1999

Endogenous monitoring and enforcement of a transferable emissions permit system

JK Stranlund

KK Dhanda

Endogenous Monitoring and Enforcement of a Transferable Emissions Permit System[1]

John K. Stranlund

Department of Resource Economics, 203 Draper Hall, University of Massachusetts-Amherst, Amherst, Massachusetts 01003
E-mail: stranlund@resecon.umass.edu

and

Kanwalroop Kathy Dhanda

School of Business Administration, University of Portland, 5000 North Willamette Boulevard, Portland, Oregon 97203

Received July 7, 1998; revised July 15, 1999

The literature on noncompliant firms in transferable emissions permit systems offers little guidance to policymakers that must determine how to commit resources to monitor firms and punish violations in such systems. We consider how a budget-constrained enforcement authority that seeks to minimize aggregate noncompliance in a transferable emissions permit system should allocate its monitoring and enforcement efforts among heterogeneous firms. With a conventional model of firm behavior in a transferable permit system, we find that differences in the allocation of monitoring and enforcement effort between any two types of firms should be independent of differences in their exogenous characteristics.

1. INTRODUCTION

Transferable emissions permit systems are gaining support in policy circles in large part because of their well-known efficiency properties, and because of the belief that efficient outcomes can be achieved more easily than with command-and-control standards. However, the efficiency properties of transferable permit system appear to hold only under quite stringent and patently unrealistic assump-tions. For example, a small but important literature has examined the effects of noncompliance on the performance of transferable permit systems. Malik [9] appears to be the first to cast doubt on the efficiency properties of transferable permit systems when firms may be noncompliant. In a comparison of a transferable permit system to uniform emissions standards with exogenous enforcement, Keeler [8] finds that noncompliance (and hence, aggregate emissions) may be greater under a transferable permit system than under uniform standards. In another comparison of a transferable permit system to a uniform emissions standard, but this time when enforcement expenditures are committed to achieve a certain degree of compliance, Malik [10] finds that a transferable permit system may be

[1] Thanks are due Barry Field, Tom Stevens, two anonymous referees, and an associate editor of the *Journal of Environmental Economics and Management* for their insightful comments and suggestions.

more costly to enforce, and hence, enforcement plus aggregate abatement costs may actually be higher for such systems than for uniform emissions standards.

Despite the importance of examining the welfare properties of transferable permit systems and comparing them to alternative policies when firms may be noncompliant, the literature offers little guidance to policymakers that are faced with the problem of how to commit resources to monitor firms and punish violations in transferable permit systems. It is toward providing some of this guidance that this paper is directed. The primary objective of this paper is to consider a fundamental policy question that has not been addressed before: How should a budget-constrained enforcement authority distribute its effort among heterogeneous, noncompliant firms in a transferable emissions permit system?

Garvie and Keeler [5] address this issue in the context of emissions standards, and we follow their approach quite closely. In particular, we model both monitoring and enforcement effort, where monitoring is captured by the probability with which a firm is audited and enforcement is a resource commitment to punish violations once they are detected. In addition, we model the regulatory choice of monitoring and enforcement as a two-stage game with complete information. In the first stage of the game, a budget-constrained enforcement authority chooses a monitoring and enforcement regime to minimize aggregate noncompliance in a transferable emissions permit system. In the second stage the firms choose their emissions, their demands for emissions permits, their consequent violations, and an equilibrium in the permit market is established.

The paper is organized in the spirit of backward induction, so we begin in Section 2 with an analysis of a firm's choices. The most important results of this section are two independence results that have significant implications for effective monitoring and enforcement. We show first that a firm's choice of emissions is independent of the monitoring and enforcement pressure applied to it. This result is not entirely new (Malik [9]), but its implications for monitoring and enforcement policy have not been explored. The most surprising result of this section is new: a firm's choice of violation, even whether it is to be in violation or not, is independent of its exogenous characteristics. This result suggests that if one observes that a firm is in violation more often than another, it is likely due to differences in monitoring and enforcement not because it employs an inferior emissions-control technology or a dirtier production process.

In Section 3 we begin by characterizing equilibria of the emissions permit market when firms are noncompliant. Taking the firm's choices and the resulting equilibria into account, we examine the enforcement authority's optimizing distribution of monitoring and enforcement pressure among the firms and find that differences in the pressure applied to different types of firms should be independent of differences in the firms' exogenous characteristics. Thus, an enforcement authority that seeks to maximize the effectiveness of its enforcement budget should not concern itself with fundamental differences among the firms to guide its decisions about distributing monitoring and enforcement efforts, even though the firms may be very different. This result stands in sharp contrast to Garvie and Keeler's finding that optimal monitoring and enforcement of emissions standards requires that greater monitoring and enforcement effort be directed at firms with parametrically higher marginal control costs.

In Section 4 we discuss a number of policy implications of our finding that differences in the application of monitoring and enforcement in a transferable emissions permit system should be independent of differences in the firms' exogenous characteristics. We note first that the result depends critically on the two independence results of Section 2—each firm's choice of emissions is independent of the monitoring and enforcement pressure applied to it, and each firm's choice of violation is independent of its exogenous characteristics. A judgement about the real-world applicability of our primary policy conclusion must rest on empirical tests of these hypotheses.

If empirically valid, our primary finding has strong implications about the value and use of information on firms' abatement costs in the design of enforcement strategies for transferable permit systems. Since the distribution of monitoring and enforcement effort should be independent of exogenous differences among the firms, if the firms face the same penalty structure and the costs of conducting audits and applying enforcement pressure do not vary across firms, a uniform monitoring and enforcement strategy that exhausts the enforcement budget minimizes aggregate noncompliance given that budget. Thus, in this case, knowledge of the firms' abatement costs is not relevant for designing an enforcement strategy. This strong conclusion needs to be qualified somewhat when penalty structures and monitoring and enforcement cost-parameters vary across firms. In these cases, a differentiated monitoring and enforcement strategy is required, and the distribution of effort depends, in part, on the equilibrium permit price. Knowledge of the firms' abatement costs is necessary to forecast the equilibrium permit price, but once again, exogenous differences of the firms' abatement costs should not affect the distribution of monitoring and enforcement effort.

