examines the factors that encourage environmental auditing at a facility as well as the impact of auditing on the facility's long term compliance with the CAA. In particular, we examine the potential for the explicit recognition of firms' differential incentives for standardization in a facility-level

²⁴ The distributions of both *Firm audit1* and *Firm audit2* are significantly different from the facility-level audit measure (both with p-value=0.00).

empirical analysis of environmental auditing to lead to different conclusions regarding the effectiveness of environmental auditing. To do so, we expand our dataset beyond the multi-facility firms included in our analysis above to include stand-alone facilities.²⁵ We extract additional facility characteristics, enforcement history, and current compliance status from the AFS. We also linked the AFS database to EPA's Enforcement and Compliance History Online (ECHO) and Integrated Data for Enforcement Analysis (IDEA) databases to obtain additional enforcement and compliance measures.

We identify an audit at a facility if the facility submitted at least one intent-to-audit notice between 1998 and 2003. In this case, the variable *Facility audit* equals one for the facility. Otherwise, we do not measure an audit at the facility and *Facility audit* equals zero. Of the 2811 facilities included in our analysis, 111 (about 4%) submitted intent-to-audit notices during the time frame we consider.²⁶

The CAA requires facilities to self-report their compliance status on an on-going basis (i.e., each quarter).²⁷ We use these data to create a binary compliance variable, *Comply*, which equals one if the facility is in compliance for each of the 12 quarters between the second quarter of 2004 and the first quarter of 2007.²⁸ *Comply* equals zero if

²⁵ Note that due to missing data for some of the variables included in our facility-level analysis not all of the 730 facilities whose parent companies are included in our firm-level analysis (Section 4) are present in our facility-level analysis. 542 facilities whose firms are included in our firm-level analysis are present in the sample used in our facility-level analysis.

²⁶ The 4 percent audit rate is roughly consistent with Potoski and Prakash's [33] estimate that approximately 4 percent of "major" CAA-regulated facilities participate in the ISO14001 certification program, a program which requires (among other things) adoption of an environmental auditing protocol.

²⁷ The self-reported nature of the compliance data introduces the potential for bias. Existing evidence on the accuracy of self-reported environmental compliance data is mixed. For example, deMarchi and Hamilton [34] find reporting irregularities in self-reported data from the Toxics Release Inventory but these data are not used by regulators for enforcement purposes. Analyses of self-reported compliance measures that are used for enforcement purposes have been unable to reject the accuracy of these data [35, 36]. The use of self-reported data for enforcement purposes provides strong incentives, such as criminal fines, to encourage truthful reporting. The CAA self-reported data are used for enforcement.

²⁸ Facilities are included in our analysis provided we observe compliance or non-compliance status for all 12 quarters. Facilities with missing compliance data during this time period are excluded from our

the facility is out of compliance for at least one of these 12 quarters. In focusing on the longer run impacts of auditing on compliance with the CAA, our analysis differs from Khanna and Widyawati [12], who examine contemporaneous compliance effects.

A key concern in conducting an empirical analysis of the effect of audits on compliance behavior is the potential for endogeneity. Specifically, because violations discovered during the course of an audit are eligible for penalty mitigation (with the possibility of a complete waiver of penalties), facilities that are concerned that they are non-compliant may be more likely to implement an environmental auditing program. If so, the audit decision would be correlated with factors that affect the compliance status of the facility. However, while Khanna and Widyawati [12] could not reject the hypothesis that their audit variable was endogenous, both Evans et al. [18] and Earnhart and Harrington [19] found no evidence that their audit measure was endogenous

Even though the endogeneity of the auditing measure may be less of a concern in this analysis given our emphasis on the longer run effects of auditing, to address the possibility for endogeneity we follow Khanna and Widyawati by estimating a simultaneous recursive bivariate probit model for the two latent binary dependent variables, *Facility audit* and *Comply*.²⁹ We follow both Evans et al. [18] and Khanna and Widyawati [12] in selecting facility- and county-level variables for inclusion in the audit and compliance equations. We can broadly classify these variables as those related to the facility's characteristics, its inspection and compliance history, the stringency of the CAA

analysis. We re-estimated our models including all facilities with at least 2 quarters of data in this time period and we obtain qualitatively similar results.

