UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 608

The Harrington Paradox Squared

by

Jessica Coria and Xiao-Bing Zhang

January 2015

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

EFMD EQUIS ACCREDITED

The Harrington Paradox Squared*

Jessica Coria[†] and Xiao-Bing Zhang[‡]

January 2015

Abstract

Harrington (1988) shows that state-dependent enforcement based on past compliance records provides an explanation to the seemingly contradictory observation that firms' compliance with environmental regulations is high despite the fact that inspections occur infrequently and fines are rare and small. This result has been labeled in the literature as the "Harrington paradox." In this paper we propose an improved transition structure for the audit framework where targeting is based not only on firms' past compliance record but also on adoption of environmentally superior technologies. We show that this transition structure would not only foster the adoption of new technology but also increase deterrence by changing the composition of firms in the industry toward an increased fraction of cleaner firms that pollute and violate less.

Key Words: imperfect compliance, state-dependent targeted enforcement, technology adoption, emission standards.

JEL classification: L51, Q55, K31, K42.

---

*Research funding from the Swedish Research Council (FORMAS) is gratefully acknowledged. Xiao-Bing Zhang gratefully acknowledges funding from the Swedish International Development Cooperation Agency (SIDA)

[†]Corresponding author: Jessica Coria, Department of Economics, University of Gothenburg, P.O. Box 640, SE 405 30 Gothenburg, Sweden. Email: Jessica.Coria@economics.gu.se.

[‡]University of Gothenburg. Email: Xiao-Bing.Zhang@economics.gu.se.

1 Introduction

Technological change is the main force improving the trade-off between economic growth and environmental quality in the long run. Therefore, the effect of environmental policies on the development and spread of new technologies is among the most important determinants of success or failure of environmental protection efforts (Aldy and Stavins 2007). Yet, environmental policy instruments impose costs on polluters. When there is room for firms to untruthfully report emissions without being caught and fined, i.e., there is imperfect enforcement, environmental policies will have lower success in creating incentives for technological development and controlling the generation of pollution than when the monitoring probability and stringency of the fines are such that truthful reporting is induced. Unfortunately, in many circumstances, the frequent monitoring and relatively high fines necessary to deter firms from under-reporting emissions are not available due to lack of accurate monitoring technology, reluctance to use high penalties, and/or budget constraints.

Harrington (1988) shows that a regulator's enforcement can be made more efficient by dividing firms into two groups according to their past compliance record. Without increasing inspection rates or fines, the regulator can lower the incidence of non-compliance by concentrating surveillance resources on firms in one of the groups (the target group), punishing violations by exile into the target group and (once there) rewarding firms found in compliance by returning them to the non-target group. This scheme generates what Harrington refers to as "enforcement leverage." Since non-compliance triggers greater future scrutiny, the expected costs of non-compliance are beyond the avoidance of immediate fines. Thus, he shows that there exists an equilibrium where firms have an incentive to comply with regulations despite the fact that the cost of compliance in each period is greater than the expected penalty.

Harrington (1988) offers an explanation for the seemingly contradictory observation that compliance rates across most industries are quite high despite the fact that inspections occur infrequently and fines are rare and small, a result labeled in the literature as the "Harrington paradox." In the present paper we propose an improved transition structure for the audit framework where targeting is based not only on firms' past compliance record but also on adoption of environmentally superior technologies. We show that this transition structure would not only foster the adoption of the new technology but it would also increase deterrence by changing the composition of firms in the industry toward an increased fraction of cleaner firms that pollute and violate less.

Harrington's work initiated a substantial amount of theoretical work analyzing the robustness of the results to alternative specifications of information and compliance cost structures (see Harford 1991, Harford and Harrington 1991, and Raymond 1999), providing alternative explanations to the "paradox" (see, e.g., Heyes and Rickman 1999, Livernois and McKenna 1999, and Nyborg and Telle 2004 and 2006)[1], and testing the empirical validity of his predictions (see, e.g., Helland 1998, Clark et al. 2004, Cason and Langadharan 2006, and Gray and Shimshack 2011 for a review of the literature).

Like our study, some previous studies have suggested alternative targeting methods.[2] For instance, in Friesen (2003), firms move randomly into the target group but escape based on observed compliance behavior (Friesen 2003). Notably, Liu and Neilson (2009) and Gilpatric et al. (2011) propose tournament-based dynamic targeting mechanisms. In their setting, a fixed number of firms are selected for inspection and those with the highest emissions are targeted with higher inspection probability, which induces dynamic rank-order tournaments among inspected firms, where enforcement leverage is enhanced by a competition effect. Similarly, in our setting, firms with the highest emissions (i.e., the firms that have not invested in more efficient abatement technologies) are also targeted with higher monitoring probability. In our model, however, firms have the option to adopt the new technology "in exchange" for a reduced monitoring probability. Since technology adoption serves the purpose of reducing emissions and increasing deterrence, the regulator can achieve the same or an increased level of compliance at a lower total enforcement cost.

The fact that the stringency of enforcement can be reduced if polluting agents show evidence of compliance-promoting activities is well documented in the literature. For example, Arguedas (2013)

[1] Heyes and Rickman (1999) show that if the environmental protection agency interacts with firms in more than one enforcement domain, it might be optimal to tolerate non-compliance in some sub-set of domains "in exchange" for compliance in others. Livernois and McKenna (1999) show that if firms self-report their emissions, lowering fines for non-compliance raises the proportion of firms that truthfully report their compliance status. Nyborg and Telle (2004) argue that if prosecution is costly, it might be optimal for the regulator just to issue a warning of some kind instead of prosecuting violators, and not to impose further penalties if violators move into compliance upon receipt of the warning.

[2] Like our paper, these studies also sorted firms into discrete groups and made use of the indefinite Markov state-switching model employed by Harrington. In constrast, some papers introduced a continuous reputation indicator (that summarizes the frequency and size of past violations) and used dynamic simulation techniques to analyze more efficient targeting of inspections (see, e.g., Hentschel and Randall 2000).

points out that in the Spanish legislation on hazardous waste, firms that invest in clean production processes associated with responsible water consumption are rarely inspected and, if inspected, they are rarely punished if found non-compliant. She also points out that penalty reductions in exchange for investment efforts by polluting firms can be found in the EPA's Audit Policy, where fines for non-compliance can be significantly reduced if firms install enhanced emission control devices that simplify regulators' monitoring processes.[3]

The paper is organized as follows. Section 2 presents the targeting scheme and the firm's compliance decisions. Section 3 presents the model of adoption and analyzes the impact of targeted state-dependent enforcement on the rate of technology adoption vis-a-vis Harrington's two-group targeting scheme. Section 4 studies the effects of the enforcement scheme on adopters', non-adopters', and aggregate emissions. Section 5 studies the effects of the enforcement scheme on the resources devoted to monitoring and enforcement. Section 6 presents some numerical simulations. The final section provides a discussion and concludes the paper.

2 The Model

Consider a competitive industry consisting of a continuum of firms of mass 1 that are risk-neutral and initially homogeneous in abatement costs. The firms are required to make two dichotomous decisions: whether to adopt a new abatement technology to reduce emissions at a lower cost and whether to comply with the emission standard $\overline{q}$. We assume that the adoption decisions made by firms are observable by the regulator. However, the emissions and compliance status of firms can only be known by the regulatory agency through costly monitoring. Like Harrington (1988), we focus on the behavior of a regulatory agency whose primary goal is enforcement and not social welfare maximization. Thus, we specify the goal of the regulatory agency as minimizing the resources devoted to monitoring and enforcement consistent with achieving a given compliance rate with the emission standard $\overline{q}$ without modeling the policy process through which the level of

[3] Arguedas (2013) analyzes whether it is socially desirable that fines for exceeding pollution standards depend on the firm's level of investment in environmentally friendly technologies. Unlike this paper, she considers a static partial equilibrium framework and focuses on the effects of fines instead of auditing. Coria and Villegas (2014) analyze the advisability of targeted enforcement of emissions taxes in a static setting. They show that the regulator can reduce aggregate emissions by engaging in a regulatory deal where a reduced monitoring probability is granted in "exchange" for adoption of new technology.

the standard is chosen.[4]

Let the abatement cost function of an individual firm be denoted $c(q)$, which is strictly convex and decreasing in the level of emissions q. The new technology allows firms to abate emissions at a lower cost $\theta c(q)$, where $\theta \in (0, 1)$ is a parameter that represents the drop in abatement cost obtained by adopting the new technology. After making the adoption decision, firms decide on the compliance or violation of the standard $\overline{q}$. We assume that after monitoring a firm, the regulator is able to perfectly determine the firm's compliance status. If the monitoring reveals that the firm is non-compliant, it faces a convex penalty $\phi(q - \overline{q}) > 0$. For zero violation, the penalty is zero $\phi(0)$, yet the marginal penalty is greater than zero $\phi'(0) > 0$.[5]

Harrington considers two groups of firms: the non-target group (G_1), which faces less stringent enforcement, and the target group (G_2), where scrutiny is high. Let π_1 and π_2 denote the probabilities that the regulator audits a firm in G_1 and G_2, respectively, where these probabilities are common knowledge among firms, and $\pi_1 < \pi_2$. Moreover, firms can move from G_1 to G_2 according to transition probabilities that depend on the adoption status, current state of the system and compliance with the emission standard (see Table 1).

Adopters				
	Comply		Violate	
	G_1^A	G_2^A	G_1^A	G_2^A
G_1^A	1	0	$1-\alpha_A$	α_A
G_2^A	γ_A	$1-\gamma_A$	0	1

Non-adopters				
	Comply		Violate	
	G_1^{NA}	G_2^{NA}	G_1^{NA}	G_2^{NA}
G_1^{NA}	1	0	$1-\alpha_{NA}$	α_{NA}
G_2^{NA}	γ_{NA}	$1-\gamma_{NA}$	0	1

Table 1: Transition matrices for adopters and non-adopters

Let G_1^A (G_1^{NA}) and G_2^A (G_2^{NA}) denote the sub-group of adopters (non-adopters) in G_1 and G_2, respectively. Furthermore, let $\alpha_A(\alpha_{NA})$ denote the probability of moving an adopter (non-adopter) to group G_2 if caught violating in group G_1, and $\gamma_A(\gamma_{NA})$ denote the probability of moving an

[4] See also Garvie and Keeler (1994).

[5] Unlike our setting, Harrington (1988) assumes a linear penalty function, implying that the decision of whether or not to comply with the emission standard is of the all-or-nothing type. Though such an assumption facilitates the modeling since it provides a clear cut-off policy where all detected violations are transferred to the group with the higher monitoring probability, it might lead to unrealistic situations where firms report zero emissions and the regulator does not monitor them.

adopter (non-adopter) back to group G_1 if discovered complying in G_2. We assume that $\alpha_A \leq \alpha_{NA}$ and $\gamma_A \geq \gamma_{NA}$. In addition, we assume that $\alpha_A \geq \gamma_A$ and $\alpha_{NA} \geq \gamma_{NA}$.

Thus, our framework is general enough to encompass Harrington's state-dependent enforcement scheme (if $\alpha_A = \alpha_{NA}$ and $\gamma_A = \gamma_{NA}$, and hence, our four-group targeting scheme converges to Harrington's two-group targeting scheme) and to allow us to analyze the effects of differentiated probabilities of transition to reward the firms that adopt the technologies (hereinafter denoted *targeted state-dependent enforcement* where $\alpha_A < \alpha_{NA}$ and $\gamma_A > \gamma_{NA}$). Finally, the framework is also general enough to analyze the effects of the allocation of adopters and non-adopters to the target and non-target groups G_1 and G_2. In particular, we analyze three different initial allocations: (1) when all firms are initially allocated to G_1, (2) when all firms are initially allocated to G_2, and (3) when adopters are initially allocated to G_1 and non-adopters to G_2, hereinafter denoted *targeted initial allocation*.[6]

As in Harrington (1988), the monitoring scheme poses a Markov decision problem to adopters and non-adopters since they move from one group to the other depending on the compliance behavior in the previous period. For each adoption status, the firm chooses among possible strategies:

- comply when in G_1 and G_2,
- comply only if in G_1,
- comply only if in G_2,
- violate in both groups.

Let the strategy f^{jklm} describe the firm's decisions to comply with (0) or violate (1) the regulation, where j and k denote the actions taken by adopters when in groups G_1 and G_2, respectively and let l and m denote the actions taken by non-adopters when in groups G_1 and G_2, respectively. In principle, we should have 16 possible strategies. However, since adopters' compliance cost is lower than non-adopters' within the same group (G_1 or G_2), it is not reasonable that non-adopters

[6] Note that firms in our model are homogeneous ex-ante and hence, should comply with the regulation to the same extent. To be consistent with this assumption, we analyze the cases where all firms are initially allocated to G_1 or G_2, yet a random initial allocation of firms to G_1 or G_2 is also feasible. Let us consider a situation where a fraction χ of the firms are initially allocated to G_1 and the remaining fraction $[1 - \chi]$ are initially allocated to G_2. In this case the results become a linear combinations of our results.

comply but adopters violate. Moreover, since the expected cost of compliance in G_2 is higher both for adopters and non-adopters, it is not reasonable that they comply in G_1 but violate in G_2. Thus, six potential strategies remain: f^{0000}, f^{0010}, f^{0011}, f^{1010}, f^{1011}, and f^{1111}. Note that the first three strategies imply full compliance by adopters (and varying levels of compliance by non-adopters) and the last three strategies imply partial or full non-compliance by adopters and non-adopters.