2. FIRM BEHAVIOR IN A TRANSFERABLE EMISSIONS PERMIT SYSTEM

2.1. *Basic Assumptions*

Throughout, we consider a fixed set of heterogeneous, risk-neutral firms. We assume a competitive emissions permit market so that the choices of a single firm have no effect on the equilibrium of the market. We wish to incorporate the fact that an enforcement authority's strategy will likely affect the permit market equilibrium, and this in turn will affect the firms' equilibrium compliance choices. Toward that end, we assume that firms are grouped by type into a set K, and that there are n^k identical firms of type k. The enforcement authority is going to choose a type-specific monitoring and enforcement strategy. We assume that there are enough firms of each type so that their aggregate choices impact the market equilibrium; hence, the enforcement authority must account for the market effects of its monitoring and enforcement strategy.[2]

The emissions-control (abatement) costs of a k-type firm are summarized by $c(e^k, \alpha^k)$, which is strictly decreasing and convex in the firm's emissions $e^k[c_e(e^k, \alpha^k) < 0$ and $c_{ee}(e^k, \alpha^k) > 0$; throughout subscripts denote partial deriva-

[2]Our assumption of identical firms of each type is not critical. A reasonable alternative would assume that the firms are all different, but that there is at least one characteristic that is common to groups of firms upon which the enforcement authority may condition its monitoring and enforcement strategy (e.g., whether a particular control device is installed or not). All of the results of this paper hold under this alternative specification.

tives in the usual manner]. Typically, a polluting firm may pursue a number of strategies to control its emissions including reducing its output, substitution toward cleaner inputs, adopting cleaner production techniques, as well as installing end-of-pipe emissions-control devices. Therefore, its emissions-control costs will depend on a number of exogenous factors including prices of outputs and inputs, and parametric characteristics of its production and emissions-control technologies.[3] For a k-type firm, these characteristics are arrayed in the vector α^k, and *exogenous firm-heterogeneity* is introduced to the model by allowing this vector of parameters to vary among types of firms. [In a slight abuse of notation, a derivative with respect to some element of α^k will usually be indicated simply by the subscript α; for example, $-c_{e\alpha}(e^k, \alpha^k)$ will denote the change in marginal abatement costs from a change in some element of α^k.]

Suppose that a total of L emissions permits are issued to the firms free-of-charge, and that possession of a permit confers the legal right to release one unit of emissions. Let l_0^k be the number of emissions permits that are initially allocated to each k-type firm, and let l^k be the number of permits each of these firms holds after trade. Competitive behavior in the permit market establishes a constant permit price p. If a k-type firm is noncompliant, its emissions exceed the number of permits it holds and the magnitude of its violation is $v^k = e^k - l^k > 0$. If a firm is compliant, $e^k - l^k \leq 0$ and $v^k = 0$.

We allow the probability with which a firm is audited (monitoring) and the commitment to penalize noncompliance (enforcement) to vary among firm-types, but not among firms of the same type. Suppose that each k-type firm is audited with constant probability π^k. We have in mind here that the enforcement authority commits to auditing $\bar{n}^k < n^k$ firms of type k at random so that $\pi^k = \bar{n}^k / n^k$. If a firm is found to be in violation, a penalty $f(v^k, \phi^k)$ is imposed by a judiciary. Following Garvie and Keeler [5], ϕ^k denotes a commitment of resources that allows the authority to bring enforcement pressure to bear on a k-type firm if necessary. Enforcement effort may include notification of a violation, bringing a civil suit against a noncompliant firm, or even developing and prosecuting a criminal case against executives of noncompliant firms. We restrict the analysis by assuming throughout that $\phi^k > 0$ for each k. Given positive enforcement expenditures, the penalty for a zero violation is zero but the marginal penalty for a zero violation is greater than zero [$f(0, \phi^k) = 0$ and $f_v(0, \phi^k) > 0$]. Assume that the penalty is increasing at an increasing rate in the level of the violation [$f_v(v^k, \phi^k) > 0$ and $f_{vv}(v^k, \phi^k) > 0$], and that the penalty and the marginal penalty are both increasing in the enforcement commitment [$f_\phi(v^k, \phi^k) > 0$ and $f_{v\phi}(v^k, \phi^k) > 0$].

Although our characterization of the firm's expected compliance costs is quite similar to those of Malik ([9, 10]) and Keeler [8], there are some important differences. None of these works explicitly distinguish between monitoring and enforcement. Malik [9] allows for nonneutral attitudes toward risk and assumes audit probabilities that depend on the firms' choices of emissions and permit-holdings, while we assume risk-neutral firms and, since audits of firms of a particular type are random, our audit probabilities do not depend a priori on the firms'

[3] In addition, a firm's control costs may be affected by regulations that are not directly related to emissions control. For example, a number of authors have noted that state public utility commission regulations may have significant impacts on the behavior and control costs of firms that trade sulfur dioxide allowances under Title IV of the 1990 Clean Air Act Amendments ([2–4]).

choices. Keeler [8] assumes that firms are risk-neutral and that expected penalties are constant for all firms, indicating implicitly that monitoring and enforcement efforts are constant across firms. He also considers the implications of expected penalties that are increasing but at constant, decreasing, and increasing rates in the level of violation. Expected penalties in our model may vary across firm-types and we choose to focus on the case in which they are increasing at an increasing rate in a firm's violation. Malik [10] assumes risk-neutral firms, audit probabilities that vary among firms, and a penalty function that is increasing and convex in the size of a violation; hence, at least on these grounds, his model is most similar to ours. It bears repeating though that we deal explicitly with the problem of designing a monitoring and enforcement program for a transferable emissions permit system, while these others do not.[4]