²⁹ Maddala [37] derives the full information maximum likelihood estimator for this model, which we estimate in Stata. According to Wilde [38], the model is identified with sufficient variation in the independent variables. In our previous work, we used exclusion restrictions for identification. Unfortunately we are unable to find valid exclusion restrictions for this application thus we adopt Khanna and Widyawati's identification strategy, which follows Wilde [38].

regulations to which the facility is subject, its regulatory exposure under non-CAA environmental programs, and county-level characteristics for the county in which the facility is located. Table 3 provides variable descriptions and summary statistics.³⁰

We minimize our discussion of these facility- and county-level variables in order to focus on the primary variables of interest for motivating our model of the firm's standardization decision. Based on our conceptual model, we expect the incentives for standardization to vary systematically across firms. Our empirical analysis in section 4 confirms that the incentives for adopting a standardized auditing policy vary between publicly traded and privately held firms, and by the degree of heterogeneity among the facilities owned by the firm. Here, we explore the potential implications of recognizing these differences in a facility-level analysis of environmental auditing by including two relevant firm-level controls. The first variable, *Multifacility*, takes the value of one if the firm owns more than one CAA-regulated facility in Michigan. The second variable, *Public*, equals one if the firm that owns the facility (i.e., the facility's parent company) is publicly-traded and zero if it is privately held. Ten percent of facilities in our sample are owned by publicly-traded firms while about 20 percent belong to a multi-facility firm.

Table 4 reports the results of our facility-level bivariate probit analysis. The results of the audit and compliance equations are given in the left-hand and right-hand panels, respectively. We report results for two specifications that differ in terms of whether or not we include the two firm-level measures. Because Model I excludes these firm-level controls, it is similar to those considered in previous analyses of environmental

³⁰ See Section 4 for a description of the regulatory program acronyms listed in Table 1.

auditing.³¹ Model II is identical to Model I other than the inclusion of *Public* and *Multifacility*.³² Model I is reported in the second and fourth columns while Model II is reported in the third and fifth columns.

The results of Models I and II are broadly consistent with each other and with past analyses with respect to the variables that measure facility- and county-level characteristics. However, the inclusion of *Public* and *Multifacility* in Model II results in two important differences between the results. First, the primary variable of interest in the compliance equation, *Audit*, is positive and significant in Model I but insignificant (and negative) in Model II. The former suggests that, controlling for the endogeneity of *Audit*, auditing facilities are more likely to be in compliance. The latter result suggests no significant impact of auditing on long-term compliance. Second, the estimate for ρ in Model I is negative and significant, which suggests a correlation between the unobserved factors that affect *Audit* and *Comply*. The estimate for ρ is insignificant in Model II. Along both of these dimensions, Model I is consistent with the findings of Khanna and Widyawati [12] while Model II aligns with Evans et al.'s [18] results. A Wald test confirms that we are able to reject the null hypothesis that *Public* and *Multifacility* are jointly equal to zero ($\chi^2(4) = 40.92$, p-value = 0.00). Thus, including these variables significantly improves the fit of the model. If, as we argue, these variables are related to

³¹ This analysis and that of Khanna and Widyawati [12] examine the effects of auditing on compliance with the CAA. However, the significant differences between the dataset we use here and the dataset used by Khanna and Widyawati make replicating their specifications unfeasible. While our data are closer to the sample used by Evans et al. [18] as both samples focus on facilities in Michigan, Evans et al.'s sample includes manufacturing facilities regulated under RCRA while our sample restricts attention to CAA-regulated facilities.