Let $E_A^{jklm}(1)$ and $E_{NA}^{jklm}(1)$ denote the present value of adopters' and non-adopters' expected cost of strategy f^{jklm} when initially allocated to G_1. By analogy, let $E_A^{jklm}(2)$ and $E_{NA}^{jklm}(2)$ denote the present value of adopters' and non-adopters' expected cost of strategy f^{jklm} when initially allocated to group G_2. As in Harrington (1988), by the stationary property, the expected present value must be the cost in this period plus the expected present value discounted one period. For instance, let us compute the present values of f^{1010} for adopters when initially allocated to G_1 and G_2, respectively. In a single play of this game, if the regulator announces beforehand that the inspection probability for an adopter is π_i ($\vee\ i = 1, 2$), the adopters' cost minimization problem corresponds to[7]:

$$Min_{q_A}\left[\theta c(q_A) + \pi_i \phi(q_A - \overline{q})\right] \text{ s.t. } q_A \leq \overline{q}.$$

The optimization problem can be represented by the Lagrangian $L = \theta c(q_A) + \pi_i \phi(q_A - \overline{q}) + \omega\left[\overline{q} - q_A\right]$, where $\omega \geq 0$ is the Lagrange multiplier. The FOC defining the optimal level of emissions is given by:

$$\theta' c(q_A) + \pi_i \phi' \left[q_A - \overline{q}\right] + \omega = 0. \tag{1}$$

Under compliance, $q_A = \overline{q}$ and $\omega = 0$. Hence, the expected cost of the regulation is equal to $\theta c(\overline{q})$. Under non-compliance (NC), adopters select an emission level $q_A^{NC} > \overline{q}$ such that $-\theta' c(q_A^{NC}) = \pi_i \phi'\left[q_A^{NC} - \overline{q}\right]$. It holds that q_A^{NC} decreases with the monitoring probability π_i, and the expected cost of the regulation is equal to $\theta c(q_A^{NC}) + \pi_i \phi(q_A^{NC}(\pi_i) - \overline{q})$. Let $0 \leq \beta < 1$ be the discount factor. Since under the strategy f^{1010} adopters violate the standard if in G_1 and comply if in G_2, the expected costs when initially allocated G_1 and G_2 are, respectively:

$$\begin{aligned} E_A^{1010}(1) &= \left[\theta c(q_A^{NC}(\pi_1)) + \pi_1 \phi(q_A^{NC}(\pi_1) - \overline{q})\right] + \beta\left[\pi_1 \alpha_A E_A^{1010}(2) + \left[1 - \pi_1 \alpha_A\right] E_A^{1010}(1)\right], \quad (2) \\ E_A^{1010}(2) &= \left[\theta c(\overline{q})\right] + \beta\left[\pi_2 \gamma_A E_A^{1010}(1) + \left[1 - \pi_2 \gamma_A\right] E_A^{1010}(2)\right]. \quad (3) \end{aligned}$$

[7] For non-adopters $\theta = 1$.

The second term in parentheses in equation (2) represents the expected present value discounted one period. It is composed of the expected cost of being caught in violation in G_1 and sent to G_2 with probability $\pi_1\alpha_A$ plus the expected cost of remaining in G_1 with probability $[1-\pi_1\alpha_A]$. By analogy, the second term in parentheses in equation (3) represents the discounted expected present value of being found in compliance in G_2 and sent to G_1 with probability $\pi_2\gamma_A$ plus the expected cost of remaining in G_2 with probability $[1-\pi_2\gamma_A]$.

Solving equations (2) and (3) simultaneously yields:

$$\begin{aligned} E_A^{1010}(1) &= \frac{\theta c(q_A^{NC}(\pi_1)) + \pi_1\phi(q_A^{NC}(\pi_1) - \overline{q})}{1-\beta} + \frac{\beta\pi_1\alpha_A\left[\theta c(\overline{q}) - \left[\theta c(q_A^{NC}(\pi_1)) + \pi_1\phi(q_A^{NC}(\pi_1) - \overline{q})\right]\right]}{[1-\beta]\left[1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A\right]}, \\ E_A^{1010}(2) &= \frac{\theta c(\overline{q})}{1-\beta} - \frac{\beta\pi_2\gamma_A\left[\theta c(\overline{q}) - \left[\theta c(q_A^{NC}(\pi_1)) + \pi_1\phi(q_A^{NC}(\pi_1) - \overline{q})\right]\right]}{[1-\beta]\left[1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A\right]}. \end{aligned}$$

Table 2 presents solutions to the sets of simultaneous equations giving the present values of each feasible strategy f^{jklm}. Note that the expected cost for those cases where firms are moved from one group to the other comprises two terms. The first term represents the expected cost if the firm remains in the initial group forever. The second term is an adjustment factor that reflects the likelihood of the firm being moved to the other group. This adjustment factor is positive if the expected cost is greater in the other group and negative otherwise.[8]

Note also that regardless of the *initial allocation*, the transition probabilities $(\alpha_{NA}, \gamma_{NA})$ only affect non-adopters' expected costs of compliance. By analogy, the transition probabilities (α_A, γ_A) only affect adopters' expected costs of compliance. Morever, like in Harrington (1988), an increased probability $\alpha_A(\alpha_{NA})$ of transiting an adopter (non-adopter) to group G_2 if caught violating in group G_1 increases adopters' (non-adopters') expected costs of compliance, while the reverse holds for the probability $\gamma_A(\gamma_{NA})$ of moving an adopter (non-adopter) back to group G_1 if discovered complying in G_2.

[8] This is consistent with Friesen (2003).

Adopters	
Initially in $\mathbf{G}_1$	
00	$\frac{\theta c(\overline{q})}{1-\beta}$
10	$\frac{\theta c(q_A^{NC}(\pi_1))+\pi_1\phi(q_A^{NC}(\pi_1)-\overline{q})}{1-\beta}+\frac{\beta\pi_1\alpha_A\left[\theta c(\overline{q})-\left[\theta c(q_A^{NC}(\pi_1))+\pi_1\phi(q_A^{NC}(\pi_1)-\overline{q})\right]\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$
11	$\frac{\theta c(q_A^{NC}(\pi_1))+\pi_1\phi(q_A^{NC}(\pi_1)-\overline{q})}{1-\beta+\beta\pi_1\alpha_A}+\frac{\beta\pi_1\alpha_A\left[\theta c(q_A^{NC}(\pi_2))+\pi_2\phi(q_A^{NC}(\pi_2)-\overline{q})\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A]}$
Initially in $\mathbf{G}_2$	
00	$\frac{\theta c(\overline{q})}{1-\beta}$
10	$\frac{\theta c(\overline{q})}{1-\beta}-\frac{\beta\pi_2\gamma_A\left[\theta c(\overline{q})-\left[\theta c(q_A^{NC}(\pi_1))+\pi_1\phi(q_A^{NC}(\pi_1)-\overline{q})\right]\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$
11	$\frac{\theta c(q_A^{NC}(\pi_2))+\pi_2\phi(q_A^{NC}(\pi_2)-\overline{q})}{1-\beta}$
Non-adopters	
Initially in $\mathbf{G}_1$	
00	$\frac{c(\overline{q})}{1-\beta}$
10	$\frac{c(q_{NA}^{NC}(\pi_1))+\pi_1\phi(q_{NA}^{NC}(\pi_1)-\overline{q})}{1-\beta}+\frac{\beta\pi_1\alpha_{NA}\left[c(\overline{q})-\left[c(q_{NA}^{NC}(\pi_1))+\pi_1\phi(q_{NA}^{NC}(\pi_1)-\overline{q})\right]\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}]}$
11	$\frac{c(q_{NA}^{NC}(\pi_1))+\pi_1\phi(q_{NA}^{NC}(\pi_1)-\overline{q})}{1-\beta+\beta\pi_1\alpha_{NA}}+\frac{\beta\pi_1\alpha_{NA}\left[c(q_{NA}^{NC}(\pi_2))+\pi_2\phi(q_{NA}^{NC}(\pi_2)-\overline{q})\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A]}$
Initially in $\mathbf{G}_2$	
00	$\frac{c(\overline{q})}{1-\beta}$
01	$\frac{c(\overline{q})}{1-\beta}-\frac{\beta\pi_2\gamma_{NA}\left[c(\overline{q})-\left[c(q_A^{NC}(\pi_1))+\pi_1\phi(q_A^{NC}(\pi_1)-\overline{q})\right]\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$
11	$\frac{c(q_{NA}^{NC}(\pi_2))+\pi_2\phi(q_{NA}^{NC}(\pi_2)-\overline{q})}{1-\beta}$

Table 2: Expected costs of compliance for adopters and non-adopters

Finally, as in Harrington (1988), there are critical probabilities π_1 and π_2 that define which strategy is optimal for the firms. In our case, let $\overline{\pi}_1^A$ ($\overline{\pi}_1^{NA}$) and $\overline{\pi}_2^A$ ($\overline{\pi}_2^{NA}$) denote the critical probabilities π_1 and π_2 that make adopters (non-adopters) indifferent between compliance and violation when in G_1 and G_2, respectively. $\overline{\pi}_1^A$ and $\overline{\pi}_2^A$ are independent of the initial allocation of adopters and non-adopters to G_1 and G_2 and are implicitly defined by the equations:

$$\theta\left[c(\overline{q})-c(q_A^{NC}(\overline{\pi}_1^A)\right] = \overline{\pi}_1^A\phi(q_A^{NC}(\overline{\pi}_1^A)-\overline{q}), \tag{4}$$

$$\theta\left[c(\overline{q})-c(q_A^{NC}(\overline{\pi}_2^A)\right]-\Gamma = \overline{\pi}_2^A\phi(q_A^{NC}(\overline{\pi}_2^A)-\overline{q}), \tag{5}$$

where $\Gamma=\frac{\beta\overline{\pi}_2^A\gamma_A\left[\theta\left[c(\overline{q})-c(q_A^{NC}(\pi_1)\right]-\pi_1\phi(q_A^{NC}(\pi_1)-\overline{q})\right]}{\left[1-\beta+\beta\pi_1\alpha_A+\beta\overline{\pi}_2^A\gamma_A\right]}$. In the case of imperfect compliance, $\pi_1<\overline{\pi}_1^A$, and hence $\Gamma > 0$. Thus, $\overline{\pi}_2^A$ is implicitly defined by equation (5), and it is a non-linear function of π_1, α_A, and γ_A. As shown in Appendix A, it holds that $\overline{\pi}_2^A$ increases when π_1 or α_A increases, and it decreases when γ_A increases. Vis-a-vis Harrington's enforcement scheme, targeted state-dependent enforcement reduces $\overline{\pi}_2^A$ since $\alpha_A < \alpha_{NA}$ and $\gamma_A > \gamma_{NA}$. From equations (4) and (5), it can also be seen that the larger the reduction in abatement costs due to the adoption of the new technology (i.e., the lower the parameter θ), the lower the critical probabilities $\overline{\pi}_1^A$ and $\overline{\pi}_2^A$. In other words, the more efficient the new technology is, the higher the incentives for adopters to comply.

Similar equations define the probabilities $\overline{\pi}_1^{NA}$ and $\overline{\pi}_2^{NA}$ for $\theta = 1$ and emission levels $q_{NA}^{NC}(\overline{\pi}_1^{NA})$ and $q_{NA}^{NC}(\overline{\pi}_2^{NA})$. Since for the same monitoring probability the expected costs of compliance are lower for non-adopters, the minimum monitoring probability necessary to ensure compliance is lower for adopters than for non-adopters, $\overline{\pi}_1^A < \overline{\pi}_1^{NA}$ and $\overline{\pi}_2^A < \overline{\pi}_2^{NA}$.

Given these critical probabilities, the optimal strategy f can be characterized as:

$$f = \begin{cases} f^{0000} \text{ if } \pi_1 > \overline{\pi}_1^{NA} \text{ and } \pi_2 > \overline{\pi}_2^{NA}, \\ f^{0010} \text{ if } \pi_1 \in \left[\overline{\pi}_1^{A}, \overline{\pi}_1^{NA}\right] \text{ and } \pi_2 > \overline{\pi}_2^{NA}, \\ f^{0011} \text{ if } \pi_1 \in \left[\overline{\pi}_1^{A}, \overline{\pi}_1^{NA}\right] \text{ and } \pi_2 \in \left[\overline{\pi}_2^{A}, \overline{\pi}_2^{NA}\right], \\ f^{1010} \text{ if } \pi_1 < \overline{\pi}_1^{A} \text{ and } \pi_2 > \overline{\pi}_2^{NA}, \\ f^{1011} \text{ if } \pi_1 < \overline{\pi}_1^{A} \text{ and } \pi_2 \in \left[\overline{\pi}_2^{A}, \overline{\pi}_2^{NA}\right] \;, \\ f^{1111} \text{ if } \pi_1 < \overline{\pi}_1^{A} \text{ and } \pi_2 < \overline{\pi}_2^{A}. \end{cases}$$

3 The Adoption Rate

We assume that buying and installing the new technology implies a fixed cost that differs among firms.[9] Let k_i denote the fixed cost of adoption for firm i, and assume that k_i is uniformly distributed on the interval $(\underline{k}, \overline{k})$. Note that the differences $E_{NA}^{jklm}(1) - E_A^{jklm}(1)$ and $E_{NA}^{jklm}(2) - E_A^{jklm}(2)$ indicate how compliance costs change with the use of new technologies when firms are initially allocated to G_1 and G_2, respectively and the difference $E_{NA}^{jklm}(2) - E_A^{jklm}(1)$ indicates how compliance costs

[9]The assumption that adoption costs differ among firms is not new in the literature analyzing the effects of choice of policy instruments on the rate of adoption of new technologies; see, e.g., Requate and Unold (2001). On the other hand, Stoneman and Ireland (1983) point out that although most theoretical and empirical literature on technological adoption focuses on the demand side alone, supply-side forces might be very important in explaining patterns of adoption in practice. Thus, e.g., costs of acquiring new technology might vary among firms due to firm characteristics, e.g., location and output, or competition among suppliers of capital goods.

change under a targeted allocation based on adoption status. Any firm whose saving in total expected cost offsets its adoption cost will adopt the new technology. For a given strategy f^{jklm} and initial allocation of non-adopters and adopters to groups y and x, respectively, where $y, x = 1, 2$, and $y \geq x$, the rate of firms $\lambda \in [0, 1]$ adopting the more efficient abatement technology is defined by:

$$\lambda^{jklm}(y \mid x) = \int_{\underline{k}}^{\widehat{k}} f(k_i)dk = F(\widehat{k}_i) = \psi \left[E_{NA}^{jklm}(y) - E_A^{jklm}(x) \right] - \varsigma, \tag{6}$$

where the RHS of equation (6) follows from the definition of the uniform cumulative distribution of k_i, $\psi = \frac{1}{\overline{k} - \underline{k}}$ and $\varsigma = \psi \underline{k}$. For simplicity, we assume hereinafter that $\varsigma \simeq 0$. Thus, the adoption rate is a function of the shift in abatement costs θ, the emission standard $\overline{q}$, the initial allocation of adopters and non-adopters to G_1 or G_2, the monitoring probabilities (π_1, π_2), and the transition probabilities (α_A, α_{NA}) and (γ_A, γ_{NA}). In addition, λ is inversely related to the length of the investment cost interval $(\overline{k} - \underline{k})$. [10]

In what follows, we analyze the impact of the targeted state-dependent enforcement strategy on the rate of adoption through comparative statics with respect to the transition probabilities $\alpha_A, \alpha_{NA}, \gamma_A$, and γ_{NA} (see Appendix B for detailed comparative statics of adopters' and non-adopters' expected costs of compliance with regard to the transition probabilities).