2.2. *A Firm's Choices of Emissions, Permits, and Violation*

At the time the firms make their choices, the enforcement authority has committed itself to a type-specific monitoring and enforcement program. We assume that each firm chooses positive emissions and permits, and never overcomplies. Then, a k-type firm's problem is to choose emissions and permits to

$$\begin{aligned} &\min c(e^k, \alpha^k) + p(l^k - l_0^k) + \pi^k f(e^k - l^k, \phi^k), \\ &\text{s.t. } e^k - l^k \geq 0. \end{aligned} \tag{1}$$

The Lagrange equation for (1) is $\theta^k = c(e^k, \alpha^k) + p(l^k - l_0^k) + \pi^k f(e^k - l^k, \phi^k) - \eta^k(e^k - l^k)$ and the Kuhn–Tucker conditions are

$$\theta_e^k = c_e(e^k, \alpha^k) + \pi^k f_v(e^k - l^k, \phi^k) - \eta^k = 0; \tag{2a}$$

$$\theta_l^k = p - \pi^k f_v(e^k - l^k, \phi^k) + \eta^k = 0; \tag{2b}$$

$$\theta_\eta^k = l^k - e^k \geq 0, \qquad \eta^k \geq 0, \qquad \eta^k \times (l^k - e^k) = 0. \tag{2c}$$

Given our assumptions about abatement costs and the penalty schedule, (2a)–(2c) are necessary and sufficient to determine the firm's optimal choices of emissions and permits uniquely.

Whether a k-type firm is compliant or not, it chooses its emissions so that the price of a permit is equal to its marginal abatement cost; that is, its emissions-choice rule is

$$e^k(\alpha^k, p) = \{e^k \mid c_e(e^k, \alpha^k) + p = 0\}. \tag{3}$$

[4] Two other papers on monitoring and enforcement of transferable permit systems deserve mention. The first paper in this literature appears to be Beavis and Walker [1]. They considered noncompliant behavior when emissions are random (emissions in our model are deterministic) and they characterized a uniform monitoring program to achieve an aggregate emissions target in a cost-effective manner. vanEgteren and Weber [15] examined noncompliance when the aggregate issuance of permits and monitoring are fixed and one firm exercises power in the permit market. Their primary result is that when market power is present, the initial distribution of permits may be used as an implicit enforcement mechanism. In our model, the permit market is competitive so that all choices are independent of the initial distribution of permits.

To establish this result, simply combine (2a) and (2b) to obtain $c_e(e^k, \alpha^k) + p = 0$. Note that, in equilibrium, the firms' marginal abatement costs are equal.

A simple inspection of (3) reveals:

RESULT 1. *A k-type firm's choice of emissions is independent of the probability that it will be audited π^k and the enforcement pressure applied to it ϕ^k.*

This is consistent with an observation by Malik ([9, p. 101]), who noted that when the probability with which a firm is audited is constant as in the case of random audits, a firm's choice of emissions is independent of the probability that it will be audited. [Harford [7] derives a similar conclusion in the case of an emissions tax]. Although not surprising, our result notes that this independence extends to enforcement pressure as well. As Malik notes, and we repeat here, this does not imply that the equilibrium distribution of emissions among the firms is independent of the choice of monitoring and enforcement policy. There will be an indirect effect on equilibrium emissions of monitoring and enforcement because the choice of a monitoring and enforcement strategy will affect the equilibrium permit price, and this in turn will affect equilibrium choices of emissions.

Turn now to the firm's demand for emissions permits. When it is compliant the number of permits it demands is simply equal to its choice of emissions; that is, $l^k(\alpha^k, p) = e^k(\alpha^k, p)$. When the firm is noncompliant its demand for emissions permits is

$$l^k(\alpha^k, \pi^k, \phi^k, p) = \{l^k \mid p - \pi^k f_v(e^k(\alpha^k, p) - l^k, \phi^k) = 0\}. \qquad (4)$$

To obtain (4) note from (2b) and (2c) that $e^k - l^k > 0$ implies $\eta^k = 0$ and $\theta_l^k = p - \pi^k f_v(e^k - l^k, \phi^k) = 0$. Substitution of the firm's choice of emissions $e^k(\alpha^k, p)$ into $p - \pi^k f_v(e^k - l^k, \phi^k) = 0$ yields (4). Note that although a noncompliant firm's choice of emissions is not directly affected by the monitoring and enforcement effort applied to it, its demand for emissions permits is.

If all firms are noncompliant, (4) implies that each firm will equate its marginal expected penalty to the permit price. Since each firm also equates its marginal abatement cost to the permit price, marginal abatement costs and marginal expected penalties are equal to each other and equal across firms. We see in Section 3 that the equilibrating nature of emissions permit markets has an important implication for the ability of an enforcement authority to exploit fundamental differences among the firms.

Having specified a firm's choice of emissions and its demand for permits, we can now turn to its choice of violation. We start with its choice of whether to be compliant or not: A k-type firm is compliant if and only if

$$p - \pi^k f_v(0, \phi^k) \leq 0.\text{[5]} \qquad (5)$$

Although this result is not new, one aspect of it has been overlooked; namely, (5) does not depend on α^k. A firm's decision to be compliant or not depends only on

[5] Malik [10] and vanEgteren and Weber [15] assert that (5) is sufficient to guarantee compliance when $f_{vv}(v^k, \phi^k) \geq 0$. When $f_{vv}(v^k, \phi^k) > 0$, (5) is also necessary. To show that $e^k - l^k = 0$ only if (5) holds, assume toward a contradiction that $e^k - l^k = 0$ while $p - \pi^k f_v(0, \phi^k) > 0$. Then, since $\eta^k \geq 0$ by (2c), $p - \pi^k f_v(0, \phi^k) > 0$ implies $p - \pi^k f_v(0, \phi^k) + \eta^k > 0$. But (2b) requires that if $e^k - l^k = 0$ is an optimal choice, $p - \pi^k f_v(0, \phi^k) + \eta^k = 0$; hence, a contradiction.

the relationship between the permit price and the marginal expected penalty of a vanishingly slight violation, not on parametric characteristics of its emissions-control costs. In fact, this independence extends further.

RESULT 2. *A k-type firm's choice of violation, including whether it is compliant or not, is independent of its exogenous characteristics* α^k.