³² We get the qualitatively same results if we use all three firm-level heterogeneity measures in place of *Multifacility*.

the firm's incentives for standardization, then our results suggest that such controls are important in analyses of environmental auditing.

6. Discussion

As external observers, researchers rely on data provided by regulated entities, through surveys or through self-reports, or by regulators to measure participation in voluntary programs and related outcomes. Often such data are only available at the firm-level even though program adoption may happen at the facility-level. When researchers use a firm-level measure of participation, they are making an assumption, which often goes unstated, regarding the nature of decision-making within the regulated entity. In particular they are assuming that the firm has a standardized adoption policy. However, our theoretical model shows that it may not be optimal for a firm to adopt a standardized policy. We find that the level of heterogeneity among a firm's facilities is inversely related to the likelihood that a standardized policy is optimal. However, at multi-facility firms with relatively high or low benefits from adoption, a firm is more likely to adopt a standardized policy.

Because our theoretical analysis suggests that assigning the same auditing outcome for all facilities owned by the same firm (i.e., assuming standardization) is more problematic for certain firms (e.g., those whose facilities exhibit more heterogeneity), we examine the implications for the model using the adoption of a voluntary environmental audit program as a case study. Using a dataset of CAA-regulated firms in Michigan, we analyze the decision to standardize the adoption of environmental audits at the firm's regulated facilities in Michigan. We find that our measures of firm heterogeneity and

proxies for the relative benefits from auditing do have statistically significant impacts on the standardization decision and that those effects are consistent with the theory.

Next we examine whether accounting for those firm-level factors that influence the standardization decision affects the results of an empirical analysis of the effect of environmental auditing on facility compliance. Using a facility-level data set on both environmental audit adoption and facility compliance with the CAA, we find that the estimated impact of auditing on compliance depends on whether or not we include controls for firm incentives to adopt a standardized auditing policy. When such controls are excluded, our empirical results suggest a positive impact of auditing on long-run compliance with the CAA. This effect disappears once these controls are added. Of course, our controls are not perfect measures of firm incentives for standardization. However, as in our dataset, many datasets are likely to contain a handful of observable firm characteristics that may be related to this decision and can therefore be used as proxies.

While our analysis examines the issue of standardization within the context of environmental auditing, the decision to adopt an environmental auditing program shares many aspects with the decision to participate in other voluntary environmental programs. Thus, we believe that the lessons from our analysis have relevance for empirical analyses of the impact on environmental performance of participation in other voluntary environmental programs. For example, the various empirical analyses of EPA's 33/50 program reach different conclusions regarding the effectiveness of participation in improving environmental performance. While these studies vary across a number of dimensions, one key dimension is whether participation is measured at the firm- or

facility-level. Interestingly, neither of the two studies that examine the effect of facility-level participation on facility-level compliance [20, 21] include firm-level controls.

While three of the four studies that use a firm-level measure of adoption (and assume that participation is standardized) include the number of facilities owned by the firm into their analyses, none of them include measures of heterogeneity or the relative benefit of the program. We believe that including more firm-level controls might help to reconcile some of the mixed results in this literature.

To help policy makers effectively assess the extent to which participation in voluntary programs can substitute for more traditional compliance and enforcement mechanisms, it is important to conduct careful studies of such programs. Given the findings of our study, we recommend that future analyses of voluntary environmental program include facility-level participation measures as well as firm-level controls, when available. If facility-level measures are not available, we urge researchers to control for the factors that make a firm more likely to standardize including the heterogeneity of the facilities in the firm's portfolio.