Proposition 1 *A targeted state-dependent enforcement scheme spurs the rate of adoption of the environmentally friendly technology.*

Recall that $\lambda^{jklm}(y \mid x) = \psi \left[E_{NA}^{jklm}(y) - E_A^{jklm}(x) \right] \vee y, x = 1, 2$, and $y \geq x$, and that the transition probabilities α_{NA} and γ_{NA} (α_A and γ_A) only affect non-adopters' (adopters') expected costs of compliance. Hence,

$$\begin{aligned} \frac{\partial \lambda^{jklm}(y \mid x)}{\partial \alpha_{NA}} &= \psi \frac{\partial E_{NA}^{jklm}(y)}{\partial \alpha_{NA}} \geq 0, \\ \frac{\partial \lambda^{jklm}(y \mid x)}{\partial \gamma_{NA}} &= \psi \frac{\partial E_{NA}^{jklm}(y)}{\partial \alpha_{NA}} \leq 0. \end{aligned}$$

Furthermore,

[10] The more heterogeneous the firms are in terms of the investment cost, the larger the interval $(\overline{k} - \underline{k})$ and the lower the rate of adoption.

$$\frac{\partial \lambda^{jklm}(y \mid x)}{\partial \alpha_A} = -\psi \frac{\partial E_{NA}^{jklm}(x)}{\partial \alpha_A} \leq 0,$$
$$\frac{\partial \lambda^{jklm}(y \mid x)}{\partial \gamma_A} = -\psi \frac{\partial E_{NA}^{jklm}(x)}{\partial \gamma_A} \geq 0.$$

Thus, *targeted state-dependent enforcement* where $\alpha_{NA} > \alpha_A$ and $\gamma_{NA} < \gamma_A$ induces a larger rate of adoption than Harrington's scheme based only on past compliance.

As shown in Appendix B, marginal variations in $(\alpha_{NA},\ \gamma_{NA})$ have a larger effect on the rate of adoption than do marginal variations in $(\alpha_A,\ \gamma_A)$ in almost all cases. Moreover, the marginal effects of (α_A, α_{NA}) on the rate of adoption are larger when firms are initially allocated to G_1. The reverse holds for (γ_A, γ_{NA}); their marginal effects on the rate of adoption are larger when firms are initially allocated to G_2.

As mentioned above, since enforcement is more stringent in G_2, it holds that $E_A^{jklm}(2) \geq E_A^{jklm}(1)$ and $E_{NA}^{jklm}(2) \geq E_{NA}^{jklm}(1)$, where equality holds only in the case where adopters/non-adopters fully comply with the regulation. Therefore, we can derive the following proposition regarding the effects of a *targeted initial allocation* on the rate of adoption.

Proposition 2 *The rate of adoption of the environmentally friendy technology under a targeted state dependent enforcement scheme is larger if the regulator also targets the initial allocation of firms based on adoption status.*

Given equation (6), the difference in adoption rate between targeted initial allocation and the allocation where all firms are initially sent to G_1 corresponds to:

$$\lambda^{jklm}(2 \mid 1) - \lambda^{jklm}(2 \mid 2) = \psi \left[E_A^{jklm}(2) - E_A^{jklm}(1) \right] \geq 0.$$

This difference is equal to zero under full compliance by non-adopters, and positive otherwise.

By analogy, the difference in adoption rate between targeted initial allocation and the allocation where all firms are initially sent to G_2 corresponds to:

$$\lambda^{jklm}(2 \mid 1) - \lambda^{jklm}(1 \mid 1) = \psi \left[E_{NA}^{jklm}(2) - E_{NA}^{jklm}(1) \right] \geq 0.$$

This difference is equal to zero under full compliance by adopters, and positive otherwise.

Hence, compared with the allocations where all firms are sent to G_1 or G_2, targeted initial allocation leads to a higher rate of adoption. Thus, our results suggest that to speed up the pace of adoption of environmentally friendly technologies, the regulator should exert a stronger monitoring pressure on non-adopters. This result goes against previous studies of targeted enforcement policy in a static setting that suggest exerting a stronger monitoring pressure on firms with lower abatement cost since their pollution levels are more responsive to the enforcement parameters than those of firms with higher abatement cost (e.g., Garvie and Keeler (1994) Macho-Stadler and Pérez-Castrillo (2006)). Since in their analysis the rate of adoption is exogenous, they do not consider that biasing the monitoring scheme against firms with lower abatement costs reduces the potential gains from investing in new technologies, and hence, discourages adoption. A similar argument applies in the case of industrial turnover. Stringent regulations that only apply to newer or cleaner firms might slow down the turnover of pollution sources, drive up the cost of environmental protection, and increase pollution levels since they provide existing sources with perverse incentives to continue operating while "taxing" newer and cleaner entrants. See, e.g., Maloney and Brady (1988).

4 Individual and Aggregate Expected Emissions

Let $\widehat{q}_A^{jklm}(y)$ and $\widehat{q}_{NA}^{jklm}(y)$ to denote the expected emissions by adopters and non-adopters under strategy $jklm$ when initially allocated to group $y = 1, 2$. Table 3 presents the summary of expected emissions by adopters and non-adopters under different strategies.

Adopters		
f	Initially in G_1	Initially in G_2
00	$\frac{\overline{q}}{1-\beta}$	$\frac{\overline{q}}{1-\beta}$
10	$\frac{[1-\beta+\beta\pi_2\gamma_A]q_A^{NC}(\pi_1)+\beta\pi_1\alpha_A\overline{q}}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$	$\frac{\overline{q}}{1-\beta}+\frac{\beta\pi_2\gamma_A\left[q_A^{NC}(\pi_1)-\overline{q}\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$
11	$\frac{q_A^{NC}(\pi_1)}{1-\beta+\beta\pi_1\alpha_A}+\frac{\beta\pi_1\alpha_A q_A^{NC}(\pi_2)}{[1-\beta][1-\beta+\beta\pi_1\alpha_A]}$	$\frac{q_A^{NC}(\pi_2)}{1-\beta}$
Non-adopters		
f	Initially in G_1	Initially in G_2
00	$\frac{\overline{q}}{1-\beta}$	$\frac{\overline{q}}{1-\beta}$
10	$\frac{[1-\beta+\beta\pi_2\gamma_{NA}]q_{NA}^{NC}(\pi_1)+\beta\pi_1\alpha_{NA}\overline{q}}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}]}$	$\frac{\overline{q}}{1-\beta}+\frac{\beta\pi_2\gamma_{NA}\left[q_{NA}^{NC}(\pi_1)-\overline{q}\right]}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}]}$
11	$\frac{q_{NA}^{NC}(\pi_1)}{1-\beta+\beta\pi_1\alpha_{NA}}+\frac{\beta\pi_1\alpha_{NA} q_{NA}^{NC}(\pi_2)}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}]}$	$\frac{q_{NA}^{NC}(\pi_2)}{1-\beta}$

Table 3: Expected emissions by adopters and non-adopters

As expected, comparing the columns of Table 3 shows that (except for the case of full compliance) expected emissions by adopters and non-adopters are larger if firms are initially allocated to G_1. Furthermore, if $\alpha_A = \alpha_{NA}$ and $\gamma_A = \gamma_{NA}$, expected emissions are higher for non-adopters than adopters in all cases, i.e., $\widehat{q}_{NA}^{jklm}(y) \geq \widehat{q}_A^{jklm}(y) \vee jklm$ and $y = 1, 2$.

For a given strategy and initial allocation of non-adopters and adopters to groups y and x, respectively, aggregate expected emissions can be represented as:

$$\widehat{Q}(y \mid x) = \lambda(y \mid x)\widehat{q}_A(x) + [1 - \lambda(y \mid x)]\, \widehat{q}_{NA}(y). \tag{7}$$

By varying the transition probabilities $(\alpha_A, \alpha_{NA}, \gamma_A, \gamma_{NA})$ we have two types of effects on aggregate emissions: a *direct effect* on adopters' or non-adopters' emissions, and an *indirect effect* on the rate of adoption. As shown in detail in Appendix C, increased probabilities (α_A, α_{NA}) have the positive effect of reducing emissions by adopters and non-adopters, respectively. In contrast, increased probabilities (γ_A, γ_{NA}) have a negative effect, leading to increased emissions. Therefore, targeted state-dependent enforcement has the positive effect of reducing emissions by means of enhancing the rate of adoption (and thus changing the composition of firms towards a larger fraction of cleaner firms). Furthermore, it has the positive (direct) effect of reducing non-adopters' emissions. Nevertheless, in some cases, this might come at the expense of increased emissions by adopters.

Proposition 3 *A targeted state-dependent enforcement scheme based on firms' past compliance and adoption of environmentally superior technologies can reduce aggregate emissions.*

Let us for a moment disregard the effects of the initial allocation of firms to G_1 or G_2. Let the supercripts T and H denote the outcomes of targeted state-dependent and Harrington's enforcement, respectively. Given equation (7), the difference in expected emissions between the two enforcement schemes corresponds to:

$$\widehat{Q}^T - \widehat{Q}^H = \left[\lambda^T - \lambda^H\right]\left[\widehat{q}_A^H - \widehat{q}_{NA}^H\right] + \left[1 - \lambda^T\right]\left[\widehat{q}_{NA}^T - \widehat{q}_{NA}^H\right] + \lambda^T\left[\widehat{q}_A^T - \widehat{q}_A^H\right]. \tag{8}$$

Note that the first term in brackets on the RHS of equation (8) is negative and corresponds to the effect of targeted state-dependent enforcement increasing the rate of adoption (vis-a-vis Harrington's enforcement), and thus reducing expected aggregate emissions as adopters emit less than non-adopters. The second term is also negative and corresponds to the reduced emissions by non-adopters, which are monitored more stringently under targeted state-dependent enforcement and hence emit less. Finally, the third term is positive and corresponds to the increased emissions by adopters, which are monitored less stringently under targeted state-dependent enforcement and hence emit more.

Regardless the initial allocation, $\widehat{q}_A^T = \widehat{q}_A^H$ under the strategies f^{0000}, f^{0010} and f^{0011}, implying that equation (8) simplifies to:

$$\widehat{Q}^T - \widehat{Q}^H = \left[\lambda^T - \lambda^H\right]\left[\widehat{q}_A^H - \widehat{q}_{NA}^H\right] + \left[1 - \lambda^T\right]\left[\widehat{q}_{NA}^T - \widehat{q}_{NA}^H\right] \leq 0.$$

This difference is equal to zero under f^{0000} and negative under f^{0010} and f^{0011}. Hence, aggregate emissions under targeted state-dependent enforcement are lower or equal to those under Harrington's enforcement. The comparison is less clear for the strategies f^{1010}, f^{1011} and f^{1111}. In what follows, let us consider how targeted state-dependent enforcement affects adopters' and non-adopters' emissions (that is, the second and third term of equation (8); see Appendix C for detailed comparative statics) under these three strategies.

- f^{1010}. We have that the marginal effects of the probabilities α_{NA} and γ_{NA} are larger than the marginal effects of the probabilities α_A and γ_A. Hence, vis-a-vis Harrington's enforcement, targeted state-dependent enforcement increases adopters' and reduces non-adopters' emissions. The overall effect is a net reduction in emissions as the reduction in non-adopters' emissions is larger than the increase in adopters's emissions regardless the initial allocation.

- f^{1011}. A targeted state-dependent enforcement increases adopters' emissions. In contrast, it has no effect on non-adopters'emissions if they are initially allocated to G_2 and reduces non-adopters'emissions if they are initially allocated to G_1. The overall effect is a net increase in emissions as the increase in adopters' emissions is larger than (any) reduction in non-adopters's emissions regardless the initial allocation.

- f^{1111}. A targeted state-dependent enforcement has no effect on adoption or on adopters' and non-adopters' emissions when firms are initially allocated to G_2. Hence $\widehat{Q}^T - \widehat{Q}^H = 0$ in such case. If firms are initially allocated to G_1, it increases adopters' and reduces non-adopters' emissions. The overall effect is a net reduction of emissions as the reduction of non-adopters' emissions is larger than the increase in adopters's emissions.

Thus, we can say that vis-a-vis Harrington's enforcement, targeted state-dependent enforcement has no effect on emissions under full compliance by adopters and non-adopters, while it unambiguously reduces emissions under the strategies f^{0010}, f^{0011}, and f^{1010}. If all firms are initially allocated to G_1, it also reduces emissions under f^{1111}. Finally, whether or not targeted dependent enforcement leads to lowered emissions under f^{1011} depends on the relative magnitude of the direct and indirect effects. Even if adopters's emissions might be larger than those under Harrington's enforcement, adopters emit less than non-adopters. Hence, aggregate emissions under targeted state-depedent enforcement can be still lower than under Harrington's enforcement due to the larger rate of adoption.