Proof of Result 2. The result that a firm's choice between a zero violation (compliance) and a positive violation (noncompliance) is independent of α^k is immediately obvious from (5). Therefore, we need only consider the effect on the firm's choice of a positive violation of a change in some element of α^k. When the firm is noncompliant, (2b) and (2c) require $p - \pi^k f_v(e^k - l^k, \phi^k) = 0$. Taking account of the firm's choice of emissions (3) and its choice of permits (4) we have

$$p - \pi^k f_v\big(e^k(\alpha^k, p) - l^k(\alpha^k, \pi^k, \phi^k, p), \phi^k\big) \equiv 0.$$

Differentiate this with respect to some element of α^k to obtain $-\pi^k f_{vv} \times (e_\alpha^k - l_\alpha^k) = 0$, which implies $v_\alpha^k = e_\alpha^k - l_\alpha^k = 0$. Q.E.D.

Result 2 is rather surprising because it reveals that, holding monitoring, enforcement and the permit price constant, a change in some parameter that affects the abatement costs of a firm has no effect on its choice of violation. To illustrate, suppose that a firm adopts a cleaner production process. This lowers its marginal abatement costs (since $-c_{e\alpha}(e^k, \alpha^k) < 0$) so it is motivated to reduce its emissions ($e_\alpha^k < 0$) because it is now cheaper to do so. What is unexpected is that the firm is also motivated to sell the corresponding number of permits ($l_\alpha^k = e_\alpha^k$) so that its level of violation remains unchanged. The intuition behind this result is as follows: The marginal expected benefit to a firm of a marginal reduction in its violation is the marginal expected penalty it avoids, which clearly does not depend on the firm's characteristics. To reduce its violation it may purchase the legal right to emit, the marginal cost of which is the equilibrium permit price, or it may reduce its emissions, the marginal cost of which is $-c_e(e^k, \alpha^k)$. But, the firm always chooses its emissions to equate its marginal abatement costs to the price of an emissions permit [see (3)]. Hence, the marginal cost of reducing its violation is simply equal to the permit price, and hence, independent of the firm's characteristics. Since the marginal costs and benefits to a firm of reducing its violation are both independent of its exogenous characteristics, so too is its choice of violation.

Result 2 suggests that a difference in the violations of any two types of firms is independent of differences in their exogenous characteristics. Thus, if two firms are audited with the same probability and the same enforcement effort is applied to each, they both should have the same level of violation even though one may employ a less advanced emissions-control technology or use a dirtier production process. The policy significance of this is that if a regulatory authority is to consider why some firms cheat more than others, the answer likely lies in differences in monitoring and enforcement, not in the firms' fundamental differences. As with Result 1, we note that Result 2 does not imply that equilibrium violations are independent of the firms' exogenous characteristics. Again, there is an indirect price effect. The firms' exogenous characteristics affect aggregate demand for emissions permits, and hence, the equilibrium permit price, which in turn affects equilibrium violations.

TABLE I
Comparative Statics of a Firm's Choices of Emissions, Permits, and Level of Violation

	Emissions (e^k)	Permits (l^k)	Violation (v^k)
α^k	$e^k_\alpha = \dfrac{-c_{e\alpha}(e^k, \alpha^k)}{c_{ee}(e^k, \alpha^k)}$	$e^k_\alpha = l^k_\alpha$	$v^k_\alpha = 0$
π^k	$e^k_\pi = 0$	$l^k_\pi = \dfrac{f_v(v^k, \phi^k)}{\pi^k f_{vv}(v^k, \phi^k)} > 0$	$v^k_\pi = -l^k_\pi < 0$
ϕ^k	$e^k_\phi = 0$	$l^k_\phi = \dfrac{f_{v\phi}(v^k, \phi^k)}{f_{vv}(v^k, \phi^k)} > 0$	$v^k_\phi = -l^k_\phi < 0$
p	$e^k_p = \dfrac{-1}{c_{ee}(e^k, \alpha^k)} < 0$	$l^k_p = e^k_p - \dfrac{1}{\pi^k f_{vv}(v^k, \phi^k)} < 0$	$v^k_p = \dfrac{1}{\pi^k f_{vv}(v^k, \phi^k)} > 0$

A noncompliant firm's choice of violation depends only on the monitoring and enforcement effort applied to it and the emissions permit price. Thus, using (3) and (4), we write the choice of violation of a noncompliant firm as

$$v^k(\pi^k, \phi^k, p) = e^k(\alpha^k, p) - l^k(\alpha^k, \pi^k, \phi^k, p). \tag{6}$$

The marginal impacts of α^k, π^k, ϕ^k, and p on the choices of emissions, permits, and violation of a noncompliant k-type firm are presented in Table I.[6] We have already discussed the marginal impacts of a change in some element of α^k on a k-type firm's choices. We have also noted that a firm's choice of emissions is independent of the monitoring and enforcement effort applied to it (Result 1); therefore, its choice of violation is affected by monitoring and enforcement only through induced changes in the number of permits it chooses to hold. For example, if a k-type firm faces a higher audit probability, then noncompliance is a relatively less attractive strategy. Hence, it is motivated to reduce its violation ($v^k_\pi < 0$) by purchasing more permits ($l^k_\pi > 0$), not by reducing its emissions ($e^k_\pi = 0$). The same qualitative effects occur when the regulator has committed itself to greater enforcement effort. A higher permit price implies that purchasing the legal right to emit is a relatively less attractive option than reducing emissions, so a firm is motivated to hold fewer permits and reduce its emissions ($e^k_p < 0$ and $l^k_p < 0$). In addition, a higher permit price makes noncompliance a relatively more attractive option so that a firm is motivated to increase its violation ($v^k_p > 0$).

3. DISTRIBUTING EFFORT AMONG NONCOMPLIANT FIRMS

In this section we endogenize the allocation of monitoring and enforcement effort among the firms. Our regulatory-choice model is the same as that of Garvie and Keeler [5], except their model is applied to the enforcement of emissions standards. Following their approach we assume that the enforcement authority has a fixed budget with which it chooses a type-specific monitoring and enforcement program to minimize aggregate violations, and we simplify the analysis by assuming that all firms are noncompliant.