REFERENCES

- [1] U.S. Government Accountability Office (GAO). 2007. "EPA-State Enforcement Partnership Has Improved, but EPA's Oversight Needs Further Enhancement." GAO-07-883.
- [2] Kotchen, Matthew. 2013. Voluntary- and Information-Based Approaches to Environmental Management: A Public Economics Perspective." *Review of Environmental Economics and Policy* 7(2): 276-295.
- [3] Khanna, Madhu and Keith Brouhle. 2008. "Effectiveness of Voluntary Environmental Initiatives." In *Governing the Environment: Interdisciplinary Perspectives*. Cambridge, England: Cambridge University Press (eds. M. Delmas and O. Young)
- [4] Prakash, Aseem and Matthew Potoski. 2012. "Voluntary Environmental Programs: A Comparative Perspective." *Journal of Policy Analysis and Management* 31: 123-138.
- [5] Lyon, Thomas and John Maxwell. 2007. "Environmental Public Voluntary Programs Reconsidered." *Policy Studies Journal* 35: 23-750.
- [6] U.S. Environmental Protection Agency (EPA). 1997. "EPA Strategic Plan." EPA/190-R-97 002.
- [7] Pfaff, Alexander S. P. and Chris William Sanchirico. 2004. "Big Field, Small Potatoes: An Empirical Assessment of EPA's Self-Audit Policy." *Journal of Policy Analysis and Management* 23: 415-432.
- [8] Short, Jodi and Michael Toffel. 2008. "Coerced Confessions: Self-Policing in the Shadow of the Regulator." *Journal of Law, Economics and Organization* 24(1): 45-71.
- [9] Stafford, Sarah. 2007. "Should You Turn Yourself In? The Consequences of Environmental Self-Policing." *Journal of Policy Analysis and Management* 26: 305-326.
- [10] Stafford, Sarah L. 2005. "Does Self-Policing Help the Environment? EPA's Audit Policy and Hazardous Waste Compliance." *Vermont Journal of Environmental Law* 6.
- [11] Stafford, Sarah L. 2006. "State Adoption of Environmental Audit Initiatives." *Contemporary Economic Policy* 24(1): 172-187.
- [12] Khanna, Madhu and Diah Widyawati. 2011. "Fostering Regulatory Compliance: The Role of Environmental Self-Auditing and Audit Policies." *Review of Law and Economics* 7: 129-163.

- [13] Mishra Birendra K., D. Paul Newman and Christopher H. Stinson. 1997. "Environmental Regulations and Incentives for Compliance Audits." *Journal of Accounting and Public Policy* 16: 187–214
- [14] Pfaff, Alexander S. P. and Chris William Sanchirico. 2000. "Environmental Self-Auditing: Setting the Proper Incentives for Discovery and Correction of Environmental Harm." *Journal of Law, Economics, and Organization* 16: 189-208.
- [15] Friesen, Lana. 2006. "The Social Welfare Implications of Industry Self-auditing." *Journal of Environmental Economics and Management* 51: 280-294.
- [16] U.S. General Accounting Office (GAO). 1995. "Environmental Auditing: A Useful Tool that Can Improve Environmental Performance and Reduce Cost Savings." GAO/RCED 95-37.
- [17] U.S. Environmental Protection Agency (EPA). 1999. "EPA/CMA Root Cause Analysis Pilot Project: An Industry Survey,"
- [18] Evans, Mary F., Lirong Liu, and Sarah Stafford. 2011. "Do Environmental Audits Improve Long-term Compliance? Evidence from Manufacturing Facilities in Michigan." *Journal of Regulatory Economics* 40: 279-302.
- [19] Earnhart, Dietrich, and Donna Ramirez Harrington. 2014. "The Role of Environmental Audits in Reducing Water Effluent Levels." *Journal of Environmental Economics and Management* 68(2):243–261.
- [20] Bi, Xiang, and Madhu Khanna. 2012. "Reassessment of the Impact of the EPA's Voluntary 33/50 Program on Toxic Releases." *Land Economics* 88(2): 341-361.
- [21] Vidovic, Martina, and Neha Khanna. 2012. "Is Voluntary Pollution Abatement in the Absence of a Carrot or Stick Effective? Evidence from Facility Participation in the EPA's 33/50 Program." *Environmental and Resource Economics* 52(3): 369-393.
- [22] Khanna, Madhu, and Lisa A. Damon. 1999. "EPA's Voluntary 33/50 Program: Impact on Toxic Releases and Economic Performance of Firms." *Journal of Environmental Economics and Management* 31(1): 1-25.
- [23] Vidovic, Martina, and Neha Khanna. 2007. "Can Voluntary Pollution Prevention Programs Fulfill Their Promises? Further Evidence from the 33/50 Program." *Journal of Environmental Economics and Management* 53(2): 180-195.
- [24] Innes, Robert, and Abdoul G. Sam. 2008. "Voluntary Pollution Reductions and the Enforcement of Environmental Law: An Empirical Study of the 33/50 Program." *Journal of Law and Economics* 51(2): 271-96.