When it comes to the expected aggregate violations, note that if firms were to always comply with the regulation, their expected emissions would be equal to $\frac{\overline{q}}{1-\beta}$. Thus, for a given strategy and initial allocation of non-adopters and adopters to groups y and x, respectively, the expected aggregate violations $\widehat{V}$ can be represented as:

$$\widehat{V}(y \mid x) = \widehat{Q}(y \mid x) - \frac{\overline{q}}{1-\beta}.$$

Hence, it is clear that if targeted state-dependent enforcement reduces expected aggregate emissions, it also reduces expected aggregate violations.

Let us now analyze the effects of a targeted initial allocation on aggregate emissions.

Proposition 4 *Expected aggregate emissions under targeted initial allocation are lower than the expected aggregate emissions under an allocation that initially sends all firms to G_1. If the increase in adoption rate due to targeted initial allocation is sufficiently large, the expected aggregate emissions under targeted initial allocation are also lower than the expected aggregate emissions under an allocation that initially sends all firms to G_2.*

Given equation (7), the difference in expected aggregate emissions between targeted initial allocation and the allocation where all firms are initially sent to G_1 corresponds to:

$$\widehat{Q}(2 \mid 1) - \widehat{Q}(1 \mid 1) = [\lambda(2 \mid 1) - \lambda(1 \mid 1)] \, [\widehat{q}_A(1) - \widehat{q}_{NA}(1)] + [1 - \lambda(2 \mid 1)] \, [\widehat{q}_{NA}(2) - \widehat{q}_{NA}(1)] \, . \quad (9)$$

Note that the first term in brackets on the RHS of equation (9) is negative and corresponds to the effect of a targeted initial allocation increasing the rate of adoption, and thus reducing expected aggregate emissions as adopters emit less than non-adopters. The second term is also negative and corresponds to the reduced emissions by non-adopters, which are monitored more stringently under targeted initial allocation and hence emit less. So, compared with the case where both adopters and non-adopters are allocated to G_1, a targeted initial allocation would not only lead to a higher adoption rate, but also to lower expected aggregate emissions.

The difference in expected aggregate emissions between a targeted initial allocation and the allocation where all firms are initially sent to G_2 corresponds to:

$$\widehat{Q}(2 \mid 1) - \widehat{Q}(2 \mid 2) = [\lambda(2 \mid 1) - \lambda(2 \mid 2)] \, [\widehat{q}_A(2) - \widehat{q}_{NA}(2)] + \lambda(2 \mid 1) \, [\widehat{q}_A(1) - \widehat{q}_A(2)] \, . \quad (10)$$

As before, the firm term in brackets on the RHS of equation (10) is negative and corresponds to the effect of targeted initiall allocation increasing the rate of adoption, and thus reducing expected aggregate emissions as adopters emit less than non-adopters. The second term is positive and corresponds to the increased emissions by adopters under targeted initial allocation: if adopters would have been initially allocated to G_2, they would have emitted $\widehat{q}_A(2)$ instead of $\widehat{q}_A(1)$. Since these two effects have different signs, the final effect of the initial allocation on expected aggregate emissions depends on their relative magnitude. Let us find the conditions for when $\widehat{Q}(2 \mid 1) - \widehat{Q}(2 \mid 2) \leq 0$. We have two cases:

- $\widehat{q}_A(1) = \widehat{q}_A(2)$, which occurs under full compliance by adopters.
- $\widehat{q}_A(1) - \widehat{q}_A(2) > 0$, and $\frac{\lambda(2|1)-\lambda(2|2)}{\lambda(2|1)} \geq \frac{\widehat{q}_A(1)-\widehat{q}_A(2)}{\widehat{q}_{NA}(2)-\widehat{q}_A(2)}$.

That is, if the increase in adoption rate due to a targeted initial allocation $\lambda(2 \mid 1) - \lambda(2 \mid 2)$ is sufficiently large, the expected aggregate emissions can be lower than when adopters and non-adopters are initially allocated to G_2.

5 Enforcement Costs

As Harrington (1988), we assume that the regulator wishes to minimize the resources devoted to monitoring and enforcement consistent with achieving a target compliance rate. For a given cost per visit for the regulatory agency equal to m (that does not differ between adopters and non-adopters)[11], we compute the expected costs of enforcing compliance among adopters and non-adopters for each strategy f^{jklm} and initial allocations to G_1 and G_2. These costs are denoted as $\widehat{m}_A^{jklm}(y)$ and $\widehat{m}_{NA}^{jklm}(y)$, respectively. Results are presented in Table 4.

[11] Millock et al. (2002) and Millock et al. (2012) analyze the incentives provided by different policy instruments for the adoption of new environmental monitoring technologies. Like in our study, in these studies the choice of installing a technology separates agents into two categories, yet their focus is on the optimal choice and stringency of policy instruments while ours is on differentiated monitoring probabilities. Furthermore, unlike our study, in these studies adoption of technological monitoring devices serves the purpose of transforming non-point sources into point sources, thus reducing the monitoring cost m.

Adopters		
f	Initially in G_1	Initially in G_2
00	$\frac{m\pi_1}{1-\beta}$	$\frac{m\pi_2}{1-\beta+\beta\pi_2\gamma_A} + \frac{\beta\pi_2\gamma_A m\pi_1}{[1-\beta][1-\beta+\beta\pi_2\gamma_A]}$
10	$\frac{m\pi_1}{[1-\beta]} + \frac{\beta\pi_1\alpha_A m[\pi_2-\pi_1]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$	$\frac{m\pi_2}{1-\beta} - \frac{\beta\pi_2\gamma_A m[\pi_2-\pi_1]}{[1-\beta][1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A]}$
11	$\frac{m\pi_1}{1-\beta+\beta\pi_1\alpha_A} + \frac{\beta\pi_1\alpha_A m\pi_2}{[1-\beta][1-\beta+\beta\pi_1\alpha_A]}$	$\frac{m\pi_2}{1-\beta}$
Non-adopters		
f	Initially in G_1	Initially in G_2
00	$\frac{m\pi_1}{1-\beta}$	$\frac{m\pi_2}{1-\beta+\beta\pi_2\gamma_{NA}} + \frac{\beta\pi_2\gamma_{NA} m\pi_1}{[1-\beta][1-\beta+\beta\pi_2\gamma_{NA}]}$
10	$\frac{m\pi_1}{[1-\beta]} + \frac{\beta\pi_1\alpha_{NA} m[\pi_2-\pi_1]}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}]}$	$\frac{m\pi_2}{1-\beta} - \frac{\beta\pi_2\gamma_{NA} m[\pi_2-\pi_1]}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}]}$
11	$\frac{m\pi_1}{1-\beta+\beta\pi_1\alpha_{NA}} + \frac{\beta\pi_1\alpha_{NA}\pi m\pi_2}{[1-\beta][1-\beta+\beta\pi_1\alpha_{NA}]}$	$\frac{m\pi_2}{1-\beta}$

Table 4: Expected cost of enforcing adopters and non-adopters

Since by construction we target surveillance resources to non-adopters, it is not surprising to see that $\widehat{m}_{NA}^{jklm}(y) \geq \widehat{m}_{A}^{jklm}(y)$, where equality holds only in the case where adopters/non-adopters fully comply with the regulation. Moreover, since enforcement is more stringent in G_2, it holds that $\widehat{m}_{A}^{jklm}(2) \geq \widehat{m}_{A}^{jklm}(1)$ and $\widehat{m}_{NA}^{jklm}(2) \geq \widehat{m}_{NA}^{jklm}(1)$.

For a given strategy and initial allocation of non-adopters and adopters to groups y and x, respectively, the total expected enforcement cost can be characterized as:

$$\widehat{M}(y \mid x) = [\lambda(y \mid x)\widehat{m}_A(x) + [1-\lambda(y \mid x)]\,\widehat{m}_{NA}(y)] \tag{11}$$

Like in the case of expected aggregate emissions, varying the transition probabilities creates two types of effects: a *direct effect* on enforcement cost, and an *indirect effect* on the rate of adoption. As shown in detail in Appendix D, increased probabilities (α_A, α_{NA}) increase the cost of enforcing compliance among adopters and non-adopters, respectively. In contrast, increased probabilities probabilities (γ_A, γ_{NA}) reduce the cost of enforcing compliance of adopters and non-adopters. Therefore, a targeted state-dependent enforcement has the positive indirect effect of reducing the total expected enforcement cost by means of enhancing the rate of adoption (and thus changing the composition of firms towards a larger fraction of firms whose cost of enforcement is lower). Nevertheless, this comes at the expense of an increased cost of enforcing compliance among non-adopters.

Proposition 5 *A targeted state-dependent enforcement scheme based on firms' past compliance and adoption of environmentally superior technologies can reduce the total expected cost of enforcing an emission standard.*

Let us disregard for the moment the effects of the initial allocation of firms to G_1 or G_2. Let the supercripts T and H denote the outcomes of targeted state-dependent and Harrington's enforcement, respectively. Given equation (11) and since under Harrington's enforcement the cost of enforcing compliance among adopters and non-adopters is the same, the difference in the total expected cost of enforcing an emission standard between the two enforcement schemes corresponds to:

$$\widehat{M}^T - \widehat{M}^H = \lambda^T \left[\widehat{m}_A^T - \widehat{m}_A^H\right] + \left[1 - \lambda^T\right] \left[\widehat{m}_{NA}^T - \widehat{m}_{NA}^H\right]. \tag{12}$$

The first term in brackets on the RHS of equation (12) is negative and corresponds to the lowered expected cost of enforcing compliance among adopters who are monitored less stringently under targeted state-dependent enforcement. The second term is positive and corresponds to the increased expected of enforcing compliance among non-adopters who are monitored more stringently under targeted state-dependent enforcement. Since these two effects have different signs, the sign of the difference in (12) depends on their relative magnitude. Let us find the conditions for when $\widehat{M}^T - \widehat{M}^H \leq 0$. We have:

$$\frac{\lambda^T}{1-\lambda^T} \geq \frac{\widehat{m}_A^H - \widehat{m}_A^T}{\widehat{m}_{NA}^T - \widehat{m}_{NA}^H}. \tag{13}$$

That is, if the adoption rate induced by targeted state-dependent enforcement is sufficiently large, this enforcement scheme can reduce the total expected cost of enforcing an emission standard vis-a-vis Harrington's enforcement. How large must the adoption rate be to induce a reduced expected cost of enforcing the standard? The answer depends on how targeted state dependent enforcement affects adopters' and non-adopters' enforcement cost. Let us analyze the effects of targeted state dependent enforcement under each feasible strategy when all firms are initially allocated to G_1 by means of comparative statics with respect to the transition probabilities (see Appendix D for detailed comparative statics).

- f^{0000}. We have that $\widehat{m}_A^T = \widehat{m}_A^H = \widehat{m}_{NA}^T = \widehat{m}_{NA}^H$ and hence, $\widehat{M}^T - \widehat{M}^H = 0$ regardless of λ^T.

- f^{0010} and f^{0011}. Targeted state dependent enforcement has no effect in the cost of enforcing compliance among adopters and increases the cost of enforcing compliance among non-adopters, and hence $\widehat{M}^T - \widehat{M}^H > 0$ regardless of λ^T.

- f^{1010}, f^{1011} and f^{1111}. Targeted state dependent enforcement reduces the cost of enforcing compliance among adopters and increases the cost of enforcing compliance among non-adopters. If condition (13) holds, the overall effect is however a net reduction in the cost of enforcement as the reduction in the enforcement cost for adopters is larger than the increase in the enforcement cost for non-adopters.

Let us assume now that all firms are initially allocated to G_2

- f^{0000}, f^{0010}, f^{0011} and f^{1010}. Targeted state dependent enforcement reduces the cost of enforcing compliance among adopters and increases the cost of enforcing compliance among non-adopters. If condition (13) holds, the overall effect is however a net reduction in the cost of enforcement as the reduction in the enforcement cost for adopters is larger than the increase in the enforcement cost for non-adopters.

- f^{1011}. Targeted state dependent enforcement reduces the cost of enforcing compliance among adopters and has no effect in the cost of enforcing compliance among non-adopters, and hence $\widehat{M}^T - \widehat{M}^H < 0$ regardless of λ^T.

- f^{1111}. We have that $\widehat{m}_A^T = \widehat{m}_A^H = \widehat{m}_{NA}^T = \widehat{m}_{NA}^H$ and hence, $\widehat{M}^T - \widehat{M}^H = 0$ regardless of λ^T.

Thus, we can say that targeted state-dependent enforcement leads to a reduced cost of enforcement under f^{1011} if all firms are initially allocated to G_2. Provided condition (13) holds, it has no effect or the positive effect of reducing the cost of enforcement under f^{0000} , f^{1010}, f^{1011} and f^{1011}. Finally, whether or not targeted state-dependent enforcement leads to lowered enforcement costs under f^{0010} and f^{0011} depends on the initial allocation. In particular, the expected cost of enforcement is not lower than Harrington's when all firms are initially allocated to G_1.

In sum, even if the cost of enforcing compliance among non-adopters might be larger under targeted state-dependent enforcement than under Harrington's enforcement, the fact that targeted state-dependent enforcement changes the composition of firms towards a larger fraction of firms for which the cost of enforcement is lower implies that its total enforcement cost can be still lower if the adoption rate is sufficiently large.