[6] To conserve space we have omitted the derivations of these comparative statics. They are available upon request.

3.1. *Permit Market Equilibrium When Firms Are Noncompliant*

To choose an effective monitoring and enforcement program, the enforcement authority must take into account how its policy affects the equilibrium in the emissions permit market. Let $\alpha = (\alpha^k)_{k \in K}$, $\pi = (\pi^k)_{k \in K}$, and $\phi = (\phi^k)_{k \in K}$. Assume that a total of L permits have been issued to the firms and the enforcement authority has committed itself to a type-specific monitoring and enforcement program $[\pi, \phi]$. Then, the equilibrium permit price when all firms are noncompliant is $\bar{p} = \bar{p}(\alpha, \pi, \phi, L)$, which from (4) must satisfy

$$\sum n^k l^k(\alpha^k, \pi^k, \phi^k, \bar{p}) \equiv L. \tag{7}$$

(Summations throughout are taken over the entire set K). Differentiate (7) with respect to π^h and ϕ^h and rearrange the results to obtain

$$\partial\bar{p}/\partial\pi^h = -n^h l_\pi^h(\alpha^h, \pi^h, \phi^h, \bar{p}) \Big/ \sum n^k l_p^k(\alpha^k, \pi^k, \phi^k, \bar{p}) > 0; \tag{8a}$$

$$\partial\bar{p}/\partial\phi^h = -n^h l_\phi^h(\alpha^h, \pi^h, \phi^h, \bar{p}) \Big/ \sum n^k l_p^k(\alpha^k, \pi^k, \phi^k, \bar{p}) > 0. \tag{8b}$$

The signs of (8a)–(8b) follow from $l_\pi^h > 0$, $l_\phi^h > 0$, and $l_p^k < 0$ (refer to Table I). Intuitively, increased monitoring of noncompliant firms of a particular type motivates them to purchase more emissions permits ($l_\pi^h > 0$) to reduce the magnitude of their violations ($v_\pi^h < 0$). This increased demand for permits then puts upward pressure on the equilibrium permit price. A similar effect occurs if the regulator directs greater enforcement expenditures at h-type firms.

3.2. *Endogenous Monitoring and Enforcement*

Suppose that the enforcement authority has a budget B, with which it chooses a type-specific monitoring and enforcement program to minimize aggregate noncompliance. Recall that the probability with which a k-type firm is audited is $\pi^k = \bar{n}^k/n^k$, where $\bar{n}^k$ is the number of random audits the regulator conducts on k-type firms and n^k is the number of these firms. Suppose that the cost of conducting an audit of any firm is a constant w. Then, the cost of establishing the audit probability π^k for k-type firms is $w\bar{n}^k = wn^k\bar{n}^k/n^k = wn^k\pi^k$. The authority backs up its audit of a k-type firm with a commitment of ϕ^k that allows it to bring enforcement pressure to bear on the firm if necessary. Since the authority audits $\bar{n}^k$ k-type firms, the cost of establishing its enforcement commitment for all of these firms is $\phi^k\bar{n}^k = \phi^k n^k \bar{n}^k/n^k = \phi^k n^k \pi^k$. Bringing all the components together, the enforcement authority's budget constraint is

$$B \geq w \sum n^k \pi^k + \sum \phi^k n^k \pi^k. \tag{9}$$

Note that the per-firm marginal cost of establishing the audit probability π^k for k-type firms is $w + \phi^k$, which consists of the audit cost and the resource commitment to punish the violation of one of these firms. The per-firm marginal cost of establishing the enforcement commitment ϕ^k for k-type firms is simply π^k.

With its limited resources the enforcement authority seeks to minimize aggregate equilibrium violations

$$\sum n^k v^k(\pi^k, \phi^k, \bar{p}), \tag{10}$$

where $v^k(\pi^k, \phi^k, \bar{p})$ is defined by (6) and $\bar{p} = \bar{p}(\alpha, \pi, \phi, L)$ is defined by (7).

An optimal monitoring and enforcement program, which we denote as $[\pi_*, \phi_*] = [(\pi^k_*)_{k\in K}, (\phi^k_*)_{k\in K}]$ minimizes (10) subject to (9), $\pi^k \in [0, 1]$ and $\phi^k \geq 0$, $\forall k \in K$.[7] Like Garvie and Keeler, we simplify the analysis further by restricting our attention to interior solutions that exhaust the enforcement budget. The Lagrange equation associated with the enforcement authority's design problem is then

$$\Lambda = \sum n^k v^k(\pi^k, \phi^k, \bar{p}) + \lambda\left[w \sum n^k \pi^k + \sum n^k \pi^k \phi^k - B\right].$$

At $[\pi_*, \phi_*]$, the following first-order conditions must hold

$$\partial\Lambda_*/\partial\pi^h = n^h v^h_\pi(\pi^h_*, \phi^h_*, \bar{p}_*) + \sum n^k v^k_p(\pi^k_*, \phi^k_*, \bar{p}_*)\left[\partial\bar{p}_*/\partial\pi^h\right] + \lambda_* n^h(w + \phi^h_*) = 0, \quad h \in K; \tag{11a}$$

$$\partial\Lambda_*/\partial\phi^h = n^h v^h_\phi(\pi^h_*, \phi^h_*, \bar{p}_*) + \sum n^k v^k_p(\pi^k_*, \phi^k_*, \bar{p}_*)\left[\partial\bar{p}_*/\partial\phi^h\right] + \lambda_* n^h \pi^h_* = 0, \quad h \in K; \tag{11b}$$

$$\partial\Lambda_*/\partial\lambda = w \sum n^k \pi^k_* + \sum n^k \pi^k_* \phi^k_* - B = 0, \qquad \lambda_* > 0. \tag{11c}$$

In (11a)–(11b), $\bar{p}_* = \bar{p}(\alpha, \pi_*, \phi_*, L)$ is the equilibrium permit price under $[\pi_*, \phi_*]$.