- [25] Gamper-Rabindran, Shanti. 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.
- [26] Sah, Raaj and Joseph Stiglitz. 1986. "The Architecture of Economic Systems: Hierarchies and Polyarchies." *The American Economic Review* 76: 716-727.
- [27] Van Zandt, Timothy. 1999. "Real-Time Decentralized Information Processing as a Model of Organizations with Boundedly Rational Agents." *The Review of Economic Studies* 66: 633-658
- [28] Chang, Myong-Hun and Joseph Harrington. 2002. "Decentralized Business Strategies in a Multi-Unit Firm." *Annals of Operations Research* 109: 77-98.
- [29] Hunnicutt, Lynn. 2001. "Mixups in the Warehouse: Centralized and Decentralized Multi-plant Firms." *Economic Inquiry* 39(4): 537-548.
- [30] Evans, Mary F., Scott M. Gilpatric, Michael McKee, and Christian A. Vossler, Managerial Incentives for Compliance with Environmental Information Disclosure Programs, in "Experimental Methods, Environmental Economics" Routledge, UK (eds. T. Cherry, J. Shogren, and S. Kroll), 244-260, 2008.
- [31] Henriques, Irene, and Perry Sadorsky. 1996. "The Determinants of an Environmentally Responsive Firm: An Empirical Approach." *Journal of Environmental Economics and Management* 30: 381-395.
- [32] Papke, Leslie E, and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics* 11: 619-632.
- [33] Potoski, Matthew and Aseem Prakash. 2005. "Green Clubs and Voluntary Governance: ISO 14001 and Firms' Regulatory Compliance." *American Journal of Political Science* 49: 235-248.
- [34] de Marchi, Scott and James T. Hamilton. 2006. "Assessing the Accuracy of Self-Reported Data: An Evaluation of the Toxics Release Inventory." *Journal of Risk and Uncertainty* 32: 57-76.
- [35] Laplante, Benoit and Paul Rilstone. 1996. "Environmental Inspections and Emissions of the Pulp and Paper Industry in Quebec." *Journal of Environmental Economics and Management* 31: 16-36.
- [36] Shimshack, Jay P. and Michael B. Ward. 2005. "Regulator Reputation, Enforcement, and Environmental Compliance." *Journal of Environmental Economics and Management* 50: 519-540.

[37] Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press.

[38] Wilde, Joachim. 2000. "Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors." *Economics Letters* 69(3): 309-312.

Table 1. Results of firm-level empirical analysis

		Probit results—dependent variable is <i>Standardization</i>		Fractional logit results—dependent variable is fraction of facilities audited
	Sample mean (standard deviation)	Estimated coefficient (robust standard error)	Conditional marginal effect (Delta-method standard error)	Estimated coefficient (robust standard error)
<i>Industry heterogeneity</i>	0.67 (0.42)	-0.89* (0.46)	-0.026 (0.014)	0.047 (0.73)
<i>Regulatory heterogeneity</i>	0.51 (0.43)	-0.88* (0.41)	-0.036* (0.015)	0.15 (0.76)
<i>Size heterogeneity</i>	0.39 (0.42)	-0.091 (0.80)	-0.0080 (0.031)	0.37 (0.64)
<i>Public</i>	0.30 (0.46)	-0.80** (0.31)	-0.050** (0.019)	0.95* (0.50)
<i># facilities</i>	4.27 (6.47)	-0.060** (0.018)	-0.0057** (0.0022)	0.045 (0.027)
Constant	--	3.25** (0.53)	--	-3.62** (1.02)

N=171, pseudo R2=0.22

For all variables except *# facilities*, conditional marginal effects are calculated as the response for the change of going from a value of zero to a value of one (at the means of the other variables). For *# facilities*, the conditional marginal effect is calculated at the means of all the variables. * indicates significance at 10% and ** indicates significance at 5%.