Proposition 6 *The expected enforcement costs under targeted initial allocation are lower than the expected enforcement costs under an allocation that initially sends all firms to G_2. If the increase in adoption rate due to targeted initial allocation is sufficiently large, the expected enforcement costs under targeted initial allocation are also lower than the expected enforcement costs under an allocation that initially sends all firms to G_1.*

Given equation (11), the difference in expected enforcement costs between targeted initial allocation and the allocation where all firms are initially sent to G_1 corresponds to:

$$\widehat{M}(2 \mid 1) - \widehat{M}(1 \mid 1) = [\lambda(2 \mid 1) - \lambda(1 \mid 1)] \, [\widehat{m}_A(1) - \widehat{m}_{NA}(1)] + [1 - \lambda(2 \mid 1)] \, [\widehat{m}_{NA}(2) - \widehat{m}_{NA}(1)] \, . \tag{14}$$

The first term in brackets on the RHS of equation (14) is negative and corresponds to the effect of targeted initial allocation increasing the rate of adoption, and thus reducing the expected cost of enforcement as adopters demand less surveillance resources than non-adopters. The second term is positive and corresponds to the increased expected of enforcing compliance among non-adopters who are monitored more stringently under targeted initial allocation. Since these two effects have different signs, the final effect of the initial allocation on the expected enforcement cost depends on their relative magnitude. Let us find the conditions for when $\widehat{M}(2 \mid 1) - \widehat{M}(1 \mid 1) \leq 0$. We have two cases:

- $\widehat{M}(2 \mid 1) - \widehat{M}(1 \mid 1) = 0$ when $\widehat{m}_A(1) = \widehat{m}_{NA}(1)$, and $\lambda(2 \mid 1) = 1$.
- $\widehat{M}(2 \mid 1) - \widehat{M}(1 \mid 1) < 0$ when $\widehat{m}_A(1) - \widehat{m}_{NA}(1) < 0$, and $\frac{\lambda(2|1) - \lambda(1|1)}{1 - \lambda(2|1)} \geq \frac{\widehat{m}_{NA}(2) - \widehat{m}_{NA}(1)}{\widehat{m}_{NA}(1) - \widehat{m}_A(1)}$.

That is, if the increase in adoption rate $\lambda(2 \mid 1) - \lambda(1 \mid 1)$ due to targeted initial allocation is sufficiently large, the total enforcement cost can be lower than when both adopters and non-adopters are initially allocated to G_1.

The difference in the expected enforcement cost between targeted initial allocation and the allocation where all firms are initially sent to G_2 corresponds to:

$$\widehat{M}(2 \mid 1) - \widehat{M}(2 \mid 2) = [\lambda(2 \mid 1) - \lambda(2 \mid 2)] \, [\widehat{m}_A(2) - \widehat{m}_{NA}(2)] + \lambda(2 \mid 1) \, [\widehat{m}_A(1) - \widehat{m}_A(2)] \, . \tag{15}$$

As before, the first term in brackets on the RHS of equation (15) corresponds to the effect of targeted initial allocation increasing the rate of adoption, and thus reducing the expected cost of enforcement as adopters demand less surveillance resources than non-adopters. The second terms corresponds to the reduction in the cost of monitoring adopters; under targeted initial allocation adopters cause an expected enforcement cost of to $\widehat{m}_A(1)$ instead of $\widehat{m}_A(2)$. Hence, compared with when both adopters and non-adopters are allocated to G_2, targeted initial allocation would not only lead to a higher adoption rate but also to a lower expected enforcement cost.

6 Numerical Simulations

In this section we present a numerical example of the effects of the targeted state-depedent enforcement on the adoption rate, aggregated emissions and total monitoring cost. In line with the assumptions of the model, let the abatement cost function be given by $c(q) = c_0 - c_1 q + \frac{c_2}{2} q^2$, where $c'(q) = c_2 q - c_1 < 0$, and $c''(q) = c_2 > 0$. The penalty function is given by $\phi(q - \overline{q}) = \varphi_1(q - \overline{q}) + \frac{\varphi_2 (q-\overline{q})^2}{2}$, where $\phi'(q - \overline{q}) = \varphi_1 + \varphi_2(q - \overline{q}) > 0$, and $\phi''(q - \overline{q}) = \varphi_2 > 0$. Then, given a monitoring probability $\pi_i \vee i = 1, 2$, the emission levels $q_A^{NC}(\pi_i)$ and $q_{NA}^{NC}(\pi_i)$ in a single play of this game are given by:[12]

$$\begin{aligned} q_A^{NC}(\pi_i) &= \overline{q} + \frac{\theta \left[c_1 - c_2 \overline{q}\right] - \pi_i \varphi_1}{\pi_i \varphi_2 + \theta c_2}, \\ q_{NA}^{NC}(\pi_i) &= \overline{q} + \frac{\left[c_1 - c_2 \overline{q}\right] - \pi_i \varphi_1}{\pi_i \varphi_2 + \theta c_2}. \end{aligned}$$

Let $c_0 = 50$, $c_1 = 10$, and $c_2 = 1$. Moreover, let $\theta = 0.65$, which implies that technology adoption allows for a 35% reduction in the abatement cost. The total number of firms is set at $n = 100$. The cost of adopting the new technology is assumed to be uniformily distributed in the interval $[20,\ 100]$. The emission standard is set at $\overline{q} = 5$. The coefficients for the penalty functions are set at $\varphi_1 = 20$ and $\varphi_2 = 1$ and the discount factor is set at $\beta = 0.95$. Finally, the unitary inspection cost m is equal to 1. Regarding the stringency of the enforcement scheme, we assume that $\pi_1 = 0.15$ and $\pi_2 = 0.5$. Moreover, under a two-group enforcement scheme, $\alpha_A = \alpha_{NA} = 0.5$, and $\gamma_A = \gamma_{NA} = 0.25$. Under a targeted state-dependent enforcement, $\alpha_A = 0.4$, $\alpha_{NA} = 0.6$, $\gamma_A = 0.35$, and $\gamma_{NA} = 0.15$.

[12] See equation (1).

Table 5 presents the adoption rate, expected aggregate emissions, and expected total enforcement cost under targeted state-dependent enforcement and a two-group enforcement scheme for each feasible strategy when all firms are initially allocated to G_1 and G_2. Table 6 compares the outcomes of both enforcement schemes under targeted initial allocation where non-adopters are initially allocated to G_2 and adopters to G_1.

		Two Groups			**Four Groups**		
f	**Initial Allocation**	λ	$\widehat{Q}$	$\widehat{M}$	λ	$\widehat{Q}$	$\widehat{M}$
0000	(1 | 1)	0.844	10000	300.00	0.844	10000	300.00
	(2 | 2)	0.844	10000	507.41	0.844	10000	469.56
0010	(1 | 1)	0.538	11130	396.00	0.589	10839	419.04
	(2 | 2)	0.629	10639	571.72	0.694	10367	552.73
0011	(1 | 1)	0.664	10481	438.02	0.683	10406	439.88
	(2 | 2)	0.844	10000	584.38	0.844	10000	545.94
1010	(1 | 1)	0.545	11352	507.81	0.596	11118	503.91
	(2 | 2)	0.633	10827	653.65	0.700	10626	629.53
1011	(1 | 1)	0.671	10766	574.71	0.691	10738	537.37
	(2 | 2)	0.849	10262	706.09	0.850	10323	638.12
1111	(1 | 1)	0.668	10648	711.34	0.688	10602	694.37
	(2 | 2)	0.844	10000	1000	0.844	10000	1000

Table 5: Targeted state-dependent enforcement vs. a two-group targeting scheme

As expected, when adopters and non-adopters fully comply with the regulation, there are no differences in adoption rate or expected aggregate emissions between targeted state-depedent enforcement and a two-group enforcement scheme regardless of the initial allocation. Nevertheless, when firms are initially allocated to G_2, the cost of enforcement is lower for the targeted state-dependent enforcement. For the remaining feasible strategies, targeted state-depedent enforcement induces a higher rate of adoption, lower emissions, lower total enforcement cost, or a combination of these changes.

		Two-Groups			Four-Groups		
f	**Initial Allocation**	λ	$\widehat{Q}$	$\widehat{M}$	λ	$\widehat{Q}$	$\widehat{M}$
0000	(2 | 1)	0.844	10000	332.41	0.844	10000	333.00
0010	(2 | 1)	0.629	10639	431.34	0.694	10367	440.42
0011	(2 | 1)	0.844	10000	409.38	0.844	10000	409.38
1010	(2 | 1)	0.635	10907	560.97	0.702	10705	539.33
1011	(2 | 1)	0.851	10374	581.34	0.851	10421	528.30
1111	(2 | 1)	0.848	10219	755.28	0.848	10248	722.51

Table 6: Initial targeted allocation under a targeted state-dependent enforcement vs. a two-group targeting scheme

Table 6 shows that, as expected, targeted initial allocation generates less emissions than an allocation that sends all firms to G_1. If all firms are initially sent to G_2, the comparison is less clear, but we can say that aggregate emissions are higher under targeted initial allocation under most feasible strategies. Finally, when it comes to total enforcement costs, as expected, targeted initial allocation generates a lower total cost of enforcement than an allocation that sends all firms to G_2. If all firms are initially sent to G_1, the comparison is less clear, but we can say that total enforcement costs are higher under targeted initial allocation under most feasible strategies.

Given our choice of parameters, the critical probabilities that define the optimal strategy are equal to $(\overline{\pi}_1^A, \overline{\pi}_2^A) = (0.163, 0.290)$ and $(\overline{\pi}_1^{NA}, \overline{\pi}_2^{NA}) = (0.250, 0.481)$. Hence, the optimal strategy corresponds to f^{1010}. Thus, with regards to a two-group enforcement scheme, targeted state-dependent enforcement induces a higher rate of adoption, lower emissions and lower total enforcement cost under all allocations of adopters and non-adopters to the target and non-target groups G_1 and G_2.

7 Conclusions

A significant fraction of the literature on environmental regulation has focused on how environmental policies are and should be enforced. Harrington (1988) shows that a suitable strategy for the regulator to deal with the budget constraints in the enforcement activity is to target enforcement. Regulators can define a monitoring schedule for firms according to their past compliance records or their potential emissions. If firms face a targeted enforcement strategy where those with higher potential emissions are monitored more closely, a plausible response may be to adopt a new and more efficient abatement technology that allows them to reduce potential emissions and thus avoid more stringent monitoring. Using a four-group targeting scheme (denoted targeted state-dependent enforcement), we have analyzed the effects of an audit framework where targeting is based only on firms' past compliance record but also on adoption of environmentally superior technologies.

The results suggest that targeted state-dependent enforcement has a deterrent effect and can help reduce total enforcement costs. Firstly, it changes the composition of firms in the industry toward an increased fraction of cleaner firms that pollute and violate less. Secondly, it reduces the minimum monitoring probability required to ensure compliance by adopters. Finally, it provides non-adopters with stronger incentives to comply since surveillance resources are targeted more heavily to non-adopters.

The fact that the technology adoption rate is influenced by monitoring strategy is good news for a regulator who wants to achieve a given level of aggregate emissions but has political constraints on the level of the emission standard to be imposed. Such a regulator may use a differentiated monitoring strategy to induce technology adoption and thereby reduce aggregate emissions for a given politically feasible emission standard. Consequently, targeted monitoring strategies should not be ruled out as a plausible enforcement policy if the interaction between monitoring probabilities and technology adoption is taken into consideration.

References

[1] Aldy, J. and R. Stavins. 2007. Architectures for Agreement, Addressing Global Climate Change in the Post-Kyoto World, Cambridge Press.

[2] Arguedas, C. (2013). Pollution standards, technology investment and fines for non-compliance. *Journal of Regulatory Economics* 44(2): 156-176.

[3] Cason, T.N. and Gangadharan, L. (2006). An experimental study of compliance and leverage in auditing and regulatory enforcement. *Economic Inquiry* 44: 352-366.

[4] Clark, J., Friesen, L. and Muller, A. (2004). The good, the bad, and the regulator: An experimental test of two conditional audit schemes. *Economic Inquiry* 42: 69-87.

[5] Coria, J. and C. Villegas-Palacio (2014). Regulatory Dealing: Technology Adoption vs. Enforcement Stringency of Emission Taxes. *Contemporary Economic Policy* 32(2): 451-473.

[6] Friesen, L. (2003). Targeting enforcement to improve compliance with environmental regulations. *Journal of Environmental Economics and Management* 46: 72-85.

[7] Garvie D. and A. Keeler. (1994). Incomplete enforcement with endogenous regulatory choice. *Journal of Public Economics* 55(1): 141-161.

[8] Gilpatric, S., C. Vossler and L. Liu (2011). Using competition to stimulate regulatory compliance: a tournament-based dynamic targeting mechanism, Working Paper.

[9] Gray, W.B. and J.P. Shimshack (2011). The Effectiveness of Environmental Monitoring and Enforcement: A Review of the Empirical Evidence. *Review of Environmental Economics and Policy* 5(1): 3-24.

[10] Harford, J.D. (1991). Measurement Error and State-Dependent Pollution Control Enforcement. *Journal of Environmental Economics and Management* 21: 67-81.

[11] Harford, J.D. and W. Harrington (1991). A Reconsideration of Enforcement Leverage when Penalties are Restricted. *Journal of Public Economics* 45(3), 391-395.

[12] Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics* 37: 29–53.

[13] Helland, E. (1998). The Enforcement of Pollution Control Laws: Inspections, Violations and Self-Reporting. *The Review of Economics and Statistics* 80(1): 141-153.

[14] Hentschel, E. and A. Randall (2000). An Integrated Strategy to Reduce Monitoring and Enforcement Costs. *Environmental and Resource Economics* 15: 57-74.

[15] Heyes, A. and N. Rickman (1999). Regulatory Dealing — Revisiting the Harrington Paradox. *Journal of Public Economics* 72: 361-378.

[16] Liu, L. and W. Neilson (2009). Enforcement leverage with fixed inspection capacity. Working Paper, Department of Economics and International Business, Sam Houston State University.

[17] Livernois, J. and C. J. McKenna (1999). Truth or Consequences: Enforcing Pollution Standards with Self-Reporting. *Journal of Public Economics* 71(3): 415 - 440.

[18] Maloney, M.T. and G.L. Brady. 1988. Capital turnover and marketable pollution rights. *Journal of Environmental Law and Economics* 31(4): 203-226.

[19] Millock, K., D. Sunding and D. Zilberman. 2002. Regulating Pollution with Endogenous Monitoring. *Journal of Environmental Economics and Management* 44(2): 221-241.