Given that the enforcement authority's budget is just exhausted, the rules for distributing monitoring and enforcement among the firms are quite straightforward.

RESULT 3. *The optimal distribution of monitoring and enforcement* $[\pi_*, \phi_*]$ *must satisfy*

$$\frac{v^h_\pi(\pi^h_*, \phi^h_*, \bar{p}_*)}{v^j_\pi(\pi^j_*, \phi^j_*, \bar{p}_*)} = \frac{w + \phi^h_*}{w + \phi^j_*} \qquad \forall h, j \in K; \tag{12a}$$

$$\frac{v^h_\phi(\pi^h_*, \phi^h_*, \bar{p}_*)}{v^j_\phi(\pi^j_*, \phi^j_*, \bar{p}_*)} = \frac{\pi^h_*}{\pi^j_*} \qquad \forall h, j \in K. \tag{12b}$$

Proof Result 3. Substitute for $\partial\bar{p}_*/\partial\pi^h$ from (8a) to rewrite (11a) as

$$\frac{v^h_\pi(\pi^h_*, \phi^h_*, \bar{p}_*) - l^h_\pi(\alpha^h, \pi^h_*, \phi^h_*, \bar{p}_*) \times A_*}{w + \phi^h_*} = -\lambda_*, \tag{13}$$

[7] The enforcement authority's design problem may be very complex, particularly because its constraint set is not convex. See Garvie and Keeler's Fig. 1 for an illustration of this nonconvexity.

where

$$A_* = \frac{\sum n^k v_p^k(\pi_*^k, \phi_*^k, \bar{p}_*)}{\sum n^k l_p^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*)}.$$

To confirm the appropriate function arguments for l_p^k, l_π^k, and v_π^k, $\forall k \in K$, see (4) and (6). Now, for any two types of firms, say h and j, (13) implies

$$\frac{v_\pi^h(\pi_*^h, \phi_*^h, \bar{p}_*) - l_\pi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*) \times A_*}{v_\pi^j(\pi_*^j, \phi_*^j, \bar{p}_*) - l_\pi^j(\alpha^j, \pi_*^j, \phi_*^j, \bar{p}_*) \times A_*} = \frac{w + \phi_*^h}{w + \phi_*^j}. \tag{14}$$

From Table I,

$$v_\pi^k(\pi_*^k, \phi_*^k, \bar{p}_*) = -l_\pi^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*) \qquad \forall k \in K. \tag{15}$$

Substitution of (15) into (14) yields (12a).

Now substitute for $\partial \bar{p}_* / \partial \phi^h$ from (8b) to rewrite (11b) as

$$\frac{v_\phi^h(\pi_*^h, \phi_*^h, \bar{p}_*) - l_\phi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*) \times A_*}{\pi_*^h} = -\lambda_*. \tag{16}$$

For any two types of firms, h and j, (16) implies

$$\frac{v_\phi^h(\pi_*^h, \phi_*^h, \bar{p}_*) - l_\phi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*) \times A_*}{v_\phi^j(\pi_*^j, \phi_*^j, \bar{p}_*) - l_\phi^j(\alpha^j, \pi_*^j, \phi_*^j, \bar{p}_*) \times A_*} = \frac{\pi_*^h}{\pi_*^j}. \tag{17}$$

From Table I,

$$v_\phi^k(\pi_*^k, \phi_*^k, \bar{p}_*) = -l_\phi^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*) \qquad \forall k \in K, \tag{18}$$

which upon substitution into (17) yields (12b). Q.E.D.

The conditions given by (12a) are derived directly from (11a); hence, they provide necessary conditions for allocating monitoring effort between every pair of firm types. They indicate that the ratio of the direct marginal impacts of monitoring on the violations of any two types of firms must equal the ratio of the per-firm marginal costs of establishing the audit probabilities for these firms. By "direct marginal impacts" we mean the marginal impacts of monitoring on the firms' violations when the permit price is held constant at $\bar{p}_*$. Similarly, the conditions given by (12b) are derived from (11b), and provide necessary conditions for allocating enforcement resources between every pair of firms. They indicate that the ratio of the direct marginal impacts of enforcement on the violations of any two types of firms must equal the ratio of the per-firm marginal costs of establishing the enforcement commitment for these firms.

Proposition 1 follows immediately from (12a)–(12b):

PROPOSITION 1. *In an optimal monitoring and enforcement program, differences in the monitoring and enforcement effort applied to any two types of firms are independent of the differences in the exogenous characteristics of these firms.*

Proof of Proposition 1. Since differences between $v_*^h = v^h(\pi_*^h, \phi_*^h, \bar{p}_*)$ and $v_*^j = v^j(\pi_*^j, \phi_*^j, \bar{p}_*)$ are independent of differences between α^h and α^j (recall Result 2), the proposition follows from a simple inspection of (12a)–(12b). Q.E.D.

Proposition 1 is the primary result of our analysis and we discuss its policy-significance at some length. Before we do, however, let us note that the proposition stems from the two independence results that we derived in Section 2. Clearly, Proposition 1 requires Result 2—each firm's choice of violation must be independent of its exogenous characteristics. What is not so clear is that Proposition 1 also requires Result 1—each firm's choice of emissions must be independent of the monitoring and enforcement effort applied to it.