Table 2. Means and standard deviations for firm-level heterogeneity measures by alignment between facility- and firm-level audit measures

Heterogeneity measure	Firm audit measure			
	Firm audit 1 (at least one facility audits)		Firm audit 2 (all facilities audit)	
	Firm-level measure equal to facility-level measure	Firm-level measure not equal to facility-level measure	Firm-level measure equal to facility-level measure	Firm-level measure not equal to facility-level measure
Industry heterogeneity	0.52 (0.016)	0.68 (0.032)	0.52 (0.016)	0.74 (0.025)
Regulatory heterogeneity	0.40 (0.015)	0.68 (0.21)	0.40 (0.015)	0.73 (0.020)
Size heterogeneity	0.30 (0.014)	0.46 (0.024)	0.31 (0.014)	0.41 (0.024)

Table reports means and standard deviations for heterogeneity measure listed in each row by whether or not the facility-level audit measure is the same as the firm-level audit measure for two different audit measures. All pairwise differences in means (i.e., between columns 2 and 3, 4 and 5) are significantly different at the 5% level.

Table 3. Variable description and summary statistics for facility-level analysis of environmental auditing

Variable name	Variable description	Mean	Standard deviation
<i>Facility-level variables</i>			
Comply	= 1 if facility reported complete compliance from 2 nd quarter 2004 through 1 st quarter 2007, = 0 if facility was out of compliance for one or more quarters	0.80	0.40
Facility Audit	= 1 if facility submitted at least one intent to audit notice between 1998 and 2003	0.040	0.19
Employees	Number of employees at facility	169.24	1013.35
Manufacturing	= 1 if facility is classified as manufacturing (SIC codes 20 – 39)	0.55	0.50
Past inspection	= 1 if facility was inspected at least once between 1994 and 1997	0.42	0.49
Count past inspection	Number of times facility was inspected between 1994 and 1998	0.97	1.62
Past violation	= 1 if the facility was cited for a violation at least once between 1994 and 1998	0.049	0.22
Past penalties	Total CAA penalties paid from 1994 to 1998, in \$1,000s	31.91	571.45
Major	= 1 if facility is classified as a major air source	0.15	0.36
MACT	= 1 if facility is subject to MACT (maximum achievable control technology).	0.071	0.26
PSD	= 1 if facility is subject to PSD (prevention of significant deterioration)	0.049	0.22
NSR	= 1 if facility is subject to NSR (new source review)	0.0068	0.082
NSPS	= 1 if facility is subject to NSPS (new source performance standards)	0.17	0.38
SIP	= 1 if facility is classified as a SIP (state implementation plan) source	0.94	0.24
CFC	= 1 if facility is subject to CFC tracking	0.052	0.22
CERCLIS	=1 if the facility is tracked in CERCLIS	0.015	0.12
ICIS	= 1 if facility is tracked ICIS	0.13	0.34
PCS	=1 if the facility is tracked in PCS	0.11	0.32
RCRA	= 1 if facility is regulated under RCRA	0.57	0.49
TRI	=1 if the facility is subject to TRI reporting	0.31	0.46
<i>County-level variables</i>			
Nonattainment	= 1 if facility is located in county that was classified as non-attainment for ozone in 2004	0.65	0.48
County conservancy	Number of individuals in the county that belonged to the Nature Conservancy in 2005 per 1000 residents	2.49	1.05
County Republicans	Percentage of the voters in the county in the 2000 Presidential election that voted Republican	47.92	10.65