[20] Millock, K., A. Xabadia and D. Zilberman. 2002. Policy for the adoption of new environmental monitoring technologies to manage stock externalities. *Journal of Environmental Economics and Management* 64(1): 102-116.

[21] Nyborg, K. and K. Telle (2006). Firms' compliance to environmental regulation: Is there really a paradox? *Environmental and Resource Economics* 35(1): 1-18.

[22] Nyborg, K. and K. Telle (2004). The Role of Warnings in Regulation: Keeping Control with Less Punishment. *Journal of Public Economics* 88(12): 2801-2816.

[23] Raymond, M. (1999). Enforcement leverage when penalties are restricted: a reconsideration under asymmetric information. *Journal of Public Economics* 73: 289-295.

[24] Requate, T. and W. Unold (2001). On the incentives created by policy instruments to adopt advanced abatement technology if firms are asymmetric. *Journal of Institutional and Theoretical Economics* 157: 536-554.

[25] Stoneman, P. and N.J. Ireland (1983). The Role of Supply Factors in the Diffusion of New Process Technology. *The Economic Journal* 93 Supplement: Conference Papers 66-78.

Appendix A

The critical probability $\bar{\pi}_2^A$ is determined by equation (5), which defines an implicit function

$$f(\bar{\pi}_2^A, \pi_1, \alpha_A, \gamma_A) = \theta[c(\bar{q}) - c(q_A^{NC}(\bar{\pi}_2^A))] - \Gamma - \bar{\pi}_2^A \varphi(q_A^{NC}(\bar{\pi}_2^A) - \bar{q}) = 0,$$

where

$$\Gamma = \beta \bar{\pi}_2^A \gamma_A \frac{\left[\theta\left[c(\bar{q}) - c(q_A^{NC}(\pi_1)\right] - \pi_1 \phi(q_A^{NC}(\pi_1) - \bar{q})\right]}{\left[1 - \beta + \beta\pi_1\alpha_A + \beta\bar{\pi}_2^A\gamma_A\right]} > 0 \text{ if } \pi_1 < \bar{\pi}_1^A.$$

By the implicit function theorem, we have:

$$\frac{\partial \bar{\pi}_2^A}{\partial \pi_1} = -\frac{\partial f / \partial \pi_1}{\partial f / \partial \bar{\pi}_2^A}.$$

Differentiating $f(.)$ with respect to π_1 yields:

$$\frac{\partial f}{\partial \pi_1} = -\frac{\partial \Gamma}{\partial \pi_1} = \frac{[1 - \beta + \beta\bar{\pi}_2^A\gamma_A]\beta\bar{\pi}_2^A\gamma_A\varphi(q_A^{NC}(\pi_1) - \bar{q}) + \beta\alpha_A\beta\bar{\pi}_2^A\gamma_A\left[\theta[c(\bar{q}) - c(q_A^{NC}(\pi_1))]\right]}{[1 - \beta + \beta\pi_1\alpha_A + \beta\bar{\pi}_2^A\gamma_A]^2} > 0.$$

Differentiating $f(.)$ with respect to π_2 yields:

$$\frac{\partial f}{\partial \bar{\pi}_2^A} = -\varphi(q_A^{NC}(\bar{\pi}_2^A) - \bar{q}) - (1 - \beta + \beta\pi_1\alpha_A)\frac{\Gamma}{\bar{\pi}_2^A(1 - \beta + \beta\pi_1\alpha_A + \beta\bar{\pi}_2^A\gamma_A)} < 0.$$

Hence, $\frac{\partial \bar{\pi}_2^A}{\partial \pi_1} > 0$.

By analogy, we differentiate $f(.)$ with respect to the transition probabilities α_A and γ_A, which yields:

$$\begin{aligned} \frac{\partial f}{\partial \alpha_A} &= -\frac{\partial \Gamma}{\partial \alpha_A} = \frac{\beta\pi_1\Gamma}{1 - \beta + \beta\pi_1\alpha_A + \beta\bar{\pi}_2^A\gamma_A} > 0, \\ \frac{\partial f}{\partial \gamma_A} &= -\frac{\partial \Gamma}{\partial \gamma_A} == -\frac{(1 - \beta + \beta\pi_1\alpha_A)\Gamma}{\gamma_A(1 - \beta + \beta\pi_1\alpha_A + \beta\bar{\pi}_2^A\gamma_A)} < 0. \end{aligned}$$

Hence, $\frac{\partial \bar{\pi}_2^A}{\partial \alpha_A} = -\frac{\partial f / \partial \alpha_A}{\partial f / \partial \bar{\pi}_2^A} > 0$ and $\frac{\partial \bar{\pi}_2^A}{\partial \gamma_A} = -\frac{\partial f / \partial \gamma_A}{\partial f / \partial \bar{\pi}_2^A} < 0$.

Appendix B

Let us compute the derivatives of adopters' and non-adopters' expected costs of compliance with respect to the probabilities (α_A, γ_A) and $(\alpha_{NA}, \gamma_{NA})$.

Effects of (α_A, γ_A)

As shown in Table 2, regardless of the *initial allocation*, the transition probabilities (α_A, γ_A) only affect adopters' expected costs of compliance. Since under the strategies f^{0000}, f^{0010} and f^{0011} adopters already comply, decreasing α_A or increasing γ_A has no effect on emissions, abatement or rate of adoption. Thus , $\frac{\partial E_A^{0000}(y)}{\partial \alpha_A} = \frac{\partial E_A^{0010}(y)}{\partial \alpha_A} = \frac{\partial E_A^{0011}(y)}{\partial \alpha_A} = 0$, and $\frac{\partial E_A^{0000}(y)}{\partial \gamma_A} = \frac{\partial E_A^{0010}(y)}{\partial \gamma_A} = \frac{\partial E_A^{0011}(y)}{\partial \gamma_A} = 0 \vee y = 1, 2$.

Instead, if the strategies f^{1010} or f^{1011} are optimal, the derivatives $\frac{\partial E_A^{1010}(y)}{\partial \alpha_A}$ and $\frac{\partial E_A^{1011}(y)}{\partial \alpha_A}$ are the same, positive, and given by:

$$\begin{aligned}
\frac{\partial E_A^{1010}(1)}{\partial \alpha_A} &= \beta\pi_1 \left[1-\beta+\beta\pi_2\gamma_A\right] \frac{\left[\theta c(\overline{q}) - \left[\theta c(q_A^{NC}(\pi_1)) + \pi_1\phi(q_A^{NC}(\pi_1) - \overline{q})\right]\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A\right]^2} > 0, \\
\frac{\partial E_A^{1010}(2)}{\partial \alpha_A} &= \frac{\beta\pi_2\gamma_A}{1-\beta+\beta\pi_2\gamma_A} \frac{\partial E_A^{1010}(1)}{\partial \alpha_A} > 0.
\end{aligned}$$

If f^{1111} is optimal, $\frac{\partial E_A^{1111}(2)}{\partial \alpha_A} = 0$. In constrast, $\frac{\partial E_A^{1111}(1)}{\partial \alpha_A} > 0$ and corresponds to:

$$\frac{\partial E_A^{1111}(1)}{\partial \alpha_A} = \beta\pi_1 \frac{\left[\left[\theta c(q_A^{NC}(\pi_2)) + \pi_2\phi(q_A^{NC}(\pi_2) - \overline{q})\right] - \left[\theta c(q_{NA}^{NC}(\pi_1)) + \pi_1\phi(q_{NA}^{NC}(\pi_1) - \overline{q})\right]\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A\right]^2} > 0.$$

The derivatives $\frac{\partial E_A^{1010}(y)}{\partial \gamma_A}$ and $\frac{\partial E_A^{1011}(y)}{\partial \gamma_A}$ are the same, negative, and given by:

$$\begin{aligned}
\frac{\partial E_A^{1010}(1)}{\partial \gamma_A} &= -\frac{\beta\pi_2\alpha_A}{1-\beta+\beta\pi_2\gamma_A} \frac{\partial E_A^{1010}(1)}{\partial \alpha_A} < 0, \\
\frac{\partial E_A^{1010}(2)}{\partial \gamma_A} &= -\frac{\pi_2\left[1-\beta+\beta\pi_1\alpha_A\right]}{\beta\pi_1\gamma_A} \frac{\partial E_A^{1010}(2)}{\partial \alpha_A} < 0.
\end{aligned}$$

If policy f^{1111} is optimal, $\frac{\partial E_A^{1111}(y)}{\partial \gamma_A} = 0, \vee y = 1, 2$.

In sum, from the analysis above it is clear that $\frac{\partial E_A^{jklm}(y)}{\partial \alpha_A} \geq 0$ and $\frac{\partial E_A^{jklm}(y)}{\partial \gamma_A} \leq 0$.

Effects of $(\alpha_{NA}, \gamma_{NA})$

Note that regardless of the *initial allocation*, the transition probabilities $(\alpha_{NA}, \gamma_{NA})$ only affect non-adopters' expected cost of compliance. Since under the policy f^{0000} non-adopters already comply, increasing α_{NA} or decreasing γ_{NA} has no effect on the rate of adoption. Thus, $\frac{\partial E_{NA}^{0000}(y)}{\partial \alpha_{NA}} = \frac{\partial E_{NA}^{0000}(y)}{\partial \gamma_{NA}} = 0 \vee y = 1, 2.$

If the strategy f^{0010} or f^{1010} is optimal, the derivatives $\frac{\partial E_{NA}^{0010}(y)}{\partial \alpha_{NA}}and\frac{\partial E_{NA}^{1010}(y)}{\partial \alpha_{NA}}$ are the same, positive and given by:

$$\begin{aligned}\frac{\partial E_{NA}^{0010}(1)}{\partial \alpha_{NA}} &= \beta\pi_1\left[1-\beta+\beta\pi_2\gamma_{NA}\right]\frac{\left[c(\overline{q})-\left[c(q_{NA}^{NC}(\pi_1))+\pi_1\phi(q_{NA}^{NC}(\pi_1)-\overline{q})\right]\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}\right]^2} > 0,\\ \frac{\partial E_{NA}^{0010}(2)}{\partial \alpha_{NA}} &= \frac{\beta\pi_2\gamma_{NA}}{1-\beta+\beta\pi_2\gamma_{NA}}\frac{\partial E_{NA}^{0010}(1)}{\partial \alpha_{NA}} > 0.\end{aligned}$$

If the strategy f^{0011} , f^{1011} or f^{1111} is optimal, $\frac{\partial E_{NA}^{0011}(y)}{\partial \alpha_{NA}}$, $\frac{\partial E_{NA}^{1011}(y)}{\partial \alpha_{NA}}$ and$\frac{\partial E_{NA}^{1111}(y)}{\partial \alpha_{NA}}$ are the same, and equal to zero if the firms are initially allocated to G_2. In contrast, the derivatives $\frac{\partial E_{NA}^{0011}(1)}{\partial \alpha_{NA}}$, $\frac{\partial E_{NA}^{1011}(1)}{\partial \alpha_{NA}}$, and$\frac{\partial E_{NA}^{1111}(1)}{\partial \alpha_{NA}}$ are positive and given by:

$$\frac{\partial E_{NA}^{0011}(1)}{\partial \alpha_{NA}} = \beta\pi_1\frac{\left[\left[c(q_{NA}^{NC}(\pi_2))+\pi_2\phi(q_{NA}^{NC}(\pi_2)-\overline{q})\right]-\left[c(q_{NA}^{NC}(\pi_1))+\pi_1\phi(q_{NA}^{NC}(\pi_1)-\overline{q})\right]\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}\right]^2} > 0.$$

The derivatives $\frac{\partial E_{NA}^{0010}(y)}{\partial \gamma_{NA}}and\frac{\partial E_{NA}^{1010}(y)}{\partial \gamma_{NA}}$ are the same, negative, and given by:

$$\begin{aligned}\frac{\partial E_{NA}^{0010}(1)}{\partial \gamma_{NA}} &= -\frac{\beta\pi_2\alpha_{NA}}{1-\beta+\beta\pi_2\gamma_{NA}}\frac{\partial E_{NA}^{0010}(1)}{\partial \alpha_{NA}} < 0,\\ \frac{\partial E_{NA}^{0010}(2)}{\partial \gamma_{NA}} &= -\frac{1-\beta+\beta\pi_1\alpha_{NA}}{\beta\pi_1\gamma_{NA}}\frac{\partial E_{NA}^{0010}(2)}{\partial \alpha_{NA}} < 0.\end{aligned}$$

By analogy, $\frac{\partial E_{NA}^{0011}(y)}{\partial \gamma_{NA}}$, $\frac{\partial E_{NA}^{1011}(y)}{\partial \gamma_{NA}}$, and$\frac{\partial E_{NA}^{1111}(y)}{\partial \gamma_{NA}}$ are the same, and equal to zero.

In sum, from the analysis above it is clear that $\frac{\partial E_{NA}^{jklm}(y)}{\partial \alpha_{NA}} \geq 0$ and $\frac{\partial E_{NA}^{jklm}(y)}{\partial \gamma_{NA}} \leq 0$.

Marginal Variations in Transition Probabilities

From the analysis above, it follows that $\left|\frac{\partial E_{NA}^{jklm}(y)}{\partial \alpha_{NA}}\right| \geq \left|\frac{\partial E_{A}^{jklm}(y)}{\partial \alpha_{A}}\right|$ and $\left|\frac{\partial E_{NA}^{jklm}(y)}{\partial \gamma_{NA}}\right| \geq \left|\frac{\partial E_{A}^{jklm}(y)}{\partial \gamma_{A}}\right|$ $\vee$ $jklm \neq 1011$, implying that at the margin, variations on (α_{NA}, γ_{NA}) have a larger effect on the rate of adoption than do marginal variations in (α_A, γ_A).