Contrary to Result 1, suppose instead that each firm's choice of emissions is not independent of the monitoring and enforcement effort applied to it. Then we would write a k-type firm's choice of emissions as $e^k(\alpha^k, \pi^k, \phi^k, p)$ instead of as (3). Doing so does not affect the essential structure of the enforcement authority's design problem, and it does not affect the first-order conditions (11a)–(11b). In addition, the marginal impacts of monitoring and enforcement effort on the equilibrium permit price (8a)–(8b) would remain the same. Therefore, Eqs. (14) and (17) in the Proof of Result 3 would also remain the same. Note in the Proof of Result 3 that (12a) is derived from (14) using (15), and (12b) is derived from (17) using (18). Now, if each firm's choice of emissions depended on the monitoring and enforcement effort applied to it, (15) and (18) would instead be

$$v_\pi^k(\pi_*^k, \phi_*^k, \bar{p}_*) = e_\pi^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*) - l_\pi^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*) \qquad \forall k \in K, \tag{15'}$$

and

$$v_\phi^k(\pi_*^k, \phi_*^k, \bar{p}_*) = e_\phi^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*) - l_\phi^k(\alpha^k, \pi_*^k, \phi_*^k, \bar{p}_*) \qquad \forall k \in K. \tag{18'}$$

Now, use (15′) to substitute for $l_\pi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*)$ and $l_\pi^j(\alpha^j, \pi_*^j, \phi_*^j, \bar{p}_*)$ in (14) and use (18′) to substitute for $l_\phi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*)$ and $l_\phi^j(\alpha_*^j, \pi_*^j, \phi_*^j, \bar{p}_*)$ in (17) to obtain

$$\frac{v_\pi^h(\pi_*^h, \phi_*^h, \bar{p}_*) \times [1 + A_*] - e_\pi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*) \times A_*}{v_\pi^j(\pi_*^j, \phi_*^j, \bar{p}_*) \times [1 + A_*] - e_\pi^j(\alpha^j, \pi_*^j, \phi_*^j, \bar{p}_*) \times A_*} = \frac{w + \phi_*^h}{w + \phi_*^j}, \tag{14'}$$

and

$$\frac{v_\phi^h(\pi_*^h, \phi_*^h, \bar{p}_*) \times [1 + A_*] - e_\phi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*) \times A_*}{v_\phi^j(\pi_*^j, \phi_*^j, \bar{p}_*) \times [1 + A_*] - e_\phi^j(\alpha^j, \pi_*^j, \phi_*^j, \bar{p}_*) \times A_*} = \frac{\pi_*^h}{\pi_*^j}. \tag{17'}$$

Equation (14′) clearly indicates that differences in the optimal monitoring of h- and j-type firms would depend on differences between $e_\pi^h(\alpha^h, \pi_*^h, \phi_*^h, \bar{p}_*)$ and $e_\pi^j(\alpha^j, \pi_*^j, \phi_*^j, \bar{p}_*)$, which in turn would depend on differences in the parametric characteristics of h- and j-type firms, α^h and α^j. In a similar fashion, Eq. (17′) indicates that differences in the optimal enforcement effort directed at h- and

j-type firms would also depend on differences in the exogenous characteristics of h- and j-type firms. Thus, Proposition 1 would not hold if, contrary to Result 1, each firm's choice of emissions depended on the monitoring and enforcement pressure applied to it.

4. POLICY IMPLICATIONS

Proposition 1 has a very strong policy implication: in a transferable permit system, an enforcement authority is not to use parametric differences among regulated firms to guide its decisions about distributing monitoring and enforcement efforts among them. We have shown that this conclusion depends critically on two independence results concerning firms' choices of emissions and violations: (1) each firm's choice of emissions is independent of the monitoring and enforcement effort applied to it, and (2) each firm's choice of violation is independent of its exogenous characteristics. With appropriate data on emissions, violations, firm characteristics, and key enforcement parameters, these hypotheses can be subjected to rigorous econometric tests. However, we doubt that these tests can be conducted at the present time. The primary problem is that the current generation of market-based initiatives have been in existence for only a short time, and thus far have achieved perfect, or very close to perfect, compliance rates.[8] There is simply not enough variation in compliance rates as yet to conduct adequate tests of these independence results.

Conducting these tests in an experimental setting is a viable second best option. Experimental tests of other aspects of the design of transferable permit systems have been conducted (for an example, see [6]), but none have focused on the design of enforcement strategies. In fact, we are not aware of any experimental analyses of enforcement strategies for environmental policies. In the absence of the appropriate data for econometric analysis, tests of our fundamental independence results in an experimental setting seems to us to be a natural way to obtain initial information about the empirical validity of out primary policy recommendation.

If empirically valid, Proposition 1 implies that there is no point to applying more intense monitoring and enforcement efforts to firms that employ less advanced emissions-control technologies, or that use dirtier production processes, or that differ in any other fundamental way. The inability of an enforcement authority to exploit exogenous differences among regulated firms to target its monitoring and enforcement effort is due to the equilibrating nature of the permit market. Recall from Section 2 that each firm chooses its emissions so that its marginal abatement cost is equal to the permit price, and it chooses the number of permits to hold so that its marginal expected penalty is also equal to the permit price. Since all firms face the same permit price, marginal abatement costs and marginal expected penalties are equal to each other and equal across firms. This will be true for any strategy the enforcement authority may choose (at least all those strategies that induce interior outcomes). Therefore, exogenous differences in abatement costs

[8]Since its inception, firms in the Sulfur Dioxide Allowance Trading program have been perfectly compliant ([12–14]). There have been a small number of violations in the Regional Economic Incentives Market program (RECLAIM) of the Los Angeles air basin, but program authorities do not attribute these violations to deliberate attempts by firms to evade their legal obligations. Rather, they seem to be the result of a few firms that lacked sufficient understanding of the required protocols [11].

cannot be exploited to enhance the productivity of a fixed enforcement budget because, in a permit market equilibrium, all marginal abatement costs are equal regardless of the enforcement authority's strategy.[9]

The inability of an enforcement authority to exploit exogenous differences among firms has important implications for the value of firm-level information in the design of a monitoring and enforcement strategy. Recall that if uniform monitoring and enforcement is applied to the firms in our model, their violations will be the same. Therefore, it is clear that a uniform monitoring and enforcement strategy will satisfy (12a)–(12b), and if it just exhausts the enforcement budget it will be an optimal program. Furthermore, a uniform monitoring and enforcement strategy suggests that knowledge of the firms' abatement costs is not relevant in the design of the monitoring and enforcement strategy. Thus, the asymmetric information problem of firms having better information about their control costs than regulatory authorities may have no bearing on an enforcement authority's choice of strategy in a transferable permit system.