County education	Percentage of the county's population aged 25 and older whose highest level of education achieved was high school in 1990	83.32	4.22
<i>Firm-level variables</i>			
Multi-facility	=1 if the firm (parent company) owns more than one CAA-regulated facility in Michigan	0.19	0.39
Public	= 1 if firm (parent company) is publicly traded	0.10	0.31

Table 4. Results of the facility-level analysis of environmental auditing

Variable name	Audit equation		Compliance equation	
	Model I	Model II	Model I	Model II
Audit	--	--	1.41** (0.25)	-0.51 (0.87)
Employees	0.00011* (0.000058)	0.000094* (0.000053)	-0.00021** (0.000078)	-0.00015 (0.000090)
Manufacturing	-0.015 (0.12)	0.0061 (0.13)	-0.015 (0.077)	-0.026 (0.082)
Past inspection	--	--	-0.088 (0.074)	-0.10 (0.082)
Past violation	--	--	-0.023 (0.41)	-0.020 (0.44)
Past inspection*Past violation	--	--	0.13 (0.44)	0.12 (0.48)
Count past inspections	-0.013 (0.027)	-0.026 (0.027)	--	--
Past penalties	-0.00062 (0.00049)	-0.00057 (0.00044)	-0.00014 (0.00017)	-0.00017 (0.00018)
Major	0.37** (0.14)	0.32** (0.15)	-1.95** (0.11)	-2.07** (0.11)
MACT	0.13 (0.16)	0.25 (0.16)	-0.13 (0.13)	-0.10 (0.14)
PSD	0.32* (0.18)	0.37* (0.19)	-0.32* (0.18)	-0.28 (0.20)
NSR	-0.93* (0.53)	-1.019 (0.56)	-0.23 (0.50)	-0.37 (0.52)
NSPS	0.078 (0.14)	0.034 (0.15)	-0.24** (0.092)	-0.21** (0.097)
SIP	0.56 (0.35)	0.52 (0.37)	0.71** (0.18)	0.83** (0.18)
CFC	0.47* (0.24)	0.38 (0.28)	-1.04** (0.20)	-1.04** (0.21)
CERCLIS	-5.08** (0.35)	-6.38** (0.93)	0.12 (0.34)	-0.13 (0.33)
ICIS	0.15 (0.11)	0.14 (0.13)	0.054 (0.11)	0.085 (0.12)
PCS	-0.041 (0.13)	-0.086 (0.14)	-0.086 (0.11)	-0.14 (0.12)
RCRA	0.44** (0.15)	0.45** (0.16)	0.077 (0.081)	0.15* (0.08)
TRI	0.65** (0.13)	0.58** (0.14)	-0.19** (0.095)	-0.038 (0.10)
Nonattainment	--	--	-0.057	-0.080

			(0.073)	(0.080)
County conservancy	-0.077 (0.072)	-0.11 (0.071)	0.0092 (0.041)	0.0040 (0.041)
County Republicans	-0.020** (0.0055)	-0.020** (0.0056)	0.016** (0.0039)	0.014** (0.0044)
County education	0.029 (0.020)	0.035* (0.021)	0.014 (0.011)	0.018 (0.012)
Public	--	0.21 (0.15)	--	0.27* (0.15)
Multi-facility	--	0.53** (0.12)	--	-0.24** (0.11)
Constant	-4.43** (1.47)	-4.99** (1.53)	-1.05 (0.84)	-1.34 (0.87)
Rho	-0.82** (0.096)	0.14 (0.41)	--	--
Log L	-1168.60	-1150.76	--	--

Table reports estimated coefficients and robust standard errors, below coefficients in parentheses. Sample size is 2811. * indicates significance at 10% and ** indicates significance at 5%.