Moreover, it follows that $\left|\frac{\partial E_{A}^{jklm}(1)}{\partial \alpha_{A}}\right| \geq \left|\frac{\partial E_{A}^{jklm}(2)}{\partial \alpha_{A}}\right|$ and $\left|\frac{\partial E_{NA}^{jklm}(1)}{\partial \alpha_{NA}}\right| \geq \left|\frac{\partial E_{NA}^{jklm}(2)}{\partial \alpha_{NA}}\right|$, implying that the marginal effects of (α_A, α_{NA}) on the rate of adoption are larger if firms are initially allocated to G_1. The reverse holds for (γ_A, γ_{NA}), where $\left|\frac{\partial E_{A}^{jklm}(2)}{\partial \gamma_{A}}\right| \geq \left|\frac{\partial E_{A}^{jklm}(1)}{\partial \gamma_{A}}\right|$, and $\left|\frac{\partial E_{NA}^{jklm}(2)}{\partial \gamma_{NA}}\right| \geq \left|\frac{\partial E_{NA}^{jklm}(1)}{\partial \gamma_{NA}}\right|$ indicating that their marginal effects on the rate of adoption are larger if firms are initially allocated to G_2.

Appendix C

Let us compute the derivatives of adopters' and non-adopters' expected emissions with respect to the probabilities (α_A, γ_A) and $(\alpha_{NA}, \gamma_{NA})$. As shown in Table 3, regardless of the *initial allocation*, the transition probabilities (α_A, γ_A) only affect adopters' expected costs of compliance. Since under the strategies f^{0000}, f^{0010} and f^{0011} adopters already comply, decreasing α_A or increasing γ_A has no effect on emissions, abatement or rate of adoption. In contrast, if the strategies f^{1010} or f^{1011} are optimal, the derivatives $\frac{\partial \widehat{q}_A^{1010}(y)}{\partial \alpha_A}$ and $\frac{\partial \widehat{q}_A^{1011}(y)}{\partial \alpha_A}$ are the same $\vee$ $y = 1, 2$, negative, and given by:

$$\begin{aligned}
\frac{\partial \widehat{q}_A^{1010}(1)}{\partial \alpha_A} &= \frac{\beta\pi_1 \left[1-\beta+\beta\pi_2\gamma_A\right]\left[\overline{q}-q_A^{NC}(\pi_1)\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_A+\beta\pi_2\gamma_A\right]^2} < 0, \\
\frac{\partial \widehat{q}_A^{1010}(2)}{\partial \alpha_A} &= \frac{\beta\pi_2\gamma_A}{1-\beta+\beta\pi_2\gamma_A}\frac{\partial \widehat{q}_A^{1010}(1)}{\partial \alpha_A} < 0.
\end{aligned}$$

If the strategy f^{1111} is optimal, $\frac{\partial \widehat{q}_A^{1111}(2)}{\partial \alpha_A} = 0$. In contrast $\frac{\partial \widehat{q}_A^{1111}(1)}{\partial \alpha_A}$ is given by:

$$\frac{\partial \widehat{q}_A^{1111}(1)}{\partial \alpha_A} = \frac{\beta\pi_1\left[q_A^{NC}(\pi_2)-q_A^{NC}(\pi_1)\right]}{\left[1-\beta+\beta\pi_1\alpha_A\right]^2} < 0.$$

The derivatives $\frac{\partial \widehat{q}_A^{1010}(y)}{\partial \gamma_A}$ and $\frac{\partial \widehat{q}_A^{1011}(y)}{\partial \gamma_A}$ are the same $\vee$ $y = 1, 2$, positive, and given by:

$$\begin{aligned}
\frac{\partial \widehat{q}_A^{1010}(1)}{\partial \gamma_A} &= \frac{\beta\pi_2\alpha_A}{1-\beta+\beta\pi_2\gamma_A}\frac{\partial \widehat{q}_A^{1010}(1)}{\partial \alpha_A} > 0, \\
\frac{\partial \widehat{q}_A^{1010}(2)}{\partial \gamma_A} &= \frac{\pi_2\left[1-\beta+\beta\pi_1\alpha_A\right]}{\pi_1\left[1-\beta+\beta\pi_2\gamma_A\right]}\frac{\partial \widehat{q}_A^{1010}(1)}{\partial \alpha_A} > 0.
\end{aligned}$$

If the policy f^{1111} is optimal, $\frac{\partial \widehat{q}_A^{1010}(2)}{\partial \gamma_A} = 0$, $\vee$ $y = 1, 2$.

Regarding the transition probabilities $(\alpha_{NA}, \gamma_{NA})$, since under f^{0000} non-adopters already comply, increasing α_{NA} or decreasing γ_{NA} has no effect on the rate of adoption. If the strategy f^{0010} or f^{1010} is optimal, the derivatives $\frac{\partial \widehat{q}_{NA}^{0010}(y)}{\partial \alpha_{NA}}$ and $\frac{\widehat{q}_{NA}^{1010}(y)}{\partial \alpha_{NA}}$ are the same $\vee$ $y = 1, 2$, positive, and given by:

$$\begin{aligned}
\frac{\partial \widehat{q}_{NA}^{0010}(1)}{\partial \alpha_{NA}} &= \frac{\beta\pi_1\left[1-\beta+\beta\pi_2\gamma_{NA}\right]\left[\overline{q}-q_{NA}^{NC}(\pi_1)\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}\right]^2} < 0, \\
\frac{\partial \widehat{q}_{NA}^{0010}(2)}{\partial \alpha_{NA}} &= \frac{\beta\pi_2\gamma_{NA}}{1-\beta+\beta\pi_2\gamma_{NA}}\frac{\partial \widehat{q}_{NA}^{0010}(1)}{\partial \alpha_{NA}} < 0.
\end{aligned}$$

If the strategy f^{0011} , f^{1011} or f^{1111} is optimal, $\frac{\partial\widehat{q}_{NA}^{0011}(y)}{\partial\alpha_{NA}}$, $\frac{\partial\widehat{q}_{NA}^{1011}(y)}{\partial\alpha_{NA}}$ and $\frac{\partial\widehat{q}_{NA}^{1111}(y)}{\partial\alpha_{NA}}$ are the same $\vee$ $y = 1, 2$ and equal to zero if the firms are initially allocated to G_2. In contrast, the derivatives $\frac{\partial\widehat{q}_{NA}^{0011}(1)}{\partial\alpha_{NA}}$, $\frac{\partial\widehat{q}_{NA}^{1011}(1)}{\partial\alpha_{NA}}$ and $\frac{\partial\widehat{q}_{NA}^{1111}(1)}{\partial\alpha_{NA}}$ are negative and given by:

$$\frac{\partial\widehat{q}_{NA}^{0011}(1)}{\partial\alpha_{NA}} = \frac{\beta\pi_1\left[q_{NA}^{NC}(\pi_2) - q_{NA}^{NC}(\pi_1)\right]}{\left[1-\beta+\beta\pi_1\alpha_{NA}\right]^2} < 0.$$

The derivatives $\frac{\partial\widehat{q}_{NA}^{0010}(y)}{\partial\gamma_{NA}}$ and $\frac{\partial\widehat{q}_{NA}^{1010}(y)}{\partial\gamma_{NA}}$ are the same $\vee$ $y = 1, 2$, positive, and given by:

$$\begin{aligned}\frac{\partial\widehat{q}_{NA}^{0010}(1)}{\partial\gamma_{NA}} &= -\frac{\beta\pi_2\alpha_{NA}}{1-\beta+\beta\pi_2\gamma_{NA}}\frac{\partial\widehat{q}_{NA}^{0010}(1)}{\partial\alpha_{NA}} > 0,\\ \frac{\partial\widehat{q}_{NA}^{0010}(2)}{\partial\gamma_{NA}} &= \frac{\pi_2\left[1-\beta+\beta\pi_1\alpha_{NA}\right]}{\pi_1\left[1-\beta+\beta\pi_2\gamma_{NA}\right]}\frac{\partial\widehat{q}_{NA}^{0010}(1)}{\partial\alpha_{NA}} > 0.\end{aligned}$$

By analogy, $\frac{\partial\widehat{q}_{NA}^{0011}(y)}{\partial\gamma_{NA}}$, $\frac{\partial\widehat{q}_{NA}^{1011}(y)}{\partial\gamma_{NA}}$, and $\frac{\partial\widehat{q}_{NA}^{1111}(y)}{\partial\gamma_{NA}}$ are the same $\vee$ $y = 1, 2$ and equal to zero.

Marginal Variations in Transition Probabilities

From the analysis above it follows:

- f^{0000}. We have that $\frac{\partial\widehat{q}_A(y)}{\partial\alpha_A} = \frac{\partial\widehat{q}_A(y)}{\partial\gamma_A} = \frac{\partial\widehat{q}_{NA}(y)}{\partial\alpha_{NA}} = \frac{\partial\widehat{q}_{NA}(y)}{\partial\gamma_{NA}} = 0$ $\vee$ $y = 1, 2$. Hence, targeted state-dependent enforcement has no effect adopters' and non-adopters' emissions.

- f^{0010}. We have that $\frac{\partial\widehat{q}_A(y)}{\partial\alpha_A} = \frac{\partial\widehat{q}_A(y)}{\partial\gamma_A} = 0$ $\vee$ $y = 1, 2$. In contrast, $\frac{\partial\widehat{q}_{NA}(y)}{\partial\alpha_{NA}} < 0 < \frac{\partial\widehat{q}_{NA}(y)}{\partial\gamma_{NA}}$ $\vee$ $y = 1, 2$. Hence, targeted state-dependent enforcement has no effect on adopters' emissions but reduces non-adopters' emissions.

- f^{0011}. We have that $\frac{\partial\widehat{q}_A(y)}{\partial\alpha_A} = \frac{\partial\widehat{q}_A(y)}{\partial\gamma_A} = 0$ $\vee$ $y = 1, 2$. In contrast, $\frac{\partial\widehat{q}_{NA}(1)}{\partial\alpha_{NA}} < 0 = \frac{\partial\widehat{q}_{NA}(1)}{\partial\gamma_{NA}}$, while $\frac{\partial\widehat{q}_{NA}(2)}{\partial\alpha_{NA}} = \frac{\partial\widehat{q}_{NA}(2)}{\partial\gamma_{NA}} = 0$. Hence, targeted state-dependent enforcement has no effect on adopters' or non-adopters' emissions if firms are initially allocated to G_2. However, it reduces non-adopters' emissions if they are initially allocated to G_1.

- f^{1010}. We have that $\frac{\partial\widehat{q}_A(y)}{\partial\alpha_A} < 0 < \frac{\partial\widehat{q}_A(y)}{\partial\gamma_A}$ and $\frac{\partial\widehat{q}_{NA}(y)}{\partial\alpha_{NA}} < 0 < \frac{\partial\widehat{q}_{NA}(y)}{\partial\gamma_{NA}}$ $\vee$ $y = 1, 2$. Moreover, $\left|\frac{\partial\widehat{q}_{NA}(y)}{\partial\alpha_{NA}}\right| > \left|\frac{\partial\widehat{q}_A(x)}{\partial\alpha_A}\right|$ and $\left|\frac{\partial\widehat{q}_{NA}(y)}{\partial\gamma_{NA}}\right| > \left|\frac{\partial\widehat{q}_A(y)}{\partial\gamma_A}\right|$ $\vee$ $y, x = 1, 2$ and $y \geq x$, implying that the marginal effects of the probabilities α_{NA} and γ_{NA} are larger than the marginal effects of the probabilities α_A and γ_A.

- f^{1011}. We have that $\frac{\partial \widehat{q}_A(y)}{\partial \alpha_A} < 0 < \frac{\partial \widehat{q}_A(y)}{\partial \gamma_A} \vee\ y = 1, 2$. In contrast, $\frac{\partial \widehat{q}_{NA}(1)}{\partial \alpha_A} < \frac{\partial \widehat{q}_{NA}(2)}{\partial \alpha_A} = 0$ and $\frac{\partial \widehat{q}_{NA}(y)}{\partial \gamma_{NA}} = 0 \vee y = 1, 2$. Moreover, $\left|\frac{\partial \widehat{q}_A(y)}{\partial \alpha_A}\right| \geq \left|\frac{\partial \widehat{q}_{NA}(y)}{\partial \alpha_{NA}}\right|$ and $\left|\frac{\partial \widehat{q}_A(y)}{\partial \gamma_A}\right| > \left|\frac{\partial \widehat{q}_{NA}(y)}{\partial \gamma_{NA}}\right| \vee y, x = 1, 2$ and $y \geq x$, implying that the the marginal effects of the probabilities α_A and γ_A are larger than the marginal effects of the probabilities α_{NA} and γ_{NA}.

- f^{1111}. We have that $\frac{\partial \widehat{q}_A(1)}{\partial \alpha_A} < \frac{\partial \widehat{q}_A(2)}{\partial \alpha_A} = 0$ and $\frac{\partial \widehat{q}_A(y)}{\partial \gamma_A} = 0 \vee y = 1, 2$. By analogy, $\frac{\partial \widehat{q}_{NA}(1)}{\partial \alpha_A} < \frac{\partial \widehat{q}_{NA}(2)}{\partial \alpha_A} = 0$ and $\frac{\partial \widehat{q}_{NA}(y)}{\partial \gamma_{NA}} = 0 \vee y = 1, 2$. Moreover, $\left|\frac{\partial \widehat{q}_{NA}(1)}{\partial \alpha_{NA}}\right| > \left|\frac{\partial \widehat{q}_A(1)}{\partial \alpha_A}\right|$, implying that the marginal effect of the probability α_{NA} is larger than the marginal effects of the probability α_A when firms are initially allocated to G_1.