However, it is clear that the recommendation of uniform monitoring and enforcement depends on the uniformity of key monitoring and enforcement parameters. If the costs of conducting audits and applying enforcement effort vary across firms, or if there are parametric differences in the penalty function that are due, for example, to regional differences in the severity of penalties imposed by courts, a uniform monitoring and enforcement strategy will not be appropriate. When a differentiated strategy is called for because of differences in monitoring and enforcement parameters, knowledge of the firms' exogenous characteristics becomes useful because these characteristics affect the optimal monitoring and enforcement program through their impacts on the equilibrium permit price. This is seen clearly in (12a)–(12b) where the firms' characteristics show up in the specification of the equilibrium permit price. In fact, determining a strategy when uniformity is not optimal requires the enforcement authority to forecast equilibrium violations. This requires a forecast of the equilibrium permit price and this requires information about the firms' exogenous characteristics.

To conclude this section, let us contrast Proposition 1 to an analogous result of Garvie and Keeler [5]. With the same regulatory-choice model but applied to firms that face emissions standards, they found that greater monitoring and enforcement effort should be directed at firms with parametrically higher marginal abatement costs. We, of course, have shown that when firms operate under a transferable emissions permit system, the distribution of optimal monitoring and enforcement is independent of exogenous differences in their abatement costs. From the asymmetric information perspective, given that audit and enforcement costs and penalty structures do not vary across firms, knowledge of firms' abatement costs is useful information in designing a compliance-maximizing enforcement strategy for emis-

[9] This issue is complicated somewhat by the fact that the enforcement authority's strategy impacts the equilibrium permit price. If monitoring and enforcement of different types of firms have differential impacts on the permit price, the enforcement authority may be able to manipulate the permit price to attain higher compliance rates. Of course, Proposition 1 suggests that it cannot. More directly, it can be shown that the differential marginal impacts of monitoring of any two types of firms (i.e., $\partial \bar{p}/\partial \pi^h - \partial \bar{p}/\partial \pi^j, \forall h \neq j$) are independent of exogenous differences of the firms. The same is true of the differential marginal impacts of enforcement pressure. Thus, even from the perspective of manipulating the equilibrium permit price to make its resources more productive, the enforcement authority is unable to exploit fundamental differences among the firms to do so.

sions standards, but this information is not relevant in the context of transferable permits. The reasons for the difference between Garvie and Keeler's result and our Proposition 1 are nearly immediate. When a firm faces an emissions standard, its choice of emissions is not independent of the monitoring and enforcement effort applied to it and its choice of violation is not independent of its exogenous characteristics; that is, our Results 1 and 2 do not apply when firms face emissions standards.

5. CONCLUSION

Market-based approaches to environmental policy have gained many adherents. Theoretical and empirical work has indicated that environmental quality standards can be achieved at much lower cost with a market-based approach than with traditional command-and-control regulations, primarily because the former provides sources of pollution more flexibility in choosing the manner in which they will control their emissions. However a critical component of such policies has not been adequately addressed; namely, how should compliance to such policies be enforced? Without enforcement strategies that are designed to achieve acceptable levels of compliance in a cost-effective manner, the efficiency-gains that can be achieved by market-based systems may not materialize.

With a conventional model of noncompliant firms in a transferable emissions permit system, we have addressed a fundamental policy question about monitoring and enforcement of these systems: How should a budget-constrained enforcement authority distribute monitoring and enforcement resources among heterogeneous noncompliant firms? And we have obtained a surprising answer: the optimal distribution of monitoring and enforcement is based on straightforward rules that indicate that differences in the application of these efforts among heterogeneous firms should be independent of differences in the firms' exogenous characteristics. Thus, the justification for applying different monitoring and enforcement pressure to different firms must come from differences in key enforcement parameters, not from fundamental differences of the firms.

REFERENCES

1. B. Beavis and M. Walker, Random wastes, imperfect monitoring and environmental quality standards, *J. Public Econom.* **21**, 377–387 (1983).
2. M. Bernstein, A. Farrell, and J. Winebrake, The impact of restricting the SO_2 allowance market, *Energy Policy* **22**, No. 9, 748–754 (1994).
3. J. Coggins and V. Smith, Some welfare effects of emissions allowance trading in a twice-regulated industry, *J. Environ. Econom. Management* **25**, 275–297 (1993).
4. D. Fullerton, S. McDermott, and J. Caulkins, Sulfur dioxide compliance of a regulated utility, *J. Environ. Econom. Management* **34**, 32–53 (1997).
5. D. Garvie and A. Keeler, Incomplete enforcement with endogenous regulatory choice, *J. Public Econom.* **55**, 141–162 (1994).
6. R. Godby, S. Mestelman, R. A. Muller, and J. D. Welland, Emissions trading with shares and coupons when control over discharges is uncertain, *J. Environ. Econom. Management* **32**, No. 3, 359–381 (1997).
7. J. D. Harford, Firm behavior under imperfectly enforceable pollution standards and taxes, *J. Environ. Econom. Management* **5**, 26–43 (1978).

8. A. Keeler, Noncompliant firms in transferable discharge permit markets: Some extensions, *J. Environ. Econom. Management* **21**, 180–189 (1991).
9. A. Malik, Markets for pollution control when firms are noncompliant, *J. Environ. Econom. Management* **18**, 97–106 (1990).
10. A. Malik, Enforcement costs and the choice of policy instruments for pollution control, *Econom. Inquiry* **30**, 714–721 (1992).
11. South Coast Air Quality Management District, "RECLAIM Program Three-Year Audit and Progress Report," South Coast Air Quality Management District, Diamond Bar, CA (1998).
12. U.S. Environmental Protection Agency, "Acid Rain Program 1995 Compliance Report," U.S. EPA Acid Rain Program, Washington DC (1996).
13. U.S. Environmental Protection Agency, "Acid Rain Program 1996 Compliance Report," U.S. EPA Acid Rain Program, Washington DC (1997).
14. U.S. Environmental Protection Agency, "Acid Rain Program 1997 Compliance Report," U.S. EPA Acid Rain Program, Washington DC (1998).
15. H. vanEgteren and M. Weber, Marketable permits, market power, and cheating, *J. Environ. Econom. Management* **30**, 161–173 (1996).