APPENDIX. Description of heterogeneity measures

Industry heterogeneity

The dataset contains up to three four-digit SIC codes for each facility. Let SIC1, SIC2, SIC3 represent these variables. Three steps are required to create *Industry heterogeneity*.

Step 1:

Sort facilities by parent company identifier, SIC1, SIC2, SIC3. Create a variable, called *Industry duplicates*, that equals zero if a particular sequence of SIC1, SIC2, SIC3 is unique among facilities within a parent company. Otherwise, *Industry duplicates* assigns a count to identify the number of times a particular sequence of SIC1, SIC2, SIC3 shows up within facilities owned by the parent company (i.e., the first facility with the sequence will have *Industry duplicates* equal to one, the second facility with that same sequence will have *Industry duplicates* equal to two, and so on). Consider a hypothetical example firm that owns five facilities with values for SIC1, SIC2, SIC3 given in the following table (-- indicates a missing value):

Facility ID	SIC1	SIC2	SIC3	Industry duplicates
1	30	37	--	0
2	20	--	--	1
3	20	24	26	0
4	20	--	--	2
5	20	--	--	3

Two sequences of SIC1, SIC2, SIC3 are unique within the facilities owned by this firm (30, 37, --; 20, 24, 26) so the facilities with these sequences (1 and 3) are assigned values of *Industry duplicates* equal to zero. One sequence, 20, --, --, occurs three times, first for facility 2 (so *Industry duplicates*=1), second for facility 4 (so *Industry duplicates*=2), and third for facility 5 (so *Industry duplicates*=3).

Step 2:

Identify the maximum value of *Industry duplicates* among facilities owned by each parent company. Call this *Max industry duplicates*. Note that if a firm has *Max industry duplicates* equal to *Num facilities* (the # of facility owned by the parent company), then all facilities owned by that parent company have the same values for the sequence SIC1, SIC2, SIC3. If *Max industry duplicates* is equal to zero, then each facility owned by the firm has a unique sequence of SIC1, SIC2, SIC3.

Step 3:

Form *Industry heterogeneity* for each firm using the following:

$$Industry\ heterogeneity = \frac{Num\ facilities - Max\ industry\ duplicates}{Num\ facilities}$$

Since the range of *Max industry duplicates* is zero to *Num facilities* for each firm, *Industry heterogeneity* is bounded between zero and one where zero indicates homogeneity (i.e., *Max industry duplicates* = *Num facilities*) and one indicates the largest degree of heterogeneity (i.e., *Max industry duplicates* = 0).

Regulatory heterogeneity

This measure is formed using an analogous three step method. However, we are interested in the values of the sequence of five variables: *CERCLIS*, *ICIS*, *PCS*, *RCRA*, *TRI*. If *Max regulatory duplicates* represent the maximum value of *Regulatory duplicates* (where this measure is formed as described above), then

$$\text{Regulatory heterogeneity} = \frac{\text{Num facilities} - \text{Max regulatory duplicates}}{\text{Num facilities}}.$$

Size heterogeneity

This measure uses only one variable, rather than a sequence of variables as the other two measures so forming it is more straightforward. We use *Employees*, the number of employees at the facility, to create a categorical variable, *Cat employees* where

$$\text{Cat employees} = \begin{cases} 1 & \text{if } \text{Employees} \leq 10 \\ 2 & \text{if } 10 < \text{Employees} \leq 100 \\ 3 & \text{if } 100 < \text{Employees} \leq 500 \\ 4 & \text{if } 500 < \text{Employees} \end{cases}$$

We then create a variable for each facility, *Employee duplicates*, formed using the same method as described above, based on the values of *Cat employees* at the facilities owned by a given parent company. Letting *Max employee duplicates* represent the maximum value of *Employee duplicates* among facilities owned by a parent company, we form *Size heterogeneity* using the same technique as with the other two heterogeneity measures:

$$\text{Size heterogeneity} = \frac{\text{Num facilities} - \text{Max employee duplicates}}{\text{Num facilities}}$$