Appendix D

Let us compute the derivatives of adopters' and non-adopters' expected enforcement cost with respect to the probabilities of transition (α_A, γ_A) and $(\alpha_{NA}, \gamma_{NA})$. As shown in Table 4, regardless of the *initial allocation*, the transition probabilities (α_A, γ_A) only affect the cost of enforcing compliance among adopters. We have that $\frac{\partial \widehat{m}_A^{0000}(y)}{\partial \alpha_A} = \frac{\partial \widehat{m}_A^{0010}(y)}{\partial \alpha_A} = \frac{\partial \widehat{m}_A^{0011}(y)}{\partial \alpha_A} \vee y = 1, 2$. Moreover, the derivatives $\frac{\partial \widehat{m}_A^{0000}(y)}{\partial \gamma_A}, \frac{\partial \widehat{m}_A^{0010}(y)}{\partial \gamma_A}$, and $\frac{\partial \widehat{m}_A^{0011}(y)}{\partial \gamma_A}$ are the same $\vee y = 1, 2$. If adopters are initially allocated to G_1, they are equal to zero. If adopters are initially allocated to G_2, they are negative and given by:

$$\frac{\partial \widehat{m}_A^{0000}(2)}{\partial \gamma_A} = -\frac{\beta m \pi_2 [\pi_2 - \pi_1]}{[1 - \beta + \beta \pi_2 \gamma_A]^2} < 0.$$

If strategy f^{1010} or f^{1011} is optimal, the derivatives $\frac{\partial \widehat{m}_A^{1010}(y)}{\partial \alpha_A}$ and $\frac{\partial \widehat{m}_A^{1011}(y)}{\partial \alpha_A}$ are the same $\vee y = 1, 2$, positive, and given by

$$\begin{aligned}
\frac{\partial \widehat{m}_A^{1010}(1)}{\partial \alpha_A} &= \frac{\beta m \pi_1 [1 - \beta + \beta \pi_2 \gamma_A] [\pi_2 - \pi_1]}{[1 - \beta] [1 - \beta + \beta \pi_1 \alpha_A + \beta \pi_2 \gamma_A]^2} > 0, \\
\frac{\partial \widehat{m}_A^{1010}(2)}{\partial \alpha_A} &= \frac{\beta \pi_2 \gamma_A}{1 - \beta + \beta \pi_2 \gamma_A} \frac{\partial \widehat{m}_A^{0010}(1)}{\partial \alpha_A} > 0.
\end{aligned}$$

The derivatives $\frac{\partial \widehat{m}_A^{1010}(y)}{\partial \gamma_A}$ and $\frac{\partial \widehat{m}_A^{1011}(y)}{\partial \gamma_A}$ are the same $\vee y = 1, 2$, negative, and given by

$$\begin{aligned}
\frac{\partial \widehat{m}_A^{1010}(1)}{\partial \gamma_A} &= -\frac{\beta \pi_2 \alpha_A}{1 - \beta + \beta \pi_2 \gamma_A} \frac{\partial \widehat{m}_A^{0010}(1)}{\partial \alpha_A} < 0, \\
\frac{\partial \widehat{m}_A^{1010}(2)}{\partial \gamma_A} &= -\frac{\beta \pi_1 \gamma_A}{1 - \beta + \beta \pi_1 \alpha_A} \frac{\partial \widehat{m}_A^{0010}(2)}{\partial \alpha_A} < 0.
\end{aligned}$$

If the strategy f^{1111} is optimal, $\frac{\partial \widehat{m}_A^{1111}(2)}{\partial \alpha_A} = 0$. In contrast,

$$\frac{\partial \widehat{m}_A^{1111}(1)}{\partial \alpha_A} = \frac{\beta m \pi_1 [\pi_2 - \pi_1]}{[1 - \beta + \beta \pi_1 \alpha_A]^2} > 0.$$

Moreover, $\frac{\partial \widehat{m}_A^{1111}(2)}{\partial \gamma_A} = 0, \vee y = 1, 2$.

Regarding the transition probabilities $(\alpha_{NA}, \gamma_{NA})$, we have that $\frac{\partial \widehat{m}_{NA}^{0000}(y)}{\partial \alpha_{NA}} = 0 \vee y = 1, 2$. Moreover, the derivatives $\frac{\partial \widehat{m}_{NA}^{0000}(y)}{\partial \gamma_{NA}}$ are the same $\vee y = 1, 2$. If non-adopters are initially allocated to G_1, they are equal to zero. If non-adopters are initially allocated to G_2, they are negative and given by:

$$\frac{\partial \widehat{m}_{NA}^{0000}(2)}{\partial \gamma_{NA}} = -\frac{\beta m \pi_2 [\pi_2 - \pi_1]}{[1 - \beta + \beta \pi_2 \gamma_{NA}]^2} < 0.$$

If the strategy f^{0010} or f^{1010} is optimal, the derivatives $\frac{\partial \widehat{m}_{NA}^{0010}(y)}{\partial \alpha_{NA}}$ and $\frac{\widehat{m}_{NA}^{1010}(y)}{\partial \alpha_{NA}}$ are the same $\vee$ $y = 1, 2$, positive, and given by:

$$\begin{aligned}
\frac{\partial \widehat{m}_{NA}^{0010}(1)}{\partial \alpha_{NA}} &= \frac{\beta m \pi_1 \left[1-\beta+\beta\pi_2\gamma_{NA}\right]\left[\pi_2-\pi_1\right]}{\left[1-\beta\right]\left[1-\beta+\beta\pi_1\alpha_{NA}+\beta\pi_2\gamma_{NA}\right]^2} > 0, \\
\frac{\partial \widehat{m}_{NA}^{0010}(2)}{\partial \alpha_{NA}} &= \frac{\beta\pi_2\gamma_{NA}}{1-\beta+\beta\pi_2\gamma_{NA}}\frac{\partial \widehat{m}_{NA}^{0010}(1)}{\partial \alpha_{NA}} > 0.
\end{aligned}$$

The derivatives $\frac{\partial \widehat{m}_{NA}^{0010}(y)}{\partial \gamma_{NA}}and\frac{\partial \widehat{m}_{NA}^{1010}(y)}{\partial \gamma_{NA}}$ are the same $\vee$ $y = 1, 2$, negative and given by:

$$\begin{aligned}
\frac{\partial \widehat{m}_{NA}^{0010}(1)}{\partial \gamma_{NA}} &= -\frac{\beta\pi_2\alpha_{NA}}{1-\beta+\beta\pi_2\gamma_{NA}}\frac{\partial \widehat{m}_{NA}^{0010}(1)}{\partial \alpha_{NA}} < 0, \\
\frac{\partial \widehat{m}_{NA}^{0010}(2)}{\partial \gamma_{NA}} &= -\frac{\beta\pi_1\gamma_{NA}}{1-\beta+\beta\pi_1\alpha_{NA}}\frac{\partial \widehat{m}_{NA}^{0010}(2)}{\partial \alpha_{NA}} < 0.
\end{aligned}$$

If the strategy f^{0011} , f^{1011} or f^{1111} is optimal, $\frac{\partial \widehat{m}_{NA}^{0011}(y)}{\partial \alpha_{NA}}$, $\frac{\partial \widehat{m}_{NA}^{1011}(y)}{\partial \alpha_{NA}}$ and$\frac{\partial \widehat{m}_{NA}^{1111}(y)}{\partial \alpha_{NA}}$ are the same $\vee$ $y = 1, 2$ and equal to zero if the firms are initially allocated to G_2. In contrast, the derivatives $\frac{\partial \widehat{m}_{NA}^{0011}(1)}{\partial \alpha_{NA}}$, $\frac{\partial \widehat{m}_{NA}^{1011}(1)}{\partial \alpha_{NA}}$ and$\frac{\partial \widehat{m}_{NA}^{1111}(1)}{\partial \alpha_{NA}}$ are positive, and given by:

$$\frac{\partial \widehat{m}_{NA}^{0011}(1)}{\partial \alpha_{NA}} = \frac{\beta m \pi_1 \left[\pi_2-\pi_1\right]}{\left[1-\beta+\beta\pi_1\alpha_{NA}\right]^2} > 0.$$

By analogy, $\frac{\partial \widehat{m}_{NA}^{0011}(y)}{\partial \gamma_{NA}}$, $\frac{\partial \widehat{m}_{NA}^{1011}(y)}{\partial \gamma_{NA}}$ and$\frac{\partial \widehat{m}_{NA}^{1111}(y)}{\partial \gamma_{NA}}$ are the same $\vee$ $y = 1, 2$ and equal to zero.

Marginal Variations in Transition Probabilities

From the analysis above it follows:

- f^{0000}. We have that $\frac{\partial \widehat{m}_A(y)}{\partial \alpha_A} = \frac{\partial \widehat{m}_{NA}(y)}{\partial \alpha_{NA}} = 0 \vee y = 1, 2$. Moreover, $\frac{\partial \widehat{m}_A(1)}{\partial \gamma_A} = \frac{\partial \widehat{m}_{NA}(1)}{\partial \gamma_{NA}} = 0$, while $\frac{\partial \widehat{m}_A(2)}{\partial \gamma_A} < \frac{\partial \widehat{m}_{NA}(2)}{\partial \gamma_{NA}} < 0$. Hence, targeted state-dependent enforcement has no effect in the cost of enforcing compliance among adopters and non-adopters if they are initially allocated to G_1. However, it reduces the enforcement cost for adopters and increases the enforcement cost for non-adopters if firms are initially allocated to G_2. Since the marginal effect of γ_A is larger than the marginal effect of γ_{NA}, the reduction in the enforcement cost for adopters is larger than the increase in the enforcement cost for non-adopters.

- f^{0010}. We have that $\frac{\partial \widehat{m}_A(y)}{\partial \alpha_A} = 0 \vee y = 1,2$. Moreover, $\frac{\partial \widehat{m}_A(1)}{\partial \gamma_A} = 0$, while $\frac{\partial \widehat{m}_A(2)}{\partial \gamma_A} < 0$. For non-adopters, $\frac{\partial \widehat{m}_{NA}(y)}{\partial \alpha_{NA}} > 0$ while $\frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}} < 0 \vee y = 1,2$. Hence, targeted state-dependent enforcement has no effect in the cost of enforcing compliance among adopters if they are initially allocated to G_1, while it reduces the enforcement cost if they are initially allocated to G_2. For non-adopters, it increases the enforcement cost for all initial allocations.

- f^{0011}. We have that $\frac{\partial \widehat{m}_A(y)}{\partial \alpha_A} = 0 \vee y = 1,2$. Moreover, $\frac{\partial \widehat{m}_A(1)}{\partial \gamma_A} = 0$, while $\frac{\partial \widehat{m}_A(2)}{\partial \gamma_A} < 0$. For non-adopters, $\frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}} = 0 \vee y = 1,2$, $\frac{\partial \widehat{m}_{NA}(2)}{\partial \alpha_{NA}} = 0$, while $\frac{\partial \widehat{m}_{NA}(1)}{\partial \alpha_{NA}} > 0$. Hence, targeted state-dependent enforcement has no effect in the cost of enforcing compliance among adopters if they are initially allocated to G_1, while it reduces the enfocement cost if they are initially allocated to G_2. For non-adopters, it increases the cost of enforcement if they are initially allocated to G_1.

- f^{1010}. We have that $\frac{\partial \widehat{m}_A(y)}{\partial \gamma_A} < 0 < \frac{\partial \widehat{m}_A(y)}{\partial \alpha_A}$ and $\frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}} < 0 < \frac{\partial \widehat{m}_{NA}(y)}{\partial \alpha_{NA}} \vee y = 1,2$. Moreover, $\left|\frac{\partial \widehat{m}_A(y)}{\partial \alpha_A}\right| > \left|\frac{\partial \widehat{m}_{NA}(x)}{\partial \alpha_{NA}}\right|$ and $\left|\frac{\partial \widehat{m}_A(y)}{\partial \gamma_A}\right| > \left|\frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}}\right| \vee y, x = 1,2$ and $y \geq x$, implying that the marginal effects of the probabilities α_A and γ_A are larger than the marginal effects of the probabilities α_{NA} and γ_{NA}. Hence, the reduction in the enforcement cost for adopters is larger than the increase in the enforcement cost for non-adopters.

- f^{1011}. We have that $\frac{\partial \widehat{m}_A(y)}{\partial \gamma_A} < 0 < \frac{\partial \widehat{m}_A(y)}{\partial \alpha_A} \vee y = 1,2$. In contrast, $\frac{\partial \widehat{m}_{NA}(1)}{\partial \alpha_A} > \frac{\partial \widehat{m}_{NA}(2)}{\partial \alpha_A} = 0$, and $\frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}} = 0 \vee y = 1,2$. Moreover, $\left|\frac{\partial \widehat{m}_A(y)}{\partial \alpha_A}\right| \geq \left|\frac{\partial \widehat{m}_{NA}(y)}{\partial \alpha_{NA}}\right|$ and $\left|\frac{\partial \widehat{m}_A(y)}{\partial \gamma_A}\right| > \left|\frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}}\right| \vee y, x = 1,2$ and $y \geq x$, implying that the marginal effects of the probabilities α_A and γ_A are larger than the marginal effects of the probabilities α_{NA} and γ_{NA}. Hence, the reduction in the enforcement cost for adopters is larger than the increase in the enforcement cost for non-adopters.

- f^{1111}. We have that $\frac{\partial \widehat{m}_A(y)}{\partial \gamma_A} = \frac{\partial \widehat{m}_{NA}(y)}{\partial \gamma_{NA}} = 0 \vee y = 1,2$, $\frac{\partial \widehat{m}_A(2)}{\partial \alpha_A} = \frac{\partial \widehat{m}_{NA}(2)}{\partial \alpha_{NA}} = 0$, while $\frac{\partial \widehat{m}_A(1)}{\partial \alpha_A} > \frac{\partial \widehat{m}_{NA}(1)}{\partial \alpha_{NA}} > 0$. Hence, targeted state-dependent enforcement has no effect in the cost of enforcing compliance among adopters and non-adopters if they are initially allocated to G_2. Instead, it reduces the enforcement cost for adopters and increases it for non-adopters if firms are initially allocated to G_1. Note that the marginal effect of α_A is larger than the marginal effect of α_{NA}. Hence, the reduction in the enforcement cost for adopters is larger than the increase in the enforcement cost for non-adopters.